# ENZYMATIC CHARACTERIZATION OF SELENOPROTEIN K AND S AND HARNESSING SELENOCYSTEINE FOR PROTEIN ENGINEERING

by

Jun Liu

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry and Biochemistry

Fall 2016

© 2016 Jun Liu All Rights Reserved ProQuest Number: 10241382

All rights reserved

INFORMATION TO ALL USERS The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10241382

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code Microform Edition © ProQuest LLC.

> ProQuest LLC. 789 East Eisenhower Parkway P.O. Box 1346 Ann Arbor, MI 48106 – 1346

# ENZYMATIC CHARACTERIZATION OF SELENOPROTEIN K AND S AND HARNESSING SELENOCYSTEINE FOR PROTEIN ENGINEERING

by

Jun Liu

Approved:

Murray V. Johnston, Ph.D. Chair of the Department of Chemistry and Biochemistry

Approved:

George H. Watson, Ph.D. Dean of the College of Arts and Sciences

Approved:

Ann L. Ardis, Ph.D. Senior Provost for Graduate and Professional Education

	I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.
Signed:	Sharon Rozovsky, Ph.D. Professor in charge of dissertation
	I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.
Signed:	Colin Thorpe, Ph.D. Member of dissertation committee
	I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.
Signed:	Tatyana Polenova, Ph.D. Member of dissertation committee
	I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.
Signed:	Abraham M. Lenhoff, Ph.D. Member of dissertation committee

#### ACKNOWLEDGMENTS

I cannot imagine finishing my Ph.D. anywhere else but in Professor Sharon Rozovsky's lab. She spent much time in training me to be independent when I came to the lab with little biochemistry background. Sharon knows how to push me to become a good researcher, and I was soon able to become self-sufficient. She unselfishly gave me a lot of opportunities to present my work in conferences, both nationally and internationally, which not only gave me the chance to communicate with other researchers in my field but also inspired new ideas. Sharon is so generous that she always made time for me whenever I needed to talk to her personally or needed scientific advice. She also helped and encouraged me to apply for awards, grants or to write my own papers. She is truly a friend to me, and also a deep lover of science, whom I appreciate very much. Even though sometimes our discussions sound more like an argument, we respect each other a great deal and, afterwards, I am always able to view my project from a different angle. I have benefitted enormously from her mentorship, which strengthened not only my understanding of science, but also shaped me into becoming a leader and allowed me to be able to manage my project efficiently.

I would like to thank my committee members, Professor Colin Thorpe, Professor Tatyana Polenova and Professor Abraham M. Lenhoff for their suggestions regarding my research project, generously taking time out of their busy schedules, and writing me supporting letters for my graduate fellowship, pre-doctoral fellowship grant, and postdoctoral applications. My first paper about selenoprotein S would not

iv

have gone smoothly if not for the help and suggestions from Professor Colin Thorpe. I have been reading publications from his research group or those he has recommended in order to learn the chemistry of thiol-disulfide exchange reaction, insulin turbidity assays, protein isomerase and oxidase assays. I very much appreciate Professor Colin Thorpe's influence on me and his mentorship, which helped me better understand redox biology.

The publications from Professor Robert J. Hondal and Professor David S. Waugh were very helpful in making my expressed selenoprotein ligation project possible. I want to recognize their special contribution towards stimulating new ideas in me.

Dr. Vidyadhar N. Daithankar has excellent molecular biology skills and is a nice person, who taught me important skills on molecular cloning. It was a great pleasure to work with him. Frank Zhang is an equally friendly guy, and he helped me considerably with the preparation of selenoprotein K (SelK) and did a wonderful job in measuring the redox potential of SelK's diselenide bond. It was a great pleasure to work with him for the past four and a half years. He also loves to play billiards with me. Maggie Chen joined the lab when I had just started the native chemical ligation project. It was a great pleasure to work with her. She helped me in many aspects in preparing various selenoprotein targets. Rujin Cheng and Oshini Ekanayake are involved in the glutathione peroxidase project with me and they are doing a wonderful job in developing it and moving it forward.

I would like to thank Jim Draper, Moiscell Robinson and John Famiglietti for their remarkable help with maintaining the FPLC, UV-vis, refrigerators and other instrumentation in the lab. I also want to thank the help from Drs. Stephen Chan and

V

Papa Nii Asare-Okai from the mass spectrometry facility; Dr. Steve Bai from the NMR facility, and Dr. Yu-Sung Wu, at the DBI protein core facility, Pat McMahon and Rick Bernard from the computer facility, all staff from the Chemistry & Biochemistry department, especially Mrs. Judy Dellose, Sue James, Susan Cheadle, Tracy Walsh, Eileen Burns, Brandon Calitree, Brenda Carboni, and David L Murray.

I also want to thank our collaborators Dr. Jodi L. Bubenik and Professor Donna M. Driscoll for stimulating collaboration on the *in vivo* function of selenoprotein S, and especially Professor Driscoll for writing a recommendation letter for my predoctoral fellowship application; Stella Chen and Professor E. James Petersson for generously giving access to stopped flow instruments; and collaborators and coauthors Dr. Eric Block, Dr. Squire J. Booker, Sonia Flores-Penalba, Dr. Graham George, Sivaji Gundala, Bradley J. Landgraf, Stephene N. Lodge, Dr. M. Jake Pushie, Dr, Abith Vattekkatte, Dr. Rama Yaghi, and Dr.Huawei Zeng.

Besides the long list of people I have acknowledged above, I also want to recognize the help I received from Yuliya Pepelyayeva, Prabhavathi Srinivasan, Diane N. Pham, Fei Li, Megan R. Carpenter, Stephanie Schaefer, Renee Rubenstein, Wayne Wilkie, Jay Subramoney, Dominic Santoleri, Jasmin Philip, Matthew Fritz, Michael Boeri, Vicki Wallace, Yuan Hsiang Cheng, Beverly Hallahan, Yuefan Li, Gulcin Unal Tosun, Yanduan Hu, Mahshid Poursaeed, Liedeke Sweitzer, Elyssa Bernfeld, Sarah Frantz, current and past members of the Rozovsky lab.

Finally, I want to thank my parents for their encouragement and constant support. I also want to thank my wife Yinyin Wang. She fully understands that science takes time, and gives me freedom to spend as much time as I would like to in lab. She makes my life outside the lab wonderful, and she is always the best part of my day.

vi

## **TABLE OF CONTENTS**

xi	LIST OF FIGURES
xxvi	LIST OF TABLES
xxvii	ABSTRACT

## Chapter

1	INT	RODU	CTION	1
	1.1	Introd	uction to Selenoproteins	1
	1.2	Selen	ocysteine in Enzymatic Catalysis	4
	1.3	The H	luman Selenoproteome	6
	1.4	Memb	brane Bound Selenoproteins	7
	1.5	Introd	uction to Selenoprotein S	. 10
	1.6	Introd	uction to Selenoprotein K	. 18
	1.7	Introd	uction to Expressed Protein Ligation	. 21
	1.8	Introd	uction to Sec-mediated Expressed Protein Ligation	. 23
R	EFER	ENCES		. 27
2	EXF	PRESSI	ON AND PURIFICATION OF THE MEMBRANE ENZYMES	
	SEL	ENOPI	ROTEIN K AND SELENOPROTEIN S	. 55
	2.1	Introd	uction	. 55
	2.2	Resul	ts	. 59
		2.2.1	Overexpression of SelK	. 59
		2.2.2	The Relation between the Length of Linker between SelK and	
			cMBP and in vivo Proteolytic Cleavage	. 61
		2.2.3	Detergent Screening	. 62
		2.2.4	Purification of SelK	. 64
		2.2.5	DDM-SelK Micelle Size	. 66
		2.2.6	SelK Secondary Structure	. 69
		2.2.7	Overexpression and Purification of SelS	. 70
	2.3	Discu	ssion	. 72
	2.4	Mater	ials and Methods	. 73

		2.4.1	Bacterial Strains, Plasmids, and Chemical Reagents	73
		2.4.2	SelK Expression Vectors	73
		2.4.3	SelS Expression Vectors	74
		2.4.4	Protein Expression Screening	74
		2.4.5	Expression and Purification of SelK From a cMBP-SelK	75
		216	Fusion	13 76
		2.4.0	Circular Dichroism Spectroscopy	יייייייייייייייייייי דד
		2.4.7	Circular Dichroisin Speciroscopy	/ /
RI	EFER	ENCES		78
3	THE	E INTRI	NSICALLY DISORDERED MEMBRANE PROTEIN	
	SEL	ENOPR	OTEIN S IS A REDUCTASE	85
	3.1	Introdu	uction	85
	3.2	Result	S	87
		3.2.1	Incorporation of Sec into SelS	87
		3.2.2	SelS' Reductase Activity Measurement	93
		3.2.3	SelS' Peroxidase Activity Measurment	96
		3.2.4	SelS' Peroxidase Substrate Specificity	98
		3.2.5	SelS' Peroxidase Kinetics Measurement	99
		3.2.6	Identification of Peroxidatic and Resolving Residues in	101
		2 2 7		101
		3.2.7	SelS is Resistant to ROS Inactivation	103
	3.3	Discus	ssion	106
	3.4	Materi	als and Methods	108
		3.4.1	Bacterial Strains, Plasmids, and Chemical Reagents	108
		3.4.2	Molecular Cloning	109
		3.4.3	Expression and Purification of cSelS and its Mutants	109
		3.4.4	Oxidase and Isomerase Assays	112
		3.4.5	Inductively Coupled Plasma (ICP) Spectrometry	113
		3.4.6	Electrospray Ionization Mass Spectroscopy	114
		3.4.7	Circular Dichroism Spectroscopy	114
		3.4.8	Insulin Reduction Assay	114
		3.4.9	Peroxidase Assays and Kinetics Measurement	115
		3.4.10	Traping of Selenenic Acid by Dimedone	116
		3.4.11	Inactivation of cSelS and cSelS C174S by H <sub>2</sub> O <sub>2</sub>	116
RI	EFER	ENCES		118

4	SELENOPROTEIN K FORMS AN INTERMOLECULAR DISELENIDE			
	BON	ID WIT	H UNUSUALLY HIGH REDOX POTENTIAL	. 129
	41	Introdu	uction	129
	4.2	Result	S	. 132
		4.2.1	Sec Incorporation into SelK	. 132
		4.2.2	SelK Forms an Intermolecular Diselenide Bond	. 133
		4.2.3	SelK's Redox Potential	. 134
		4.2.4	SelK can Reduce Lipid Hydroperoxides	. 138
		4.2.5	Substrate Specificity	. 139
		4.2.6	Steady State Kinetics	. 140
	4.3	Discus	ssion	. 144
	4.4	Materi	als and Methods	. 147
		4.4.1	Bacterial Strains, Plasmids, and Chemical Reagents	. 147
		4.4.2	Expression and Purification of SelK and its Mutants	. 147
		4.4.3	Electrospray Ionization Mass Spectrometry	. 150
		4.4.4	Determination of Redox Potentials	. 150
		4.4.5	Peroxidase Activity Assays	. 152
		4.4.6	Steady State Kinetics	. 152
RE	FERI	ENCES		. 154
5	EXP	RESSE	D SELENOPROTEIN LIGATION HARNESS	
	SEL	ENOCY	STEINE REACTIVITY TO EXPAND SCOPE OF NATIVE	
	CHE	MICAI	L LIGATION	. 164
	5.1	Introdu	uction	. 164
		511	Preparation of Selenoproteins is Challenging	164
		512	Preparation of Selenoproteins whose Sec is Close to the C-	. 104
		5.1.2	terminus	165
		5.1.3	Preparation of Selenoproteins Using a Cysteine (Cys)	. 100
			Auxotrophic Strain	. 165
		5.1.4	Incorporation of Sec with an Engineered tRNA <sup>Sec</sup>	. 166
		5.1.5	Incorporation of Sec through an Orthogonal tRNA/aaRS Pair	. 167
		5.1.6	Sec-mediated Native Chemical Ligation	. 167
		5.1.7	Expressed Selenoprotein Ligation	. 169
	5.2	Result	S	. 170

		5.2.1	Incorporation of Sec into MBP-SelM <sup>CT</sup> and TEV Protease	
			Cleavage.	170
		5.2.2	MBP-SelM <sup>NT</sup> Thioester Preparation	173
		5.2.3	Expressed Selenoprotein Ligation of SelM	176
		5.2.4	Expressed Selenoprotein Ligation of Selenoprotein W	177
		5.2.5	Expressed Selenoprotein Ligation Followed by Deselenization	n
			of Trx	184
		5.2.6	Site Selective Modification through Sec Conjugation	184
		5.2.7	Selective Generation of Dehydroalanine for Site Specific	
			Labeling.	196
	5.3	Discus	sion	199
	5.4	Materi	als and Methods	201
		5.4.1	Molecular Cloning	201
		5.4.2	Protein Expression and Purification	212
		5.4.3	Expressed Selenoprotein Ligation	217
		5.4.4	Deselenization of Trx A94U to Trx	219
		5.4.5	Enzymatic Activity of Trx	219
		5.4.6	Conjugation of Thioglucose and Thiophosphate to Trx A94U	J.219
		5.4.7	Conjugation of Ubiquitin G76C or Ubiquitin G76U to Trx	220
		510	Selective Alleylation of Ubiquitin C76U	220
		5.4.8 5.4.0	Selective Aikylation of Oblydroalaning	220
		5.4.9	Mass Sportrometry	221
		5.4.10	Mass Spectrometry	221
RE	FERE	NCES		223
6	OUTI	LOOK	AND FUTURE PERSPECTIVES	234
	6.1	The In	teraction between SelS and p97	234
	6.2	Explor	ation of Selenoprotein K Auto-proteolytic Cleavage	236
	6.3	Future	Expansions of Expressed Selenoprotein Ligation for Protein	
		Engine	ering	238
RE	FERE	NCES		243
Append	dix			

### **LIST OF FIGURES**

- Figure 1.5. Schematic representation of the conformational changes accompanying SelS' reductase activity. In step 1 the selenylsulfide bond of SelS is reduced by a cellular redox system such as the thioredoxin/thioredoxin reductase system. In step 2 the selenolate in the reduced segment attacks the substrate's disulfide bond, forming a mixed selenylsulfide. In step 3 SelS' selenylsulfide is reformed and the substrate and SelS are released.
- Figure 1.6. SelS interacts with multiple protein complexes. (A) SelS' protein partners are best documented in the ERAD pathway. SelS associates with the

central channel composed of several transmembrane proteins (here abbreviated as Derlin) that is responsible for retro-translocation of misfolded or misassembled proteins from the ER to the cytoplasm where they are degraded by the proteasome. (B) SelS also interacts extensively with various membrane-embedded protein complexes. Hence, it was proposed that it is responsible for maintaining these complexes in the ER membrane<sup>71</sup>. Interactions between SelS and its multiple protein partners occur through its transmembrane helix as well as a coiled-coil region. While the cytoplasmic disordered C-terminal was demonstrated to have a reductase activity *in vitro*, its substrates and physiological roles are still unknown.

- Figure 1.8. Methodology of Cys-mediated expressed protein ligation. Expressed protein ligation is a class of NCL reactions that involves at least one protein fragment made by recombinant expression in a host organism. .22
- Figure 2.1. Schematic representations of SelK and SelS. (A) Human SelK has a predicted single-pass transmembrane helix; shown here is a prediction by the TMHMM program<sup>15</sup>. It also has two potential SH3 binding elements and a putative phosphorylation site. The reactive Sec residue resides at the C-terminal, near a conserved Arg, and is exposed to the cytoplasm.
  (B) Human SelS is also predicted to have a single-pass transmembrane helix. Similar to SelK the C-terminal domain faces the cytoplasm. In addition to the transmembrane helix it has a p97/Valosin-containing Protein (VCP)-interacting Motif (VIM), a coiled coil dimerization interface and a disordered C-terminal region with an internal selenylsulfide bond.
- Figure 2.2. Schematic representations of the expression constructs tested in this study. His-tag, hexahistidine sequence; SUMO, small ubiquitin-like modifier protein; cMBP, cytoplasmic maltose-binding protein; pMBP, periplasmic

maltose-binding protein; short linker, NSSS; long linker, NSSSNNNNNNNNLG; TEV, tobacco etch virus protease cleavage site (ENLYFQ/G); Strep-tag, streptavidin binding tag (WSHPQFEK); SelK U92C, Homo sapiens SelK with a U92C mutation at the active site. 

- Figure 2.3. Proteolytic cleavage of cMBP-SelK fusion in *E. coli* depends on the length of the linker between MBP and SelK and the expression temperature. (A) Protein expression at 18 °C of cMBP-SelK fusion with short (NSSS) and long (NSSSNNNNNNNNLG) linker length. Lanes 1 to 4: SelK with a short linker at 0, 4, 8, and 22 h post-induction. Lanes 5 to 8: SelK with a long linker at 0, 4, 8, and 22 h post-induction. M: Protein molecular weights standard: 20, 25, 30, 40, 50, 60, 80, 100, 150, 250 kDa where the 25 and 80 kDa bands triple the intensity. The arrow points to the cMBP-SelK fusion at 54 kDa and cMBP at 42 kDa. (B) cMBP-SelK with short and long linker induced at 37 °C. Lanes 1 to 6: cMBP-SelK short linker 0, 2, 4, 6, 8, and 22 h post-induction. Lanes 7 to 12: long linker 0, 2, 4, 6, 8, and 22 h post-induction. M: protein molecular weights standard as listed above. The arrow points to the cMBP-SelK fusion at 53 kDa and
- Figure 2.4. Cleavage of SelK from its fusion partner cMBP by TEV protease in the presence of different detergents. Lane 1: TEV protease prior to incubation with cMBP-SelK. Lane 2: cMBP-SelK prior to incubation with TEV protease. Lane 3: Cleavage without detergent. Lane 4: Cleavage in the presence of 40 µM MNG-3. Lane 5: Cleavage in the presence of 0.48 mM DDM. Lane 6: Cleavage in the presence of 32 mM CHAPS. Lane 7: Cleavage in the presence of 0.144 mM LMPC. Lane 8: Cleavage in the presence of 0.06% Triton X-100. M: Protein molecular weights standard (the molecular mass in kDa is noted on the right). The
- Figure 2.5. Expression and purification of SelK-E (A) SDS-PAGE analysis. Lane 1: Crude cell extract prior to induction. Lane 2: Crude cell extract after overnight expression at 18 °C. Lane 3: Amylose column elute. Lane 4: Following exchange of detergent from 0.1% Triton X-100 to 0.067% DDM by IMAC. Lane 5: TEV protease cleavage mixture. Lane 6: SelK after the His6-tagged TEV and MBP were removed by IMAC. The arrow points to SelK. M: Protein molecular weights standard (the molecular mass in kDa is noted on the right). (B) The molecular weight of SelK detected by mass spectroscopy is 10654 Da (predicted molecular weight

- Figure 2.6. Characterization of the SelK-DDM complex by size exclusion chromatography. (A) SelK DDM complex elution profile on Superdex 200 in the presence (solid line) and absence of the reducing agent DTT (dotted line). (B) The molecular weight of the SelK DDM complex was determined by measuring its elution volume relative to standard proteins. The black triangle overlaid on the calibration curve corresponds to the obtained molecular weight of 91 kDa for the SelK DDM complex.......68
- Figure 2.7. SelK posses a α-helical secondary structure in various detergent micelles. The helical structure is reduced in the presence of SDS. CD spectra of SelK in the presence of 0.24 mM DDM (circles), 40 mM β-OG (squares), 0.072 mM LMPC (triangles) and 25 mM SDS (solid line)......70
- Figure 2.8. Expression and purification of SelS. (A) Schematic representations of SelS expression construct. His-tag, hexahistidines sequence; cMBP, cytoplasmic maltose-binding protein; short linker, NSSS; TEV, tobacco etch virus protease cleavage site (ENLYFQ/S); SelS U188C, *Homo sapiens* SelS with a U188C mutation at the active site. (B) SDS-PAGE analysis. Lane 1: Crude cell extract prior to induction. Lane 2: Crude cell extract after overnight expression at 18 °C. Lane 3: Amylose column elute with buffer containing DDM detergent. Lane 4: TEV protease cleavage mixture. Lane 5: SelS after the His6-tagged TEV and MBP were removed by IMAC. The arrow points to SelS dimer and monomer (from top to bottom). M: Protein molecular weights standard (the molecular mass in kDa is noted on the right). (C) CD spectrum of SelS in the presence of 0.24 mM DDM.
- Figure 3.2. Characterization of cSelS U188C by size-exclusion chromatography. (A) cSelS U188C eluted from Sephacryl S-100 column in the presence (solid line) and absence of the reducing agent DTT (dotted line). (B) The apparent molecular weight of the cSelS U188C was determined by measuring its elution volume relative to standard proteins. The black

Figure 3.3. Enrichment of Sec-containing proteins from mixtures of cSelS and cSelS 188Δ. (A) Separation of cSelS and cSelS 188Δ by size exclusion chromatography. cSelS (solid line) and cSelS U188C (dotted line) elution profile from a Sephacryl 100 under non-reducing condition. Peak 2, eluted last, contains the selenium-rich form of the enzyme. (B) SDS-PAGE analysis of cSelS / cSelS 188Δ mixture and cSelS U188C oligomerization pattern. Lanes 1, 2, 3: cSelS U188C, cSelS first peak and cSelS second peak under non-reducing conditions. M: Protein molecular weight standards (the molecular mass in kDa is noted on the right). Lanes 4, 5, 6: cSelS U188C, cSelS first peak and cSelS second peak under reducing conditions.

- Figure 3.8. Substrate specificity of cSelS peroxidase's activity. (A) Reaction scheme. (B) cSelS substrate specificity. Reaction with cSelS and hTrx (black bars), reaction excluding hTrx (striped bars), and reaction excluding

cSelS (dotted bars). Only H <sub>2</sub> O <sub>2</sub> and COOH have significant rates	
compared to the control ( $\times \times$ , P<0.001, n=3). The affinity for H <sub>2</sub> O <sub>2</sub> is th	e
highest among the four substrates (****, P<0.0001, n=3)	99

- Figure 3.11. Proposed reaction mechanism of SelS peroxidase activity......103
- Figure 3.12. cSelS resistance to inactivation by H<sub>2</sub>O<sub>2</sub>. The ability of cSelS to reduce insulin's intermolecular disulfide bond following incubations with H<sub>2</sub>O<sub>2</sub> is monitored by recording the increasing turbidity caused by insulin's chain B aggregation. (A) cSelS reductase activity assay after incubations with H<sub>2</sub>O<sub>2</sub>. (B) Percentage of remaining activity of wild type cSelS after treatment with H<sub>2</sub>O<sub>2</sub>. (C) cSelS C174S reductase activity assay after incubations with H<sub>2</sub>O<sub>2</sub>. (D) Percentage of remaining activity of cSelS C174S after treatment with H<sub>2</sub>O<sub>2</sub>. (D) Percentage of remaining activity of cSelS C174S after treatment with H<sub>2</sub>O<sub>2</sub>. The error bars represent the range of measurements (that is, highest and lowest values) among two repetitions, using two independent protein preparations.
- Figure 4.1. Schematic representation of SelK as well as the mutants used in this study.
  (A) SelK's domain organization. In blue, the N terminal domain is located in the ER lumen; In purple, the hydrophobic segment formed by residues 20-42 is predicted to contain a single transmembrane helix (TM); In green, the intrinsically disordered domain facing the cytoplasm. This domain contains an SH3 binding site as well as the catalytic selenocysteine (Sec). The location of the Sec is shown in red. SelK has no additional cysteines or selenocysteines besides the one outlined. (B) Proteins used in this study. SelK U92 was prepared by misloading of cysteyl-tRNA with Sec by a cysteine auxotrophic strain and a media with Sec. About 10% of this sample is SelK U92C generated from remaining

pools of Cys in the cell. SelK U92C and SelK U92S were prepared by conventional heterologous expression and are purely the mutant form.131

- Figure 4.3. The redox potential of SelK's intermolecular diselenide bond is -257 mV. Gel shift assays based on alkylation with fluorescein-5-maleimide. 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 (given in kDa); i TEV protease (bands with higher molecular weights are complexes of TEV protease with SelK); ii oxidized SelK present as a dimer; iii alkylated SelK present as a monomer (bands with lower molecular weights are truncated SelK generated by an *E. coli* protease). (A) Percent of reduced and oxidized wild-type SelK. Visualized by coomassie brilliant blue stain using white light imaging. (B) Alkylation pattern of wild-type SelK. Visualized by fluorescence imaging. Only alkylated protein is detected. (C) Same as A for SelK U92C. (D) Same as B for SelK U92C. (E) Wild-type SelK redox potential titration curve (reduction potential of -257 mV) and (F) SelK U92C redox potential titration curve (reduction potential of  $-207 \pm$ 1 mV). The fraction of reduced protein is plotted against the buffer redox
- Figure 4.4. SelK reductase assay. Insulin reductase activity was determined in 100 mM potassium phosphate (pH 7), 2 mM EDTA, 0.067% DDM, 0.13 mM bovine insulin, and 0.33 mM dithiothreitol (DTT). The reaction was initiated by the addition of SelK U92 (blue) or hTrx1 (red). The reaction without an addition of an enzyme served as negative control. The reaction's progress was monitored by recording the increase in turbidity at 650 nm versus time. The temperature was kept constant at 25 °C....136
- Figure 4.5. SelK peroxidase activity assays. (A) NADPH consumption is monitored in the presence of 8 or 64 nM hTrxR, with and without 5  $\mu$ M hTrx, 200  $\mu$ M H<sub>2</sub>O<sub>2</sub>, and 5  $\mu$ M of SelK U92. The reaction with 5  $\mu$ M SelK U92, 5  $\mu$ M

- Figure 5.2. MBP-SelM<sup>CT</sup> characterization by mass spectrometry. Deconvoluted electrospray ionization mass spectrum of reduced MBP-SelM<sup>CT</sup>. The calculated molecular mass of MBP-SelM<sup>CT</sup> is 53674 Da. The 53933 Da

form is due to a 6-phosphogluconolactone modification, most likely on  $His_6$ -tag at the N-terminal of MBP<sup>60</sup>. Inset shows the M/Z spectrum...171

Figure 5.4. Cleavage of a Sec-containing TEV recognition sequence in MBP-SelM<sup>CT</sup> by TEV protease in the absence of reductant. MBP-SelM<sup>CT</sup> was not cleaved by TEV protease in the absence of reductant. The failure of TEV protease to cleave is most likely due to the formation of intermolecular diselenide bonds in MBP-SelM<sup>CT</sup>, which masked the cleavage site. MBP-SelM<sup>CT</sup> was incubated with TEV protease at a molar ratio of 5:1 in 50 mM sodium phosphate (pH 7.5), 200 mM NaCl, and 2 mM EDTA. Proteins are identified on a 16% Tris-Glycine SDS-PAGE with reducing agents (right of the protein markers) or without (left of the protein markers) reducing agents.

Figure 5.5. MBP-SelM<sup>NT</sup> thioester characterization by mass spectrometry. (A) Scheme demonstrating MBP-SelM<sup>NT</sup> thioester production. MBP-SelM<sup>NT</sup>-VMA fusion was first purified by amylose column. It was then bound to a chitin column due to the presence of a chitin-binding domain fused at the C-terminal of the intein. During cleavage of the intein, the thioester form of MBP-SelM<sup>NT</sup> (MBP-SelM<sup>NT</sup>-MES) was formed with the addition of 75 mM MESNA. (B) The thioester form of MBP-SelM<sup>NT</sup> purified by a chitin column was detected by mass spectrometry. The theoretical mass of the MBP-SelM<sup>NT</sup>-MES is 45322 Da. The 42809 Da peak arises from a truncation. Inset shows the M/Z spectrum.

Figure 5.6. SelM U48C purification and characterization by mass spectrometry. (A) The purification of SelM U48C prepared by heterologous expression analyzed by 16% Tris-Glycine SDS-PAGE. Lane 1: cell lysate; lane 2: following purification by amylose affinity chromatography; lane 3: TEV

Figure 5.7. (A) Principle of expressed selenoprotein ligation. The seleno-fragment of interest is expressed in E. coli as a fusion with MBP in defined growth medium enriched with selenocystine. The complementary thioester protein fragment is prepared by established intein technology. Immediately prior to ligation with the thioester-fragment, the selenofragment is cleaved from MBP by TEV protease. The selenolate attacks the thioester and the native protein is generated spontaneously through an amide bond formation. "POI" stands for protein of interest. (B) Downstream applications of Sec-mediated ligation: Sec is readily deselenized into an alanine (anaerobically, blue) or serine (aerobically, pink) by TCEP generating a non-selenoprotein; oxidative elimination generates dehydroalanine, an electrophilic tag for chemical labelling; reactions with thiols and selenolates introduce chemical probes and posttranslational modifications via the selenylsulfide, diselenide or selenoether bonds. Panels C-H) ESL preparation of SelM. (C) The minimal Trx fold of Mus musculus SelM (PDB entry 2A2P). Red: thioester fragment; Green: seleno-fragment; Yellow: sulfur atom; Orange: selenium atom. (D) The design of thioester (SelM<sup>NT</sup>) and seleno (SelM<sup>CT</sup>) fragments. SelM residues 1 to 24 constitute an ER targeting sequence that is cleaved in vivo and thus were omitted here. Sec48 was mutated to Cys to allow Sec incorporation using the cysteinyl-tRNA misloading protocol. (E) SelM ligation monitored by SDS-PAGE. Following the generation of the SelM<sup>CT</sup> fragment, the thioester-fragment MBP-SelM<sup>NT</sup> is added and the reaction is initiated as shown in panel A. Since the TEV protease was not removed from the SelM<sup>CT</sup> mixture, it releases the SelM<sup>NT</sup> fragment as the ligation progresses. The ligation was monitored using the appearance of a 14 kDa band. Lane 1: MBP-SelM<sup>CT</sup>; lane 2: same sample as in lane 1 following cleavage by TEV protease; lane 3: MBP-SelM<sup>NT</sup>-VMA; lane 4: MBP-SelM<sup>NT</sup> thioester following cleavage of MBP-SelM<sup>NT</sup>-VMA with thiols and subsequent purification; lanes 5-7 the ligation monitored on days 0, 1, and 2, respectively. The ligation mixture contained MBP-SelM<sup>NT</sup> thioester and MBP-SelM<sup>CT</sup> TEV cleavage mixture. Lane M: protein molecular mass standards (the molecular mass in kDa is noted on the right). SelM<sup>NT</sup> is too small to be detected. (F) Deconvoluted ESI-MS of intact SelM (the calculated molecular mass is 14168 Da). (G) ESI-MS spectrum of the

- Figure 5.8. MBP-SelW<sup>NT</sup> thioester characterization by mass spectrometry. Deconvoluted mass spectrum of MBP-SelW<sup>NT</sup> thioester after chitin affinity chromatography. The calculated molecular mass of the thioester form of MBP-SelW<sup>NT</sup> (MBP-SelW<sup>NT</sup>-MES) is 43934 Da. Inset shows the M/Z spectrum. 178
- Figure 5.9. MBP-SelW<sup>CT</sup> characterization by mass spectrometry. Deconvoluted electrospray ionization mass spectrum of reduced MBP-SelW<sup>CT</sup>. The calculated molecular mass of MBP-SelW<sup>CT</sup> is 50510 Da. The calculated molecular mass for the protein with Cys is 50463 Da (denoted with \*). The 50767 Da form is due to a 6-phosphogluconolactone modification, most likely on the His6-tag at the N-terminal of MBP<sup>60</sup>. Inset shows the M/Z spectrum. 179

- Figure 5.14. Expressed selenoprotein ligation can be used to generate nonselenoproteins and to introduce a unique site for labeling. Panels A-D) chemical ligation of E. coli Trx with Sec at position 94 and subsequent deselenization to form wild-type Trx. (A) The structure of E. coli Trx A94U based on PDB entry 2TRX. Red: thioester fragment; Green: seleno-fragment; Yellow: sulfur atoms; Orange: selenium atom. (B) The design of thioester (Trx<sup>NT</sup>) and seleno (Trx<sup>CT</sup>) fragments. (C) Deconvoluted ESI-MS of intact Trx A94U (the calculated molecular mass is 11754 Da). (D) Deconvoluted ESI-MS of wild-type Trx generated by deselenization of Sec to Ala (the calculated molecular mass 11675 Da; the \* denotes Trx A94S). (E) The enzymatic activity of Trx generated by deselenization and that of Trx prepared by heterologous expression are similar as measured by insulin turbidity assays. (F) Labelling of Trx A94U with a thiophosphate (the calculated molecular mass is 11864 Da; the peak marked with \*\* is a contaminant from the commercial thiophosphate). (G) Conjugation of Trx A94U to ubiquitin G76C via the formation of a selenylsulfide bond (the calculated molecular mass is 20521 Da). (H) Conjugation of Trx A94U to ubiquitin G76U via the formation of a diselenide bond (the calculated molecular mass is 20568 Da). (I-J) Due to Sec's low  $pK_a$  it can be selectively alkylated. (I) A mixture of ubiquitin G76C and ubiquitin G76U prior to alkylation (the calculated molecular mass for ubiquitin G76C is 8769 Da and that of ubiquitin G76U is 8816 Da). (J) The same sample from panel I alkylated with MM(PEG)<sub>24</sub> (the calculated molecular mass for MM(PEG)<sub>24</sub>-alkylated ubiquitin G76U is 10055 Da). Only ubiquitin

- Figure 5.20. Characterization of Trx prepared by heterologous expression in *E. coli*. The calculated molecular mass of Trx is 11675 Da. Inset shows the M/Z spectrum. 192

- Figure 5.24. Chemoselective dehydroalanine formation from Sec. (A) Sec deselenization into dehydroalanine can be promoted through two successive alkylation steps. The reaction can be carried out selectively at low pH, even when cysteines are present, due to Sec's low pK<sub>a</sub> (5.2) and high nucleophilicity.
- Figure 5.25. Conjugations of thiol nucleophiles to SelM Sec48dehydroalanine. Deconvoluted mass spectra of (A) Conjugation of SelM Sec48dehydroalanine to 1-thio-β-D-glucose. Calculated molecular weights are 14087 Da for SelM Sec48dehydroalanine and 14283 Da after 1-thio-β-D-glucose incorporation. (B) SelM incubated with DBBA at pH 8.0. A SelM-DBAA adduct formed with both Cys45 and Sec48. Calculated molecular weight of SelM with a DBAA adduct is 14308 Da. (C) Sample from B was reacted with 20 mM idoacetamide. No alkylation was observed confirming that Cys45 and Sec48 are not exposed for modifications. (D) SelM prior to incubation with DBAA is fully alkylated when the procedure at panel C is applied. Calculated molecular weight for SelM with two idoacetamide is 14282 Da.
- Figure 6.1. SelK auto-proteolytic cleavage may be inhibited by its binding partners.
- Figure 6.2. Sequential ligation for preparation of selenoprotein with multiple cysteines. The challenge here is how to prepare a thioester for the second ligation step. A) A proposed sequential ligation from C-terminal to Nterminal. B) A proposed sequential ligation from N-terminal to C-

terminal. T refers to the TEV protease cleavage site and X refers to factor
Xa protease's cleavage site. In both cases, the Sec fragment contains a N-
terminal split intein providing a thioester precursor. The N-terminal split
intein fragment will only generate a thioester when incubated with the
other complementary C-terminal split intein fragment
Figure 6.3. Sec-mediated protein trans-splicing. Both the intein's first residue and the
first residue of the C-extein (POI <sup>C1</sup> ) contain conserved residues such as
Cys, Ser, or Thr. We hypothesized that their substitution into Sec may
enhance protein trans-splicing because selenolate is a stronger
nucleophile than thiolate, and also because the selenoester is more
reactive than thioester

## LIST OF TABLES

Table 1.1. Membrane-bound human selenoproteins	8
Table 1.2. Selenoproteins made by EPL where the N-terminal Sec peptide was only synthesized by SPPS	24
Table 4.1. Kinetic parameters of SelK activity with PLPCOOH as substrate	43
Table 5.1. Primers used for cloning	01
Table 5.2. Defined growth medium.   2	.16

#### ABSTRACT

Selenoproteins contain the 21<sup>st</sup> amino acid selenocsyteine (Sec). Twenty-five selenoproteins have been identified in the human selenoproteome, and they have been implicated in various diseases, such as cancer, cardiovascular diseases, and aging. However their biochemical function, in contrast to their biological study, is less characterized due to the challenge to make sufficient amounts of them.

Selenoprotein K (SelK) and selenoprotein S (SelS) are two membrane selenoproteins that belong to a SelK/SelS family that contains intrinsically disordered regions. Both of them were shown to participate in antioxidant defense and to be involved in the endoplasmic reticulum associated protein degradation pathway. Nevertheless, their precise function is unknown. The work described in this thesis established methods to incorporate Sec into SelK and SelS, and to characterize their biochemical properties. SelS in isolation possesses an intra selenylsulfide bond that has disulfide reductase activity relying on the presence of Sec. SelS is also found to possess weak peroxidase activity toward  $H_2O_2$ . For the first time, we have shown evidence of trapping a selenenic acid intermediate, and the mutagenesis study also established that the contribution of selenylsulfide rendered SelS resistant to inactivation by reactive oxygen species (ROS). The other member of the SelK/SelS family, SelK, was identified to contain a diselenide bond, with an apparent redox potential of -254 mV. SelK was also characterized to have weak lipid peroxidase activity. In addition, SelK was found to have auto-proteolytic cleavage, which may play a regulatory role in vivo.

In a separate research effort, we developed a new method to prepare selenoproteins, which we termed expressed selenoprotein ligation (ESL). The method incorporates Sec at any position in the protein, with no constraints on the fragment's size and properties using expressed protein ligation. Sec's high reactivity was exploited to facilitate challenging ligation reactions. Furthermore, the selenocysteine can be converted in a straightforward and selective fashion to an alanine or serine offering great flexibility in the selection of the ligation site. Alternatively, the selenocysteine can be used to introduce bioconjugates in a site-specific fashion. Thus, this method facilitates complicated ligations, increases overall yield, and introduces site-specific chemistry into the protein scaffold.

### Chapter 1

#### **INTRODUCTION**

#### **1.1 Introduction to Selenoproteins**

Selenoproteins are distinct in that they rely on the use of selenium to expand their chemical versatility<sup>1</sup>. Selenium, in contrast to metals or organic cofactors, is covalently incorporated into these proteins in the form of the rare amino acid selenocysteine (Sec, U). Sec resembles cysteine (Cys) structurally, but with selenium instead of sulfur. Unlike the conventional 20 amino acids, Sec's codon UGA is also a termination codon. In addition, Sec is biosynthesized exclusively on its own dedicated tRNA<sup>2</sup>, and its incorporation into proteins requires ancillary proteins, as well as a selenocysteine inserting sequence (SECIS) in its mRNA. In prokaryotes, SECIS is in the open reading frame, and immediately downstream of the Sec UGA codon<sup>3</sup>, while for eukaryotes and archaea, the SECIS is located in the 3' untranslated region<sup>2</sup>. After serylation, in prokaryotes it is directly converted to Sec, while in eukaryotes there is an additional phosphorylation step (Figures 1.1 and 1.2). More in-depth information on the selenoprotein synthesis can be found in a number of excellent reviews<sup>4-6</sup>.



Figure 1.1. Sec incorporation in the prokaryotic organism. The tRNA<sup>Sec</sup> is first serylated by seryl-tRNA synthetase and converted to seryl-tRNA<sup>Sec</sup>. Subsequently selenocysteine synthase catalyzes the formation of the selenocystely moiety that uses selenophosphate as a selenium donor. The decoding of the UGA codon to Sec instead of a stop signal also requires interaction of a quaternary complex composed of a dedicated translation factor SelB, GTP, selenocystyl-tRNA<sup>Sec</sup>, and SECIS.



Figure 1.2. Sec incorporation in the eukaryotic organism. The tRNA<sup>Sec</sup> is first serylated by seryl-tRNA synthetase and converted to seryl-tRNA<sup>Sec</sup>. Subsequently the seryl-tRNA<sup>Sec</sup> is phosphorylated through an ATP dependent catalysis by O-phosphoseryl tRNA kinase. The activated phosphoseryl-tRNA<sup>Sec</sup> is then converted to selenocystely moiety catalyzed by selenocysteine synthase, that uses selenophosphate as a selenium donor. The decoding of the UGA codon to Sec also requires interaction of a quaternary complex composed of a dedicated translation factor SelB, GTP, selenocystyl-tRNA<sup>Sec</sup>, and SECIS.

Although selenoproteins are found in all three domains of life, the use of selenium is not universal<sup>7</sup>. While the number of selenoproteins in a given genome is low, their contribution is often disproportionally important and even critical for the organism's survival under stress. Defining the physiological and biochemical properties of the selenoproteome is essential for identifying the relationships between the specialized chemistry of selenium and the health benefits it produces<sup>1, 8-11</sup>. The majority of selenoproteins are enzymes, mostly oxidoreductases, i.e., they mediate the transfer of electrons. Their specific cellular roles in mammals can be broadly divided into three classes: i) regulation of sulfur-based redox pathways; ii) cellular protection against oxidative stress by direct detoxification of reactive species; and iii) signal transduction, particularly the regulation of the oxidative stress response and calcium influx. In the past, progress in the biochemical characterization of selenoproteins was slow due to the difficulty of producing sufficient quantities<sup>12, 13</sup>. Fortunately, chemical biology has recently supplied novel methods for selenoprotein preparation that dramatically open new possibilities for biochemical characterization of the selenoproteome<sup>14-16</sup>.

#### **1.2** Selenocysteine in Enzymatic Catalysis

As only a small fraction of the selenoproteome has been characterized to date, the chemical diversity of Sec's reactivity still remains to be fully captured and described. It seems that Sec utilization may provide a functional advantage since it was not lost during evolution<sup>17</sup>. Yet, that advantage is difficult to pinpoint as many selenoproteins have Cys-containing orthologues<sup>18</sup>. The selenium and sulfur atoms are very similar; they are in the same group of the periodic table, with selenium radius 15 pm bigger than sulfur. Sec resembles Cys in many of its physicochemical properties,

except that the redox potential of Sec is 200 mV lower and the pK<sub>a</sub> is 3.1 units lower than Cys (Figure 1.3). The catalytic properties of such naturally-occurring Sec- and Cys-containing orthologues from different organisms have been compared and are often at par<sup>12</sup>. (In contrast Cys mutants of naturally Sec-occurring enzymes are 10-1000 times less efficient<sup>12</sup>.) Indeed, researchers in the field continue to investigate which features are inherently unique to the chemistry of selenium and could explain its pattern of utilization in nature<sup>19, 20</sup>. For one, Sec is a stronger nucleophile than Cys at physiological pH, because it is fully deprotonated and because selenium is more polarized than sulfur (see full discussion and additional references in<sup>19, 20</sup>). The high nucleophilicity offers fast reaction rates with electrophiles, which could not be achieved using Cys with its low pKa. Another consequence of Sec's high nucleophilicity is that the selenylsulfide bond is reformed more quickly than the disulfide bond. This feature can help protect the active site residues from being damaged by reactive oxygen species as well as fine-tune the selenoproteins' response time to cellular events. At the same time, also in part due to its polarizability, selenium is not only a highly reactive nucleophile but also a better electrophile and a better leaving group than sulfur. The combination of these three factors endows selenoproteins with high reactivity. These factors, along with the stability of different redox states, render selenoproteins particularly resistant to inactivation by reactive oxygen species<sup>21, 22</sup>.

HSe OH NH <sub>2</sub>	HS	O NH <sub>2</sub> OH
Properties	selenocysteine	cysteine
atomic radius (pm) bond length (Se-Se or S-S) (pm) redox potential (mV) Electronegativity pKa	115 232 -381 2.55 5.2	100 205 -180 2.58 8.3
oxidation state	-2, 0, +4, +6	-2, 0, +4, +6

Figure 1.3. Physicochemical properties of selenocysteine and cysteine<sup>23-27</sup>.

#### **1.3** The Human Selenoproteome

Since the identification of the 25 genes that encode selenoproteins in the human genome in 2003, significant progress has led to the description of the structure, expression, binding partners, localization and interaction networks of many of these selenoproteins (most recently reviewed in<sup>1, 28, 29</sup>). The reaction mechanisms of thioredoxin reductase<sup>30, 31</sup>, glutathione peroxidase<sup>32</sup>, methionine-R-sulfoxide reductase<sup>33</sup> and deiodinase<sup>34</sup> have been characterized in depth. Biochemical evidence further supports enzymatic activity for additional members such as selenoprotein P (SelP)<sup>35</sup>, selenoprotein H (SelH)<sup>36</sup>, 15-kDa selenoprotein (Sep15)<sup>37</sup>, and selenoprotein S (SelS)<sup>38</sup>. While the biochemical function of Sec continues to be investigated for these and other members, it is interesting to note that the redox motifs of several human selenoproteins were demonstrated to be essential for the formation of protein

complexes such as selenoprotein O (SelO)<sup>39</sup>, SelH<sup>36</sup>, and selenoprotein W (SelW)<sup>40</sup>. Finally, while Sec is predominantly involved in catalytic functions, it can in rare cases be responsible for storage and structural roles, such as in SelP<sup>41</sup>.

#### 1.4 Membrane Bound Selenoproteins

Even as knowledge about selenoproteins is increasing, the membrane-bound members of this group remain considerably less understood. This lacuna is analogous to the general knowledge gap regarding the structure and functions of membrane proteins compared to soluble proteins<sup>42, 43</sup>. This is largely due to the fact that a specialized, membrane-mimicking environment is needed in order for these proteins to be properly characterized. Eight out of the 25 human selenoproteins (isoforms not counted) are believed to be inserted into cellular membranes (Table 1.1). This percentage mirrors the overall cellular distribution of soluble and membrane-bound proteins in humans. A recent map of the human proteome, based on quantitative transcriptomics, classified 23-27% of human genes (5,500 out of the total 20,344) as encoding membrane-bound proteins<sup>44</sup>. Most of the human membrane-bound selenoproteins are bitopic, i.e., they contain a single transmembrane helix that spans the lipid bilayer only once<sup>45</sup>. In comparison, over 40% of mammalian membrane proteins, and thus by far the largest class, are bitopic proteins<sup>44, 46, 47</sup>. Bitopic proteins, which are prominently involved in signal transduction<sup>48-51</sup>, are often found to oligomerize, where the process of homo- or hetero-oligomerization turns them into a biologically active complex<sup>52-55</sup>.

In general, anchoring proteins to the lipid bilayer increases the rate at which they encounter protein partners by restricting the search space from three to two dimensions<sup>56</sup>. Productive encounters can be further facilitated by fixing the orientation
of the respective proteins. The membrane can thus be considered as a reactive surface that is constructed to recruit, organize and connect multiple proteins into protein complexes<sup>57</sup>. Another obvious advantage is access to lipophilic substrates or to membrane-bound protein partners. Accordingly, many of the membrane-bound selenoproteins are part of membrane-embedded protein complexes or are involved in signal transduction.

Selenoprotein K	Selenoprotein S	Protein (Abbreviation)
(SelK)	(SelS/VIMP)	
Q9Y6D0	Q9BQE4	Uniprot ID
NP_067060.2	NP_060915.2	Accession Code
92 (94)	188 (189)	Sec location of (total AAs)
ER membrane	ER membrane	Membrane Location
Single pass (bitopic)	Single pass (bitopic)	Putative Membrane
		Attachment
Palmitoylation	Multi protein	Biological Function
	complexes targeting	
Deficient calcium flux	Not available	Phenotype of Knockout
in immune cells <sup>58</sup>		Mice

Table 1.1. Membrane-bound human selenoproteins

Deiodinase 1 (D1/DIO1)	Selenoprotein N	Selenoprotein I	Selenoprotein T
	(SelN/SepN1)	(Sell/hEPT1)	(SelT)
P49895	57ZV5	Q9C0D9	P62341
NP_000783.2	NP_996809.1	NP_277040.1	NP_057359.2
126 (249)	428 (556)	387 (397)	36 (182)
Plasma membrane	ER membrane	Putative ER	ER membrane
Single pass	Not known	Multi pass	Amphipathic
(bitopic)		(polytopic)	(monotopic)
Thyroid hormone	Muscle development	Lipid biosynthesis	Calcium regulation
activation			
Normal thyroid gland	Affected muscle function,	Not available	Conditional knockout
function (euthyroid) <sup>63</sup>	abnormal lung		in β-cell impaired
	development <sup>60-62</sup>		glucose tolerance <sup>59</sup>

Deiodinase 3 (D3/DIO3)	Deiodinase 2 (D2/DIO2)
P55073	Q92813
NP_001353.4	NP_054644.1
170 (304)	133, 266 (273)
Plasma membrane	ER membrane
Single pass (bitopic)	Single pass
	(bitopic)
Thyroid hormone deactivation	Thyroid hormone activation
Hypothyroid, impaired reproductive capacity <sup>68</sup>	Euthyroid, abnormal motor phenotype, hearing impairment <sup>64-67</sup>

# 1.5 Introduction to Selenoprotein S

SelS (also called VIMP (VCP-interacting membrane protein), SEPS1 and Tanis) was originally shown to recruit protein partners in the ER-associated degradation (ERAD) pathway of misfolded proteins or misassembled protein machineries<sup>69, 70</sup>. SelS, however, not only is associated with the ERAD pathway but also contributes to the transport and maintenance of multicomplexes<sup>71</sup>. Interestingly, Sec does not appear to be required for all of these functions, even though it is essential for SelS' enzymatic competency. The intrinsically disordered Sec-containing C- terminal domain has reductase activity<sup>38</sup>, but neither the redox substrates nor the cellular role of SelS' enzymatic functions has been unequivocally identified.



Figure 1.4. Domain architecture of the SelS/SelK family adapted from Shchedrina et al.<sup>72</sup>. Members of the family do not necessarily share homology but are all predicted to be single-pass (bitopic) intrinsically disordered membrane proteins.

SelS is a member of a recently identified eukaryotic family of membrane proteins named the SelS/SelK family (Figure 1.4)<sup>72</sup>. This family contains over 108 proteins, and 33 members have been identified in humans alone. Members share several commonalities: they contain fewer than 300 amino acids, possess a sole transmembrane helix within 60 amino acids from the N-terminus, and hold a relatively high fraction of glycine, proline and polar residues in their C-terminus, which adds to a high isoelectric point (>8.3). This composition is a hallmark of intrinsically disordered proteins (IDPs)<sup>73, 74</sup>. Finally, either a Cys or Sec is positioned within the terminal five residues. Because the majority of selenoproteins are enzymes, members of the SelS/SelK family are expected to have enzymatic activity as well. However, IDPs commonly function in regulation and signaling pathways, and only a handful were shown to be enzymes<sup>75, 76</sup>. Therefore, this unusual feature of the family stands out.

Multiple sequence alignment demonstrated that SelS has at least three different classes: in class I, the penultimate Sec is separated by 12-13 residues from a partnering Cys; in class II the penultimate Sec has no obvious binding partner; and class III relies on a penultimate Cys separated by 4-5 residues from a partner Cys<sup>77, 78</sup>. All three classes share extensive homology in their C-terminal segments despite the fact that disordered regions typically acquire mutations faster than ordered regions. Only *Homo sapiens* and *Plasmodium falciparum* SelS have been characterized so far<sup>38, 77-79</sup>. Human SelS has two isoforms: a 187 amino acid long protein without a Sec and a 189 amino acid long protein that has a Sec<sup>80</sup>. The Sec-less variant comprises 5-16% of the population of total SelS transcripts.

Human SelS is a single-pass transmembrane protein with a short segment in the ER lumen and an extended cytoplasmic region<sup>81</sup>. The cytoplasmic segment contains a coiled-coil region (residues 52-122) followed by a disordered segment (residues 123-189 out of 189)<sup>78</sup> that includes the Sec at position 188. The coiled-coil region carries two extended  $\alpha$ -helices (PDB entry 2Q2F) and is essential for homo and hetero-oligomerization as well as associations with protein partners<sup>71</sup>. This segment also contains the p97 ATPase binding site named the valosin-containing protein (VCP) interacting motif (residues 78-88). p97 ATPase is a general contributor of energy for the disassembly of protein machineries and the factor responsible for pulling misfolded proteins out of the lumen<sup>82, 83</sup>. Pro178 and Pro183 of the C-terminal segment are also implicated in p97 ATPase binding<sup>84</sup>.

NMR spectroscopy studies of SelS demonstrated that region 123-189 has little structural order in the oxidized and reduced states<sup>78</sup>. Varying the oxidation state led to a significant change in the chemical shifts of residues 173-189, suggesting local structural rearrangement in the active site. The region did not gain order in the presence of p97 ATPase's N-domain. The disordered domain contains the active site residues Cys174 and Sec188 (Figure 1.5), which were shown to form an intramolecular selenylsulfide bond<sup>38</sup>. Pro178 and Pro183 are important for binding of protein partners, presumably by bringing the two catalytic residues in close spatial proximity (Figure 1.6). That a conformation switch is relevant for function is further supported by the finding that there is a Sec-less variant of SelS<sup>80</sup>. This form, despite being catalytically inactive, would be permanently in the "reduced" conformation. Cys174 and Sec188 must be in close spatial proximity even in the reduced state due to the speed with which the selenylsulfide bond between them is reformed once reduced. Consequently, it was impossible to trap the selenenic acid (SeOH), formed when U188 reacted with H<sub>2</sub>O<sub>2</sub>, using the alkylating reagent dimedone. In contrast, the selenenic acid was trapped by dimedone once the resolving Cys174 was mutated to a Ser<sup>79</sup>. The selenylsulfide bond between Cys174 and Sec188 was shown to be reduced by the human enzymes thioredoxin and thioredoxin reductase in vitro and its redox potential was determined to be  $-234 \text{ mV}^{38}$ .

Our group has recently shown that *in vitro* human SelS has disulfide reductase activity similar to that of the well-known reductases thioredoxin and protein disulfide isomerase (Figure 1.5); see chapter 3 for detailed experiments<sup>38</sup>. In the absence of

known redox substrates, the activity was measured by monitoring turbidity due to aggregation of the insulin chains. Once the biological substrate(s) of SelS are identified, more specific reductase assays are needed before the catalytic efficiency of SelS can be assessed. SelS was also found to have weak peroxidase activity, suggesting that this function is at most secondary<sup>79</sup>.



Figure 1.5. Schematic representation of the conformational changes accompanying SelS' reductase activity. In step 1 the selenylsulfide bond of SelS is reduced by a cellular redox system such as the thioredoxin/thioredoxin reductase system. In step 2 the selenolate in the reduced segment attacks the substrate's disulfide bond, forming a mixed selenylsulfide. In step 3 SelS' selenylsulfide is reformed and the substrate and SelS are released.

Both the reductase and peroxidase activities were contingent upon the presence of Sec. Furthermore, SelS resisted inactivation by high concentrations of  $H_2O_2$ . Therefore, SelS is sensitive to low levels of reactive oxygen species, yet cannot be inactivated by them even at high concentrations.

SelS is localized to the ER as well as the perinuclear speckles<sup>80, 85</sup>. There have been no reports yet of a knockout model in either cells or animals. SelS was initially identified as a member of the ERAD machinery<sup>70, 86</sup>. This pathway is responsible for transporting misfolded proteins from the ER to the cytoplasm for degradation by the proteasome<sup>87-90</sup>. Accordingly, SelS strongly binds p97 ATPase, derlin 1, and derlin 2<sup>72, 84</sup>. A study surveying protein networks in the ERAD pathway reported that, in addition to those partners, SelS associates with various ERAD partners, such as SelK, UBE4A, UBXD6, UBXD8 and KLHDC2<sup>91</sup>. An earlier report listed the protein partners Herp and Hrd among others<sup>92</sup>. Some of these targets may potentially interact with SelS via its partner p97 ATPase. The abundance of SelS' binding partners was reinforced by an exhaustive analysis of SelS interactions that identified over 200 protein partners, particularly of multiprotein complexes, including oligosaccharyltransferase, multisynthetase, anaphase-promoting, and nuclear pore complexes<sup>71</sup>. This study showed that SelS is associated not merely with the ERAD path but more generally with the transport and maintenance of protein machineries. Considering the complexity of SelS' network of interactions, the large number of reports regarding SelS' function in a variety of pathways comes as no surprise. For example, SelS has been implicated in pathways regulating ER shape<sup>93</sup>, lipid metabolism<sup>94, 95</sup> and the management of lipid droplets<sup>95, 96</sup>. SelS' contribution to so many protein complexes may also explain its many connections to conditions such as cardiovascular diseases<sup>97-99</sup>, metabolic disorders<sup>100</sup>, diabetes<sup>101</sup>, inflammation<sup>102, 103</sup>, autoimmune thyroid diseases<sup>104</sup>, preeclampsia<sup>105</sup>, and cancer<sup>106-108</sup>.

Overall, the emerging picture is that SelS is a diverse scaffolding protein whose transmembrane helix and coiled coil region are responsible for mediating protein-protein interactions, designating SelS to the appropriate membrane-residing multiprotein complex, while its C-terminal disordered region may possibly have additional functionalities that are not yet clearly delineated (Figure 1.6). One of these functions may involve signaling. Several studies have shown that SelS is involved in modulating the expression of kinases such as in the PI3K/Akt signaling pathways<sup>109,</sup> <sup>110</sup>, but the paths have not yet been fully charted<sup>111, 112</sup>. Nevertheless, a perplexing question persists: if SelS participates in signaling who are its signaling pathers?

16

Likewise, are there proteins that act as SelS' substrates? No such targets have yet emerged, even in a meticulous pull-down survey involving several mutants of the active site<sup>71</sup>. Despite these ongoing challenges, it appears that SelS' network of interactions is gradually emerging and will contribute toward a more a coherent picture of how all these seemingly disconnected physiological functions are related.



Figure 1.6. SelS interacts with multiple protein complexes. (A) SelS' protein partners are best documented in the ERAD pathway. SelS associates with the central channel composed of several transmembrane proteins (here abbreviated as Derlin) that is responsible for retro-translocation of misfolded or misassembled proteins from the ER to the cytoplasm where they are degraded by the proteasome. (B) SelS also interacts extensively with various membrane-embedded protein complexes. Hence, it was proposed that it is responsible for maintaining these complexes in the ER membrane<sup>71</sup>. Interactions between SelS and its multiple protein partners occur through its transmembrane helix as well as a coiled-coil region. While the cytoplasmic disordered C-terminal was demonstrated to have a reductase activity *in vitro*, its substrates and physiological roles are still unknown.

# 1.6 Introduction to Selenoprotein K

SelK is a member of the SelK/SelS family (Figure 1.4) but shares little homology with its other members<sup>72</sup>. Based on predictions of transmembrane helices in proteins, SelK has a short N-terminal segment in the ER lumen, a single-pass transmembrane helix<sup>85</sup> embedded in the ER membrane, and a cytoplasmic C-terminal disordered region that is particularly rich in Gly, Pro and basic residues and contains at least one sequence that is capable of binding Src-homology 3 (SH3) domains<sup>113, 114</sup>. The predicted transmembrane helix at residues 20-42 has an uncommon composition as it contains both an aspartic acid and a glutamic acid, which are not frequently encountered in transmembrane segments<sup>115</sup>. Yet, such marginally hydrophobic helices do efficiently integrate into membranes, particularly when flanked by positive residues<sup>116</sup>. Because the transmembrane helix seems insufficiently hydrophobic<sup>117</sup>. SelK will likely be present in a complex with itself or a protein partner. Membrane proteins with single-pass transmembrane helices, like SelK, are typically part of oligomeric complexes<sup>118</sup>. Recognition between partners is often mediated by transmembrane polar amino acids. Experimentally, we found that in the absence of protein partners, SelK is at a minimum a homodimer, regardless of its oxidation state<sup>119</sup>. In our view, SelK's oligomerization state is likely to change as function necessitates, as is often the case with bitopic proteins (see Membrane selenoproteins).

The Sec, at position 92 of 94 residues, is located in the disordered region facing the cytoplasm and is not in a recognizable redox motif. In many selenoproteins, a vicinal cysteine forms a selenenylsulfide bond with the Sec<sup>1, 120, 121</sup>. Remarkably, SelK contains an intermolecular diselenide bond as a homodimer<sup>119</sup>. While intramolecular diselenide bonds were reported for SelL<sup>122</sup> and most likely also occur in SelP<sup>123</sup>, an intermolecular diselenide has not previously been described for a

selenoprotein. The redox potential of the intermolecular diselenide bond was determined to be -257 mV, which implies that it can be effectively reduced by cellular proteins. Indeed, we have shown that it is reduced by thioredoxin reductase<sup>119</sup>. We proposed that the diselenide bond might be present under conditions where SelK is a homooligomer, as SelK putatively alternate oligomerization states.

In contrast to SelS, SelK has no obvious oxidoreductase activity in the absence of protein partners. It is capable of reducing lipid peroxides in detergent micelles even though under these conditions its catalytic efficiency is low<sup>119</sup>. The catalytic efficiency of membrane enzymes is often affected by the choice of detergents or physicochemical properties of the membrane in which they reside. Therefore, SelK's membrane environment or protein partners will likely influence its activity. Furthermore, intrinsically disordered proteins are rarely efficient enzymes, possibly because the conformational search for the appropriate catalytic conformation is timeconsuming<sup>76</sup>. Nonetheless, SelK's ability to process lipid peroxides is a significant observation since only a handful of peroxidases accept hydrophobic substrates. Interestingly, among them are the two selenoproteins glutathione peroxidase 4 and SelP<sup>35</sup>. Regardless, peroxidase activity does not appear to be SelK's central function considering its association with ERAD and palmitoylation (see below). We proposed that SelK can, among other functions, be a sensor of lipid peroxides<sup>119</sup>.



Figure 1.7. SelK associates with at least two membrane complexes: the ERAD machinery<sup>72</sup> and the palmitoyl transferase DHHC6<sup>124</sup>. The association between SelK and DHHC6, mediated via a SH3 domain, stabilizes DHHC6 and possibly activates it. The role of SelK in the ERAD machinery is still unknown. The oligomerization state of SelK in any of these complexes remains also undetermined. SelK was shown to be regulated by the protease calpain and may also be capable of autoproteolysis.

SelK knockout mice were viable but exhibited impaired immune responses<sup>58</sup>. Subsequently, SelK was shown to be a novel target of m-calpain, a cysteine protease with diverse functions<sup>125</sup>. Calpain cleavage between R<sub>81</sub> and G<sub>82</sub> generates a truncated form that lacks the Sec<sup>126</sup>. The cleavage is regulated by Toll-like receptor-induced calpastatin in macrophages. Interestingly, we found that SelK is capable of undergoing autoproteolysis when purified in detergent micelles (unpublished data). The autoproteolysis cleavage site resides between S<sub>55</sub> and S<sub>56</sub>, again generating a Seccontaining peptide and a membrane embedded Sec-less protein. It is possible that the Sec-containing peptides generated by both calpain and SelK itself play physiological roles<sup>127</sup>.

Recently, SelK's role in palmitoylation was defined by reports that it associates with the DHHC palmitoyl transferase DHHC6 (DHHC-type containing 6, where DHHC refers to the active site residues Asp, His, His and Cys) and is responsible for modulating its activity<sup>124</sup>. S-palmitoylation is the reversible posttranslational attachment of a fatty acid onto cysteine residues, which moderates proteins' cellular localization<sup>128, 129</sup>. The full range of DHHC6 substrates remains unknown<sup>130</sup>, but among those that require the presence of SelK for their palmitoylation are the CD36 receptor<sup>131</sup> and the inositol 1,4,5-triphosphate receptor (IP3R)<sup>124</sup>. Ancillary proteins were previously shown to contribute to the activation and stabilization of DHHC proteins<sup>132, 133</sup>, but SelK appears to be the only one so far that is bitopic.

Palmitoylation may not be the only path by which SelK functions since its other protein partners are members of the ERAD pathway (Figure 1.7). SelK and SelS co-immunoprecipitated with each other as well as with the cytoplasmic p97 ATPase<sup>83, <sup>84</sup>, derlin 1 and derlin 2<sup>71, 72, 91</sup>, which are responsible for recognition of misfolded proteins in the membrane environment and facilitate their extraction from the lipid bilayer<sup>134, 135</sup>. Overall, it appears that SelK – like most small membrane proteins<sup>136-138</sup> – acts to stabilize and modulate the activity of larger protein complexes. It will be particularly helpful to learn whether it can also contribute directly to the chemical reactivity of DHHC6 as suggested by Hoffmann and colleagues<sup>124</sup>.</sup>

## 1.7 Introduction to Expressed Protein Ligation

Native chemical ligation (NCL) refers to Cys-mediated chemical ligation<sup>139</sup>, where a N-terminal Cys of one fragment attacks a C-terminal thioester of the complementary fragment, followed by an S $\rightarrow$ N shift to spontaneously form the native amide bond (Figure 1.8). Both fragments are unprotected peptides synthesized through solid phase peptide synthesis (SPPS). NCL is most efficient in synthesizing challenging peptides, which exceed the length or properties permitted by SPPS<sup>140</sup>. An extension of NCL is expressed protein ligation (EPL), during which an N-terminal Cys fragment synthesized by SPPS or by recombinant methods can be chemoselectively joined with the complementary recombinant fragment thioester, typically prepared by intein technology<sup>141, 142</sup>. An intein is a protein segment able to cleave itself out of a precursor protein and join the remaining portions (the exteins) with a native amide bond<sup>143</sup>. It has been engineered for diverse applications such as protein purification, native chemical ligation, and splicing<sup>144</sup>.



Figure 1.8. Methodology of Cys-mediated expressed protein ligation. Expressed protein ligation is a class of NCL reactions that involves at least one protein fragment made by recombinant expression in a host organism.

EPL is a powerful chemical tool for protein semisynthesis that can be used to assemble proteins from their respective peptide fragment building blocks. It has a wide range of applications including semisynthesis of protein targets that are toxic to *E. coli* 

or face challenges of protein folding<sup>145</sup>, site selective labeling for proteomic applications<sup>146-149</sup> and biosynthesis of cyclic proteins to improve biophysical properties<sup>150</sup>. EPL can also be used for segmental isotopic labeling, which allows the fragments to be examined by NMR spectroscopy and helps reduce NMR spectral complexity when analyzing macromolecules or intrinsically disordered proteins<sup>148, 151-<sup>154</sup>. EPL is also useful for the incorporation of non-conventional amino acids, especially for those that are challenging for genetic incorporation with bioorthogonal tRNA/tRNA synthetase pair through nonsense suppression<sup>155-157</sup>. Similarly, EPL is useful for introducing selective posttranslational modifications, an example of which is the Muir group generation of synthetic histones and modifications of chromatin<sup>158</sup>.</sup>

The N-terminal Cys fragment can be made either through SPPS, or recombinant expression while a protease recognition sequence can be introduced after a protein tag<sup>146, 159, 160</sup>. It can also be expressed using a leader sequence such as MC-POI<sup>CT</sup> (methionine-cysteine followed by protein of interest C terminus) with the Cys at position two, where the methionine at the first position will be completely removed by endogenous methionine aminopeptidase<sup>161</sup>.

The C-terminal thioester fragment can also be synthesized by SPPS<sup>162</sup> or by the widely used intein technology through thiolysis of the engineered intein fusion<sup>163,</sup> <sup>164</sup>. In addition, peptide or protein hydrazide can be used to generate thioesters in a controlled fashion, especially useful for sequential ligation<sup>165, 166</sup>.

## 1.8 Introduction to Sec-mediated Expressed Protein Ligation

As discussed in section 1.2, Sec has similar physicochemical properties to Cys, except that Sec has a lower pK<sub>a</sub>, and the selenolate is a stronger nucleophile than thiolate (Figure 1.3). Sec-mediated NCL relies on the selenolate of an N-terminal Sec

in one fragment attacking the C-terminal thioester of the complementary fragment, followed by a Se $\rightarrow$ N acyl shift to form a native amide bond<sup>167-169</sup>. Because of its lower pK<sub>a</sub> and stronger nucleophilicity, Sec-mediated NCL was shown to have a faster ligation rate and a wider working pH range than Cys-mediated ligations<sup>167</sup>. For example, at pH 5.0, and pH 6.5, Sec-mediated NCL was 1000 and 100 fold faster than Cys, respectively<sup>167</sup>.

It was originally found that Sec peptides tend to dimerize, and need to be efficiently reduced for chemical ligation. Exogenous reducing reagent, such as thiol or *tris*(2-carboxyethyl)phosphine (TCEP), can be used, but in the process of using TCEP, deselenization takes place<sup>169</sup>. This traceless deselenization leads to expansion of the ligation sites for NCL at alanine, proline, and phenylalanine through Sec derivatives<sup>170-174</sup>. Recent studies have identified the mechanism of deselenization of Sec in peptide or protein. TCEP treatment under anaerobic or aerobic conditions leads to the formation of alanine or serine respectively<sup>175, 176</sup>. In addition, Sec-mediated NCL was used to prepare dehydroalanine-containing peptides, providing an electrophilic center for protein labeling with nucleophiles<sup>177</sup>.

Selenoprotein by EPL	N-terminal Sec peptide preparation,
	length, in total protein length
RNase A <sup>167</sup>	SPPS, 15, (124)
Azurin <sup>178</sup>	SPPS, 17, (148)
Sortase A <sup>179</sup>	SPPS. 22 (106)
TrxR <sup>180</sup>	SPPS. 3 (490)

Table 1.2. Selenoproteins made by EPL where the N-terminal Sec peptide was only synthesized by SPPS

For preparation of selenoproteins, the thioester was made by standard intein technology<sup>163, 181</sup>, while the N-terminal Sec peptide was prepared by SPPS (Table 1.2) (Figure 1.9). For instance, Sec substituted ribonuclease A (RNase A) was prepared by Sec-mediated ligation technology, in which a 109 amino acid recombinant thioester, made using intein technology, was ligated to a synthetic 15 residue Sec peptide<sup>167</sup>. The Sec displaced RNase A had comparable biological activity to the native one, suggesting that the synthetic enzyme is in the correct fold. A similar semisynthesis strategy has been used to engineer Sec-containing azurin, a blue copper protein<sup>178</sup> and sortase A, a bacterial transpeptidase<sup>179</sup>.

The significance of Sec-mediated EPL would be to extend its use in the preparation of human selenoproteins for functional characterization. Sec-mediated NCL was proposed as an efficient way to make human selenoproteins for biochemical characterization, due to its independence of Sec genetic incorporation machinery<sup>182,</sup> <sup>183</sup>. However, in the past decade, only thioredoxin reductase (TrxR), whose Sec is located in the penultimate position, has been semisynthesized by EPL<sup>180</sup>. The EPL TrxR was produced by ligating a TrxR (1-487) recombinant thioester to a tripeptide synthesized CUG. The EPL TrxR compared well to the natively purified one in terms of catalytic efficiency. Despite its many advantages, the wide usage of Sec-mediated EPL is hampered due to the challenge of synthesizing longer N-terminal Sec peptide<sup>140, 184</sup>. Chapter 5 describes a novel method to prepare N-terminal Sec peptides without size limitations and irrespective of the fragment's properties, for use in expressed selenoprotein ligation.



Figure 1.9. Sec-mediated expressed protein ligation. Sec-mediated EPL is a powerful tool to make selenoproteins. The thioester is produced through intein technology, and the N-terminal Sec peptide is prepared by SPPS.

### REFERENCES

- Hatfield, D.L., Tsuji, P.A., Carlson, B.A. & Gladyshev, V.N. Selenium and selenocysteine: roles in cancer, health, and development. *Trends Biochem Sci* 39, 112-120 (2014).
- Bulteau, A.L. & Chavatte, L. Update on selenoprotein biosynthesis.
   Antioxidants & redox signaling 23, 775-794 (2015).
- 3. Yoshizawa, S. & Bock, A. The many levels of control on bacterial selenoprotein synthesis. *Biochim Biophys Acta* **1790**, 1404-1414 (2009).
- 4. Turanov, A.A., Xu, X., Carlson, B.A., Yoo, M., Gladyshev, V.N., & Hatfield,
  D.L. Biosynthesis of Selenocysteine, the 21st Amino Acid in the Genetic
  Code, and a Novel Pathway for Cysteine Biosynthesis. *Advances in Nutrition*2, 122-128 (2011).
- 5. Driscoll, D.M. & Copeland, P.R. Mechanism and regulation of selenoprotein synthesis. *Annual Review of Nutrition* **23**, 17-40 (2003).
- Yoshizawa, S. & Böck, A. The many levels of control on bacterial selenoprotein synthesis. *Biochim. Biophys. Acta* 1790, 1404-1414 (2009).
- Lobanov, A.V., Hatfield, D.L. & Gladyshev, V.N. Eukaryotic selenoproteins and selenoproteomes. *Biochim. Biophys. Acta* 1790, 1424-1428 (2009).

- Pillai, R., Uyehara-Lock, J.H. & Bellinger, F.P. Selenium and Selenoprotein Function in Brain Disorders. *IUBMB life* 66, 229-239 (2014).
- Pitts, M.W., Byrns, C.N., Ogawa-Wong, A.N., Kremer, P. & Berry, M.J. Selenoproteins in Nervous System Development and Function. *Biological Trace Element Research* 161, 231-245 (2014).
- Roman, M., Jitaru, P. & Barbante, C. Selenium biochemistry and its role for human health. *Metallomics* 6, 25-54 (2014).
- Huang, Z., Rose, A.H. & Hoffmann, P.R. The role of selenium in inflammation and immunity: from molecular mechanisms to therapeutic opportunities. *Antioxid. Redox Signal.* 16, 705-743 (2012).
- Johansson, L., Gafvelin, G. & Arnér, E.S.J. Selenocysteine in proteins properties and biotechnological use. *Biochim. Biophys. Acta* 1726, 1-13 (2005).
- Rozovsky, S. in Biochalcogen Chemistry: The Biological Chemistry of Sulfur, Selenium, and Tellurium. (eds. C.A. Bayse & J.L. Brumaghim) 127-142 (ACS Press, 2013).
- Aldag, C., Bröcker, M.J., Hohn, M.J., Prat, L., Hammond, G., Plummer, A. & Söll, D. Rewiring translation for elongation factor Tu-dependent selenocysteine incorporation. *Angew. Chem.* 52, 1441-1445 (2013).
- Thyer, R., Robotham, S.A., Brodbelt, J.S. & Ellington, A.D. Evolving tRNA(Sec) for Efficient Canonical Incorporation of Selenocysteine. *Journal of the American Chemical Society* 137, 46-49 (2015).

- Rakauskaitė, R., Urbanaviciute, G., Ruksenaite, A., Liutkeviciute, Z., Juskenas, R., Masevicius, V. & Klimasauskas, S. Biosynthetic selenoproteins with genetically-encoded photocaged selenocysteines. *Chem Commun* 51, 8245-8248 (2015).
- Castellano, S. On the unique function of selenocysteine Insights from the evolution of selenoproteins. *Biochimica Et Biophysica Acta-General Subjects* 1790, 1463-1470 (2009).
- Romagne, F., Santesmasses, D., White, L., Sarangi, G. K., Mariotti, M., Hubler, R., Weihmann, A., Parra, G., Gladyshev, V. N., Guigo, R. & Castellano, S. SelenoDB 2.0: annotation of selenoprotein genes in animals and their genetic diversity in humans. *Nucleic Acids Research* 42, D437-D443 (2014).
- Arnér, E.S.J. Selenoproteins-what unique properties can arise with selenocysteine in place of cysteine? *Exp. Cell Res.* **316**, 1296-1303 (2010).
- Hondal, R.J., Marino, S.M. & Gladyshev, V.N. Selenocysteine in thiol/disulfide-like exchange reactions. *Antioxid. Redox Signal.* 18, 1675-1689 (2013).
- Ruggles, E.L., Snider, G.W. & Hondal, R.J. in Selenium: Its Molecular Biology and Role in Human Health. (eds. D.L. Hatfield, M.J. Berry & V.N. Gladyshev) 73-83 (Springer, New York; 2012).

- Snider, G.W., Ruggles, E., Khan, N. & Hondal, R.J. Selenocysteine confers resistance to inactivation by oxidation in thioredoxin reductase: Comparison of selenium and sulfur enzymes. *Biochemistry* 52, 5472-5481 (2013).
- Huber, R.E. & Criddle, R.S. Comparison of the chemical properties of selenocysteine and selenocystine with their sulfur analogs. *Archives of biochemistry and biophysics* 122, 164-173 (1967).
- Besse, D., Siedler, F., Diercks, T., Kessler, H. & Moroder, L. The redox potential of selenocystine in unconstrained cyclic peptides. *Angew Chem Int Edit* 36, 883-885 (1997).
- Malins, L.R., Mitchell, N.J. & Payne, R.J. Peptide ligation chemistry at selenol amino acids. *J Pept Sci* 20, 64-77 (2014).
- Muttenthaler, M. & Alewood, P.F. Selenopeptide chemistry. *Journal of Peptide Science* 14, 1223-1239 (2008).
- Arnold, A.P., Tan, K.S. & Rabenstein, D.L. Nuclear magnetic resonance studies of the solution chemistry of metal complexes. 23. Complexation of methylmercury by selenohydryl-containing amino acids and related molecules. *Inorganic Chemistry* 25, 2433-2437 (1986).
- Kasaikina, M.V., Hatfield, D.L. & Gladyshev, V.N. Understanding selenoprotein function and regulation through the use of rodent models. *Biochimica Et Biophysica Acta-Molecular Cell Research* 1823, 1633-1642 (2012).

- Labunskyy, V.M., Hatfield, D.L. & Gladyshev, V.N. Selenoproteins: molecular pathways and physiological roles. *Physiol Rev* 94, 739-777 (2014).
- Arner, E.S.J. Focus on mammalian thioredoxin reductases Important selenoproteins with versatile functions. *Biochimica Et Biophysica Acta-General Subjects* 1790, 495-526 (2009).
- Lothrop, A.P., Snider, G.W., Ruggles, E.L. & Hondal, R.J. Why Is Mammalian Thioredoxin Reductase 1 So Dependent upon the Use of Selenium? *Biochemistry* 53, 554-565 (2014).
- Toppo, S., Flohé, L., Ursini, F., Vanin, S. & Maiorino, M. Catalytic mechanisms and specificities of glutathione peroxidases: Variations of a basic scheme. *Biochim. Biophys. Acta* 1790, 1486-1500 (2009).
- 33. Kryukov, G.V., Kumar, R.A., Koc, A., Sun, Z.H. & Gladyshev, V.N. Selenoprotein R is a zinc-containing stereo-specific methionine sulfoxide reductase. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 4245-4250 (2002).
- Schweizer, U., Schlicker, C., Braun, D., Köhrle, J. & Steegborn, C. Crystal structure of mammalian selenocysteine-dependent iodothyronine deiodinase suggests a peroxiredoxin-like catalytic mechanism. *Proc. Natl. Acad. Sci. U S* A 111, 10526-10531 (2014).
- 35. Takebe, G., Yarimizu, J., Saito, Y., Hayashi, T., Nakamura, H., Yodoi, J.,Nagasawa, S. & Takahashi, K. A comparative study on the hydroperoxide and

thiol specificity of the glutathione peroxidase family and selenoprotein P. J. *Biol. Chem.* **277**, 41254-41258 (2002).

- Novoselov, S.V., Kryukov G.V., Xu, X., Carlson, B.A., Hatfield, D.L. & Gladyshev, V.N. Selenoprotein H is a nucleolar thioredoxin-like protein with a unique expression pattern. *Journal of Biological Chemistry* 282, 11960-11968 (2007).
- 37. Ferguson, A.D., Labunskyy, V. M., Fomenko, D. E., Arac, D., Chelliah, Y., Amezcua, C. A., Rizo, J., Gladyshev, V. N. & Deisenhofer, J. NMR structures of the selenoproteins Sep15 and SelM reveal redox activity of a new thioredoxin-like family. *J. Biol. Chem.* 281, 3536-3543 (2006).
- 38. Liu, J., Li, F. & Rozovsky, S. The intrinsically disordered membrane protein selenoprotein S is a reductase in vitro. *Biochemistry* **52**, 3051-3061 (2013).
- Han, S.J., Lee, B.C., Yim, S.H., Gladyshev, V.N. & Lee, S.R. Characterization of mammalian selenoprotein O: a redox-active mitochondrial protein. *PLoS One* 9, e95518 (2014).
- 40. Dikiy, A., Novoselov S.V., Fomenko, D.E., Sengupta, A., Carlson, B.A., Cerny, R.L., Ginalski, K., Grishin, N.V., Hatfield, D.L. & Gladyshev, V.N. SelT, SelW, SelH, and Rdx12: genomics and molecular insights into the functions of selenoproteins of a novel thioredoxin-like family. *Biochemistry* 46, 6871-6882 (2007).
- 41. Burk, R.F. & Hill, K.E. Selenoprotein P-expression, functions, and roles in mammals. *Biochim. Biophys. Acta* **1790**, 1441-1447 (2009).

- Pieper, U., Schlessinger, A., Kloppmann, E., Chang, G. A., Chou, J. J., Dumont, M. E., Fox, B. G., Fromme, P., Hendrickson, W. A., Malkowski, M. G., Rees, D. C., Stokes, D. L., Stowell, M. H., Wiener, M. C., Rost, B., Stroud, R. M., Stevens, R. C. & Sali, A. Coordinating the impact of structural genomics on the human alpha-helical transmembrane proteome. *Nature Structural & Molecular Biology* 20, 135-138 (2013).
- 43. Babcock, J.J. & Li, M. Deorphanizing the human transmembrane genome: A landscape of uncharacterized membrane proteins. *Acta Pharmacologica Sinica* 35, 11-23 (2014).
- Uhlén, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., Alm, T., Edqvist, P. H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J. M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J. & Ponten, F. Tissue-based map of the human proteome. *Science* 347, 1260419 (2015).
- Blobel, G. Intracellular protein topogenesis. *Proc. Natl. Acad. Sci. U S A* 77, 1496-1500 (1980).
- 46. Fagerberg, L., Jonasson, K., von Heijne, G., Uhlen, M. & Berglund, L.
  Prediction of the human membrane proteome. *Proteomics* 10, 1141-1149 (2010).

- Almén, M.S., Nordstöm, K.J.V., Fredriksson, R. & Schiöh, H.B. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol.* 7 (2009).
- Worch, R., Bokel, C., Hofinger, S., Schwille, P. & Weidemann, T. Focus on composition and interaction potential of single-pass transmembrane domains. *Proteomics* 10, 4196-4208 (2010).
- 49. Gayen, S., Li, Q.X. & Kang, C.B. Solution NMR study of the transmembrane domain of single-span membrane proteins: opportunities and strategies.
   *Current Protein & Peptide Science* 13, 585-600 (2012).
- von Heijne, G. Membrane-protein topology. *Nature Reviews Molecular Cell Biology* 7, 909-918 (2006).
- 51. von Heijne, G. The membrane protein universe: what's out there and why bother? *Journal of Internal Medicine* **261**, 543-557 (2007).
- Arkin, I.T. Structural aspects of oligomerization taking place between the transmembrane alpha-helices of bitopic membrane proteins. *Biochim. Biophys. Acta* 1565, 347-363 (2002).
- Zviling, M., Kochva, U. & Arkin, I.T. How important are transmembrane helices of bitopic membrane proteins? *Biochim. Biophys. Acta* 1768, 387-392 (2007).
- Langosch, D. & Arkin, I.T. Interaction and conformational dynamics of membrane-spanning protein helices. *Protein Science* 18, 1343-1358 (2009).

- Langosch, D., Herrmann, J.R., Unterreitmeier, S. & Fuchs, A. in Structural Bioinformatics of Membrane Proteins. (ed. D. Frishman) 165-186 (2010).
- Axelrod, D. & Wang, M.D. Reduction-of-dimensionality kinetics at reactionlimited cell-surface receptors. *Biophysical Journal* 66, 588-600 (1994).
- 57. Vereb, G., Szollosi, J., Matko, J., Nagy, P., Farkas, T., Vigh, L., Matyus, L., Waldmann, T. A. & Damjanovich, S. Dynamic, yet structured: the cell membrane three decades after the Singer-Nicolson model. *Proc. Natl. Acad. Sci. U S A* 100, 8053-8058 (2003).
- Verma, S., Hoffmann, F. W., Kumar, M., Huang, Z., Roe, K., Nguyen-Wu, E., Hashimoto, A. S. & Hoffmann, P. R. Selenoprotein K knockout mice exhibit deficient calcium flux in immune cells and impaired immune responses. *J. Immunol.* 186, 2127-2137 (2011).
- 59. Prevost, G., Arabo, A., Jian, L., Quelennec, E., Cartier, D., Hassan, S., Falluel-Morel, A., Tanguy, Y., Gargani, S., Lihrmann, I., Kerr-Conte, J., Lefebvre, H., Pattou, F. & Anouar, Y. The PACAP-regulated gene selenoprotein T Is abundantly expressed in mouse and human beta-cells and its targeted inactivation impairs glucose tolerance. *Endocrinology* **154**, 3796-3806 (2013).
- Moghadaszadeh, B., Rider B.E., Lawlor, M.W., Childers, M.K., Grange, R.W., Gupta, K., Boukedes, S.S., Owen, C.A. & Beggs, A.H. Selenoprotein N deficiency in mice is associated with abnormal lung development. *FASEB J* 27, 1585-1599 (2013).

- 61. Castets, P., Bertrand, A. T., Beuvin, M., Ferry, A., Le Grand, F., Castets, M., Chazot, G., Rederstorff, M., Krol, A., Lescure, A., Romero, N. B., Guicheney, P. and Allamand, V. Satellite cell loss and impaired muscle regeneration in selenoprotein N deficiency. *Hum Mol Genet* 20, 694-704 (2011).
- Rederstorff, M., Castets, P., Arbogast, S., Laine, J., Vassilopoulos, S., Beuvin, M., Dubourg, O., Vignaud, A., Ferry, A., Krol, A., Allamand, V., Guicheney, P., Ferreiro, A. & Lescure, A. Increased muscle stress-sensitivity induced by selenoprotein N inactivation in mouse: a mammalian model for SEPN1-related myopathy. *PLoS One* 6 (2011).
- Schneider, M.J., Fiering, S. N., Thai, B., Wu, S. Y., St Germain, E., Parlow, A. F., St Germain, D. L. & Galton, V. A. Targeted disruption of the type 1 selenodeiodinase gene (Dio1) results in marked changes in thyroid hormone economy in mice. *Endocrinology* 147, 580-589 (2006).
- 64. Schneider, M.J., Fiering, S.N., Pallud, S.E., Parlow, A.F., St Germain, D.L. & Galton, V.A. Targeted disruption of the type 2 selenodeiodinase gene (D102) results in a phenotype of pituitary resistance to T4. *Mol. Endocrinol.* 15, 2137-2148 (2001).
- de Jesus, L.A., Carvalho, S.D., Ribeiro, M.O., Schneider, M., Kim, S.W.,
  Harney, J.W., Larsen, P.R. & Bianco, A.C. The type 2 iodothyronine
  deiodinase is essential for adaptive thermogenesis in brown adipose tissue. *J. Clin. Invest.* 108, 1379-1385 (2001).

- Ng, L., Goodyear, R.J., Woods, C.A., Schneider, M.J., Diamond, E.,
  Richardson, G.P., Kelley, M.W., Germain, D.L., Galton, V.A. & Forrest, D.
  Hearing loss and retarded cochlear development in mice lacking type 2
  iodothyronine deiodinase. *Proceedings of the National Academy of Sciences of the United States of America* 101, 3474-3479 (2004).
- 67. Bárez-López, S., Bosch-García, D., Gómez-Andrés, D., Pulido-Valdeolivas, I., Montero-Pedrazuela, A., Obregon, M.J. & Guadaño-Ferraz, A. Abnormal motor phenotype at adult stages in mice lacking type 2 deiodinase. *PLoS One* 9, e103857 (2014).
- Hernandez, A., Martinez, M.E., Fiering, S., Galton, V.A. & St Germain, D. Type 3 deiodinase is critical for the maturation and function of the thyroid axis. *J. Clin. Invest.* 116, 476-484 (2006).
- Lilley, B.N. & Ploegh, H.L. A membrane protein required for dislocation of misfolded proteins from the ER. *Nature* 429, 834-840 (2004).
- Ye, Y.H., Shibata, Y., Yun, C., Ron, D. & Rapoport, T.A. A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol. *Nature* 429, 841-847 (2004).
- Turanov, A.A., Shchedrina, V.A., Everley, R.A., Lobanov, A.V., Yim, S.H., Marino, S.M., Gygi, S.P., Hatfield, D.L. & Gladyshev, V.N. Selenoprotein S is involved in maintenance and transport of multiprotein complexes. *Biochem. J.* 462, 555-565 (2014).

- Shchedrina, V.A., Everley, R.A., Zhang, Y., Gygi, S.P., Hatfield, D.L. & Gladyshev, V.N. Selenoprotein K binds multiprotein complexes and is involved in the regulation of endoplasmic reticulum homeostasis. *J. Biol. Chem.* 286, 42937-42948 (2011).
- 73. Uversky, V.N., Davé, V., Iakoucheva, L.M., Malaney, P., Metallo, S.J.,
  Pathak, R.R. & Joerger, A.C. Pathological unfoldomics of uncontrolled chaos: Intrinsically disordered proteins and human diseases. *Chemical Reviews* 114, 6844-6879 (2014).
- Wright, P.E. & Dyson, H.J. Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology* 16, 18-29 (2015).
- 75. Vendruscolo, M. Enzymatic activity in disordered states of proteins. *Current Opinion in Chemical Biology* **14**, 671-675 (2010).
- 76. Schulenburg, C. & Hilvert, D. in Dynamics in Enzyme Catalysis, Vol. 337.(eds. J. Klinman & S. HammesSchiffer) 41-67 (2013).
- Röseler, A., Prieto, J.H., Iozef, R., Hecker, B., Schirmer, R.H., Külzer, S.,
  Przyborski, J., Rahlfs, S. & Becker, K. Insight into the selenoproteome of the malaria parasite plasmodium falciparum. *Antioxid. Redox Signal.* 17, 534-543 (2012).
- 78. Christensen, L.C., Jensen, N.W., Vala, A., Kamarauskaite, J., Johansson, L., Winther, J.R., Hofmann, K., Teilum, K. & Ellgaard, L. The human selenoprotein VCP-interacting membrane protein (VIMP) is non-globular and

harbors a reductase function in an intrinsically disordered region. *J. Biol. Chem.* **287**, 26388-26399 (2012).

- Liu, J. & Rozovsky, S. Contribution of selenocysteine to the peroxidase activity of selenoprotein S. *Biochemistry* 52, 5514-5516 (2013).
- Bubenik, J.L., Miniard, A.C. & Driscoll, D.M. Alternative transcripts and
   3'UTR elements govern the incorporation of selenocysteine into selenoprotein
   S. *PLoS One* 8, e62102 (2013).
- Shchedrina, V.A., Zhang, Y., Labunskyy, V.M., Hatfield, D.L. & Gladyshev,
   V.N. Structure-function relations, physiological roles, and evolution of
   mammalian ER-resident selenoproteins. *Antioxid. Redox Signal.* 12, 839-849
   (2010).
- Haenzelmann, P. & Schindelin, H. The structural and functional basis of the p97/Valosin-containing protein (VCP)-interacting motif (VIM). *Journal of Biological Chemistry* 286, 38679-38690 (2011).
- Meyer, H. & Weihl, C.C. The VCP/p97 system at a glance: connecting cellular function to disease pathogenesis. *Journal of Cell Science* 127, 3877-3883 (2014).
- Lee, J. H., Kwon, J. H., Jeon, Y. H., Ko, K. Y., Lee, S. R. & Kim, I. Y. Pro178 and Pro183 of selenoprotein S are essential residues for interaction with p97(VCP) during endoplasmic reticulum-associated degradation. *J. Biol. Chem.* 289, 13758-13768 (2014).

- Kryukov, G.V., Castellano, S., Novoselov, S.V., Lobanov, A.V., Zehtab, O.,
   Guigó, R. & Gladyshev, V.N. Characterization of mammalian selenoproteomes. *Science* 300, 1439-1443 (2003).
- Lilley, B.N. & Ploegh, H.L. Multiprotein complexes that link dislocation, ubiquitination, and extraction of misfolded proteins from the endoplasmic reticulum membrane. *Proc. Natl. Acad. Sci. U S A* 102, 14296-14301 (2005).
- 87. Smith, M.H., Ploegh, H.L. & Weissman, J.S. Road to ruin: Targeting proteins for degradation in the endoplasmic reticulum. *Science* **334**, 1086-1090 (2011).
- Christianson, J.C. & Ye, Y. Cleaning up in the endoplasmic reticulum: ubiquitin in charge. *Nature Structural & Molecular Biology* 21, 325-335 (2014).
- Lemus, L. & Goder, V. Regulation of Endoplasmic Reticulum-Associated
   Protein Degradation (ERAD) by Ubiquitin. *Cells* 3, 824-847 (2014).
- Bernasconi, R. & Molinari, M. ERAD and ERAD tuning: disposal of cargo and of ERAD regulators from the mammalian ER. *Current Opinion in Cell Biology* 23, 176-183 (2011).
- 91. Christianson, J. C., Olzmann, J. A., Shaler, T. A., Sowa, M. E., Bennett, E. J., Richter, C. M., Tyler, R. E., Greenblatt, E. J., Harper, J. W. & Kopito, R. R. Defining human ERAD networks through an integrative mapping strategy. *Nat. Cell Biol.* 14, 93-105 (2011).
- 92. Schulze, A., Standera, S., Buerger, E., Kikkert, M., van Voorden, S., Wiertz,E., Koning, F., Kloetzel, P.M. & Seeger, M. The ubiquitin-domain protein

HERP forms a complex with components of the endoplasmic reticulum associated degradation pathway. *Journal of Molecular Biology* **354**, 1021-1027 (2005).

- 93. Noda, C., Kimura, H., Arasaki, K., Matsushita, M., Yamamoto, A., Wakana,
  Y., Inoue, H. & Tagaya, M. Valosin-containing Protein-interacting Membrane
  Protein (VIMP) Links the Endoplasmic Reticulum with Microtubules in
  Concert with Cytoskeleton-linking Membrane Protein (CLIMP)-63. *Journal of Biological Chemistry* 289, 24304-24313 (2014).
- 94. Kim, C.Y. & Kim, K.H. Dexamethasone-induced selenoprotein S degradation is required for adipogenesis. *Journal of Lipid Research* **54**, 2069-2082 (2013).
- 95. Olzmann, J.A., Richter, C.M. & Kopito, R.R. Spatial regulation of UBXD8 and p97/VCP controls ATGL-mediated lipid droplet turnover. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 1345-1350 (2013).
- Kim, H. & Ye, J. Cellular responses to excess fatty acids: focus on ubiquitin regulatory X domain-containing protein 8. *Current Opinion in Lipidology* 25, 118-124 (2014).
- 97. Alanne, M., Kristiansson, K., Auro, K., Silander, K., Kuulasmaa, K., Peltonen, L., Salomaa, V. & Perola, M. Variation in the selenoprotein S gene locus is associated with coronary heart disease and ischemic stroke in two independent Finnish cohorts. *Hum. Genet.* 122, 355-365 (2007).

- 98. Gao, Y., Xie, X., Ma, Y. T., Yang, Y. N., Li, X. M., Fu, Z. Y., Zheng, Y. Y., Ma, X., Chen, B. D., Liu, F. & Huang, Y. Genetic variation in Tanis was associated with elevating plasma triglyceride level in Chinese nondiabetic subjects. *Lipids Health Dis* 12, 97 (2013).
- Cox, A. J., Lehtinen, A. B., Xu, J. Z., Langefeld, C. D., Freedman, B. I., Carr, J. J. & Bowden, D. W. Polymorphisms in the Selenoprotein S gene and subclinical cardiovascular disease in the Diabetes Heart Study. *Acta Diabetol* 50, 391-399 (2013).
- 100. Olsson, M., Olsson, B., Jacobson, P., Thelle, D.S., Björkegren, J., Walley, A., Froguel, P., Carlsson, L.M. & Sjöholm, K. Expression of the selenoprotein S (SELS) gene in subcutaneous adipose tissue and SELS genotype are associated with metabolic risk factors. *Metabolism* 60, 114-120 (2011).
- Walder, K., Kantham, L., McMillan, J. S., Trevaskis, J., Kerr, L., de Silva, A.,
  Sunderland, T., Godde, N., Gao, Y., Bishara, N., Windmill, K., Tenne-Brown,
  J., Augert, G., Zimmet, P. Z. & Collier, G. R. Tanis: A link between type 2
  diabetes and inflammation? *Diabetes* 51, 1859-1866 (2002).
- 102. Curran, J. E., Jowett, J. B., Elliott, K. S., Gao, Y., Gluschenko, K., Wang, J., Abel Azim, D. M., Cai, G., Mahaney, M. C., Comuzzie, A. G., Dyer, T. D., Walder, K. R., Zimmet, P., MacCluer, J. W., Collier, G. R., Kissebah, A. H. & Blangero, J. Genetic variation in selenoprotein S influences inflammatory response. *Nature Genetics* 37, 1234-1241 (2005).

- 103. Seiderer, J., Dambacher, J., Kühnlein, B., Pfennig, S., Konrad, A., Török, H.P., Haller, D., Göke, B., Ochsenkühn, T., Lohse, P. & Brand, S. The role of the selenoprotein S (SELS) gene 105G > A promoter polymorphism in inflammatory bowel disease and regulation of SELS gene expression in intestinal inflammation. *Tissue Antigens* 70, 238-246 (2007).
- Santos, L,R., Durães, C., Mendes, A., Prazeres, H., Alvelos, M.I., Moreira, C.S., Canedo, P., Esteves, C., Neves, C., Carvalho, D., Sobrinho-Simões, M., Soares, P. A polymorphism in the promoter region of the selenoprotein S gene (SEPS1) contributes to Hashimoto's thyroiditis susceptibility. *J. Clin. Endocrinol. Metab.* **99**, E719-E723 (2014).
- Moses, E.K., Johnson, M.P., Tømmerdal, L., Forsmo, S., Curran, J.E.,
  Abraham, L.J., Charlesworth, J.C., Brennecke, S.P., Blangero, J. & Austgulen,
  R. Genetic association of preeclampsia to the inflammatory response gene
  SEPS1. *Am. J. Obstet. Gynecol.* **198**, 336.e331–336.e335 (2008).
- Sutherland, A., Kim, D.H., Relton, C., Ahn, Y.O. & Hesketh, J.
  Polymorphisms in the selenoprotein S and 15-kDa selenoprotein genes are associated with altered susceptibility to colorectal cancer. *Genes Nutr.* 5, 215-223 (2010).
- 107. Meplan, C. & Hesketh, J. in Advances in Nutrition and Cancer, Vol. 159. (eds. V. Zappia, S. Panico, G.L. Russo, A. Budillon & F. DellaRagione) 145-166 (2014).
- Shibata, T., Arisawa, T., Tahara, T., Ohkubo, M., Yoshioka, D., Maruyama, N., Fujita, H., Kamiya, Y., Nakamura, M., Nagasaka, M., Iwata, M., Takahama, K., Watanabe, M. & Hirata, I. Selenoprotein S (SEPS1) gene 105G>A promoter polymorphism influences the susceptibility to gastric cancer in the Japanese population. *BMC Gastroenterol* 9, 2 (2009).
- 109. Zhao, Y., Li, H., Men, L. L., Huang, R. C., Zhou, H. C., Xing, Q., Yao, J. J., Shi, C. H. & Du, J. L. Effects of selenoprotein S on oxidative injury in human endothelial cells. *J Transl Med* 11, 287 (2013).
- Du, X. A., Wang, H. M., Dai, X. X., Kou, Y., Wu, R. P., Chen, Q., Cao, J. L., Mo, X. Y. & Xiong, Y. M. Role of selenoprotein S (SEPS1)-105G>A polymorphisms and PI3K/Akt signaling pathway in Kashin-Beck disease. *Osteoarthritis and Cartilage* 23, 210-216 (2015).
- 111. Speckmann, B., Gerloff, K., Simms, L., Oancea, I., Shi, W., McGuckin, M.A., Radford-Smith, G. & Khanna, K.K. Selenoprotein S is a marker but not a regulator of endoplasmic reticulum stress in intestinal epithelial cells. *Free Radic. Biol. Med.* 67, 265-277 (2014).
- 112. Touat-Hamici, Z., Legrain, Y., Bulteau, A.L. & Chavatte, L. Selective Upregulation of Human Selenoproteins in Response to Oxidative Stress. *Journal* of Biological Chemistry 289, 14750-14761 (2014).
- Liu, J., Srinivasan, P., Pham, D.N. & Rozovsky, S. Expression and purification of the membrane enzyme selenoprotein K. *Protein Expression and Purification* 86, 27-34 (2012).

- Hoffmann, P.R. in Selenium: Its Molecular Biology and Role in Human Health, Third Edition. (ed. D.B. Hatfield, MJ Gladyshev, VN) 335-344 (Springer, New York; 2012).
- 115. Ulmschneider, M.B. & Sansom, M.S.P. Amino acid distributions in integral membrane protein structures. *Biochim. Biophys. Acta* **1512**, 1-14 (2001).
- 116. von Heijne, G. Membrane proteins: from bench to bits. *Biochemical Society Transactions* 39, 747-750 (2011).
- Cymer, F., von Heijne, G. & White, S.H. Mechanisms of Integral Membrane Protein Insertion and Folding. *J Mol Biol* 427, 999-1022 (2015).
- 118. Schneider, D., Finger, C. & Prodöhl, A. From interactions of single transmembrane helices to folding of alpha-helical membrane proteins: analyzing transmembrane helix-helix interactions in bacteria. *Curr. Protein Pept. Sci.* 8, 45-61 (2007).
- Liu, J., Zhang, Z. Q. & Rozovsky, S. Selenoprotein K form an intermolecular diselenide bond with unusually high redox potential. *FEBS Lett* 588, 3311-3321 (2014).
- Li, F., Lutz, P. B., Pepelyayeva, Y., Arner, E. S. J., Bayse, C. A. & Rozovsky,
  S. Redox active motifs in selenoproteins. *Proc. Natl. Acad. Sci. U S A* 111,
  6976-6981 (2014).
- 121. Lu, J. & Holmgren, A. Selenoproteins. J. Biol. Chem. 284, 723-727 (2009).
- 122. Shchedrina, V.A., Novoselov, S.V., Malinouski, M.Y. & Gladyshev, V.N.Identification and characterization of a selenoprotein family containing a

diselenide bond in a redox motif. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 13919-13924 (2007).

- 123. Ma, S.G., Hill, K.E., Burk, R.F. & Caprioli, R.M. Mass spectrometric determination of selenenylsulfide linkages in rat selenoprotein P. *Journal of Mass Spectrometry* 40, 400-404 (2005).
- 124. Fredericks, G. J., Hoffmann, F. W., Rose, A. H., Osterheld, H. J., Hess, F. M., Mercier, F. & Hoffmann, P. R. Stable expression and function of the inositol 1,4,5-triphosphate receptor requires palmitoylation by a DHHC6/selenoprotein K complex. *Proc Natl Acad Sci U S A* 111, 16478-16483 (2014).
- Goll, D.E., Thompson, V.F., Li, H.Q., Wei, W. & Cong, J.Y. The calpain system. *Physiological Reviews* 83, 731-801 (2003).
- 126. Huang, Z., Hoffmann, F.W., Norton, R.L., Hashimoto, A.C. & Hoffmann, P.R. Selenoprotein K is a novel target of m-calpain, and cleavage is regulated by toll-like receptor-induced calpastatin in macrophages. *J. Biol. Chem.* 286, 34830-34838 (2011).
- Lemberg, M.K. Intramembrane Proteolysis in Regulated Protein Trafficking. *Traffic* 12, 1109-1118 (2011).
- Linder, M.E. & Jennings, B.C. Mechanism and function of DHHC Sacyltransferases. *Biochem Soc Trans* 41, 29-34 (2013).
- 129. Greaves, J. & Chamberlain, L.H. DHHC palmitoyl transferases: substrate interactions and (patho)physiology. *Trends Biochem Sci* **36**, 245-253 (2011).

- 130. Korycka, J., Lach, A., Heger, E., Boguslawska, D. M., Wolny, M.,
  Toporkiewicz, M., Augoff, K., Korzeniewski, J. & Sikorski, A. F. Human
  DHHC proteins: a spotlight on the hidden player of palmitoylation. *Eur. J. Cell Biol.* 91, 107-117 (2012).
- Meiler, S., Baumer, Y., Huang, Z., Hoffmann, F. W., Fredericks, G. J., Rose,
  A. H., Norton, R. L., Hoffmann, P. R. & Boisvert, W. A. Selenoprotein K is required for palmitoylation of CD36 in macrophages: implications in foam cell formation and atherogenesis. *J. Leukocyte Biol.* 93, 771-780 (2013).
- 132. Mitchell, D. A., Hamel, L. D., Ishizuka, K., Mitchell, G., Schaefer, L. M. & Deschenes, R. J. The Erf4 subunit of the yeast Ras palmitoyl acyltransferase is required for stability of the Acyl-Erf2 intermediate and palmitoyl transfer to a Ras2 substrate. *The Journal of biological chemistry* 287, 34337-34348 (2012).
- 133. Swarthout, J. T., Lobo, S., Farh, L., Croke, M. R., Greentree, W. K., Deschenes, R. J. & Linder, M. E. DHHC9 and GCP16 constitute a human protein fatty acyltransferase with specificity for H- and N-Ras. *The Journal of biological chemistry* 280, 31141-31148 (2005).
- 134. Greenblatt, E.J., Olzmann, J.A. & Kopito, R.R. Derlin-1 is a rhomboid pseudoprotease required for the dislocation of mutant alpha-1 antitrypsin from the endoplasmic reticulum. *Nat Struct Mol Biol* 18, 1147-1152 (2011).
- Mehnert, M., Sommer, T. & Jarosch, E. Der1 promotes movement of misfolded proteins through the endoplasmic reticulum membrane. *Nature Cell Biology* 16, 77-86 (2014).

- Kemp, G. & Cymer, F. Small membrane proteins elucidating the function of the needle in the haystack. *Biological Chemistry* 395, 1365-1377 (2014).
- 137. Ramamurthi, K.S. & Storz, G. The small protein floodgates are opening; now the functional analysis begins. *Bmc Biology* **12** (2014).
- 138. Storz, G., Wolf, Y.I. & Ramamurthi, K.S. Small proteins can no longer be ignored. *Annual review of biochemistry* **83**, 753-777 (2014).
- Dawson, P., Muir, T., Clark-Lewis, I. & Kent, S. Synthesis of proteins by native chemical ligation. *Science* 266, 776-779 (1994).
- 140. Dawson, P.E. & Kent, S.B. Synthesis of native proteins by chemical ligation.*Annu Rev Biochem* 69, 923-960 (2000).
- Muir, T.W., Sondhi, D. & Cole, P.A. Expressed protein ligation: a general method for protein engineering. *Proc Natl Acad Sci U S A* 95, 6705-6710 (1998).
- 142. David, R., Richter, M.P. & Beck-Sickinger, A.G. Expressed protein ligation.Method and applications. *Eur J Biochem* 271, 663-677 (2004).
- Anraku, Y., Mizutani, R. & Satow, Y. Protein splicing: its discovery and structural insight into novel chemical mechanisms. *IUBMB life* 57, 563-574 (2005).
- 144. Shah, N.H. & Muir, T.W. Inteins: Nature's Gift to Protein Chemists. *Chemical science (Royal Society of Chemistry : 2010)* **5**, 446-461 (2014).

- 145. Malins, L.R. & Payne, R.J. Recent extensions to native chemical ligation for the chemical synthesis of peptides and proteins. *Current opinion in chemical biology* 22, 70-78 (2014).
- 146. Tolbert, T.J. & Wong, C.H. New methods for proteomic research: preparation of proteins with N-terminal cysteines for labeling and conjugation. *Angew Chem Int Ed Engl* **41**, 2171-2174 (2002).
- Yuan, L., Lin, W., Xie, Y., Zhu, S. & Zhao, S. A native-chemical-ligationmechanism-based ratiometric fluorescent probe for aminothiols. *Chemistry* 18, 14520-14526 (2012).
- 148. De Rosa, L., Cortajarena, A.L., Romanelli, A., Regan, L. & D'Andrea, L.D. Site-specific protein double labeling by expressed protein ligation: applications to repeat proteins. *Org Biomol Chem* **10**, 273-280 (2012).
- Yi, L., Sun, H., Itzen, A., Triola, G., Waldmann, H., Goody, R. S. & Wu, Y.
  W. One-pot dual-labeling of a protein by two chemoselective reactions. *Angew Chem Int Ed Engl* 50, 8287-8290 (2011).
- Camarero, J.A. & Muir, T.W. Biosynthesis of a Head-to-Tail Cyclized Protein with Improved Biological Activity. *Journal of the American Chemical Society* 121, 5597-5598 (1999).
- Cowburn, D. & Muir, T.W. Segmental isotopic labeling using expressed protein ligation. *Methods in enzymology* 339, 41-54 (2001).

- Nabeshima, Y., Mizuguchi, M., Kajiyama, A. & Okazawa, H. Segmental isotope-labeling of the intrinsically disordered protein PQBP1. *FEBS Lett* 588, 4583-4589 (2014).
- 153. Xu, R., Ayers, B., Cowburn, D. & Muir, T.W. Chemical ligation of folded recombinant proteins: segmental isotopic labeling of domains for NMR studies. *Proc Natl Acad Sci U S A* 96, 388-393 (1999).
- 154. Minato, Y., Ueda, T., Machiyama, A., Shimada, I. & Iwai, H. Segmental isotopic labeling of a 140 kDa dimeric multi-domain protein CheA from Escherichia coli by expressed protein ligation and protein trans-splicing. *J Biomol NMR* 53, 191-207 (2012).
- 155. Lu, W., Gong, D., Bar-Sagi, D. & Cole, P.A. Site-specific incorporation of a phosphotyrosine mimetic reveals a role for tyrosine phosphorylation of SHP-2 in cell signaling. *Molecular cell* 8, 759-769 (2001).
- Tam, J.P. & Yu, Q. Methionine ligation strategy in the biomimetic synthesis of parathyroid hormones. *Biopolymers* 46, 319-327 (1998).
- Muir, T.W. Semisynthesis of proteins by expressed protein ligation. *Annu Rev Biochem* 72, 249-289 (2003).
- 158. Holt, M. & Muir, T. Application of the protein semisynthesis strategy to the generation of modified chromatin. *Annu Rev Biochem* **84**, 265-290 (2015).
- Erlanson, D.A., Chytil, M. & Verdine, G.L. The leucine zipper domain controls the orientation of AP-1 in the NFAT.AP-1.DNA complex. *Chem Biol* 3, 981-991 (1996).

- Liu, D., Xu, R., Dutta, K. & Cowburn, D. N-terminal cysteinyl proteins can be prepared using thrombin cleavage. *FEBS Lett* 582, 1163-1167 (2008).
- Gentle, I.E., De Souza, D.P. & Baca, M. Direct production of proteins with Nterminal cysteine for site-specific conjugation. *Bioconjug Chem* 15, 658-663 (2004).
- Raibaut, L., El Mahdi, O. & Melnyk, O. Solid phase protein chemical synthesis. *Top Curr Chem* 363, 103-154 (2015).
- Evans, T.C., Jr. & Xu, M.Q. Intein-mediated protein ligation: harnessing nature's escape artists. *Biopolymers* 51, 333-342 (1999).
- 164. Muralidharan, V. & Muir, T.W. Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nature methods* 3, 429-438 (2006).
- 165. Thom, J., Anderson, D., McGregor, J. & Cotton, G. Recombinant protein hydrazides: application to site-specific protein PEGylation. *Bioconjug Chem*22, 1017-1020 (2011).
- 166. Fang, G.M., Wang, J.X. & Liu, L. Convergent chemical synthesis of proteins by ligation of peptide hydrazides. *Angew Chem Int Ed Engl* 51, 10347-10350 (2012).
- 167. Hondal, R.J., Nilsson, B.L. & Raines, R.T. Selenocysteine in native chemical ligation and expressed protein ligation. *J Am Chem Soc* 123, 5140-5141 (2001).
- Quaderer, R., Sewing, A. & Hilvert, D. Selenocysteine-Mediated Native Chemical Ligation. *Helvetica Chimica Acta* 84, 1197-1206 (2001).

- Gieselman, M.D., Xie, L. & van Der Donk, W.A. Synthesis of a selenocysteine-containing peptide by native chemical ligation. *Org Lett* 3, 1331-1334 (2001).
- Dawson, P.E. Native Chemical Ligation Combined with Desulfurization and Deselenization: A General Strategy for Chemical Protein Synthesis. *Israel Journal of Chemistry* 51, 862-867 (2011).
- Metanis, N., Keinan, E. & Dawson, P.E. Traceless Ligation of Cysteine
   Peptides Using Selective Deselenization. *Angewandte Chemie-International Edition* 49, 7049-7053 (2010).
- Townsend, S. D., Tan, Z. P., Dong, S. W., Shang, S. Y., Brailsford, J. A. & Danishefsky, S. J. Advances in proline ligation. *J Am Chem Soc* 134, 3912-3916 (2012).
- Malins, L.R. & Payne, R.J. Synthesis and utility of β-selenol-phenylalanine for native chemical ligation-deselenization chemistry. *Org Lett* 14, 3142-3145 (2012).
- Metanis, N., Keinan, E. & Dawson, P.E. Synthetic seleno-glutaredoxin 3 analogues are highly reducing oxidoreductases with enhanced catalytic efficiency. *J Am Chem Soc* 128, 16684-16691 (2006).
- 175. Dery, S., Reddy, P. S., Dery, L., Mousa, R., Dardashti, R. N. & Metanis, N.
  Insights into the deselenization of selenocysteine into alanine and serine. *Chemical Science* 6, 6207-6212 (2015).

- Malins, L.R., Mitchell, N.J., McGowan, S. & Payne, R.J. Oxidative Deselenization of Selenocysteine: Applications for Programmed Ligation at Serine. *Angew Chem Int Ed Engl* 54, 12716-12721 (2015).
- 177. Levengood, M.R. & van der Donk, W.A. Dehydroalanine-containing peptides: preparation from phenylselenocysteine and utility in convergent ligation strategies. *Nat Protoc* 1, 3001-3010 (2006).
- Berry, S.M., Gieselman, M.D., Nilges, M.J., van Der Donk, W.A. & Lu, Y. An engineered azurin variant containing a selenocysteine copper ligand. *J Am Chem Soc* 124, 2084-2085 (2002).
- Schmohl, L., Wagner, F.R., Schumann, M., Krause, E. & Schwarzer, D.
   Semisynthesis and initial characterization of sortase A mutants containing selenocysteine and homocysteine. *Bioorganic & medicinal chemistry* 23, 2883-2889 (2015).
- Hondal, R.J. Using chemical approaches to study selenoproteins focus on thioredoxin reductases. *Biochimica et biophysica acta* 1790, 1501 (2009).
- 181. Southworth, M.W., Amaya, K., Evans, T.C., Xu, M.Q. & Perler, F.B.
  Purification of proteins fused to either the amino or carboxy terminus of the
  Mycobacterium xenopi gyrase A intein. *Biotechniques* 27, 110-114, 116, 118-120 (1999).
- Hondal, R.J. & Raines, R.T. Semisynthesis of proteins containing selenocysteine. *Methods Enzymol* 347, 70-83 (2002).

- Hondal, R.J. Incorporation of selenocysteine into proteins using peptide ligation. *Protein Pept Lett* 12, 757-764 (2005).
- 184. Muttenthaler, M. & Alewood, P.F. Selenopeptide chemistry. *J Pept Sci* 14, 1223-1239 (2008).

## Chapter 2

# EXPRESSION AND PURIFICATION OF THE MEMBRANE ENZYMES SELENOPROTEIN K AND SELENOPROTEIN S

#### 2.1 Introduction

Selenoproteins form a specialized family of enzymes that contain the genetically encoded amino acid selenocysteine (Sec, U). Members of the family participate in regulation of redox pathway, signaling, and synthesis of lipids and hormones, and they also protect the cells against oxidative stress<sup>1, 2</sup>. Selenoprotein K (SelK) was identified in 2003 as one of the 25 selenoproteins encoded in the human genome<sup>3</sup>. It is a membrane selenoprotein, with a critical role in development<sup>4</sup>, protection against oxidative stress<sup>5, 6</sup>, and longevity<sup>7</sup>. Information about SelK is scant-it is almost certainly an enzyme, as the vast majority of the selenoproteome utilizes Sec to catalyze reactions<sup>8</sup>. It partakes in oxidative defense, since overexpression of SelK in cardiomyocytes decreased the level of intracellular reactive oxygen species (ROS)<sup>5</sup>. SelK was also shown to protect the cells against exogenously imposed oxidative stress<sup>5, 6</sup>. It is expressed in all tissues but is most abundant in spleen and immune cells<sup>9</sup> and is coexpressed at high levels in four brain regions, along with SelP, GPx4, SelM, SelW, and Sep15 – all selenoproteins involved in protecting cells from oxidative stress<sup>10</sup>. Nevertheless, the precise enzymatic reaction and the physiological function(s) of SelK are unknown. Several studies suggest that SelK is involved in calcium regulation and in the ER-assisted protein degradation (ERAD) pathway-a quality control system responsible for the dislocation of misfolded

proteins from the ER for degradation in the cytoplasm<sup>11</sup>. A recent paper demonstrated co-immunoprecipitation of SelK with components of the ERAD protein machinery<sup>12</sup>. In a different study, SelK knockout mice exhibited deficiencies in calcium-related functions in the immune system<sup>9</sup>. A follow up study has shown that SelK undergoes proteolytic cleavage by m-calpain following Toll-like receptor activation<sup>13</sup> and is involved in regulation of inositol 1,4,5-tris-phosphate receptors<sup>14</sup>.



Figure 2.1. Schematic representations of SelK and SelS. (A) Human SelK has a predicted single-pass transmembrane helix; shown here is a prediction by the TMHMM program<sup>15</sup>. It also has two potential SH3 binding elements and a putative phosphorylation site. The reactive Sec residue resides at the C-terminal, near a conserved Arg, and is exposed to the cytoplasm. (B) Human SelS is also predicted to have a single-pass transmembrane helix. Similar to SelK the C-terminal domain faces the cytoplasm. In addition to the transmembrane helix it has a p97/Valosin-containing Protein (VCP)-interacting Motif (VIM), a coiled coil dimerization interface and a disordered C-terminal region with an internal selenylsulfide bond.

SelK is predicted to have a single-pass transmembrane helix (Figure 2.1A) and localizes *in vivo* to the endoplasmic reticulum (ER)<sup>5</sup>. The predicted transmembrane helix does not have any of the canonic motifs that lead to oligomerization of membrane proteins<sup>16, 17</sup>. However, its relatively high content of polar residues supports potential oligomerization<sup>18</sup>. The reactive Sec, located at position 92 of 94 residues in a conserved M(A/G)GGUGR sequence, is exposed to the cytoplasm<sup>19, 20</sup>. Uncharacteristically for selenoproteins, the Sec is not paired with a nearby Cys, Ser or Thr. In other selenoproteins, such a neighboring residue protects the easily oxidized Sec by forming a selenenylsulfide or hydrogen bond. It is possible that a hydrogen bond donor is not near in the primary sequence but is in close proximity to the Sec in the three dimensional structure or is provided by a yet-to-be-identified protein partner(s). Indeed, SelK has several motifs responsible for interactions with signaling proteins: a Src homology 3 (SH3) binding sequence<sup>21</sup>, a second atypical SH3 domain<sup>22</sup>, and a putative phosphorylation site at Ser 51. Pull down assays identified the ERAD components Derlin-1, Derlin-2 and Selenoprotein S (SelS, also known as VIMP) as SelK's binding partners. SelS, which belongs to the same family of membrane proteins, was proposed to be a reductase<sup>23</sup>. SelK and SelS were recently classified as members of a novel eukaryotic SelK/SelS family of proteins [12]. Members of this family have a short N-terminal ER luminal sequence; an N-terminal single pass transmembrane helix; a region rich in Gly, Pro, and charged residues; and a C-terminal active site (with either Sec or Cys). Their role is not well understood but could be broadly related to oxidative stress.

Biophysical characterization of SelK remains limited, owing to difficulties in preparation of selenoproteins<sup>24</sup> and membrane proteins<sup>25</sup>. In this study, we have successfully developed an efficient protocol for overexpression and purification of the full length human SelK, in which the active site Sec was substituted with Cys (U92C). A Sec to Cys substitution in selenoproteins is commonly employed for the high-level protein production that is necessary for biophysical and structural characterization<sup>24</sup>. This substitution typically reduces enzymatic activity by 10 - 1000 fold but does not otherwise interfere with function or structural integrity. We show that by employing this mutation, it is possible to overexpress SelK as a fusion protein, purify it to homogeneity, and stabilize it in various detergents. This work is essential for establishing successful structural and functional characterization of SelK and for determining its mechanism of action.

We also demonstrate that the purification strategy for SelK might be generally applicable to other members of this emerging protein family. To test this hypothesis we have employed the procedures described for SelK on its protein partner SelS. Even though SelK and SelS belong to the same family of membrane enzymes their transmembrane segments differ significantly (Figure 2.1). The SelK transmembrane helix has, rather unusually, three residues that could potentially be charged at physiological pH (Glu, Asp, and Lys) while SelS having only one (Cys). Their cytoplasmic portions are disparate with the dimeric SelS has an extended coil coiled region and a stabilizing intramolecular selenylsulfide bond while SelK has a proline rich short segment that does appear to be stabilized by intramolecular bonds. Hence, SelS provides a suitable example to test the generality of the procedure described for

58

SelK for other members of the family. We demonstrate that this expression and purification strategy can also be applied to SelK's binding partner SelS.

## 2.2 Results

## 2.2.1 Overexpression of SelK

To facilitate structural and biophysical characterization, we have developed bacterial overexpression and purification strategies for SelK. SelK is predicted to be a single pass transmembrane protein: a class of proteins that have a propensity to be in inclusion bodies, rather than insert in the E. coli plasma membrane. The expression level and solubility of membrane proteins are enhanced by the use of fusion proteins that can be cleaved after production $^{26, 27}$ . The selected affinity tags, a hexahistidine, SUMO, pMBP, and cMBP tags, were previously reported to be successful for the expression of membrane proteins or difficult targets<sup>28, 29</sup>. Since SelK's active site is located at the C-terminal, we constructed N-terminal fusions that could be separated from SelK following expression using TEV protease (see bacterial expression vectors in Figure 2.2). In each case, the Sec was mutated to a Cys (U92C) to assist the protein production. A Sec to Cys substitution is necessary to circumvent the unique requirements for Sec incorporation in proteins (it shares the TGA codon with a stop signal<sup>28</sup>). Since sulfur and selenium share many physicochemical properties<sup>30</sup>, the substitution does not otherwise interfere with structural integrity, oligomerization state, or protein stability $^{24}$ .



Figure 2.2. Schematic representations of the expression constructs tested in this study. His-tag, hexahistidine sequence; SUMO, small ubiquitin-like modifier protein; cMBP, cytoplasmic maltose-binding protein; pMBP, periplasmic maltose-binding protein; short linker, NSSS; long linker, NSSSNNNNNNNNNLG; TEV, tobacco etch virus protease cleavage site (ENLYFQ/G); Strep-tag, streptavidin binding tag (WSHPQFEK); SelK U92C, *Homo sapiens* SelK with a U92C mutation at the active site.

SelK fusion proteins were expressed in BL21(DE3) cells, and the expression levels using the different expression vectors were assessed by SDS-PAGE and Western blot, following small-scale purifications (data not shown). The identity of the protein was further corroborated using MS/MS sequencing. The best expression was observed in the case of the cMBP fusion (also called the mature MBP), a 42 kDa protein that increases solubility and acts as a chaperone assisting protein folding. cMBP-fusion has been reported to be a particularly successful strategy for increasing expression of membrane proteins<sup>28</sup>. This chapter describes optimization of expression and purification using the cMBP-SelK fusion.

# 2.2.2 The Relation between the Length of Linker between SelK and cMBP and *in vivo* Proteolytic Cleavage

Most MBP fusion proteins are constructed with a flexible, Asn-rich linker to improve binding of the fusion protein to the affinity column and to increase accessibility to the cleavage site by the designated protease<sup>31</sup>. In the case of cMBP-SelK fusion protein, we noticed significant cleavage at the linker region between the two proteins, as evidenced by the appearance of untagged cMBP (Figure 2.3). Cleavage appeared to occur *in vivo*, as the degree of truncation did not depend on the manner in which the cells were lysed and prepared for SDS-PAGE. Cleavage was present even when the cells were lysed using SDS to denature potential proteases. Furthermore, cleavage was strongly temperature dependent, with the lowest level attained during expression at 18 °C. To reduce the occurrence of the undesired *in vivo* proteolytic cleavage and the resulting reduction in cMBP-SelK yield, we tested the effect of shortening the linker between the two proteins. The original long linker, NSSSNNNNNNNNLG (SelK-D), was changed to a shorter tetrapeptide linker SNNN (SelK-E). As Figure 2.3 and subsequent figures demonstrate, we found that the resulting fusion proteins, with short and long linkers respectively, behaved identically during purification. Furthermore, the TEV cleavage site was equally accessible to the TEV protease. Nevertheless, the shorter linker (SelK-E) did reduce the extent of in *vivo* proteolytic cleavage significantly and was thus used exclusively. Interestingly, additional affinity tags introduced before the TEV cleavage site to assist in the purification did not increase the in vivo proteolytic cleavage rate (Figure 2.2 SelK-F and SelK-G). All data shown in subsequent figures was obtained using SelK-E, since it introduces only one non-native amino acid (glycine) in SelK's sequence and hence is least likely to perturb its function and structure.

61



Figure 2.3. Proteolytic cleavage of cMBP-SelK fusion in *E. coli* depends on the length of the linker between MBP and SelK and the expression temperature. (A) Protein expression at 18 °C of cMBP-SelK fusion with short (NSSS) and long (NSSSNNNNNNNNNNCG) linker length. Lanes 1 to 4: SelK with a short linker at 0, 4, 8, and 22 h post-induction. Lanes 5 to 8: SelK with a long linker at 0, 4, 8, and 22 h post-induction. M: Protein molecular weights standard: 20, 25, 30, 40, 50, 60, 80, 100, 150, 250 kDa where the 25 and 80 kDa bands triple the intensity. The arrow points to the cMBP-SelK fusion at 54 kDa and cMBP at 42 kDa. (B) cMBP-SelK with short and long linker induced at 37 °C. Lanes 1 to 6: cMBP-SelK short linker 0, 2, 4, 6, 8, and 22 h post-induction. Lanes 7 to 12: long linker 0, 2, 4, 6, 8, and 22 h post-induction. M: protein molecular weights standard as listed above. The arrow points to the cMBP-SelK fusion at 53 kDa and cMBP at 42 kDa

## 2.2.3 Detergent Screening

Membrane proteins are difficult study targets because invariably they require a specialized environment in which to maintain their structural and functional integrity. Key to any functional characterization of SelK is that the protein is maintained in a stable form and that aggregation is prevented by selecting the most suitable detergent, ionic strength, and stabilizing agents (i.e. glycerol, lipids, etc.). Detergents affect membrane proteins' conformation, enzymatic activity, and the ability to properly interact with affiliated proteins. The choice of detergent may vary between different applications: Crystallization requires detergents that do not completely cover the protein area, allowing it to form specific protein-protein contacts<sup>32</sup>. Nuclear magnetic resonance (NMR) spectroscopy benefits from detergents in which the spectra are well

resolved<sup>33-35</sup>. Enzymatic activity, however, is often favored in nonionic detergents, as opposed to zwitterionic detergents<sup>36</sup>, and certainly compared to ionic detergents. Thus, it is essential to characterize SelK in a selection of detergents. Accordingly, we surveyed the efficiency of proteolytic cleavage of SelK from its fusion partner, cMBP, by TEV protease in the presence of different detergents. They include several mild and several non-denaturing detergents: 2,2-didecylpropane-1, 3-bis- $\beta$ -D-maltopyranoside (MNG-3), DDM, 3-[(3-cholamidopropyl) dimethylammonio]-1- propanesulfonate (CHAPS), 1-myristoyl-2-hydroxy-sn-glycero-3-phosphocholine (LMPC), and Triton X-100 (Figure 2.4, all at concentrations of 4xCMC). These detergents are frequently used in functional studies of membrane proteins, due their ability to preserve the native structure and sustain function<sup>37</sup>. TEV has been previously shown to be active in these detergents<sup>38, 39</sup>. Successful cleavage by TEV protease was observed under all the conditions tested. SelK can also be cleaved in the absence of detergents but aggregates subsequently. Similarly, we found that the protein gradually aggregated in the presence of CHAPS and MNG-3.



Figure 2.4. Cleavage of SelK from its fusion partner cMBP by TEV protease in the presence of different detergents. Lane 1: TEV protease prior to incubation with cMBP-SelK. Lane 2: cMBP-SelK prior to incubation with TEV protease. Lane 3: Cleavage without detergent. Lane 4: Cleavage in the presence of 40  $\mu$ M MNG-3. Lane 5: Cleavage in the presence of 0.48 mM DDM. Lane 6: Cleavage in the presence of 32 mM CHAPS. Lane 7: Cleavage in the presence of 0.144 mM LMPC. Lane 8: Cleavage in the presence of 0.06% Triton X-100. M: Protein molecular weights standard (the molecular mass in kDa is noted on the right). The arrow points to SelK at 10.6 kDa.

## 2.2.4 Purification of SelK

Usually purification strategies for cMBP fused to transmembrane proteins do not utilize detergents during extraction from cells, since the fusion protein is typically soluble<sup>28</sup>. MBP-SelK is soluble and localized to the cytoplasm but it aggregated slowly in the absence of detergents. This may be due to the unusual nature of SelK's

transmembrane helix, which has several ionizible residues (Figure 2.1). To purify MBP-SelK, it was necessary to extract it in the presence of Triton X-100 or DDM (other detergents were not tested). The protein was then purified using amylose affinity chromatography. The detergent used for extraction was exchanged to the target detergent using IMAC, and SelK was cleaved from its partner cMBP by the TEV protease. Following cleavage, SelK was separated from the hexahistidine-tagged TEV and cMBP by rebinding them to IMAC beads. The final purity of the protein was above 95%. An example of a successful expression of the SelK in *E. coli* and subsequent purification is shown in Figure 2.5A. In several purifications, two minor truncation products, in addition to the desired full-length protein, were detected. This cleavage is most likely taking place in vivo between positively charged residues, sites that are prone to proteolytic cleavage in bacterial systems.

The amino acid sequence of the 53 kDa fusion protein band was verified by LC MS/MS mass spectrometry analysis to be that of the full length SelK. LC-ESI-TOF mass spectrometry was employed to confirm the expected molecular weight for intact cMBP-SelK and SelK (Figure 2.5B). The yield for the 53 kDa fusion protein is  $\sim$ 50 mg per liter of culture. The yield of the purified, 10.6 kDa SelK is  $\sim$ 10 mg per liter of culture ( $\sim$ 1 µmol).



Figure 2.5. Expression and purification of SelK-E (A) SDS-PAGE analysis. Lane 1: Crude cell extract prior to induction. Lane 2: Crude cell extract after overnight expression at 18 °C. Lane 3: Amylose column elute. Lane 4: Following exchange of detergent from 0.1% Triton X-100 to 0.067% DDM by IMAC. Lane 5: TEV protease cleavage mixture. Lane 6: SelK after the His6-tagged TEV and MBP were removed by IMAC. The arrow points to SelK. M: Protein molecular weights standard (the molecular mass in kDa is noted on the right). (B) The molecular weight of SelK detected by mass spectroscopy is 10654 Da (predicted molecular weight is 10655 Da).

## 2.2.5 DDM-SelK Micelle Size

DDM is frequently used in structural characterization of membrane proteins<sup>32,</sup> <sup>33</sup> and is reported here to successfully solubilize and stabilize SelK. To better evaluate its suitability for future structural studies, the molecular weight of the micelle-protein complex was measured using size exclusion chromatography. We also compared the behavior of the protein in buffer, with or without the reducing agent DTT (Figure 2.6A), to remove potential intermolecular disulfide bonds. The micelle-protein complex elutes at the same location in the presence or absence of DTT. The molecular weight of the complex is 91 kDa (Figure 2.6B). This value is in agreement with previous reports of the DDM detergent micelle size falling between 58 kDa to 86 kDa<sup>40</sup>. It is not possible to evaluate from the elution profile whether SelK elutes as a monomer or dimer since SelK is small (the molecular weight is 10655 Da). Hence, the presence of one or two copies of SelK per micelle may not cause a noticeable shift in the elution profile of the size exclusion chromatography.

To further facilitate detergent exchange, we generated a construct of SelK in which a streptavidin binding tag (WSHPQFEK, StrepII tag) was inserted after the TEV cleavage site. The SelK retains the Strep tag after cleavage from the fusion partner, which allows its specific binding to a Strep-tactin column and exchange of detergents while bound to the column. The Strep-tagged SelK (SelK-G) behaved identically to SelK-E during purification and had a similar circular dichroism (CD) spectrum (discussed in the next section).



Figure 2.6. Characterization of the SelK-DDM complex by size exclusion chromatography. (A) SelK DDM complex elution profile on Superdex 200 in the presence (solid line) and absence of the reducing agent DTT (dotted line). (B) The molecular weight of the SelK DDM complex was determined by measuring its elution volume relative to standard proteins. The black triangle overlaid on the calibration curve corresponds to the obtained molecular weight of 91 kDa for the SelK DDM complex.

## 2.2.6 SelK Secondary Structure

SelK sequence is abundant in Arg, Gly, Pro, and Ser. This pattern is indicative of intrinsically disordered proteins, which are abundant in hydrophilic amino acids (Arg, Gly, Gln, Ser, Pro, Glu, and Lys)<sup>41</sup>. Proline-rich regions can also form left-handed polyproline II helices (PPII)<sup>42</sup> that have CD spectroscopy absorptions near 220 nm<sup>43</sup>. To probe whether SelK possesses a secondary structure, we employed CD spectroscopy. SelK CD spectrum was acquired in three detergents—DDM,  $\beta$ -OG, and LMPC–as they are compatible with future structural characterization and have considerably different properties<sup>33, 44</sup>. The detergent content was exchanged from DDM to the detergent of choice, using size exclusion as described above. As can be seen in Figure 2.7, the presence of a positive band at 195 nm and negative bands at 208 and 222 nm indicates that SelK possesses an  $\alpha$ -helix secondary structure in the three detergents. The  $\alpha$ -helix signature is most pronounced in DDM micelles. In the presence of SDS the CD spectrum has a reduction in helical structure and an increase in random coil.



Figure 2.7. SelK posses a  $\alpha$ -helical secondary structure in various detergent micelles. The helical structure is reduced in the presence of SDS. CD spectra of SelK in the presence of 0.24 mM DDM (circles), 40 mM  $\beta$ -OG (squares), 0.072 mM LMPC (triangles) and 25 mM SDS (solid line).

## 2.2.7 Overexpression and Purification of SelS

In order to test whether the purification strategy based on extraction of MBP fusions in the presence of detergents might be generally applicable to other members of the SelK/SelS family we have employed the procedures described for preparation of SelK on its protein partner SelS. Figure 2.8 demonstrates the successful expression and purification of *Homo sapiens* SelS U188C using the procedure described above for SelK. The full length protein elutes as a dimer in size exclusion chromatography

(data not shown) in agreement with reports on dimerization of its soluble portion<sup>23</sup>. In addition to the transmembrane segment SelS has an extended coiled coil region as well as a disordered C-terminal domain. Hence, the CD spectrum has a signal that is mostly  $\alpha$ -helical. The spectrum is similar to that recently published for the soluble portion but the ratio between the two negative bands at 208 and 202 nm suggests a higher degree of  $\alpha$ -helix content. SelS was equally well behaved in DDM micelles as SelK.



Figure 2.8. Expression and purification of SelS. (A) Schematic representations of SelS expression construct. His-tag, hexahistidines sequence; cMBP, cytoplasmic maltosebinding protein; short linker, NSSS; TEV, tobacco etch virus protease cleavage site (ENLYFQ/S); SelS U188C, *Homo sapiens* SelS with a U188C mutation at the active site. (B) SDS-PAGE analysis. Lane 1: Crude cell extract prior to induction. Lane 2: Crude cell extract after overnight expression at 18 °C. Lane 3: Amylose column elute with buffer containing DDM detergent. Lane 4: TEV protease cleavage mixture. Lane 5: SelS after the His6-tagged TEV and MBP were removed by IMAC. The arrow points to SelS dimer and monomer (from top to bottom). M: Protein molecular weights standard (the molecular mass in kDa is noted on the right). (C) CD spectrum of SelS in the presence of 0.24 mM DDM.

## 2.3 Discussion

This paper presents an expression and purification strategy for human SelK in *E. coli*. We find that even though SelK is rich in Gly, Pro, and charged residues, it adopts an  $\alpha$ -helical secondary structure. This supports a view of SelK not as an intrinsically disordered protein, but rather as possessing a well-defined 3D architecture. This would agree with the presence of two SH3-binding sequences that typically have a specific geometry. Evidence of an  $\alpha$ -helical secondary structure further supports the feasibility of structural characterization by X-ray crystallography or NMR spectroscopy. Our studies indicate that SelK can be stabilized in a variety of detergents and acquires an  $\alpha$ -helical secondary structure in all the detergents tested in this study. The flexibility to explore structural studies in different detergents should facilitate its characterization. We are currently in the process of exploring sample conditions for solution-state NMR studies.

A second finding reported here is that a similar expression and purification strategy is successful in purifying human SelS. Both enzymes belong to the same family of eukaryotic membrane proteins but differ substantially in length and in the polar content of their transmembrane helices. Nevertheless, the two proteins could be purified with same procedure and were stable in the presence of DDM. We suggest that it is possible that the purification strategy based on extraction of MBP fusions in the presence of detergents might be generally applicable to other members of this redox-related membrane protein family.

#### 2.4 Materials and Methods

## 2.4.1 Bacterial Strains, Plasmids, and Chemical Reagents

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 (cMBP)<sup>45</sup> was purchased from the Protein Structure Initiative: Biology Materials Repository<sup>46</sup>. Chromatography media was supplied by GE Healthcare Bio-Sciences Corporation (Pittsburgh, PA), New England Biolabs, and Qiagen (Hilden, Germany). Detergents were acquired from Affymetrix (Santa Clara, CA). All other chemicals and reagents were supplied by Sigma-Aldrich (St. Louis, MO), Acros Organics (Geel, Belgium), and GoldBio (St. Louis, MO). All reagents and solvents were at least analytical grade and were used as supplied.

#### 2.4.2 SelK Expression Vectors

*Homo sapiens* SelK gene (GenBank accession no. BC013162.1) was codon optimized for expression in *E. coli* and the gene synthesized by DNA2.0 (Menlo Park, CA). Several SelK constructs were generated by cloning the codon-optimized SelK into various expression vectors. The gene was cloned into a pProExHTa expression vector (Life Technologies, NY) with an N-terminal hexahistidine-tag using SfoI and EcoRI, a pET28a (Agilent, Santa Clara, CA) as a fusion with the small ubiquitin-like modifier (SUMO) protein using BamHI and HindIII, and into pMAL-p4X (New England Biolabs) as a fusion with the periplasmic maltose binding protein (pMBP) using SacI and BamHI. pMAL-C5X (New England Biolabs) was employed to generate a series of fusions of cMBP with SelK with different linking sequence and affinity tags (Figure 2.2). In all constructs, we have introduced a hexahistidine tag between I3 and E4 of cMBP. SacI and BamHI were used to generate a construct with a 16 amino acid linker, NSSSNNNNNNNNNNNNLG, between cMBP and SelK. AvaI and BamHI were used to generate the construct with the short linker, NSSS, between the two proteins. All constructs included a TEV protease cleavage site between SelK and the fusion tag as well as a U92C mutation in SelK. Following cleavage with TEV protease, a Gly was present before the first native amino acid in all the constructs. All the clones were verified by DNA sequencing.

## 2.4.3 SelS Expression Vectors

*Homo sapiens* SelS gene (GenBank accession no. GI: 45439348) was codon optimized for expression in *E. coli* and the gene synthesized by genescript (Piscataway, NJ). The gene was cloned into a pMAL-C5X (New England Biolabs) as a fusion with the cytoplasmic maltose binding protein (cMBP). A hexahistidine tag was introduced between residues I3 and E4 of cMBP. A short linker NSSS and a TEV cleavage site, ENLYFQS, was used to connect the two proteins. Following cleavage with TEV protease, a Ser was present before the first native amino acid. The protein contained a U188C mutation.

## 2.4.4 Protein Expression Screening

For protein expression, the plasmids were transformed into *E. coli* BL21(DE3) strain. Cells were grown in LB at 37 °C, with good aeration and the relevant antibiotic selection. When the optical density (OD) at 600 nm reached 0.5, the temperature was lowered to 18 °C, and the cells were allowed to shake at the lower temperature for an additional hour. Protein expression was induced with 0.5 mM isopropyl-1-thio- $\beta$ -D-galactopyranoside (IPTG) when the OD at 600 nm reached 0.7. Time points were

taken periodically up to 24 h of expression and protein expression was visualized by SDS-PAGE and western blots. The identity of the fusion proteins was verified by western blots using anti-MBP monoclonal (New England Biolabs) or His-probe (Santa Cruz Biotechnology, Santa Cruz, CA) antibodies.

## 2.4.5 Expression and Purification of SelK From a cMBP-SelK Fusion

Cells were grown in LB, supplemented with 0.2% glucose at 37 °C, with good aeration and an ampicillin selection of 100 µg/mL. The temperature was changed to 18 °C about half an hour prior to induction. Protein expression was induced with 0.5 mM IPTG for 16 h when the OD reached 0.7. The cells were harvested, and the cell paste (6 g/L) was resuspended in 50 mM sodium phosphate, 200 mM sodium chloride, 0.1% Triton X-100, pH 7.5 (amylose buffer), supplemented with 0.5 mM benziamidine, 1 mM phenylmethylsulfonyl fluoride (PMSF), and 1 mM EDTA. Cells were lysed using a high-pressure homogenizer (EmulsiFlex-C5, Avestin, Ottawa, Canada). Cell debris was removed by centrifugation at 20,000 g for 1 h. The supernatant was loaded on an amylose column, and the column was washed with the amylose buffer. The fusion cMBP-SelK was eluted using amylose buffer containing 20 mM maltose. The purity of the eluted cMBP-SelK was about 80%, and the most abundant contamination was cMBP. The protein was then loaded onto an immobilized metal affinity chromatography (IMAC) HisTrap FF column. The column was washed with 50 mM sodium phosphate, 200 mM NaCl, 0.067% n-dodecyl-β-D-maltopyranoside (DDM), and pH 7.5 (TEV cleavage buffer). cMBP-SelK was then eluted with an imidazole linear gradient from 0-1 M imidazole in cleavage buffer. Fractions containing the cMBP-SelK fusion protein were combined and the buffer exchanged to the TEV cleavage buffer by dialysis. Cleavage of the fusion partner cMBP was carried out by

adding hexahistidine-tagged TEV protease to the dialysis bag, at 4 °C for 12 h<sup>45, 47</sup>. The TEV protease was added at a molar ratio of 1:2 relative to the fusion protein, but concentrations as low as 1:10 were also successful. Following cleavage, the mixture was loaded again to a HisTrap FF column to remove the hexahistidine tagged cMBP and TEV. The flowthrough containing the purified SelK was concentrated to 5 mg/L and stored at 4 °C. Protein purity, as determined by 15% SDS-PAGE Tris-glycine and 16% Tris-tricine gels, was higher than 95%. Protein concentration was determined using an extinction coefficient of 15470 M<sup>-1</sup>cm<sup>-1</sup> at 280 nm.

Size exclusion chromatography was used to exchange detergents. Approximately 200 µL of 2 mg/mL SelK was loaded on a Superdex 200 10/30 GL column (GE Healthcare) and eluted at 0.4 mL/min with 50 mM sodium phosphate, 200 mM sodium chloride, pH 7.5, and the detergent of choice. The column was calibrated using GE Healthcare Bio-Sciences gel filtration protein standards: aldolase (158 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa). The void volume was measured using blue dextran 2000.

Expression and purification of SelS U188C based on a cMBP-SelS fusion was identical to the procedure above. The protein behaves as a dimer on superdex 200 10/30 GL column (data not shown).

## 2.4.6 Mass Spectroscopy

Mass spectrometry spectra were obtained using a QTOF Ultima (Waters, MA), operating under positive electrospray ionization (+ESI) mode, 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% with 0.1% TFA as the mobile phase, in 25 min, at a flow rate of 0.2 mL/min at room temperature. Data were acquired from m/z 350 to 3,000, at a rate of 1 sec/scan.

## 2.4.7 Circular Dichroism Spectroscopy

CD spectra of SelK and SelS were measured by using a J-810 circular dichroism spectropolarimeter (Jasco, Essex, UK) that had been calibrated using camphor sulfonic acid for optical rotation and benzene vapor for wavelength. Far-UV spectra were recorded using a 1 mm path-length cell for the 190-250 nm region at 20 °C. Samples for CD spectroscopy were prepared in 10 mM potassium phosphate, 50 mM Na<sub>2</sub>SO<sub>4</sub>, pH 7.5, and either 0.24 mM DDM, 40 mM  $\beta$ -OG, 25 mM SDS, or 0.072 mM LMPC. Three accumulation scans were collected for baseline, and eight accumulation scans were taken for each sample.

## REFERENCES

- 1. Lobanov, A.V., Hatfield, D.L. & Gladyshev, V.N. Eukaryotic selenoproteins and selenoproteomes. *Biochim. Biophys. Acta* **1790**, 1424-1428 (2009).
- 2. Lu, J. & Holmgren, A. Selenoproteins. J. Biol. Chem. 284, 723-727 (2009).
- Kryukov, G. V., Castellano, S., Novoselov, S. V., Lobanov, A. V., Zehtab, O., Guigo, R. & Gladyshev, V. N. Characterization of mammalian selenoproteomes. *Science* 300, 1439-1443 (2003).
- Morozova, N., Forry, E. P., Shahid, E., Zavacki, A. M., Harney, J. W., Kraytsberg, Y. & Berry, M. J. Antioxidant function of a novel selenoprotein in Drosophila melanogaster. *Genes Cells* 8, 963-971 (2003).
- Lu, C. L., Qiu, F. C., Zhou, H. J., Peng, Y., Hao, W., Xu, J. L., Yuan, J. G., Wang, S. Z., Qiang, B. Q., Xu, C. M. & Peng, X. Z. Identification and characterization of selenoprotein K: An antioxidant in cardiomyocytes. *FEBS Lett.* 580, 5189-5197 (2006).
- Du, S.Q., Zhou, J., Jia, Y. & Huang, K.X. SelK is a novel ER stress-regulated protein and protects HepG2 cells from ER stress agent-induced apoptosis. *Arch. Biochem. Biophys.* 502, 137-143 (2010).
- Lai, C.Q., Parnell, L.D., Lyman, R.F., Ordovas, J.A. & Mackay, T.F.C.
   Candidate genes affecting Drosophila life span identified by integrating

microarray gene expression analysis and QTL mapping. *Mech. Ageing Dev.* **128**, 237-249 (2007).

- Castellano, S., Lobanov, A. V., Chapple, C., Novoselov, S. V., Albrecht, M., Hua, D. M., Lescure, A., Lengauer, T., Krol, A., Gladyshev, V. N. & Guigo, R. Diversity and functional plasticity of eukaryotic selenoproteins: Identification and characterization of the SelJ family. *Proc. Natl. Acad. Sci. USA* 102, 16188-16193 (2005).
- Verma, S., Hoffmann, F. W., Kumar, M., Huang, Z., Roe, K., Nguyen-Wu, E., Hashimoto, A. S. & Hoffmann, P. R. Selenoprotein K knockout mice exhibit deficient calcium flux in immune cells and impaired immune responses. *J. Immunol.* 186, 2127-2137 (2011).
- Zhang, Y., Zhou, Y., Schweizer, U., Savaskan, N. E., Hua, D., Kipnis, J., Hatfield, D. L. & Gladyshev, V. N. Comparative analysis of selenocysteine machinery and selenoproteome gene expression in mouse brain identifies neurons as key functional sites of selenium in mammals. *J. Biol. Chem.* 283, 2427-2438 (2008).
- Smith, M.H., Ploegh, H.L. & Weissman, J.S. Road to ruin: Targeting proteins for degradation in the endoplasmic reticulum. *Science* 334, 1086-1090 (2011).
- Shchedrina, V. A., Everley, R. A., Zhang, Y., Gygi, S. P., Hatfield, D. L. & Gladyshev, V. N. Selenoprotein K binds multiprotein complexes and is involved in the regulation of endoplasmic reticulum homeostasis. *J. Biol. Chem.* 286, 42937-42948 (2011).
- Huang, Z., Hoffmann, F.W., Norton, R.L., Hashimoto, A.C. & Hoffmann, P.R. Selenoprotein K Is a novel target of m-calpain, and cleavage is regulated by toll-like receptor-induced calpastatin in macrophages. *J. Biol. Chem.* 286, 34830-34838 (2011).
- Ben, S. B., Wang, Q. Y., Xia, L., Xia, J. Z., Cui, J., Wang, J., Yang, F., Bai,
  H., Shim, M. S., Lee, B. J., Sun, L. G. & Chen, C. L. Selenoprotein dSelK in
  Drosophila elevates release of Ca2+ from endoplasmic reticulum by
  upregulating expression of inositol 1,4,5-tris-phosphate receptor. *Biochemistry- Moscow* 76, 1030-1036 (2011).
- Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E.L.L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* 305, 567-580 (2001).
- Rath, A. & Deber, C.A. Surface recognition elements of membrane protein oligomerization. *Proteins-Structure Function and Bioinformatics* 70, 786-793 (2008).
- Walters, R.F.S. & DeGrado, W.F. Helix-packing motifs in membrane proteins.
   *Proc. Natl. Acad. Sci. USA* 103, 13658-13663 (2006).
- Li, E., Wimley, W.C. & Hristova, K. Transmembrane helix dimerization: Beyond the search for sequence motifs. *Biochim. Biophys. Acta, Biomembr.* 1818, 183-193 (2012).
- Chen, C. L., Shim, M. S., Chung, J., Yoo, H. S., Ha, J. M., Kim, J. Y., Choi, J.,
   Zang, S. L., Hou, X., Carlson, B. A., Hatfield, D. L. & Lee, B. J. G-rich, a

Drosophila selenoprotein, is a Golgi-resident type III membrane protein. *Biochem. Biophys. Res. Commun.* **348**, 1296-1301 (2006).

- Shchedrina, V.A., Zhang, Y., Labunskyy, V.M., Hatfield, D.L. & Gladyshev, V.N. Structure-function relations, physiological roles, and evolution of mammalian ER-resident selenoproteins. *Antioxid. Redox Signal.* 12, 839-849 (2010).
- Li, S.S.C. Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. *Biochem. J* 390, 641-653 (2005).
- Carducci, M., Perfetto, L., Briganti, L., Paoluzi, S., Costa, S., Zerweck, J., Schutkowski, M., Castagnoli, L. & Cesareni, G. The protein interaction network mediated by human SH3 domains. *Biotechnol. Adv.* 30, 4-15 (2012).
- Christensen, N.W.J., Andrea Vala1,§, Jurate Kamarauskaite1, Linda & Johansson, J.R.W., Kay Hofmann2, Kaare Teilum1 and Lars Ellgaard The human selenoprotein VCP-interacting membrane protein (VIMP) is nonglobular and harbors a reductase function in an intrinsically disordered region. *J. Biol. Chem.* 287, 26388-26399 (2012).
- Johansson, L., Gafvelin, G. & Arner, E.S.J. Selenocysteine in proteins properties and biotechnological use. *Biochim. Biophys. Acta* 1726, 1-13 (2005).

- Junge, F., Schneider, B., Reckel, S., Schwarz, D., Dotsch, V. & Bernhard, F. Large-scale production of functional membrane proteins. *Cell. Mol. Life Sci.* 65, 1729-1755 (2008).
- 26. Esposito, D. & Chatterjee, D.K. Enhancement of soluble protein expression through the use of fusion tags. *Curr. Opin. Biotechnol.* **17**, 353-358 (2006).
- 27. Sorensen, H.P. & Mortensen, K.K. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microb. Cell Fact.* **4** (2005).
- Hu, J., Qin, H.J., Gao, F.P. & Cross, T.A. A systematic assessment of mature MBP in membrane protein production: Overexpression, membrane targeting and purification. *Protein Expr. Purif.* 80, 34-40 (2011).
- Marblestone, J. G., Edavettal, S. C., Lim, Y., Lim, P., Zuo, X. & Butt, T. R. Comparison of SUMO fusion technology with traditional gene fusion systems: Enhanced expression and solubility with SUMO. *Protein Sci.* 15, 182-189 (2006).
- Wessjohann, L.A., Schneider, A., Abbas, M. & Brandt, W. Selenium in chemistry and biochemistry in comparison to sulfur. *Biol. Chem.* 388, 997-1006 (2007).
- Sachdev, D. & Chirgwin, J.M. Fusions to maltose-binding protein: Control of folding and solubility in protein purification. *Methods Enzymol.* 326, 312-321 (2000).
- Newstead, S., Ferrandon, S. & Iwata, S. Rationalizing alpha-helical membrane protein crystallization. *Protein Sci.* 17, 466-472 (2008).

- 33. Kim, H.J., Howell, S.C., Van Horn, W.D., Jeon, Y.H. & Sanders, C.R. Recent advances in the application of solution NMR spectroscopy to multi-span integral membrane proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* 55, 335-360 (2009).
- Tamm, L.K. & Liang, B.Y. NMR of membrane proteins in solution. *Prog. Nucl. Magn. Reson. Spectrosc.* 48, 201-210 (2006).
- Sanders, C.R. & Sonnichsen, F. Solution NMR of membrane proteins: practice and challenges. *Magn. Reson. Chem.* 44, S24-S40 (2006).
- 36. le Maire, M., Champeil, P. & Moller, J.V. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim. Biophys. Acta, Biomembr.* 1508, 86-111 (2000).
- Prive, G.G. Detergents for the stabilization and crystallization of membrane proteins. *Methods* 41, 388-397 (2007).
- Lundback, A.-K., van den Berg, S., Hebert, H., Berglund, H. & Eshaghi, S.
   Exploring the activity of tobacco etch virus protease in detergent solutions.
   *Anal. Biochem.* 382, 69-71 (2008).
- Mohanty, A.K., Simmons, C.R. & Wiener, M.C. Inhibition of tobacco etch virus protease activity by detergents. *Protein Expr. Purif.* 27, 109-114 (2003).
- Strop, P. & Brunger, A.T. Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. *Protein Sci.* 14, 2207-2211 (2005).

- Uversky, V.N. Intrinsically disordered proteins from A to Z. Int. J. Biochem. Cell Biol. 43, 1090-1103 (2011).
- Stapley, B.J. & Creamer, T.P. A survey of left-handed polyproline II helices.
   *Protein Sci.* 8, 587-595 (1999).
- 43. Bochicchio, B. & Tamburro, A.M. Polyproline II structure in proteins: Identification by chiroptical spectroscopies, stability, and functions. *Chirality* 14, 782-792 (2002).
- 44. Poget, S.F. & Girvin, M.E. Solution NMR of membrane proteins in bilayer mimics: Small is beautiful, but sometimes bigger is better. *Biochim. Biophys. Acta, Biomembr.* 1768, 3098-3106 (2007).
- Blommel, P.G. & Fox, B.G. A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expr. Purif.* 55, 53-68 (2007).
- 46. Cormier, C. Y., Mohr, S. E., Zuo, D. M., Hu, Y. H., Rolfs, A., Kramer, J., Taycher, E., Kelley, F., Fiacco, M., Turnbull, G. and LaBaer, J. Protein Structure Initiative Material Repository: an open shared public resource of structural genomics plasmids for the biological community. *Nucleic Acids Res.* 38, D743-D749 (2010).
- Kapust, R. B., Tozser, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D. & Waugh, D. S. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.* 14, 993-1000 (2001).

### Chapter 3

# THE INTRINSICALLY DISORDERED MEMBRANE PROTEIN SELENOPROTEIN S IS A REDUCTASE

#### 3.1 Introduction

Selenoprotein S (SelS) belongs to a family of human enzymes that contains the genetically encoded amino acid selenocysteine (Sec)<sup>1, 2</sup>. Selenoproteins typically act as oxidoreductases in redox regulation and the management of oxidative species and SelS belong to a eukaryotic protein family that is related to redox stress<sup>3</sup>. Accumulating evidence suggests that it is involved in inflammation and management of oxidative stress<sup>4, 5</sup>. It was identified as a member of the ER-assisted protein degradation (ERAD) machinery<sup>6, 7</sup> - a pathway responsible for transporting misfolded proteins from the ER to the cytoplasm for degradation by the proteasome<sup>8</sup>.

This work focuses on the enzymatic function of human SelS (also known by the alternative names SEPS1, Tanis, VIMP and SELENOS) is a single-pass transmembrane protein<sup>9</sup> with a short segment in the ER lumen and an extended cytoplasmic region<sup>10</sup>. The cytoplasmic segment contains a disordered segment (residues 123-189 out of 189) that includes the Sec at position 188. The Sec forms a selenylsulfide bond with a nearby Cys (Cys174)<sup>11</sup>. SelS was shown to dimerize through a coiled-coil region and contains a valosin-containing protein (VCP, p97) interacting motif<sup>11, 12</sup>. It mediates the interactions of the ERAD component Derlin-1 with p97, an AAA ATPase that pulls the protein targets to the cytoplasm, where they are broken down by the proteasome<sup>13</sup>. In addition to p97, SelS also interacts with

Derlin 1 and 2, components of the putative ERAD channel<sup>6</sup>; selenoprotein K (SelK), a Sec-containing enzyme with unknown function<sup>3</sup>; the VCP accessory protein ubiquitin conjugation factor E4 A (UBE4A), a contributor to multiubiquitin chain extension<sup>14</sup>; the UBX domain–containing protein 8 (UBXD8), a sensor for unsaturated fatty acids and regulator of triglyceride synthesis<sup>15</sup> that together with derlin-1 controls degradation of lipidated apolipoprotein B-100<sup>16</sup>; the UBX domain–containing protein 6 (UBXD6, Rep), which tethers p97 to the ER membrane<sup>17</sup> and the kelch containing protein 2 (KLHDC2, HCLP-1), a regulator of the LZIP transcription factor<sup>18</sup>. As mentioned above, these interactions take place in the cytoplasm, in the context of the ERAD complex. However, the actual function of SelS in the ERAD machinery and its relevance to management of oxidative stress is unknown.

Previously, the cytosolic segment of a SelS U188C mutant (cSelS U188C) was characterized by NMR spectroscopy and biochemical methods<sup>11</sup>. It was found that SelS' C-terminal domain is unstructured, classifying it as an intrinsically disordered protein. This is a class of proteins that are rich in charged and polar residues and typically adopt a stable tertiary structure only in the presence of their protein partner<sup>19</sup>. The enzymatic activity of cSelS U188C, however, was not characterized. This information is necessary to tie together not only the role of SelS in the ERAD machinery but also its strong relation to redox stress. Here, we set out to characterize *in vitro* the enzymatic activity of human SelS that contains its native Sec.

Characterization of the Sec-containing form requires incorporation of Sec at position 188. However, the specific incorporation of Sec into selenoproteins is typically the bottleneck in their characterization. Selenoproteins are produced *in vivo* by synthesizing Sec on its dedicated tRNA. The Sec codon, which is shared with the

opal stop codon, UGA, is recognized by the dual action of a mRNA loop termed the SEC Insertion Sequence (SECIS) and ancillary proteins<sup>20</sup>. A major challenge for efficient production of selenoproteins is that protein translation is prematurely terminated at the UGA codon due to its misreading as a stop codon. Hence, heterologous expression contains a high ratio of truncated protein. For SelS, the separation of the truncated and full-length Sec-containing forms, which differ by only two amino acids, is complicated by the strong dimerization of truncated and full-length, Sec-containing form, we developed an expression system that allows high yield preparation of both the full-length and cytoplasmic portions of SelS (cSelS)<sup>21</sup>. *E. coli* selenium expression machinery was employed to insert a genetically encoded Sec at the active site. We have purified the full-length cSelS by relying on the tendency of the truncated protein to aggregate in the absence of reducing agents. With a sample enriched with the native Sec-containing protein at hand, we demonstrate that only the Sec-containing enzyme has reductase and peroxidase activities.

### 3.2 Results

### 3.2.1 Incorporation of Sec into SelS

The incorporation of Sec in proteins at a unique position remains an experimental challenge and a barrier for their characterization<sup>22</sup>. To develop preparation strategies for the Sec-containing cSelS, we have first optimized strategies for obtaining high yield and purity of both the full length and the cytoplasmic portion of SelS U188C in *E. coli*<sup>23</sup>. We focus on the characterization of the cytoplasmic portion of SelS (residues 52-189), abbreviated as cSelS. cSelS was fused to a binding

partner, cytoplasmic maltose binding protein (cMBP), and expressed at 18 °C, to aid solubility and folding. Using this strategy, it is possible to prepare cSelS U188C with over 98% purity (Figure 3.1A) at a yield of 10 mg/L (0.6 mmol). The expected molecular weight was confirmed by electrospray mass spectrometry (Figure 3.1B). cSelS U188C dimerized in the presence and absence of reducing agents, even though the transmembrane region was deleted (Figure 3.2). Hence, the coiled-coil region is sufficient for dimerization, as was previously shown<sup>11</sup>.

With the high-yield expression and purification strategy at hand, the production of the native Sec-containing cSelS was undertaken by utilizing E. coli's innate Sec incorporation machinery<sup>24</sup>. Incorporation of Sec into SelS in a genetically encoded fashion required numerous modifications to the expression vector and growth conditions. E. coli makes use of an mRNA structure termed the SEC Insertion Sequence (SECIS). This stem-loop architecture coordinates the recruitments to the ribosome proteins essential for Sec synthesis on its dedicated tRNA. In E. coli, those include: SelB, an elongation factor; SelC, the Sec tRNA; SelA, a selenocysteine synthase and SelD, a selenophosphate synthase. Arnér and Böck demonstrated that it is possible to introduce Sec into recombinant proteins in E. coli by utilizing the SECIS element borrowed from its formate dehydrogenase H and coexpressing SelA, B and  $C^{24}$ . We have introduced the UGA codon at position 188, as well as *E. coli* formate dehydrogenase H SECIS element past the stop codon of the cMBP-cSelS coding DNA. The pSUABC vector expressing SelA, B and C under the control of their own natural promoter was a gift from Prof. Arnér. The incorporation of Sec is further augmented by inducing expression at a late exponential phase, in which the level of

88

release factor 2 is decreased <sup>25</sup>. The incorporation ratio of Sec was about 3% of the total expressed protein.

We have tested several methods to separate the full-length, seleniumcontaining protein from the truncated enzyme cSelS 188 $\Delta$  (188 $\Delta$  indicates that residues 188 and 189 are missing due to misreading of the Sec codon as a stop codon). The separation of cSelS and cSelS 188∆ is challenging since they are identical for all but two amino acids and form a dimer. To obtain a sample rich in cSelS, we have relied on the tendency of the cSelS 188 $\Delta$ , which has only one cysteine, to aggregate in the absence of reducing agents. We have extracted and purified the protein expressed with a UGA codon using the procedures optimized for the cSelS U188C mutant but deliberately excluding reducing agents. The truncated enzyme, cSelS  $188\Delta$ , aggregated during purification, due to the formation of intermolecular disulfide bonds. The last step relied on size exclusion chromatography to separate the tetramer and higher oligomeric forms of SelS from the dimeric Sec-containing forms. Subsequently, only Sec-rich dimeric fractions were retained. Using this approach, samples were enriched from a starting ratio of 3% cSelS to about 50%. Figure 3.3A displays the gel filtration elution profile from a typical purification, in which the first peak eluted contains tetramer and higher order aggregates while the peak eluted last contains the dimeric form of cSelS and cSelS 188A (Figure 3.3B). On average, 50% of the protein in that fraction contained Sec, as detected by inductively coupled plasmon (ICP) spectroscopy. The yield for the selenium-rich fraction was about 0.6 mg per liter media growth. The resulting protein is pure of contaminants, other than the truncated cSelS 188A. The CD spectra of cSelS and cSelS U188C are nearly identical (Figure

89

3.4). However, as detailed in the following sections, only cSelS has enzymatic activity.



Figure 3.1. Expression and purification of cSelS U188C. (A) SDS-PAGE analysis. Lane 1: Crude cell extract prior to induction. Lane 2: Crude cell extract after overnight expression at 18 °C. Lane 3: IMAC elute. Lane 4: TEV protease cleavage mixture. Lane 5: cSelS U188C after TEV protease and MBP were removed by cation exchange chromatography. Lane 6: cSelS U188C (indicated by an arrow) following purification with sephacryl S-100. M: Unstained protein molecular weights standard (the molecular mass in kDa is noted on the right). (B) The molecular weight of cSelS U188C detected by electrospray ionization mass spectrometry is 15252 Da (predicted molecular weight is 15253 Da).



Figure 3.2. Characterization of cSelS U188C by size-exclusion chromatography. (A) cSelS U188C eluted from Sephacryl S-100 column in the presence (solid line) and absence of the reducing agent DTT (dotted line). (B) The apparent molecular weight of the cSelS U188C was determined by measuring its elution volume relative to standard proteins. The black triangle overlaid on the calibration curve corresponds to the apparent molecular weight of 37 kDa for cSelS U188C. Hence, SelS U188C is a dimer.



Figure 3.3. Enrichment of Sec-containing proteins from mixtures of cSelS and cSelS 188 $\Delta$ . (A) Separation of cSelS and cSelS 188 $\Delta$  by size exclusion chromatography. cSelS (solid line) and cSelS U188C (dotted line) elution profile from a Sephacryl 100 under non-reducing condition. Peak 2, eluted last, contains the selenium-rich form of the enzyme. (B) SDS-PAGE analysis of cSelS / cSelS 188 $\Delta$  mixture and cSelS U188C oligomerization pattern. Lanes 1, 2, 3: cSelS U188C, cSelS first peak and cSelS second peak under non-reducing conditions. M: Protein molecular weight standards (the molecular mass in kDa is noted on the right). Lanes 4, 5, 6: cSelS U188C, cSelS first peak and cSelS second peak under reducing conditions.



Figure 3.4. CD spectra of cSelS (dotted line) and cSelS U188C (solid line). CD spectra were acquired at 20 °C in 10 mM potassium phosphate buffer (pH 7.5) and 50 mM sodium sulfate.

### 3.2.2 SelS' Reductase Activity Measurement

To test SelS' potential reductase activity, we compared two activity assays that are based on reducing the disulfide bond between insulin's chains A and B: by monitoring insulin reduction by NADPH consumption in glutathione reductase coupled assays<sup>26</sup> and by following the process of reduction of interdisulfide bond between chain A and chain B, and the aggregation of the B chains by turbidity.<sup>27</sup> The former assay relies on GSH as a redox couple, while the latter is compatible with other reducing agents. A reductase activity was detected only in insulin turbidity assays that were carried out with dithiothreitol (DTT) as a reducing agent. GSH is not an efficient electron donor. Interestingly, the reductase activity is present only when Sec is incorporated and is absent for the Cys mutant (Figure 3.5). When the amount of enzyme was doubled, the interval for the onset of turbidity halved (absorbance at 650 nm A650 > 0.02) and the reaction rate in the linear region doubled. cSelS 188 $\Delta$  causes rapid aggregation in the absence of reducing agents. This form can only be isolated as a tetramer or higher order oligomers prone to aggregation. The reductase activity did not depend on metals or cofactors. As the insulin turbidity assays cannot be used to extract the catalytic rate, we used hTrx and human protein disulfide isomerase (hPDI) to evaluate cSelS efficiency. For comparable concentrations of hTrx, cSelS and hPDI the turbidity onset was 12, 16 and 20 minutes, respectively. The reaction rates in the linear region were 0.027, 0.025 and 0.014  $\Delta A650 \text{ min}^{-1} \mu \text{M}^{-1}$  respectively. Based on the turbidity assays, cSelS reductase activity is comparable to that of human thioredoxin 1 (hTrx)<sup>28</sup> but slightly lower. Evidently, cSelS is an effective reductase *in* vitro.

94



Figure 3.5. Reductase activity assays. The ability of cSelS to reduce insulin's intermolecular disulfide bond is monitored by recording the increasing turbidity due to protein aggregation. The reaction was initiated with 1  $\mu$ M cSelS (triangle), 2  $\mu$ M cSelS (square), 10  $\mu$ M cSelS U188C (solid line), 2  $\mu$ M hTrx (circle) or without enzyme (dotted line).



Figure 3.6. cSelS oxidase and isomerase activities were measured by monitoring the absorbance at 296 nm of RNase A-catalyzed hydrolysis of cCMP. Human protein disulfide isomerase (PDI) is used as a positive control. (A) Oxidase activity using reduced and denatured RNase A. (B) Isomerase activity using scrambled RNase A. cSelS has no oxidase or isomerase activities.

### 3.2.3 SelS' Peroxidase Activity Measurment

Other plausible enzymatic functions for SelS are oxidase, isomerase and peroxidase. cSelS had neither an oxidase nor an isomerase activity for ribonuclease A (RNase A) (Figure 3.6). This is not surprising, since the selenylsulfide active site of SelS is located in the cytosol, while disulfide bond formation and oxidative folding processes take place in the ER. A peroxidase activity is more plausible since several selenoproteins possess peroxidase activity<sup>1</sup>. Specifically, there are at least two selenoprotein families that specialize in the removal of oxidative species: glutathione peroxidases and selenoproteins with minimal thioredoxin fold<sup>29-32</sup>. cSelS itself does not accept electrons from GSH, like the former, nor does it have a known fold. As we show in this section, cSelS is also capable of reducing hydrogen peroxide efficiently - but not hydrophobic peroxides substrates.

We have measured the ability of SelS to reduce hydrogen peroxide using the rat thioredoxin reductase (rTrxR) / hTrx coupled enzymatic assays. cSelS efficiently utilized reducing equivalents from the rTrxR / hTrx system (Figures 3.7).



Figure 3.7. Peroxidase activity assays. (A) cSelS catalyze the reduction of  $H_2O_2$  by coupling with thioredoxin system. The reaction with both 5  $\mu$ M cSelS and 5  $\mu$ M hTrx is shown in circles, in the absence of cSelS in squares, and in the absence of hTrx in triangles. (B) cSelS peroxidase activity depends on the presence of Sec (\*\*\*, P<0.001). The activity of different mutants was compared: 5  $\mu$ M cSelS (WT), 5  $\mu$ M cSelS U188C (U188C), 5  $\mu$ M cSelS U188S (U188S), and 5  $\mu$ M cSelS 188 $\Delta$  (188 $\Delta$ ).

# 3.2.4 SelS' Peroxidase Substrate Specificity

While cSelS can accept electrons directly from rTrxR, it is a low affinity substrate. The peroxidase activity was solely dependent on the presence of Sec (Figure 3.7B). The specificity of cSelS was tested for: hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>), tert-butyl hydroperoxide (tBOOH), cumene hydroperoxide (COOH), 15S-hydroperoxy-6E, 8Z, 11Z, 14Z-eicosatetraenoic acid (HpETE) and 15S-hydroxy-6E, 8Z, 11Z, 14Z-eicosatetraenoic acid (HpETE) and 15S-hydrogen peroxide and cumene hydroperoxide have turnover rates significantly higher than the control (p<0.001). Hence, cSelS is not a broad specificity peroxidase. Still, the specificity may depend on cSelS interactions with binding partners, which are likely to influence the structure of the active site.



Figure 3.8. Substrate specificity of cSelS peroxidase's activity. (A) Reaction scheme. (B) cSelS substrate specificity. Reaction with cSelS and hTrx (black bars), reaction excluding hTrx (striped bars), and reaction excluding cSelS (dotted bars). Only  $H_2O_2$  and COOH have significant rates compared to the control (xx, P<0.001, n=3). The affinity for  $H_2O_2$  is the highest among the four substrates (\*\*\*\*, P<0.0001, n=3).

# 3.2.5 SelS' Peroxidase Kinetics Measurement

Next, we measured the steady-state kinetic parameters of cSelS' peroxidase activity (Figure 3.9). cSelS exhibits a two substrate ping-pong mechanism with

Michaelis-Menten-type saturable kinetics (Scheme 1).

$$cSelS_{ox} + hTrx_{red} \Leftrightarrow cSelS_{red} + hTrx_{ox}$$
$$cSelS_{red} + H_2O_2 \Leftrightarrow cSelS_{ox} + H_2O$$
Scheme 1

With H<sub>2</sub>O<sub>2</sub> as substrate, cSelS exhibited a Michaelis constant (K<sub>m</sub>) of  $58\pm 5$   $\mu$ M, a catalytic constant (k<sub>cat</sub>) of 0.110 $\pm$ 0.003 s<sup>-1</sup>, and a catalytic efficiency (k<sub>cat</sub>/K<sub>m</sub>) of (2.1 $\pm$ 0.8)×10<sup>3</sup> M<sup>-1</sup> s<sup>-1</sup> (Figure 3.9). We have also determined the kinetic parameters with hTrx as a substrate: K<sub>m</sub> = 1.1 $\pm$ 0.1  $\mu$ M, k<sub>cat</sub> = 0.098 $\pm$ 0.003 s<sup>-1</sup>, and k<sub>cat</sub>/K<sub>m</sub> = (9 $\pm$ 1)×10<sup>4</sup> M<sup>-1</sup> s<sup>-1</sup> (Figure 3.9B). The catalytic efficiency of cSelS with H<sub>2</sub>O<sub>2</sub> as substrate is lower than that reported for GPx and peroxiredoxin. The k<sub>cat</sub>/K<sub>m</sub> is of the order of 10<sup>7</sup> M<sup>-1</sup> s<sup>-1</sup> for the selenium-containing GPx<sup>33</sup>, while that of peroxiredoxins is of the order of 10<sup>4</sup>-10<sup>5</sup> M<sup>-1</sup> s<sup>-134</sup>. However, k<sub>cat</sub>/K<sub>m</sub> was determined for cSelS in isolation. Since it is an intrinsically disordered enzyme, this value may change in the presence of its binding partner.



Figure 3.9. Kinetic parameters of cSelS peroxidase activity. (A) With H<sub>2</sub>O<sub>2</sub> as substrate:  $K_m = 52\pm5 \ \mu M$ ,  $k_{cat} = 0.110\pm0.003 \ s^{-1}$  and  $k_{cat}/K_m = (2.1\pm0.8)\times10^3 \ M^{-1} \ s^{-1}$ . (B) With hTrx as substrate:  $K_m = 1.1\pm0.1 \ \mu M$ ,  $k_{cat} = 0.098\pm0.003 \ s^{-1}$  and  $k_{cat}/K_m = (9\pm1)\times10^4 \ M^{-1} \ s^{-1}$ . Means  $\pm$  SD of three independent experiments are shown.

#### 3.2.6 Identification of Peroxidatic and Resolving Residues in SelS

We established that for SelS, the Sec188 is the peroxidatic residue, while Cys174 is the resolving residue, by examining the peroxidase activity of a series of mutants (Figure 3.10A). In all of the experiments described here, we employ a form of SelS (cSelS) in which the transmembrane helix was deleted. Sec is inserted into the wild type cSelS employing *E. coli* genetic insertion machinery<sup>35</sup>. cSelS is a dimer in which Sec188 and Cys174 form an intramolecular selenenylsulfide bond. That Sec188 is the peroxidase residue is proved by the inactivity of the U188S and U188 $\Delta$  mutants (in U188 $\Delta$  translation is terminated after residue 187). Hence, Cys174 must be the resolving residue. It can be reduced by both hTrxR and hTrx in vitro (Figure 3.10A). Due to the proposed similarity of SelS' reaction mechanism to that of atypical 2-Cys Prxs, we hypothesized that selenenic acid (Se-OH) is transiently formed during the reaction. Indeed, it is possible to trap that intermediate with dimedone (5,5-dimethyl-1,3-cyclohexanedione) if the resolving Cys is mutated into Ser to prevent the rapid formation of the selenylsulfide bond. Reduced cSelS C174S was incubated with H<sub>2</sub>O<sub>2</sub> and dimedone for 20 min and examined by electrospray ionization mass spectrometry. A mass gain of 139 Da indicates labeling with dimedone (Figure 3.10B). Dimedone reacts exclusively with sulfenic acid or selenenic acid in proteins and since cSelS C174S has only one Sec and no Cys the adduct must reside on Sec188. This intermediate cannot be trapped in the wild type, most likely because the intermolecular reaction with dimedone is considerably slower than the intramolecular reaction with the resolving Cys174. Thus the peroxidase reaction mechanism was proposed as illustrated in Figure 3.11.



Figure 3.10. Reaction mechanism exploration. (A) Identification of the peroxidatic and resolving residues in SelS using peroxidase assays. NADPH consumption is monitored in the presence of 8 nM hTrxR, 5  $\mu$ M hTrx and 200  $\mu$ M H<sub>2</sub>O<sub>2</sub> and with and without cSelS, with and without thioredoxin, and with C174S, U188S, and U188 $\Delta$  mutants. (B) Trapping of the reaction intermediate, selenenic acid, in cSelS C174S. Mass spectrum showing cSelS C174S (15284 Da) and its dimedone adduct (15423 Da).



Figure 3.11. Proposed reaction mechanism of SelS peroxidase activity.

# 3.2.7 SelS is Resistant to ROS Inactivation

Thiol-dependent peroxidases are inactivated by  $H_2O_2$ , via the irreversible modification of Cys residues to sulfinic and sulfonic forms. While some peroxidases are rather robust, tolerating the presence of as much as 30 mM H<sub>2</sub>O<sub>2</sub>, others are hypersensitive, becoming fully inactivated at considerably lower exposures.<sup>36, 37</sup> Even the Sec-dependent GPx1 is inactivated by H<sub>2</sub>O<sub>2</sub>, though this was proposed to be a result of selenium elimination.<sup>38</sup> To gain further insight into the potentially advantageous role of Sec in SelS, we tested the ability of cSelS to maintain activity after exposure to high concentrations of H<sub>2</sub>O<sub>2</sub>. We have chosen to assay the reductase – not the peroxidase - activity of SelS since in our hands hTrxR was partially inactivated at high concentrations of H<sub>2</sub>O<sub>2</sub> (data not shown). cSelS reductase activity was assayed using insulin turbidity tests<sup>27</sup>, following 20 min incubations with various concentrations of H<sub>2</sub>O<sub>2</sub> up to 50 mM and its subsequent removal by catalase. Figure 3.12A,B shows that for the wild type the reaction rate in the linear region (A<sub>650</sub>> 0.2) did not significantly change irrespective of the  $H_2O_2$  concentrations used during the incubations. In contrast, C174S, which also has a reductase activity, is more prone for inactivation (Figure 3.12C,D). Nevertheless, it still maintains about 80% of its activity after incubation at 50 mM  $H_2O_2$ . The increased resistance of the wild type to inactivation might provide an advantage for SelS, allowing it to sustain functionality under oxidative stress.



Figure 3.12. cSelS resistance to inactivation by  $H_2O_2$ . The ability of cSelS to reduce insulin's intermolecular disulfide bond following incubations with  $H_2O_2$  is monitored by recording the increasing turbidity caused by insulin's chain B aggregation. (A) cSelS reductase activity assay after incubations with  $H_2O_2$ . (B) Percentage of remaining activity of wild type cSelS after treatment with  $H_2O_2$ . (C) cSelS C174S reductase activity assay after incubations with  $H_2O_2$ . (D) Percentage of remaining activity of cSelS C174S after treatment with  $H_2O_2$ . (D) Percentage of remaining activity of cSelS C174S after treatment with  $H_2O_2$ . The error bars represent the range of measurements (that is, highest and lowest values) among two repetitions, using two independent protein preparations.

### 3.3 Discussion

The experiments establish that human SelS acts primarily as a reductase. This is consistent with the observation that most selenoproteins are oxidoreductases. SelS does not have isomerase or oxidase activities in RNase assays but it does have a significant peroxidase activity toward hydrogen peroxide, albeit lower than that of GPx or peroxiredoxins. SelS does not, however, efficiently reduce peroxides compared to GPx or peroxiredoxins and is not able to reduce complex peroxides. Since it is neither an efficient nor a broad-spectrum peroxidase, SelS peroxidase activity (at least in the absence of its protein partners) should be seen as an extension of SelS' activity as a reductase. This is analogous to the other selenium-containing reductase TrxR<sup>39</sup>. TrxR reduces, in addition to protein disulfides, hydrogen peroxide, selenite, dehydroascorbate, a-lipoic acid, and lipid hydroperoxides<sup>40</sup>. In this way, it provides a broad antioxidative capacity.

The rare amino acid Sec is essential for enzymatic activity; both the reductase and peroxidase activities require the presence of Sec and are not present when Sec is substituted with Cys. Why the selenium is critical for activity is not easy to explain. The enzymatic activity of only a small subset of selenoproteins had been characterized<sup>2</sup> and the unique role of Sec in catalysis still awaits a unified description<sup>41, 42</sup>. Differences in pK<sub>a</sub> may be a potential explanation for the reliance on Sec, but pK<sub>a</sub> is easily fine tuned by the protein microenvironment and can be shifted as needed to introduce activity for the cSelS U188C. Certainly other members of the SelS/SelK eukaryotic family do not rely on Sec<sup>3, 43</sup>. Perhaps a better explanation would invoke the formation of a particularly short-lived selenenic acid (Se-OH) or seleninic acid (SeO<sup>2-</sup>) as part of the mechanism. Indeed, we were able to trap the Se-OH on a cSelS C174S mutant, the mechanism of the peroxidase activity appears to be related to atypical 2-cysteine peroxiredoxin where the resolving cysteine is in the same subunit and in which a sulfenic acid (S-OH) intermediate was identified<sup>44</sup>.

The finding that SelS is primarily a reductase suggests future directions in delineating its biological mechanism of action. SelS is a member of the ERAD machinery and presumably may take part in unfolding and transporting target proteins. However, SelS, whose active site is in the cytoplasm, is unlikely to remove disulfide bonds in ER-residing soluble misfolded proteins as several ER residing reductases are involved in that process prior to the transport in the channel<sup>45, 46</sup>. Similarly, management of disulfide bonds in misfolded transmembrane proteins is not well understood, but SelS does not appear to be a critical component of the machinery<sup>47</sup>. A unique feature of SelS is that it is an intrinsically disordered enzyme, a class that relies on order to disorder transition or conformational selection to restrict the conformational space upon binding the substrate<sup>48</sup>. Most intrinsically disordered proteins take part in signaling or regulation. This suggests that SelS' physiological partner(s) is a signaling protein, possibly a member of the unfolded protein response (UPR)<sup>49, 50</sup>, a pathway responsible for combating oxidative stress. The ERAD path is linked to the UPR response via amplification loops in the cytoplasm<sup>51</sup>. Many of the proteins in the path are also single-pass transmembrane proteins, raising the possibility that they interact with a membrane-bound reductase, such as SelS. Indeed, several selenoproteins act as stress sensors in regulatory pathways<sup>52, 53</sup>. One notable binding partner of SelS that was also implicated in signaling is selenoprotein K (SelK), a single-pass transmembrane enzyme with unknown function<sup>3, 54</sup>. SelK does not contain a conventional seleno-redox motif where the Sec is coupled with a Cys, Ser or Thr in close proximity, and the reactive Sec may be able to interact with SelS' active site. We are currently testing whether SelK binding influences SelS' structure and enzymatic activity.

In summary, here we describe the enzymatic function of SelS. We demonstrate that it is enzymatically active in the absence of its binding partner and that the Sec is critical to cSelS' function. The selenylsulfide bond renders SelS resistant to inactivation by ROS.

### **3.4 Materials and Methods**

#### 3.4.1 Bacterial Strains, Plasmids, and Chemical Reagents

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 (cMBP)<sup>55</sup> was purchased from the Protein Structure Initiative: Biology Materials Repository<sup>56</sup>. Chromatography media was supplied by GE Healthcare Bio-Sciences Corporation (Pittsburgh, PA) and New England Biolabs. The expression construct of human thioredoxin 1 (hTrx) was generously provided by Prof. Marletta<sup>57</sup>. Human protein disulfide isomerase (PDI) was a gift from Prof. Thorpe<sup>58</sup>. The expression construct for human glutaredoxin 1 (hGrx) was a gift from Prof. Mieyal<sup>59</sup>. 15S- hydroperoxy- 5Z, 8Z, 11Z, 13E- eicosatetraenoic acid was from Cayman (Ann Arbor, MI). Bovine insulin was from Cell Application (San Diego, CA). Rat thioredoxin reductase 1 (rTrxR) was from Cayman (recombinant form) or Sigma (purified from rat liver). The pSUABC plasmid was generously provided by Prof. Arnér from the Karolinska Institutet<sup>24</sup>. All other chemicals and reagents were supplied by Sigma-Aldrich (St. Louis, MO), Acros Organics (Geel, Belgium) and GoldBio (St. Louis, MO). All reagents and solvents were at least analytical grade and were used as supplied.

## 3.4.2 Molecular Cloning

The *Homo sapiens* SelS gene (GenBank accession no. GI: 45439348) was codon optimized for expression in *E. coli* and the gene synthesized by genscript (Piscataway, NJ). The soluble portion of the gene (residues 52-189) was cloned into a pMAL-C5X (New England Biolabs) as a fusion with the cytoplasmic maltose binding protein (cMBP). To enable selenocysteine insertion, a non-encoding *E. coli* formate dehydrogenase SECIS element was inserted immediately after the stop codon<sup>60</sup>. A short linker NSSS and a TEV protease cleavage site, ENLYFQS, was used to connect the two proteins. Following cleavage with TEV protease, no non-native residues were retained.

### 3.4.3 Expression and Purification of cSelS and its Mutants

For protein expression of cSelS mutants (cSelS U188C and cSelS U188S), the plasmids were transformed into an *E. coli* BL21(DE3) strain. Cells were grown in LB, supplemented with 0.2% glucose at 37 °C, with good aeration and the relevant antibiotic selection (100  $\mu$ g/ml). When the optical density (OD) at 600 nm reached 0.5, the temperature was lowered to 18 °C, and the cells were allowed to shake at the lower temperature for an additional hour. Protein expression was induced with 0.5 mM isopropyl-1-thio-b-D-galactopyranoside (IPTG) when the optical density (OD) at 600 nm reached 0.7. The cells were harvested after 18-20 h, and the cell paste (7 g/L) was resuspended in 50 mM sodium phosphate, 200 mM sodium chloride, pH 7.5 (amylose buffer), supplemented with 0.5 mM benzamidine, 1 mM

phenylmethylsulfonyl fluoride (PMSF), and 1 mM ethylenediaminetetraacetic acid (EDTA). Cells were lysed using a high-pressure homogenizer (EmulsiFlex-C5, Avestin, Ottawa, Canada) on ice, and all subsequent procedures were carried out at 4 °C. Cell debris was removed by centrifugation at 20000 g for 1 h. The supernatant was loaded on an amylose column, and the column was washed with the amylose buffer. The fusion cMBP-cSelS was eluted using an amylose buffer containing 20 mM maltose. The purity of the eluted cMBP-cSelS was about 90%. Cleavage of the fusion partner cMBP was carried out by adding a hexahistidine-tagged TEV protease to the dialysis bag for 12 h. The TEV protease was added at a molar ratio of 1:10 relative to the fusion protein. Following cleavage, the protein was then loaded on a 5-ml HiTrap SP HP and eluted with a salt gradient between 200 and 1000 mM sodium chloride, over 20 column volumes. The fractions containing SelS were concentrated to 20 mg/ml, loaded on a HiPrep<sup>™</sup> Sephacryl S-100 column (GE Healthcare) and eluted at 0.4 ml/min with 50 mM sodium phosphate and 200 mM sodium chloride (pH 7.5). The column was calibrated using GE Healthcare Bio-Sciences gel filtration protein standards: bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa). The void volume was measured using blue dextran 2000. Protein purity, as determined by 16% Tris-tricine SDS-PAGE gel was higher than 99%. Protein concentration was determined using an extinction coefficient of 14105 M-1 CM-1 for cSelS U188C and 13980 M-1 CM-1 for both cSelS U188S and cSelS 188<sup>Δ</sup>. The successful incorporation of selenium was confirmed by mass and inductively coupled plasma (ICP) spectroscopy.

For protein expression of the cSelS and cSelS 188 $\Delta$ , which was generated by mutating UGA to UAA, the plasmid was cotransformed into *E. coli* BL21(DE3), along

with the pSUABC plasmid expressing E. coli SelA, SelB, and SelC under the control of their endogenous promoters <sup>24</sup>. Cells were grown in Studier's MDAG defined media at 37 °C<sup>61</sup>, with good aeration and the relevant antibiotic selection (100  $\mu$ g/ml ampicillin +  $34 \mu g/ml$  chloramphenicol). When the OD at 600 nm reached 2.4, the temperature was lowered to 18 °C, and the cells were allowed to shake at the lower temperature for an additional hour. Protein expression was induced with 0.5 mM IPTG. At induction, 5 mM Na<sub>2</sub>SeO<sub>3</sub> and 100 µg/ml L-cysteine were added to the media. The remaining steps are identical to those described above. It was possible to distinguish two well-separated peaks on the SP column, eluted at 400 mM sodium chloride and 600 mM sodium chloride. The latter contains aggregates and was not included in further purifications. The fraction eluted at lower ionic strength was concentrated down to 20 mg/ml and further purified on a Sephacryl S-100 column. The dimeric form of the protein was separated from other fractions and assayed for activity. Other fractions also contain full-length protein but at a much lower ratio compared to the truncated form. The successful incorporation of selenium was confirmed by ICP spectroscopy.

For protein expression and purification of cSelS C174S, it was expressed with the cysteine auxotrophic *E. coli* host cell BL21(DE3) selB::kan cys51E<sup>62</sup>. Selenocystine was incorporated according to the protocol developed by the Böck group with minor modifications<sup>63</sup>. Briefly, cells were grown in modified LB media (10 g tryptone, 5 g NaCl, 5 g yeast extract, 2 g glucose per liter), at 37 °C, with good aeration and the relevant antibiotic selection (100 µg/ml ampicillin + 25 µg/ml kanamycin), and 50 µg/ml L-cysteine. When the optical density (OD) at 600 nm reached 1.1, the cells were harvested by centrifugation at 4 °C, 4000 g for 10 min. The pellet was washed in sterile ice cold buffer (20 mM phosphate, 50 mM NaCl (pH 7.5)) and reharvested by centrifugation. The pellet was resuspended in equal volumes of Studier's MDAG defined media<sup>61</sup> supplemented with 100 µg/ml ampicillin + 25 µg/ml kanamycin. The cells were allowed to shake at 37 °C for an additional 0.5 h to consume residual free cysteine. The temperature was lowered to 18 °C, and the cells were allowed to shake at the lower temperature for an additional 0.5 h. Protein expression was induced with 0.5 mM isopropyl-1-thio-b-D-galactopyranoside (IPTG). At induction, 200 µM L-selenocystine was added to the media. The cells were harvested after 18-20 h. The remaining steps and the purification procedure were identical to those reported<sup>35</sup>. The percent of selenium incorporation in a given protein sample was measured by mass spectrometry, as it was previously shown to be in agreement with values obtained by Inductively Coupled Plasma (ICP) spectrometry. Selenium incorporation in all batches of wild type cSelS used in this study was at least 50% while incorporation in all batches of cSelS C174S was at least 90%.

### 3.4.4 Oxidase and Isomerase Assays

The ability of cSelS to act as either an oxidase or isomerase was followed by detecting the recovery of either reduced and denatured or scrambled RNase activity<sup>64</sup>. The reaction was monitored at 25 °C, by following the hydrolysis of cCMP by newly oxidized and properly folded RNase A at 296 nm<sup>65</sup>. For the oxidase assays, reduced and denatured RNase A was prepared by adding 100 molar excess of DTT to RNase A in 50 mM Tris-HCl (pH 7.5), 6 M guanidine hydrochloride and 1 mM EDTA. After an overnight incubation at 4 °C, the excess DTT and guanidine were removed by Nap-5 column and the buffer exchanged into oxidase assay buffer: 50 mM HEPES-NaOH (pH 7.0), 150 mM sodium chloride, 75 mM imidazole, 2 mM EDTA and 0.5% Tween

20. The free thiol count of fully reduced RNase A, determined by thiol titer using Ellman's reagent 5,5' –dithiolbis-2-nitrobenzoic acid (DTNB)<sup>66</sup>, was close to 8 thiols per protein, demonstrating that RNase was fully reduced. The assay buffer included 50 mM HEPES-NaOH (pH 7.0), 150 mM sodium chloride, 75 mM imidazole, 2 mM EDTA, 0.5% Tween 20, 4.5 mM cCMP, 4  $\mu$ M reduced and denatured RNase A, 0.4 mM GSSG and 1.2 mM GSH as a redox buffer ([GSSG]: [GSH] = 1:3). The reaction was initiated by adding 1  $\mu$ M cSelS. Human protein disulfide isomerase (hPDI) served as a positive control.

cSelS isomerase activity was measured using scrambled RNase A, prepared as previously described by Thorpe<sup>58</sup>. Briefly, native RNase A was added to 50 mM Tris-HCl (pH 7.5), 6 M guanidine hydrochloride with an equimolar amount of DTT for 30 h under anaerobic conditions. This solution was then exposed to air and allowed to slowly reoxidize in the dark, at room temperature. The reaction was monitored until less than 1% free thiol per RNase molecule remained. The buffer was then exchanged to 50 mM HEPES-NaOH (pH 7.0), 150 mM sodium chloride, 75 mM imidazole, 2 mM EDTA and 0.5% Tween 20. Scrambled RNase was stored at -20 °C until use. The assay mixture contained 50 mM HEPES-NaOH (pH 7.0), 150 mM sodium chloride, 75 mM imidazole, 2 mM EDTA, 0.5% Tween 20, 4.5 mM cCMP, 0.4 mM GSSG, and 1.2 mM GSH as a redox buffer and 8 µM scrambled RNase A. The reaction was initiated by adding 1 µM cSelS and monitored at 25 °C. hPDI served as a positive control.

### 3.4.5 Inductively Coupled Plasma (ICP) Spectrometry

ICP spectrometry (Thermo IRIS Intrepid II XSP Dual View) was used to determine the ratio of sulfur and selenium in 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.

### **3.4.6 Electrospray Ionization Mass Spectroscopy**

Mass spectra were obtained using a QTOF Ultima (Waters, MA), operating under positive electrospray ionization (+ESI) mode, 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%, with 0.1% formic acid as the mobile phase, in 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. The obtained spectra were deconvoluted using maximum entropy in MassLynx (Waters, MA).

### 3.4.7 Circular Dichroism Spectroscopy

CD spectra of SelS were measured by using a J-810 circular dichroism spectropolarimeter (Jasco, Essex, UK) that had been calibrated using camphor sulfonic acid for optical rotation and benzene vapor for wavelength. Far-UV spectra were recorded using a 1 mm path-length cell for the 190-250 nm region at 20 °C. Samples for CD spectroscopy were prepared in 10 mM potassium phosphate (pH 7.5) and 50 mM sodium sulfate. Three accumulation scans were collected for baseline, and eight accumulation scans were taken for each sample.

## 3.4.8 Insulin Reduction Assay

The insulin reduction assay was performed as previously described by Holmgren<sup>27</sup>. Briefly, the reaction mixture included 100 mM sodium phosphate buffer (pH 7.0), 2 mM EDTA, 0.13 mM bovine insulin and 0.33 mM dithiothreitol (DTT).

The reaction was initiated by addition of hTrx, hPDI, cSelS, or cSelS mutants. The reaction's progress was monitored by recording the increase in turbidity at 650 nm against time. The temperature was kept constant at 25 °C.

### 3.4.9 Peroxidase Assays and Kinetics Measurement

Peroxidase activity was measured via a coupled reaction with rTrxR and  $hTrx^{67}$ . The reaction mixture contained 100 mM potassium phosphate, pH 7.0, 2 mM EDTA, 150  $\mu$ M NADPH, 2 nM rTrxR, 5  $\mu$ M hTrx and either no enzyme or 5  $\mu$ M of cSelS, cSelS U188C, cSelS U188S, or cSelS 188 $\Delta$ . Enzymatic oxidation of NADPH was initiated by adding 200  $\mu$ M H<sub>2</sub>O<sub>2</sub> after incubation for 3 minutes to the sample cuvette and monitoring the consumption of NADPH spectroscopically at 340 nm. The rate of NADPH oxidation for each condition was calculated from 3 repeats.

Steady-state kinetic analysis was performed using the assay described above. To measure the kinetic parameters of cSelS peroxidase activity,  $H_2O_2$  was added at concentrations of 12.5-1600  $\mu$ M to a reaction mixture containing 8 nM rTrxR, 5  $\mu$ M cSelS, 5  $\mu$ M hTrx and 150  $\mu$ M NADPH. For measurements with hTrx as a substrate, hTrx at concentrations of 0.625-20  $\mu$ M was added to a reaction mixture containing 8 nM rTrxR, 5  $\mu$ M cSelS, 150  $\mu$ M NADPH and 200  $\mu$ M H<sub>2</sub>O<sub>2</sub>. Each condition was repeated three times.

To test the substrate specificity of cSelS' peroxidase activity, a variety of potential peroxide substrates were tested in a similar assay as described above. The soluble  $H_2O_2$ , cumene hydroperoxide (COOH) and tert-Butyl hydroperoxide (tBuOOH) were dissolved in 100 mM potassium phosphate (pH 7.0) and employed at a concentration of 200  $\mu$ M. To prepare the hydrophobic 15-HETE and 15-HpETE, ethanol was evaporated by purging with a N<sub>2</sub> stream, dissolved in 100 mM potassium
phosphate (pH 7.0), and added to the reaction buffer at a concentration of 30  $\mu$ M. The rate of the initial NADPH oxidation was calculated by monitoring the absorbance at 340 nm. Experiments were repeated three times for each substrate.

# 3.4.10 Traping of Selenenic Acid by Dimedone

 $70 \ \mu\text{M}$  of the relevant protein (cSelS, cSelS U188C, cSelS C174S, cSelS U188S, and cSelS U188 $\Delta$ ) was incubated with 10 mM DTT at 25 °C for 2 h. The buffer was exchanged for 100 mM potassium phosphate, 2 mM EDTA (pH 7.0) using a Nap-5 desalting column. 200  $\mu$ M H<sub>2</sub>O<sub>2</sub> and 10 mM dimedone were added immediately after the desalting step and the mixture incubated at 25 °C for 20 min. Following incubation, the buffer was exchanged for deionized distilled water, using polyacrylamide spin desalting column and the samples were examined using mass spectrometry. A mass shift of 138 Da (the covalent adduct with dimedone) demonstrated successful trapping of the selenenic acid species. Experiments were preformed in triplicates.

#### 3.4.11 Inactivation of cSelS and cSelS C174S by H<sub>2</sub>O<sub>2</sub>

To investigate cSelS and cSelS C174S' resistance to inactivation by  $H_2O_2$ , 140  $\mu$ M cSelS or cSelS C174S was first reduced by incubation with 20 mM DTT at 25 °C for 2 h. The sample buffer was exchanged for 100 mM potassium phosphate, 2 mM EDTA (pH 7.0), using a Nap-5 desalting column and the volume was quickly aliquoted and  $H_2O_2$  was added to a final concentration of 0, 1, 2, 4, 8, 16, 32, and 50 mM  $H_2O_2$ . The reaction mixtures were incubated at 25 °C for 20 min. Subsequently, catalase was added to the incubation mixtures to a final 100 U for 10 min to consume

the remaining  $H_2O_2$ . The complete consumption of  $H_2O_2$  was confirmed by measuring the disappearance of the absorbance at 240 nm (about 1 min).<sup>68</sup>

A modified insulin turbidity  $assay^{27}$  was used to check the remaining activity of cSelS and cSelS C174S. Briefly, the reaction mixture included 100 mM potassium phosphate (pH 7.0), 2 mM EDTA, 0.13 mM bovine insulin and 1 mM DTT. The reaction was initiated by the addition of 8  $\mu$ L of a given incubation mixture (final protein concentration in the cuvette was 8  $\mu$ M). The reaction's progress was monitored by recording the increase in turbidity at 650 nm over time. The temperature was kept constant at 25 °C. Values and error bars are the mean and standard deviation of three independent experiments.

# REFERENCES

- 1. Lu, J. & Holmgren, A. Selenoproteins. J. Biol. Chem. 284, 723-727 (2009).
- Kasaikina, M.V., Hatfield, D.L. & Gladyshev, V.N. Understanding selenoprotein function and regulation through the use of rodent models. *Biochim. Biophys. Acta Mol. Cell Res.* 1823, 1633-1642 (2012).
- Shchedrina, V. A., Everley, R. A., Zhang, Y., Gygi, S. P., Hatfield, D. L. & Gladyshev, V. N. Selenoprotein K binds multiprotein complexes and is involved in the regulation of endoplasmic reticulum homeostasis. *J. Biol. Chem.* 286, 42937-42948 (2011).
- Curran, J. E., Jowett, J. B. M., Elliott, K. S., Gao, Y., Gluschenko, K., Wang, J. M., Azim, D. M. A., Cai, G. W., Mahaney, M. C., Comuzzie, A. G., Dyer, T. D., Walder, K. R., Zimmet, P., MacCluer, J. W., Collier, G. R., Kissebah, A. H. & Blangero, J. Genetic variation in selenoprotein S influences inflammatory response. *Nat. Genet.* 37, 1234-1241 (2005).
- Du, S.Q., Liu, H.M. & Huang, K.X. Influence of SelS gene silence on beta-Mercaptoethanol-mediated endoplasmic reticulum stress and cell apoptosis in HepG2 cells. *Biochim. Biophys. Acta* 1800, 511-517 (2010).

- Ye, Y.H., Shibata, Y., Yun, C., Ron, D. & Rapoport, T.A. A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol. *Nature* 429, 841-847 (2004).
- Lilley, B.N. & Ploegh, H.L. Multiprotein complexes that link dislocation, ubiquitination, and extraction of misfolded proteins from the endoplasmic reticulum membrane. *Proc. Natl. Acad. Sci. USA* 102, 14296-14301 (2005).
- 8. Smith, M.H., Ploegh, H.L. & Weissman, J.S. Road to ruin: Targeting proteins for degradation in the endoplasmic reticulum. *Science* **334**, 1086-1090 (2011).
- Kryukov, G. V., Castellano, S., Novoselov, S. V., Lobanov, A. V., Zehtab, O., Guigo, R. & Gladyshev, V. N. Characterization of mammalian selenoproteomes. *Science* 300, 1439-1443 (2003).
- Shchedrina, V.A., Zhang, Y., Labunskyy, V.M., Hatfield, D.L. & Gladyshev, V.N. Structure-function relations, physiological roles, and evolution of mammalian ER-resident selenoproteins. *Antioxid. Redox Signal.* 12, 839-849 (2010).
- Christensen, L. C., Jensen, N. W., Vala, A., Kamarauskaite, J., Johansson, L., Winther, J. R., Hofmann, K., Teilum, K. & Ellgaard, L. The human selenoprotein VCP-interacting membrane protein (VIMP) is non-globular and harbors a reductase function in an intrinsically disordered region. *J. Biol. Chem.* 287, 26388-26399 (2012).

- Haenzelmann, P. & Schindelin, H. The structural and functional basis of the p97/Valosin-containing protein (VCP)-interacting motif (VIM). *J. Biol. Chem.* 286, 38679-38690 (2011).
- Lilley, B.N. & Ploegh, H.L. A membrane protein required for dislocation of misfolded proteins from the ER. *Nature* 429, 834-840 (2004).
- Christianson, J. C., Olzmann, J. A., Shaler, T. A., Sowa, M. E., Bennett, E. J., Richter, C. M., Tyler, R. E., Greenblatt, E. J., Harper, J. W. & Kopito, R. R. Defining human ERAD networks through an integrative mapping strategy. *Nat. Cell Biol.* 14, 93-U176 (2012).
- Lee, J. N., Kim, H., Yao, H. B., Chen, Y., Weng, K. & Ye, J. Identification of Ubxd8 protein as a sensor for unsaturated fatty acids and regulator of triglyceride synthesis. *Proc. Natl. Acad. Sci. USA* 107, 21424-21429 (2010).
- Suzuki, M., Otsuka, T., Ohsaki, Y., Cheng, J. L., Taniguchi, T., Hashimoto,
  H., Taniguchi, H. & Fujimoto, T. Derlin-1 and UBXD8 are engaged in
  dislocation and degradation of lipidated ApoB-100 at lipid droplets. *Mol. Biol. Cell* 23, 800-810 (2012).
- Madsen, L., Kriegenburg, F., Vala, A., Best, D., Prag, S., Hofmann, K., Seeger, M., Adams, I. R. & Hartmann-Petersen, R. The tissue-specific Rep8/UBXD6 tethers p97 to the endoplasmic reticulum membrane for degradation of misfolded proteins. *Plos One* 6 (2011).

- Chin, K.T., Xu, H.T., Ching, Y.P. & Jin, D.Y. Differential subcellular localization and activity of kelch repeat proteins KLHDC1 and KLHDC2. *Mol. Cell. Biochem.* 296, 109-119 (2007).
- Uversky, V.N. Intrinsically disordered proteins from A to Z. Int. J. Biochem. Cell Biol. 43, 1090-1103 (2011).
- 20. Yoshizawa, S. & Bock, A. The many levels of control on bacterial selenoprotein synthesis. *Biochim. Biophys. Acta* **1790**, 1404-1414 (2009).
- Liu, J., Srinivasan, P., Pham, D.N. & Rozovsky, S. Expression and purification of the membrane enzyme selenoprotein K. *Protein Expr. Purif.* 86, 27-34 (2012).
- Johansson, L., Gafvelin, G. & Arner, E.S.J. Selenocysteine in proteins properties and biotechnological use. *Biochim. Biophys. Acta* 1726, 1-13 (2005).
- Watson, W. H., Pohl, J., Montfort, W. R., Stuchlik, O., Reed, M. S., Powis, G. & Jones, D. P. Redox potential of human thioredoxin 1 and identification of a second dithiol/disulfide motif. *J. Biol. Chem.* 278, 33408-33415 (2003).
- Arner, E.S.J., Sarioglu, H., Lottspeich, F., Holmgren, A. & Bock, A. Highlevel 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, 1003-1016 (1999).

- Rengby, O., Johansson, L., Carlson, L. A., Serini, E., Vlamis-Gardikas, A., Karsnas, P. & Arner, E. S. J. Assessment of production conditions for efficient use of Escherichia coli in high-yield heterologous recombinant selenoprotein synthesis. *Appl. Environ. Microbiol.* **70**, 5159-5167 (2004).
- Lambert, N. & Freedman, R.B. The latency of rat-liver microsomal protein disulfide isomerase. *Biochem. J* 228, 635-645 (1985).
- 27. Holmgren, A. Thioredoxin catalyzes the reduction of insulin disulfides by dithiothreitol and dihydrolipoamide. *J. Biol. Chem.* **254**, 9627-9632 (1979).
- Ren, X.L., Bjornstedt, M., Shen, B., Ericson, M.L. & Holmgren, A. Mutagenesis of structural half-cysteine residues in human thioredoxin and effects on the regulation of activity by selenodiglutathione. *Biochemistry* 32, 9701-9708 (1993).
- Flohe, L., Toppo, S., Cozza, G. & Ursini, F. A comparison of thiol peroxidase mechanisms. *Antioxid. Redox Signaling* 15, 763-780 (2011).
- Flohe, L. & Ursini, F. Peroxidase: A term of many meanings. *Antioxid. Redox Signaling* 10, 1485-1490 (2008).
- Saito, Y., Hayashi, T., Tanaka, A., Watanabe, Y., Suzuki, M., Saito, E. & Takahashi, K. Selenoprotein P in human plasma as an extracellular phospholipid hydroperoxide glutathione peroxidase - Isolation and enzymatic characterization of human selenoprotein P. *J. Biol. Chem.* 274, 2866-2871 (1999).

- 32. Novoselov, S. V., Kryukov, G. V., Xu, X. M., Carlson, B. A., Hatfield, D. L. & Gladyshev, V. N. Selenoprotein H is a nucleolar thioredoxin-like protein with a unique expression pattern. *J. Biol. Chem.* 282, 11960-11968 (2007).
- Toppo, S., Flohe, L., Ursini, F., Vanin, S. & Maiorino, M. Catalytic mechanisms and specificities of glutathione peroxidases: Variations of a basic scheme. *Biochim. Biophys. Acta* 1790, 1486-1500 (2009).
- Trujillo, M., Ferrer-Sueta, G., Thomson, L., Flohe, L. & Radi, R. Kinetics of peroxiredoxins and their role in the decomposition of peroxynitrite. *Subcell. Biochem.* 44, 83-113 (2007).
- 35. Liu, J., Li, F. & Rozovsky, S. The intrinsically disordered membrane protein selenoprotein S is a reductase in vitro. *Biochemistry* **52**, 3051-3061 (2013).
- 36. Wood, Z.A., Poole, L.B. & Karplus, P.A. Peroxiredoxin evolution and the regulation of hydrogen peroxide signaling. *Science* **300**, 650-653 (2003).
- Pigeolet, E., Corbisier, P., Houbion, A., Lambert, D., Michiels, C., Raes, M., Zachary, M. D. & Remacle, J. Glutathione-peroxidase, superoxide-dismutase, and catalase inactivation by peroxides and oxygen derived free-radicals. *Mech. Ageing Dev.* 51, 283-297 (1990).
- Cho, C. S., Lee, S., Lee, G. T., Woo, H. A., Choi, E. J. & Rhee, S. G. Irreversible Inactivation of Glutathione Peroxidase 1 and Reversible Inactivation of Peroxiredoxin II by H2O2 in Red Blood Cells. *Antioxid. Redox Signaling* 12, 1235-1246 (2010).

- Arner, E.S.J. Focus on mammalian thioredoxin reductases Important selenoproteins with versatile functions. *Biochim. Biophys. Acta* 1790, 495-526 (2009).
- 40. Nordberg, J. & Arner, E.S.J. Reactive oxygen species, antioxidants, and the mammalian thioredoxin system. *Free Radical Biol. Med.* 31, 1287-1312 (2001).
- 41. Arner, E.S.J. Selenoproteins-What unique properties can arise with selenocysteine in place of cysteine? *Exp. Cell Res.* **316**, 1296-1303 (2010).
- 42. Ruggles, E.L., Snider, G.W. & Hondal, R.J. in Selenium: Its Molecular Biology and Role in Human Health. (eds. D.L. Hatfield, M.J. Berry & V.N. Gladyshev) 73-83 (Springer, New York; 2012).
- Roseler, A., Prieto, J. H., Iozef, R., Hecker, B., Schirmer, R. H., Kulzer, S., Przyborski, J., Rahlfs, S. and Becker, K. Insight into the selenoproteome of the malaria parasite plasmodium falciparum. *Antioxid. Redox Signaling* 17, 534-543 (2012).
- Rhee, S.G. & Woo, H.A. Multiple functions of peroxiredoxins: Peroxidases, sensors and regulators of the intracellular messenger H2O2, and protein chaperones. *Antioxid. Redox Signaling* 15, 781-794 (2011).
- Hagiwara, M. & Nagata, K. Redox-dependent protein quality control in the endoplasmic reticulum: Folding to degradation. *Antioxid. Redox Signaling* 16, 1119-1128 (2012).

- 46. Hetz, C. The unfolded protein response: controlling cell fate decisions under ER stress and beyond. *Nature Reviews Molecular Cell Biology* 13, 89-102 (2012).
- 47. Houck, S.A. & Cyr, D.M. Mechanisms for quality control of misfolded transmembrane proteins. *Biochim. Biophys. Acta, Biomembr.* 1818, 1108-1114 (2012).
- 48. Li, E., Wimley, W.C. & Hristova, K. Transmembrane helix dimerization: Beyond the search for sequence motifs. *Biochim. Biophys. Acta, Biomembr.* 1818, 183-193 (2012).
- Claessen, J.H.L., Kundrat, L. & Ploegh, H.L. Protein quality control in the ER: balancing the ubiquitin checkbook. *Trends Cell Biol.* 22, 22-32 (2012).
- Walter, P. & Ron, D. The unfolded protein response: From stress pathway to homeostatic regulation. *Science* 334, 1081-1086 (2011).
- Higa, A. & Chevet, E. Redox signaling loops in the unfolded protein response.
   *Cell. Signal.* 24, 1548-1555 (2012).
- Hawkes, W.C. & Alkan, Z. Regulation of redox signaling by selenoproteins.
   *Biol. Trace Elem. Res.* 134, 235-251 (2010).
- Brigelius-Flohe, R. & Flohe, L. Basic principles and emerging concepts in the redox control of transcription factors. *Antioxid. Redox Signaling* 15, 2335-2381 (2011).
- 54. Verma, S., Hoffmann, F. W., Kumar, M., Huang, Z., Roe, K., Nguyen-Wu, E.,Hashimoto, A. S. & Hoffmann, P. R. Selenoprotein K knockout mice exhibit

deficient calcium flux in immune cells and impaired immune responses. *J. Immunol.* **186**, 2127-2137 (2011).

- Blommel, P.G. & Fox, B.G. A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expr. Purif.* 55, 53-68 (2007).
- 56. Cormier, C. Y., Mohr, S. E., Zuo, D. M., Hu, Y. H., Rolfs, A., Kramer, J., Taycher, E., Kelley, F., Fiacco, M., Turnbull, G. & LaBaer, J. Protein Structure Initiative Material Repository: an open shared public resource of structural genomics plasmids for the biological community. *Nucleic Acids Res.* 38, D743-D749 (2010).
- 57. Mitchell, D.A., Morton, S.U., Fernhoff, N.B. & Marletta, M.A. Thioredoxin is required for S-nitrosation of procaspase-3 and the inhibition of apoptosis in Jurkat cells. *Proc. Natl. Acad. Sci. USA* **104**, 11609-11614 (2007).
- 58. Rancy, P.C. & Thorpe, C. Oxidative protein folding in vitro: A study of the cooperation between quiescin-sulfhydryl oxidase and protein disulfide isomerase. *Biochemistry* **47**, 12047-12056 (2008).
- Jao, S. C., Ospina, S. M. E., Berdis, A. J., Starke, D. W., Post, C. B. & Mieyal,
  J. J. Computational and mutational analysis of human glutaredoxin (thioltransferase): Probing the molecular basis of the low pK(a) of cysteine 22 and its role in catalysis. *Biochemistry* 45, 4785-4796 (2006).

- 60. Cheng, Q., Stone-Elander, S. & Arner, E.S.J. Tagging recombinant proteins with a Sel-tag for purification, labeling with electrophilic compounds or radiolabeling with C-11. *Nature Protocols* **1**, 604-613 (2006).
- 61. Studier, F.W. Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **41**, 207-234 (2005).
- Muller, S., Senn, H., Gsell, B., Vetter, W., Baron, C. & Bock, A. The formation of diselenide bridges in proteins by incorporation of selenocysteine residues biosynthesis and characterization of (Se)(2)-thioredoxin. *Biochemistry* 33, 3404-3412 (1994).
- Strub, M. P., Hoh, F., Sanchez, J. F., Strub, J. M., Bock, A., Aumelas, A. & Dumas, C. Selenomethionine and selenocysteine double labeling strategy for crystallographic phasing. *Structure* 11, 1359-1367 (2003).
- 64. Lyles, M.M. & Gilbert, H.F. Catalysis of the oxidative folding of ribonucleaseA by protein disulfide isomerase pre-steady state kinetics and the utilization
  of the oxidizig equivalents of the isomerase. *Biochemistry* 30, 619-625 (1991).
- 65. Crook, E.M., Mathias, A.P. & Rabin, B.R. Spectrophotometric assay of bovine pancereatic ribonuclease-A by the use of cytisine 2'-3'-phosphate. *Biochem. J* 74, 234-238 (1960).
- 66. Ellman, G.L. Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* 82, 70-77 (1959).
- 67. Chae, H. Z., Robison, K., Poole, L. B., Church, G., Storz, G. & Rhee, S. G.Cloning and sequencing of thiol-specific antioxidant from mammalian brain -

alkyl hydroperoxide reductase and thiol-specific antioxidant define a large family of antioxidant enzymes. *Proc. Natl. Acad. Sci. USA* **91**, 7017-7021 (1994).

 Hildebrandt, A.G. & Roots, I. Reduced nicotinamide adenine-dinucleotide phosphate (NADPH)-dependent formation and breakdown of hydrogenperoxide during mixed-function oxidation reactions in liver-microsomes. *Arch. Biochem. Biophys.* 171, 385-397 (1975).

### Chapter 4

# SELENOPROTEIN K FORMS AN INTERMOLECULAR DISELENIDE BOND WITH UNUSUALLY HIGH REDOX POTENTIAL

## 4.1 Introduction

Selenoprotein K (SelK) is a membrane protein linked to development, immunity<sup>1</sup>, longevity<sup>2, 3</sup>, and susceptibility to prostate cancer<sup>4</sup>, but its precise function is unknown<sup>1, 5</sup>. It is a member of a eukaryotic membrane protein family that contributes to the management of oxidative stress<sup>5</sup>, that is the detoxification of reactive oxygen and nitrogen species that can harm cellular components<sup>6</sup>. Members of the family are single-span transmembrane proteins with intrinsically disordered Cterminal domains. SelK is a selenoprotein, a class of enzymes that employ the rare amino acid selenocysteine (Sec, U)<sup>7</sup>. It is localized to the endoplasmic reticulum (ER) membrane, and the Sec (at position 92 of 94 amino acids, scheme 1) is exposed to the cytoplasm<sup>8</sup>. Selenoproteins are often found either to combat reactive oxygen species (ROS) directly or to regulate the cellular response to ROS-related stress<sup>9</sup>. SelK itself has an antioxidant function since it can mitigate the toxicity of ROS both *in vitro*<sup>10</sup> and in live cells<sup>11, 12</sup>.

In addition to its antioxidant function, or perhaps as part of this function, SelK is likely to engage in a signaling pathway. Many bitopic proteins, membrane proteins that transverse the bilayer once, are involved in signaling<sup>13</sup>, and SelK is particularly rich in Pro (15%) and positively charged residues, a composition common in signaling proteins<sup>14</sup>. Furthermore, it contains at least one, possibly two, SH3 recognition

sequences that identify it as a probable signaling protein. Recent reports tied SelK to several binding partners, including palmitoyl acyl transferase ZDHHC6<sup>15</sup> and components of the ER-associated degradation (ERAD) pathway, such as Derlin-1, p97 ATPase, and selenoprotein S (SelS, VIMP)<sup>5, 16, 17</sup>. Finally, SelK mediates Ca<sup>2+</sup> flux in immune cells and is the target of calpain proteases in myeloid cells, a class of proteases involved in regulation of inflammation and immune response<sup>1, 18</sup>.

In the search for SelK's biological function, it is necessary to reconcile two aspects. The majority of selenoproteins are enzymes that operate independently<sup>19</sup>; hence, the presence of Sec suggests an enzymatic function. However, SelK is an intrinsically disordered protein. Such proteins typically function in regulation and signaling but rarely as enzymes<sup>20, 21</sup>. Enzymatic catalysis is typically associated with a well-organized active site in which the degrees of motion are tightly controlled. Nevertheless, folding intermediates of several enzymes were shown to still catalyze their reactions though at lower efficiency, and a few intrinsically disordered proteins were also shown to function in the absence of a protein partner<sup>21, 22</sup>. Indeed, SelK's protein partner SelS, which is also disordered, is an efficient reductase even in isolation<sup>23</sup>. In contrast to SelS, however, SelK's Sec is not in a recognizable redox motif. In most selenoproteins, a vicinal cysteine typically forms a selenenylsulfide bond with the Sec. In SelK the sole Sec is not in a recognizable pattern.

Here we examine whether SelK has similar oxidoreductase activity as that displayed by other selenoproteins: an oxidase, a reductase, an isomerase, or a peroxidase function. Since the *in vivo* system responsible for acting as an electron donor to SelK is unknown, we tested the two most abundant cytoplasmic systems: thioredoxin reductase/thioredoxin (TrxR/Trx) and glutathione reductase/glutaredoxin (GR/Grx)<sup>24</sup> for their ability to reduce SelK. In addition, the SelK redox potential was characterized. Finally, we measured the rate at which it can interact with hydroperoxides in isolation.



Figure 4.1. Schematic representation of SelK as well as the mutants used in this study. (A) SelK's domain organization. In blue, the N terminal domain is located in the ER lumen; In purple, the hydrophobic segment formed by residues 20-42 is predicted to contain a single transmembrane helix (TM); In green, the intrinsically disordered domain facing the cytoplasm. This domain contains an SH3 binding site as well as the catalytic selenocysteine (Sec). The location of the Sec is shown in red. SelK has no additional cysteines or selenocysteines besides the one outlined. (B) Proteins used in this study. SelK U92 was prepared by misloading of cysteyl-tRNA with Sec by a cysteine auxotrophic strain and a media with Sec. About 10% of this sample is SelK U92C generated from remaining pools of Cys in the cell. SelK U92C and SelK U92S were prepared by conventional heterologous expression and are purely the mutant form.

# 4.2 Results

#### 4.2.1 Sec Incorporation into SelK

Sec is the catalytic residue in selenoproteins, and its presence is essential for studies of their enzymatic activity. Incorporation of Sec into proteins can be carried out by several methods<sup>25</sup>, but the most frequently employed approach is heterologous expression in E. coli that harnesses its dedicated selenium incorporation machinery<sup>26</sup>. However, as Sec is encoded by a UGA codon<sup>27</sup>, which is also a stop codon, this approach leads not only to the desired selenoprotein but also to its truncated form $^{28}$ . The truncated form in our preparations still dimerized with the full-length form and cannot be separated. Since that may lead to structural and functional changes, preparation methods based on genetic incorporation of Sec are not a viable procedure for SelK. In its place Sec was incorporated by an alternative method in which Cys in the protein of interest is substituted for Sec when an E. coli cysteine auxotroph strain is grown on media supplemented with Sec<sup>29, 30</sup>. Since Cys and Sec share similar physicochemical properties, Sec will be misloaded onto the cysteyl-tRNA and, thus, will get incorporated into proteins as Sec moiety. SelK has no native Cys residues, which makes it possible to use a SelK U92C mutant to incorporate Sec at position 92 (Figure 4.1A) and obtain the Sec-containing SelK. Since it is impossible to prevent Cys recycling in the cell the sample does contain a small percent of the Cys-containing SelK. However, it is possible to characterize the Cys-containing protein separately to isolate its potential contribution to enzymatic activity. The dominant form in our samples is the homodimer wild-type SelK (Figure 4.2). We label this selenium-rich preparation of the enzyme SelK U92 (Figure 4.1B).



Figure 4.2. SelK contains an intermolecular diselenide bond. (A) SDS–PAGE analysis of SelK, SelK U92C, and SelK U92S run under reducing and non-reducing conditions. Lanes 1–3 were run under non-reducing conditions: lane 1, SelK; lane 2, SelK U92C; lane 3, SelK U92S; lane M, protein molecular mass standards (the molecular mass is noted on the right of the corresponding band). Lanes 4–6 were run under reducing conditions: lane 4, SelK; lane 5, SelK U92C; lane 6, SelK U92S. The arrows point to the monomer and dimer forms of SelK. (B) SelK ionizes as a dimer in electrospray ionization mass spectrum. Molecular masses: 23132 Da for homodimer SelK U92C, and 23225 Da for homodimer SelK.

# 4.2.2 SelK Forms an Intermolecular Diselenide Bond

In DDM micelles, SelK U92 was at minimum a homodimer containing an intermolecular diselenide bond. As can be observed from SDS-PAGE run under non-reducing conditions (Figure 4.2A), SelK U92 migrated as a dimer of 23 kDa, SelK U92C migrated mostly as a dimer, and SelK U92S migrated mostly as a monomer. All three proteins migrated as monomers of 11.5 kDa under reducing conditions. A thiol count using Ellman's reagent confirmed that SelK U92 contained no exposed thiols/selenols. Furthermore, when examined by electrospray mass spectrometry, SelK U92 ionized as a dimer under non-reducing conditions (Figure 4.2B).

#### 4.2.3 SelK's Redox Potential

Numerous bitopic membrane proteins change oligomerization states depending on the protein functional state. Therefore, SelK is likely to engage in self-interactions as well as interactions with protein partners to fulfill its role. This suggests that the intermolecular diselenide bond is physiologically relevant and that a cytoplasmic protein is capable of reducing it. To assess which proteins can potentially reduce SelK, we measured the redox potential of SelK by gel shift assays using maleimide alkylation of U92. A fluorescein tagged maleimide was employed to validate that the protein is indeed alkylated.

SelK U92 was equilibrated in buffers with known redox potentials, the reaction was quenched with acid and SelK U92 was then alkylated with fluorescein-5maleimide and analyzed by SDS-PAGE. It was feasible to measure the ratio of oxidized and reduced (and subsequently alkylated) proteins by SDS-PAGE analysis since the alkylated protein run as a monomer while the oxidized protein run as a dimer (Figure 4.3A). Furthermore, only the alkylated protein will be fluorescent (Figure 4.3B). The ratios were subsequently used to determine the redox potentials of the Secand Cys-containing SelK using the Nernst equation (see Materials and Methods). Using this assay the redox potential of SelK U92 was determined to be -257 mV. We have also validated the results with alkylation with both N-ethylmaleimide (NEM) and iodoacetamide, IAA. The results were independent of the alkylating reagent. Note that a small amount of TEV forms an oligomer with SelK that dissociated as the redox potential is lowered. Also, due to the method employed for Sec incorporation, a small percent of the protein in these samples is the Cys-containing protein (in Figure 4.3A and Figure 4.3B the percent of the Cys-containing SelK is less than 5%). However, because SelK U92C redox potential is -207 mV (Figure 4.3C and Figure 4.3D), all of

it is reduced at redox buffers used to determine the redox potential of the SelK U92. Hence, it does not bias the measurements.



Figure 4.3. The redox potential of SelK's intermolecular diselenide bond is -257 mV. Gel shift assays based on alkylation with fluorescein-5-maleimide. 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 (given in kDa); i TEV protease (bands with higher molecular weights are complexes of TEV protease with SelK); ii oxidized SelK present as a dimer; iii alkylated SelK present as a monomer (bands with lower molecular weights are truncated SelK generated by an *E. coli* protease). (A) Percent of reduced and oxidized wild-type SelK. Visualized by coomassie brilliant blue stain using white light imaging. (B) Alkylation pattern of wild-type SelK. Visualized by fluorescence imaging. Only alkylated protein is detected. (C) Same as A for SelK U92C. (D) Same as B for SelK U92C. (E) Wild-type SelK redox potential titration curve (reduction potential of -257 mV) and (F) SelK U92C redox potential titration curve (reduction potential of  $-207 \pm 1 \text{ mV}$ ). The fraction of reduced protein is plotted against the buffer redox potential.



Figure 4.4. SelK reductase assay. Insulin reductase activity was determined in 100 mM potassium phosphate (pH 7), 2 mM EDTA, 0.067% DDM, 0.13 mM bovine insulin, and 0.33 mM dithiothreitol (DTT). The reaction was initiated by the addition of SelK U92 (blue) or hTrx1 (red). The reaction without an addition of an enzyme served as negative control. The reaction's progress was monitored by recording the increase in turbidity at 650 nm versus time. The temperature was kept constant at 25 °C.



Figure 4.5. SelK peroxidase activity assays. (A) NADPH consumption is monitored in the presence of 8 or 64 nM hTrxR, with and without 5  $\mu$ M hTrx, 200  $\mu$ M H<sub>2</sub>O<sub>2</sub>, and 5  $\mu$ M of SelK U92. The reaction with 5  $\mu$ M SelK U92, 5  $\mu$ M hTrx and 8 nM hTrxR is shown in circles, with 5  $\mu$ M SelK U92 and 64 nM hTrxR is shown in diamonds, with 5  $\mu$ M SelK U92 and 8 nM hTrxR in triangles, and with 5  $\mu$ M hTrx and 8 nM hTrxR but in the absence of SelK in squares. (B) The activity of different mutants was compared using the TrxR/Trx system: 5  $\mu$ M SelK U92 (U92), 5  $\mu$ M SelK U92C (U92C), 5  $\mu$ M SelK U92S (U92S) and in the absence of SelK (No SelK). (C) NADPH consumption is monitored in the presence of 1 mU/ $\mu$ L yGR, 5  $\mu$ M hGrx, 1 mM GSH, with or without 5  $\mu$ M of SelK U92. The reaction with both 5  $\mu$ M SelK and 5  $\mu$ M hGrx is shown in circles, in the absence of SelK in squares, and in the absence of hGrx in triangles. (D) Same as B, but activity was tested using the GR/Grx system. In either system, SelK peroxidase activity depends on the presence of Sec (\*\*\*, P<0.001, n=3).

# 4.2.4 SelK can Reduce Lipid Hydroperoxides

Selenoproteins possess enzymatic activity in practically all characterized cases, with the majority being oxidoreductases. Thus, we tested SelK for reductase and peroxidase activities. SelK does not have disulfide reductase activity in an insulin reduction assay (Figure 4.4)<sup>31</sup>. This assay is based on the aggregation of insulin chain B following the reduction of insulin's intermolecular disulfide bond. It is a general assay for reductase activity as DTT is included as the reducing reagent to generate SelK. SelK also does not have an isomerase or oxidase activity as expected from its redox potential. SelK did exhibit peroxidase activity as assayed by the consumption of NADPH in the presence of human thioredoxin reductase (hTrxR) and human thioredoxin 1 (hTrx) or yeast glutathione reductase (yGR) and human glutaredoxin 1 (hGrx) (Figure 4.5)<sup>32</sup>. We have elected to test the activity using both the Trx and Grx systems as both are highly abundant in the cytoplasm and can theoretically couple with SelK. Indeed, both systems efficiently reduced SelK's diselenide bond, allowing SelK to act as a peroxidase. hTrxR, which is known to exhibit a broad substrate selectivity<sup>33</sup>, reduced SelK directly and does not require Trx (Figure 4.5A). As will be discussed later, hTrxR is the most efficient electron donor. In the GR/Grx system, the NADPH oxidation rate in the presence of SelK and yGR but without hGrx is the same as that of the control, i.e. NADPH oxidation in buffer (Figure 4.5C). This implies that on the time scale of this experiment, SelK U92 is reduced more efficiently by hGrx than by GSH. Since the Grx system and hTrxR alone efficiently reduce SelK U92, we use both systems interchangeably in the remaining characterization of the enzymatic activity.

In both systems, the peroxidase activity was dependent on the presence of Sec, and SelK U92C and SelK U92S mutants did not exhibit significant activity at pH 7.0 (Figure 4.5B and Figure 4.5D). All assays were carried out in DDM since in our previous publication we concluded that SelK behaved best in the presence of this detergent<sup>34</sup>.

# 4.2.5 Substrate Specificity

Because SelK is a membrane protein, we tested whether it is able to reduce hydrophobic substrates. Among the selenium-containing peroxidases, only GPx4, GPx3, and SelP can catalyze the reduction of phospholipid hydroperoxides<sup>35, 36</sup>. We used similar substrates as those employed for the characterization of SelP's substrate specificity<sup>35, 36</sup>. Those include the hydrophilic substrates hydrogen peroxide, tertiary butyl hydroperoxide (tBOOH), cumene hydroperoxide (COOH), and the hydrophobic substrates 15(S)-hydroperoxy-6(E),8(Z),11(Z),14(Z)-eicosatetraenoic acid (HpETE), and 2-linoleoyl-1-palmitoyl-sn-glycero-3-phosphocholine hydroperoxide (PLPCOOH). As Figure 4.6 shows the hydrophobic HpETE and PLPCOOH are the preferred substrates.



Figure 4.6. Substrate specificity of SelK's peroxidase activity. (A) Reaction scheme. (B) NADPH for the reaction with SelK and hGrx (black bars), the reaction excluding hGrx (striped bars), and the reaction excluding SelK (dotted bars). Of the five substrates tested, HpETE and PLPCOOH have the highest reaction rates (\*\*\*, p<0.001, n=3).

# 4.2.6 Steady State Kinetics

To evaluate the catalytic efficiency of SelK we have measured its reaction rate with lipid hydroperoxides. We utilized a steady-state kinetic analysis, as described by  $Dalziel^{37}$ , because peroxidases often exhibit a bi-substrate mechanism without an apparent saturation (see Table 1 of<sup>38</sup>). In this case there is no enzyme-substrate complex and  $V_{max}$  and  $K_m$  are infinite.

To evaluate SelK's catalytic efficiency we have recorded the initial rate for the reduction of HpETE or PLPCOOH as a function of either the concentration of the substrate or the concentration of the reducing substrate (either Grx or TrxR). The primary Dalziel plots, which are the double reciprocal plots of the initial rate versus substrate concentration, displayed characteristic parallel lines of a ping-pong mechanism (Figure 4.7A, C, and E). The reciprocal slope corresponds to the rate constant  $k_1$  for the reduction of hydroperoxides. The rate of the reaction between reduced SelK U92 and the hydrophobic substrates is on the order of  $10^3 \text{ M}^{-1} \text{ s}^{-1}$ , which is on the low range reported for lipid hydroperoxides' peroxidases (see Discussion). The kinetic parameters are listed in Table 4.1.

To evaluate whether the reaction follows saturation or non-saturation kinetics, secondary Dalziel plots were employed. That is the concentration of the enzyme divided by the apparent maximal velocity is plotted against the second substrate, which was either hGrx or hTrxR. Secondary Dalziel plots of unstarurated kinetics have an intercept with the Y axis of zero. SelK displayed saturation kinetics in both Grx/GR and TrxR systems, as shown in Figure 4.7B, D, and F. When the kinetics is saturated, the reciprocal intercept gives the value for the maximum velocity ( $k_{cat}$ ) and the reciprocal slope corresponds to the rate constant  $k_2$  for the reduction of SelK by Grx or TrxR. The rate constant  $k_2$  for hGrx and hTrxR is  $10^5 \text{ M}^{-1} \text{ s}^{-1}$  and  $10^6 \text{ M}^{-1} \text{ s}^{-1}$ , respectively. This shows that both the Grx/GR system and hTrxR alone are proficient at reducing SelK, with the later being more efficient.



Figure 4.7. Dalziel plots for reduction of PLPCOOH catalyzed by SelK by coupling Grx/GR or TrxR reduction system. (A)  $[E_0]/V_0$  vs. 1/[PLPCOOH] at various concentrations of Grx: 0.1  $\mu$ M ( $\bullet$ ), 0.4  $\mu$ M ( $\bullet$ ), 1  $\mu$ M ( $\blacktriangle$ ), and 5  $\mu$ M ( $\bigtriangledown$ ). (B)  $[E_0]/V_0$  vs. 1/[PLPCOOH] at various concentrations of TrxR: 20 nM ( $\odot$ ), 30 nM ( $\Box$ ), 40 nM ( $\triangle$ ), and 80 nM ( $\bigtriangledown$ ). (C)  $[E_0]/V_0$  vs. 1/[HpETE] at various concentrations of TrxR: 10 nM ( $\bullet$ ), 40 nM ( $\blacksquare$ ), 80 nM ( $\blacktriangle$ ), and 120 nM ( $\bigtriangledown$ ). (D) Secondary plot of the intercepts of the primary plot A versus the reciprocal Grx concentrations. (E) Secondary plot of the intercepts of the primary plot B versus the reciprocal TrxR concentrations. (F) Secondary plot of the intercepts of the primary plot C versus the reciprocal TrxR concentrations.

Reducing	ROOH	$k_1$ (ROOH)	$k_2$ (RS)	k <sub>cat</sub>	$K_m$	$k_{\rm cat}(\rm RS)$	K <sub>m</sub>
substrate		(M's')x	$(M^{\prime}s^{\prime})$	(ROOH)	ROOH	(s <sup>-1</sup> )	(RS)
		$10^{3}$	x 10 <sup>6</sup>	$(s^{-1})$	(µM)		(µM)
Grx	PLPCOOH	1.624±	0.193±	0.066±	41.0±	0.066±	0.343±
		0.060	0.021	0.013	8.3	0.009	0.028
TrxR	PLPCCOH	2.556±	1.18±	0.076±	30±	0.076±	0.0640±
		0.047	0.23	0.031	12	0.016	0.0046
	HpETE	1.984±	3.06±	0.201±	101±	0.201±	0.0660±
	_	0.063	0.36	0.080	41	0.027	0.0040

Table 4.1. Kinetic parameters of SelK activity with PLPCOOH as substrate.

# 4.3 Discussion

Here we report that, in isolation, SelK's oligomerization state is at least a homodimer containing an intermolecular diselenide bond. SelK has a short N-terminal single-pass transmembrane helix that resides in the ER membrane, and its transmembrane helix is unusually rich in polar residues (including a Glu and an Asp). This composition is a strong indication that SelK oligomerizes and is likely to associate in the membrane with a protein partner<sup>39</sup>. It is also quite possible that, like other bitopic proteins, it is capable of forming both homo- and hetero-oligomers. Many bitopic proteins dynamically exchange between different oligomerization states<sup>40</sup>. The close proximity of Sec from opposing monomers will facilitate the formation of diselenide bonds across the dimer interface. However, it is also possible that SelK pairs with a protein partner forming a selenylsulfide (Se-S) bond during function.

If such homo-oligomerization takes place under physiological conditions, then SelK represents the first example of a human selenoprotein possessing a diselenide bond. An intramolecular diselenide linkage (UXXU) was shown to occur in selenoprotein L (SelL)<sup>41</sup>, which does not occur in humans. It is also possible, though not yet corroborated, that internal diselenide bonds can form in selenoproteins that have more than one Sec: the selenium storage and delivery enzyme selenoprotein P<sup>42</sup> and *Metridium senile*'s methionine sulfoxide reductase<sup>43</sup>. However, these examples of multiple Sec residues are rare as the majority of selenoproteins utilize only one Sec residue. Thus, SelK's intermolecular diselenide bond stands out as an exception. The redox potential of SelK's intermolecular diselenide bond, -257 mV, is considerably higher than that reported for selenoglutathione (-407 mV<sup>44</sup>); selenocystine (-383 mV<sup>45</sup>); a diselenide bond in a peptide (-381 mV<sup>46</sup>); or selenoproteins with an intramolecular disulfide bond. For example, *Danio rerio* SelL's native UXXU redox motif was not reduced by DTT ( $E^{\circ'} = -327 \text{ mV}$ )<sup>41</sup>. Similarly, the redox potential of non-native *E. coli* Grx3 and Trx, into which the diselenide bond was introduced, had a redox potential of -309 mV (Grx3)<sup>47</sup> and between -280 mV and -327 mV (Trx)<sup>29,47</sup>. Thus, the redox potential of SelK's diselenide bond is higher than values reported so far for a diselenide bond in biological molecules. We propose that the protein molecular environment tailors it to be in the range accessible to the cytoplasmic proteins TrxR, Trx, and Grx. SelK was most efficiently reduced by hTrxR. The observation that hTrxR is capable of reducing a diselenide bond is in agreement with previous reports that it reduces a diselenide bond in ebselen, a GPx mimic<sup>48</sup>.

In regard to SelK's enzymatic activity, our data demonstrate that SelK in micelles has no obvious oxidoreductase function other than a weak peroxidase activity. The peroxidase activity relies on the formation of an intermolecular diselenide bond. In addition, SelK most efficiently reduces lipid hydroperoxides. This is notable, as polyunsaturated fatty acids are a major target of ROS, which leads to disruption of membrane organization and related diseases<sup>49</sup>. However, when the peroxidase activity of SelK is contrasted with other Sec-containing peroxidases, its efficiency is comparatively low. Human selenoproteins with peroxidase activity range from the highly efficient glutathione peroxidases (GPxs) with the Sec reacting with hydroperoxides at a rate of  $10^5$ - $10^7$  M<sup>-1</sup> s<sup>-150, 51</sup>, to moderate peroxidases such as

145

SelS<sup>23</sup>, and TrxR<sup>52, 53</sup>, whose bimolecular rate constant for the reaction of the reduced protein with hydroperoxides are lower than 10<sup>5</sup> M<sup>-1</sup> s<sup>-1</sup>. Accordingly, at least for SelS and TrxR, the peroxidase activity is considered secondary to their primary functions as reductases. Non-selenium peroxidases, like peroxiredoxin 6<sup>54, 55</sup>, were also reported to react with lipid hydroperoxides at a rate of  $10^6 \text{ M}^{-1} \text{ s}^{-1}$ . The reaction resolution (that is regeneration of the enzyme after the reaction with the hydroperoxide) typically takes place at a bimolecular rate of  $10^4$  to  $10^5$  M<sup>-1</sup> s<sup>-150</sup>. The rate of SelK's oxidation by PLPCOOH is  $10^3 \text{ M}^{-1} \text{ s}^{-1}$ . While this value is still within the range exhibited by peroxidases, it calls into question whether SelK's peroxidase activity is its primary enzymatic function. It is possible that since this moderate catalytic efficiency was measured using the isolated SelK in detergent micelles, the activity might increase if SelK interacts with a protein partner or in the presence of its native membrane environment. The mechanisms of lipid hydroperoxides sensing in ROS signaling are not well understood <sup>56</sup> but are clearly critical for health. However, the link between SelK's ability to reduce lipid hydroperoxides and its physiological function remains an open question.

Overall, our work established the redox properties of SelK, the bonding interactions of the sole Sec and that dimerization does not require the presence of the intermolecular diselenide bond. It demonstrated that it is premature to conclusively assign an enzymatic activity in the absence of protein partners and/or the membrane environment. However, SelK can interact with lipid hydroperoxides. Future experiments will examine the role of known protein partners on SelK's oligomerization and activity.

## 4.4 Materials and Methods

#### 4.4.1 Bacterial Strains, Plasmids, and Chemical Reagents

The Homo sapiens SelK (UniProtKB Q9Y6D0) construct used in this study was modified from the SelK-E expression vector described in our previous publication<sup>34</sup>. SelK-E contains a maltose binding protein (MBP) fused to SelK with a Tabaco Etch Virus (TEV) protease cleavage site (ENLYFQG) between the two proteins to release SelK. Here, to assist in purification, an eight amino acid StrepII tag (WSHPQFEK) was inserted between the TEV protease cleavage site and SelK. The first Met of SelK was deleted to reflect the fact that it is proteolytically cleaved in vivo. Following cleavage of the fusion protein by TEV protease, SelK retains in its N-terminal a GWSHPQFEK peptide. Site-directed mutagenesis was used to prepare mutants of this construct. All other sources of bacterial strains, plasmids, chemical reagents and enzymes were as previously described<sup>34, 57</sup>, with the addition of 5,5-dimethyl-1,3cyclohexanedione (dimedone), 5-methyl-2-oxo-4-imidazolidinehexanoic acid (desthiobiotin), sodium deoxycholate, N-ethylmaleimide (NEM), *Glycine max* lipoxidase, yeast glutathione reductase, 2-linoleoyl-1-palmitoyl-sn-glycero-3phosphocholine (PLPC) supplied by Sigma-Aldrich and fluorescein-5-maleimide supplied by Life Technologies. Homo sapiens thioredoxin 1, thioredoxin reductase and glutaredoxin were prepared in house<sup>57</sup>. The expression construct for human glutaredoxin 1 (hGrx) was a gift from J. J. Mieyal<sup>58</sup> and that for human thioredoxin 1 (hTrx) was provided by M. A. Marletta<sup>59</sup>.

#### 4.4.2 Expression and Purification of SelK and its Mutants

SelK U92 was prepared by incorporating Sec into position 92 of a SelK U92C mutant using the cysteine auxotrophic *E. coli* host cell BL21(DE3) *selB::kan* 

 $cys51E^{29}$ . The protocol was similar to that developed by the Böck group with minor modifications<sup>30</sup>. Cells were grown in modified LB media (10 g tryptone, 5 g NaCl, 5 g yeast extract, 2 g glucose per liter), at 37 °C, with good aeration and the relevant antibiotic selection (100  $\mu$ g/ml ampicillin + 25  $\mu$ g/ml kanamycin), and 50  $\mu$ g/ml Lcysteine. When the optical density (OD) at 600 nm reached 1.0, the cells were harvested by centrifugation at 4 °C, 4000g for 10 min. The pellet was washed in sterile buffer (20 mM phosphate, 50 mM NaCl (pH 7.5)) that was chilled on ice, and centrifuged once again. The pellet was resuspended in Studier's MDAG defined media<sup>60</sup> supplemented with 100  $\mu$ g/ml ampicillin and 25  $\mu$ g/ml kanamycin. The volume for the media was the same as the volume of the modified LB media used at the beginning. The cells were allowed to shake at 37 °C for an additional 30 min to consume residual cysteine. The temperature was lowered to 18 °C, and the cells were allowed to shake at the lower temperature for an additional 30 min. Protein expression was induced with 0.5 mM isopropyl-1-thio-b-D-galactopyranoside (IPTG), and 200 uM L-selenocystine was added to the growth media. Accumulation of SelK in the cells stalled after 10 h at 18 °C; hence, the cells were harvested after 14-16 h. The pellet was resuspended in 50 mM sodium phosphate (pH 7.5), 200 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), 0.5% Triton X-100, 1 mM benzamidine, and 1 mM phenylmethylsulfonyl fluoride (PMSF) (amylose Buffer); frozen in liquid nitrogen and stored at -80 °C. Protein expression of SelK U92C and SelK U92S mutants was carried out with LB media as previously described<sup>34</sup>.

Protein purification of SelK U92, SelK U92C, and SelK U92S was as formerly detailed with the additional step of Strep-Tactin affinity chromatography<sup>34</sup>. In short, cells were defrosted and sonicated in amylose buffer. Cell debris was removed by

centrifugation at 20000 g for 1 h. The supernatant was loaded on an amylose column, which was washed with the amylose buffer for 5 column volumes (CV) followed by a wash with 50 mM sodium phosphate (pH 7.5), 200 mM NaCl, 1 mM EDTA, and 0.067% n-dodecyl-b-D-maltopyranoside (DDM) (exchange buffer) for 5 CV. The fusion cMBP-SelK was eluted using 5 CV exchange buffer supplemented with 20 mM maltose. Cleavage of the fusion partner, MBP, was carried out by adding a hexahistidine-tagged TEV protease to the dialysis bag overnight at 4 °C. The TEV protease was added at a molar ratio of 1:10 relative to the fusion MBP-SelK protein. Following cleavage, TEV and MBP were removed using immobilized metal ion affinity chromatography. The protein was then loaded on a 5 ml Strep Trap HP (GE Healthcare) equilibrated with exchange buffer. SelK was eluted in the exchange buffer, with a linear gradient of 0-2.5 mM desthiobiotin. The fractions containing SelK were collected, concentrated to 2 mg/mL, loaded onto a sephacryl 200 (GE Healthcare), and eluted with exchange buffer. Protein purity was analyzed using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and coomassie blue staining. The protein concentrations of SelK U92, SelK U92C, and SelK U92S were determined using an extinction coefficient of 20970 M<sup>-1</sup> cm<sup>-1</sup> or by bicinchoninic acid (BCA) protein assay. The free thiol/selenol count of SelK U92 and SelK U92C was determined using Ellman's reagent 5,5' -dithiobis-(2-nitrobenzoic acid) (DTNB)<sup>61</sup>. The extent of selenium incorporation in a given protein sample was measured by electrospray mass spectrometry as described below. When values were compared to measurements by Inductively Coupled Plasma (ICP) spectroscopy, the agreement was within a factor of 3. We believe that in this specific case mass spectrometry more accurately reflects the actual percent incorporation because ICP

also reported sulfur from residual TEV protease. TEV protease contains several Met residues and, hence, can bias the ratio of sulfur to selenium. The incorporation ratio as judged from mass spectrometry varied from 80% to 90% between batches. We refer to this form as SelK U92. The concentrations given in the paper are those of the selenium containing SelK in a given preparation.

#### 4.4.3 Electrospray Ionization Mass Spectrometry

Mass spectra were obtained using a quadrupole time-of-flight mass spectrometer QTOF Ultima (Waters), operating under positive electrospray ionization (+ESI) mode and connected to an LC-20AD (Shimadzu). 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%, with 0.1% formic acid as the mobile phase, in 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. The obtained spectra were deconvoluted using maximum entropy in MassLynx (Waters).

#### 4.4.4 Determination of Redox Potentials

The redox potential of SelK U92 and SelK U92C were determined by redox titrations with the reduced and oxidized forms of DTT. 25  $\mu$ L SelK U92 (34  $\mu$ M) was added to aliquots of 25  $\mu$ L redox buffer containing 100 mM potassium phosphate (pH 7.0), 1 mM EDTA, 0.067% DDM, and various amounts of reduced and oxidized DTT (to a total DTT concentration of 2 mM). SelK U92 runs as a dimer on SDS-PAGE when oxidized and as a monomer when reduced. The reaction of SelK with DTTred/DTTox redox pair (Reaction I) and the corresponding equilibrium constant

 $(K_{eq})$ , as well as the redox potential (E) of the various DTT buffers are given as follows:

$$\begin{split} & 2 \text{SelK}_{\text{red}} + \text{DTT}_{\text{ox}} \rightleftarrows (\text{SelK}_{\text{ox}})_2 + \text{DTT}_{\text{red}} & \text{Reaction 1} \\ & \text{K}_{\text{eq}} = \frac{[\text{SelK}_{\text{ox}}][\text{DTT}_{\text{red}}]}{[\text{SelK}_{\text{red}}]^2[\text{DTT}_{\text{ox}}]} & \text{Eq. 1} \\ & \text{E} = \text{E}_0(\text{DTT}) - \frac{\text{RT}}{\text{nF}} \times \ln \text{K}_{\text{eq}} & \text{Eq. 2} \end{split}$$

where  $E_0 (DTT_{red}/DTT_{ox}) = -327$  mV at pH 7.0 and 298 K<sup>62</sup>, R is the gas constant (8.315 J K<sup>-1</sup> mol<sup>-1</sup>), T is the absolute temperature (298 K), and n=2. The reaction mixture was equilibrated at 25 °C for 2 h after which the reaction was quenched by adding ice-cold 100% W/V trichloroacetic acid to a final concentration of 20%. The quenched reaction mixture was then spun for 10 min at 16110g. 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 re-suspended in sample loading buffer (50 mM Tris-HCl (pH 6.8), 2% SDS, 12.5 mM EDTA, 10% glycerol) containing 10 mM NEM and analyzed by 16% Tris-glycine SDS-PAGE. Densitometry of coomassie-stained non-reducing gels was performed using ImageJ. The fraction of reduced SelK to oxidized SelK was calculated based on the band intensity (where the dimer's band intensity was halved to obtain the fraction of the dimer). The midpoint redox potential of SelK U92 was calculated by fitting the known redox potential of the buffer and the measured ratio of reduced SelK U92 to oxidized SelK U92. The successful alkylating of SelK U92 was also validated using fluorescein-5-maleimide.

The redox potential measurement of SelK U92C was similar to that described above but with reduced and oxidized glutathione with a total concentration of 2 mM where  $E_0$  (*GSH/GSSG*)= -240 mV<sup>63</sup>.
#### 4.4.5 Peroxidase Activity Assays

Peroxidase activity was measured via a coupled reaction with human thioredoxin reductase/human thioredoxin (TrxR/Trx) or yeast glutathione reductase/human glutaredoxin 1 (GR/Grx) as previously described<sup>32</sup>. The reaction mixture contained 100 mM potassium phosphate (pH 7.0), 0.067% DDM, 2 mM EDTA, 150  $\mu$ M NADPH, and 5  $\mu$ M SelK. For the TrxR/Trx assays, the reaction also included 8 nM hTrxR and 5  $\mu$ M hTrx. For the GR/Grx assays, the reaction also included 1 mU/ $\mu$ L yGR, 1 mM GSH, and 5  $\mu$ M hGrx. The temperature was kept constant at 25 °C. The reaction mixture was incubated for three min, and the reaction was initiated by adding 200  $\mu$ M H<sub>2</sub>O<sub>2</sub> to the cuvette. The enzymatic oxidation of NADPH was monitored spectroscopically by recording its consumption at 340 nm.

To determine the substrate specificity of SelK's peroxidase activity, a variety of potential peroxide substrates were tested in GR/Grx reaction mixture with 0.067% DDM as described above. The reaction was triggered by addition of 60  $\mu$ M H<sub>2</sub>O<sub>2</sub>, cumene hydroperoxide (COOH), tert-Butyl hydroperoxide (tBuOOH), HpETE, or PLPCOOH respectively. HpETE and PLPCOOH were dissolved in the buffer used for the activity assays. All experiments were repeated three times, and the size of the error bar is one standard deviation.

# 4.4.6 Steady State Kinetics

Steady-state kinetic analysis was performed using the GR/Grx peroxidase assay described above. Briefly, the assays were carried out with varying concentrations of  $H_2O_2$  (0 to 600  $\mu$ M), PLPCOOH (0 to 60  $\mu$ M), and hGrx (0 to 5  $\mu$ M). The reaction was initiated by the addition of  $H_2O_2$  or PLPCOOH. The rate of the initial NADPH oxidation was measured by monitoring the absorbance at 340 nm. The reaction mixture without SelK served as a blank. Double reciprocal plots were prepared by plotting  $[E_0]/V_0$  versus 1/[S] and fitted to the transformed Lineweaver-Burk equation (Eq. 3). The plots were used to determine apparent kinetic parameters  $K_m$ ,  $k_{cat}$ . Each condition was repeated three times, and the size of the error bar is one standard deviation.

$$\frac{[E]_0}{V_0} = \frac{K_{Grx}[H_2O_2] + K_{H_2O_2}[Grx] + [Grx][H_2O_2]}{k_{cat}[Grx][H_2O_2]}$$
Eq. 3

#### REFERENCES

- Verma, S., Hoffmann, F. W., Kumar, M., Huang, Z., Roe, K., Nguyen-Wu, E., Hashimoto, A. S. & Hoffmann, P. R. Selenoprotein K knockout mice exhibit deficient calcium flux in immune cells and impaired immune responses. *J. Immunol.* 186, 2127-2137 (2011).
- Morozova, N., Forry, E. P., Shahid, E., Zavacki, A. M., Harney, J. W., Kraytsberg, Y. & Berry, M. J. Antioxidant function of a novel selenoprotein in Drosophila melanogaster. *Genes Cells* 8, 963-971 (2003).
- Lai, C.Q., Parnell, L.D., Lyman, R.F., Ordovas, J.A. & Mackay, T.F.C.
   Candidate genes affecting Drosophila life span identified by integrating microarray gene expression analysis and QTL mapping. *Mech. Ageing Dev.* 128, 237-249 (2007).
- Meplan, C., Rohrmann, S., Steinbrecher, A., Schomburg, L., Jansen, E., Linseisen, J. & Hesketh, J. Polymorphisms in thioredoxin reductase and selenoprotein k genes and selenium status modulate risk of prostate cancer. *Plos One* 7 (2012).
- Shchedrina, V. A., Everley, R. A., Zhang, Y., Gygi, S. P., Hatfield, D. L. & Gladyshev, V. N.. Selenoprotein K binds multiprotein complexes and is

involved in the regulation of endoplasmic reticulum homeostasis. *J. Biol. Chem.* **286**, 42937-42948 (2011).

- Finkel, T., Serrano, M. & Blasco, M.A. The common biology of cancer and ageing. *Nature* 448, 767-774 (2007).
- 7. Lu, J. & Holmgren, A. Selenoproteins. J. Biol. Chem. 284, 723-727 (2009).
- Chen, C. L., Shim, M. S., Chung, J., Yoo, H. S., Ha, J. M., Kim, J. Y., Choi, J., Zang, S. L., Hou, X., Carlson, B. A., Hatfield, D. L. & Lee, B. J. G-rich, a Drosophila selenoprotein, is a Golgi-resident type III membrane protein. *Biochem. Biophys. Res. Commun.* 348, 1296-1301 (2006).
- Hatfield, D.L., Tsuji, P.A., Carlson, B.A. & Gladyshev, V.N. Selenium and selenocysteine: roles in cancer, health, and development. *Trends Biochem Sci* 39, 112-120 (2014).
- Roseler, A., Prieto, J. H., Iozef, R., Hecker, B., Schirmer, R. H., Kulzer, S., Przyborski, J., Rahlfs, S. & Becker, K. Insight into the selenoproteome of the malaria parasite plasmodium falciparum. *Antioxidants & redox signaling* 17, 534-543 (2012).
- Lu, C. L., Qiu, F. C., Zhou, H. J., Peng, Y., Hao, W., Xu, J. L., Yuan, J. G., Wang, S. Z., Qiang, B. Q., Xu, C. M. & Peng, X. Z. Identification and characterization of selenoprotein K: An antioxidant in cardiomyocytes. *FEBS Lett.* 580, 5189-5197 (2006).

- Du, S.Q., Zhou, J., Jia, Y. & Huang, K.X. SelK is a novel ER stress-regulated protein and protects HepG2 cells from ER stress agent-induced apoptosis. *Arch. Biochem. Biophys.* 502, 137-143 (2010).
- Hubert, P., Sawma, P., Duneau, J. P., Khao, J., Henin, J., Bagnard, D. & Sturgis, J. Single-spanning transmembrane domains in cell growth and cell-cell interactions More than meets the eye? *Cell Adh. Migr.* 4, 313-324 (2010).
- Shchedrina, V.A., Zhang, Y., Labunskyy, V.M., Hatfield, D.L. & Gladyshev, V.N. Structure-function relations, physiological roles, and evolution of mammalian ER-resident selenoproteins. *Antioxid. Redox Signal.* 12, 839-849 (2010).
- Meiler, S., Baumer, Y., Huang, Z., Hoffmann, F. W., Fredericks, G. J., Rose,
   A. H., Norton, R. L., Hoffmann, P. R. & Boisvert, W. A. Selenoprotein K is required for palmitoylation of CD36 in macrophages: implications in foam cell formation and atherogenesis. *J. Leukocyte Biol.* 93, 771-780 (2013).
- Christianson, J. C., Olzmann, J. A., Shaler, T. A., Sowa, M. E., Bennett, E. J., Richter, C. M., Tyler, R. E., Greenblatt, E. J., Harper, J. W. & Kopito, R. R. Defining human ERAD networks through an integrative mapping strategy. *Nature Cell Biology* 14, 93-105 (2012).
- Turanov, A. A., Shchedrina, V. A., Everley, R. A., Lobanov, A. V., Yim, S. H., Marino, S. M., Gygi, S. P., Hatfield, D. L. & Gladyshev, V. N. Selenoprotein S is involved in maintenance and transport of multiprotein complexes. *Biochem J* 462, 555-565 (2014).

- Huang, Z., Hoffmann, F.W., Norton, R.L., Hashimoto, A.C. & Hoffmann, P.R. Selenoprotein K Is a novel target of m-calpain, and cleavage is regulated by toll-like receptor-induced calpastatin in macrophages. *J. Biol. Chem.* 286, 34830-34838 (2011).
- Labunskyy, V.M., Hatfield, D.L. & Gladyshev, V.N. Selenoproteins: molecular pathways and physiological roles. *Physiol Rev* 94, 739-777 (2014).
- 20. Uversky, V.N. A decade and a half of protein intrinsic disorder: Biology still waits for physics. *Protein Sci.* **22**, 693-724 (2013).
- Vendruscolo, M. Enzymatic activity in disordered states of proteins. *Current Opinion in Chemical Biology* 14, 671-675 (2010).
- Zambelli, B., Cremades, N., Neyroz, P., Turano, P., Uversky, V. N. & Ciurli,
  S. Insights in the (un)structural organization of Bacillus pasteurii UreG, an intrinsically disordered GTPase enzyme. *Molecular Biosystems* 8, 220-228 (2012).
- 23. Liu, J., Li, F. & Rozovsky, S. The intrinsically disordered membrane protein selenoprotein S is a reductase in vitro. *Biochemistry* **52**, 3051-3061 (2013).
- Holmgren, A., Johansson, C., Berndt, C., Lonn, M. E., Hudemann, C. & Lillig,
  C. H. Thiol redox control via thioredoxin and glutaredoxin systems. *Biochem Soc Trans* 33, 1375-1377 (2005).
- Rozovsky, S. in Biochalcogen Chemistry: The Biological Chemistry of Sulfur, Selenium, and Tellurium. (eds. C.A. Bayse & J.L. Brumaghim) 127-142 (ACS Press, 2013).

- Johansson, L., Gafvelin, G. & Arner, E.S.J. Selenocysteine in proteins properties and biotechnological use. *Biochim. Biophys. Acta* 1726, 1-13 (2005).
- Yoshizawa, S. & Bock, A. The many levels of control on bacterial selenoprotein synthesis. *Biochim. Biophys. Acta* 1790, 1404-1414 (2009).
- 28. Arner, E.S.J., Sarioglu, H., Lottspeich, F., Holmgren, A. & Bock, A. Highlevel 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. *Journal* of Molecular Biology 292, 1003-1016 (1999).
- Muller, S., Senn, H., Gsell, B., Vetter, W., Baron, C. & Bock, A. The formation of diselenide bridges in proteins by incorporation of selenocysteine residues biosynthesis and characterization of (Se)(2)-thioredoxin. *Biochemistry* 33, 3404-3412 (1994).
- Strub, M. P., Hoh, F., Sanchez, J. F., Strub, J. M., Bock, A., Aumelas, A. & Dumas, C. Selenomethionine and selenocysteine double labeling strategy for crystallographic phasing. *Structure* 11, 1359-1367 (2003).
- 31. Holmgren, A. Thioredoxin catalyzes the reduction of insulin disulfides by dithiothreitol and dihydrolipoamide. *J. Biol. Chem.* **254**, 9627-9632 (1979).
- 32. Chae, H. Z., Robison, K., Poole, L. B., Church, G., Storz, G. & Rhee, S. G. Cloning and sequencing of thiol-specific antioxidant from mammalian brain alkyl hydroperoxide reductase and thiol-specific antioxidant define a large

family of antioxidant enzymes. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 7017-7021 (1994).

- Gromer, S., Urig, S. & Becker, K. The thioredoxin system From science to clinic. *Med. Res. Rev.* 24, 40-89 (2004).
- Liu, J., Srinivasan, P., Pham, D.N. & Rozovsky, S. Expression and purification of the membrane enzyme selenoprotein K. *Protein Expression and Purification* 86, 27-34 (2012).
- Saito, Y., Hayashi, T., Tanaka, A., Watanabe, Y., Suzuki, M., Saito, E. & Takahashi, K. Selenoprotein P in human plasma as an extracellular phospholipid hydroperoxide glutathione peroxidase - Isolation and enzymatic characterization of human selenoprotein P. *Journal of Biological Chemistry* 274, 2866-2871 (1999).
- Takebe, G., Yarimizu, J., Saito, Y., Hayashi, T., Nakamura, H., Yodoi, J., Nagasawa, S. & Takahashi, K. A comparative study on the hydroperoxide and thiol specificity of the glutathione peroxidase family and selenoprotein P. *J. Biol. Chem.* 277, 41254-41258 (2002).
- Dalziel, K. Initial steady state velocities in the evaluation of enzymecoenzyme-substrate reaction mechanisms. *Acta Chem. Scand.* 11, 1706-1723 (1957).
- Toppo, S., Flohe, L., Ursini, F., Vanin, S. & Maiorino, M. Catalytic mechanisms and specificities of glutathione peroxidases: Variations of a basic scheme. *Biochim. Biophys. Acta* 1790, 1486-1500 (2009).

- Gayen, S., Li, Q.X. & Kang, C.B. Solution NMR study of the transmembrane domain of single-span membrane proteins: opportunities and strategies. *Current Protein & Peptide Science* 13, 585-600 (2012).
- 40. Vostrikov, V.V., Mote, K.R., Verardi, R. & Veglia, G. Structural dynamics and topology of phosphorylated phospholamban homopentamer reveal its role in the regulation of calcium transport. *Structure* **21**, 2119-2130 (2013).
- Shchedrina, V.A., Novoselov, S.V., Malinouski, M.Y. & Gladyshev, V.N. Identification and characterization of a selenoprotein family containing a diselenide bond in a redox motif. *Proceedings of the National Academy of Sciences of the United States of America* 104, 13919-13924 (2007).
- Ma, S.G., Hill, K.E., Burk, R.F. & Caprioli, R.M. Mass spectrometric determination of selenenylsulfide linkages in rat selenoprotein P. *Journal of Mass Spectrometry* 40, 400-404 (2005).
- 43. Lee, B. C., Lobanov, A. V., Marino, S. M., Kaya, A., Seravalli, J., Hatfield, D. L. & Gladyshev, V. N. A 4-Selenocysteine, 2-Selenocysteine Insertion Sequence (SECIS) element methionine sulfoxide reductase from metridium senile reveals a non-catalytic function of selenocysteines. *Journal of Biological Chemistry* 286, 18747-18755 (2011).
- 44. Beld, J., Woycechowsky, K.J. & Hilvert, D. Selenoglutathione: Efficient oxidative protein folding by a diselenide. *Biochemistry* 46, 5382-5390 (2007).
- 45. Nauser, T., Dockheer, S., Kissner, R. & Koppenol, W.H. Catalysis of electron transfer by selenocysteine. *Biochemistry* **45**, 6038-6043 (2006).

- Besse, D., Siedler, F., Diercks, T., Kessler, H. & Moroder, L. The redox potential of selenocystine in unconstrained cyclic peptides. *Angew. Chem.* 36, 883-885 (1997).
- Metanis, N., Keinan, E. & Dawson, P.E. Synthetic seleno-glutaredoxin 3 analogues are highly reducing oxidoreductases with enhanced catalytic efficiency. J. Am. Chem. Soc. 128, 16684-16691 (2006).
- Zhao, R. & Holmgren, A. A novel antioxidant mechanism of ebselen involving ebselen diselenide, a substrate of mammalian thioredoxin and thioredoxin reductase. *J. Biol. Chem.* 277, 39456-39462 (2002).
- Niki, E. Lipid peroxidation: Physiological levels and dual biological effects.
   *Free Radical Biol. Med.* 47, 469-484 (2009).
- Saito, Y., Hayashi, T., Tanaka, A., Watanabe, Y., Suzuki, M., Saito, E. & Takahashi, K. Selenoprotein P in human plasma as an extracellular phospholipid hydroperoxide glutathione peroxidase. Isolation and enzymatic characterization of human selenoprotein p. *The Journal of biological chemistry* 274, 2866-2871 (1999).
- 51. Takebe, G., Yarimizu, J., Saito, Y., Hayashi, T., Nakamura, H., Yodoi, J., Nagasawa, S. & Takahashi, K. A comparative study on the hydroperoxide and thiol specificity of the glutathione peroxidase family and selenoprotein P. *The Journal of biological chemistry* 277, 41254-41258 (2002).
- 52. Zhong, L.W. & Holmgren, A. Essential role of selenium in the catalytic activities of mammalian thioredoxin reductase revealed by characterization of

recombinant enzymes with selenocysteine mutations. *J. Biol. Chem.* **275**, 18121-18128 (2000).

- 53. Bjornstedt, M., Hamberg, M., Kumar, S., Xue, J. & Holmgren, A. Human thioredoxin reductase directly reduces lipid hydroperoxides by nadph and selenocystine strongly stimulates the reaction via catalytically generated selenols. *J. Biol. Chem.* **270**, 11761-11764 (1995).
- 54. Fisher, A.B., Dodia, C., Manevich, Y., Chen, J.W. & Feinstein, S.I.
  Phospholipid hydroperoxides are substrates for non-selenium glutathione
  peroxidase. *The Journal of biological chemistry* 274, 21326-21334 (1999).
- 55. Manevich, Y., Shuvaeva, T., Dodia, C., Kazi, A., Feinstein, S. I. & Fisher, A.
  B. Binding of peroxiredoxin 6 to substrate determines differential phospholipid hydroperoxide peroxidase and phospholipase A(2) activities. *Arch. Biochem. Biophys.* 485, 139-149 (2009).
- Girotti, A.W. Translocation as a means of disseminating lipid hydroperoxideinduced oxidative damage and effector action. *Free Radic. Biol. Med.* 44, 956-968 (2008).
- Liu, J. & Rozovsky, S. Contribution of selenocysteine to the peroxidase activity of selenoprotein S. *Biochemistry* 52, 5514-5516 (2013).
- Jao, S. C., Ospina, S. M. E., Berdis, A. J., Starke, D. W., Post, C. B. & Mieyal,
  J. J. Computational and mutational analysis of human glutaredoxin (thioltransferase): Probing the molecular basis of the low pK(a) of cysteine 22 and its role in catalysis. *Biochemistry* 45, 4785-4796 (2006).

- 59. Mitchell, D.A., Morton, S.U., Fernhoff, N.B. & Marletta, M.A. Thioredoxin is required for S-nitrosation of procaspase-3 and the inhibition of apoptosis in Jurkat cells. *Proceedings of the National Academy of Sciences of the United States of America* 104, 11609-11614 (2007).
- 60. Studier, F.W. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* **41**, 207-234 (2005).
- Ellman, G.L. Tissue sulfhydryl groups. *Archives of Biochemistry and Biophysics* 82, 70-77 (1959).
- Lees, W.J. & Whitesides, G.M. Equilibrium-constants for thiol disulfide interchange reactions - a coherent, corrected set. *J. Org. Chem.* 58, 642-647 (1993).
- 63. Aslund, F., Berndt, K.D. & Holmgren, A. Redox potentials of glutaredoxins and other thiol-disulfide oxidoreductases of the thioredoxin superfamily determined by direct protein-protein redox equilibria. *Journal of Biological Chemistry* 272, 30780-30786 (1997).

#### Chapter 5

# EXPRESSED SELENOPROTEIN LIGATION HARNESS SELENOCYSTEINE REACTIVITY TO EXPAND SCOPE OF NATIVE CHEMICAL LIGATION

# 5.1 Introduction

### 5.1.1 Preparation of Selenoproteins is Challenging

Selenium is an essential micronutrient for mammals. It is present in selenoproteins where it exerts its role in the form of selenocysteine (Sec). Selenoproteins are a special class of enzymes that serve various biological functions<sup>1</sup>, such as thyroid hormone regulation, calcium regulation, redox homeostasis, and neuroprotective functions. Thus they have been implicated in many human diseases including cardiovascular diseases, immune dysfunction, cancer, and neurological disorders<sup>2</sup>. The biosynthesis of selenoproteins requires dedicated genetic machinery that is unique compared to canonical protein translation. It requires its own tRNA<sup>sec</sup> that is first charged with serine by seryl-tRNA synthetase and then converted to SectRNA<sup>Sec</sup> by Sec synthase. The decoding of the Sec codon (UGA, also a stop codon) also requires the help of a Sec inserting sequence (SECIS) in its mRNA and ancillary proteins<sup>3</sup>. 25 genes, encoding selenoproteins, have been identified in the human genome<sup>4</sup>. In contrast to their established biological study, the biochemical and structural characterization of selenoproteins has lagged due to the challenge of preparing sufficient amounts of them<sup>5</sup>. Fortunately, great progress has been made in

the past two decades in preparing selenoproteins<sup>6</sup>, which will be discussed in the following sections.

### 5.1.2 Preparation of Selenoproteins whose Sec is Close to the C-terminus

Selenoproteins can be prepared by utilizing Sec's specific genetic incorporation machinery<sup>7</sup>. The selenoprotein gene can be expressed by fusing it with the bacterial selenoprotein formate dehydrogenase F SECIS element, and can then be cotransformed into *E. coli* with the SelA, SelB, and SelC genes (coding for Sec synthase (SelA), Sec-specific elongation factor (SelB), and tRNA<sup>sec</sup> (SelC) respectively)<sup>7</sup>. The genetic incorporation method is used routinely for producing selenoproteins whose Sec is located close to the C-terminus of the protein of interest (POI)<sup>8, 9</sup>. However, additional mutations have to be introduced for the production of selenoproteins whose Sec is in the middle or close to the N-terminus due to the fidelity of the *E. coli* genetic machinery<sup>10, 11</sup>. The Sec incorporation efficiency is also quite low because the Sec codon, UGA, is also decoded as a termination signal. In addition, separating the full length from the truncated form can be demanding considering that typically there is only a difference of several amino acids in length between them. Thus, achieving a homogeneous sample for further characterization of the protein is challenging.

#### 5.1.3 Preparation of Selenoproteins Using a Cysteine (Cys) Auxotrophic Strain

Selenoproteins can also be prepared by using a Cys auxotrophic strain supplied with Sec instead of Cys during protein translation<sup>12</sup>. In bacteria, L-Cys is converted from L-serine by the enzyme serine acetyltransferase (SAT) that is coded by the cysE gene. A Cys auxotrophic strain with a CysE knocked out cannot produce Cys but only absorb Cys from the surrounding medium<sup>13, 14</sup>. Due to the biochemical similarities between Sec and Cys, Sec can over-compete Cys for charging on cysteinyl-tRNA<sup>Cys</sup>. In this way, Sec can be randomly misloaded into any Cys position. Due to the simplicity of operation and the good yield, this method has been used to engineer selenoenzymes<sup>15</sup> as well as for production of mutant glutathione peroxidases (GPxs), where all Cys are mutated to serine so that Sec is only incorporated into the catalytic site<sup>16-19</sup>. A similar method that uses a non-Cys auxotrophic BL21(DE3) strain, with similar efficiency, has been adapted to incorporate Sec or both Sec and selenomethione into selenoproteins for crystallography determination<sup>20-22</sup>. But, incorporation with a Cys auxotrophic strain is not suitable for specific biochemical characterization of selenoproteins that have extra Cys residues. Most human selenoproteins normally contain more than one Cys residue in addition to Sec. The additional Cys residues might be part of the active site and essential for enzymatic activities<sup>23-25</sup>, maintaining structural stability<sup>26</sup>, or interacting with its protein partners<sup>27</sup>. Thus this method is not generally used for functional and structural characterization of the majority of naturally found selenoproteins.

# 5.1.4 Incorporation of Sec with an Engineered tRNA<sup>Sec</sup>

Recently, novel protein engineering methods have also been reported. In one study, a synthetically evolved tRNA<sup>Sec</sup> was engineered by random mutagenesis that could be misacylated with serine by the endogenous seryl-tRNA synthetase (SerRS), followed by its conversion to Sec with the catalysis by Sec synthase from the selenium donor, selenophosphate. The engineered tRNA<sup>Sec</sup> can also be recognized by elongation factor Tu (EF-Tu) for incorporation of Sec through the canonical translational pathway<sup>28, 29</sup>. Thus, using this method eliminates the burden of expressing SelB and

SECIS element, and is very specific. However, the amber codon UAG or opal codon UGA (also a Sec codon) used in this study can also be decoded as Ser or stop signal, and thus the sample is a mixture. Therefore, the yield is poor and requires further optimization by site directed evolution of tRNA<sup>Sec</sup> to better interface it with SerRS, and EF-Tu.

# 5.1.5 Incorporation of Sec through an Orthogonal tRNA/aaRS Pair

In a different study, a Sec analogue photocaged Sec precursor had been genetically incorporated into green fluorescent protein (GFP) with an orthogonal tRNA/aminoacyl tRNA synthetase (tRNA/aaRS) pair through amber suppression codon in yeast cells<sup>30</sup>. The photocaged Sec containing GFP can be converted to Sec containing GFP upon *in situ* UV illumination. Though this method opens new possibilities for biochemical study of selenoproteins, synthesis of photocaged Sec can be a challenge, and it is not clear whether the photocaged Sec can interfere with proper folding in additional targets. The study needs to be expanded to demonstrate its general application.

### 5.1.6 Sec-mediated Native Chemical Ligation

In contrast to the biochemical method of selenoprotein production, selenoproteins can be prepared with chemical approaches by Sec-mediated native chemical ligation (NCL). Sec-mediated NCL is a method that resembles Cys-mediated NCL due to the similarities between Sec and Cys<sup>31</sup>. It relies on an N-terminal Sec from one fragment attacking the C-terminal thioester from the complementary fragment and is subsequently followed by a Se $\rightarrow$ N acyl shift to generate the native amide bond<sup>32</sup>(Figure 5.1). The ligation rate of Sec-mediated NCL is faster and more pH tolerant than Cys-mediated ligation<sup>32</sup>. Sec-mediated NCL was proposed to be an efficient way to make selenoprotein because it does not need Sec genetic incorporation machinery<sup>33, 34</sup>. A few artificial selenoproteins, as well as selenopeptides, were made in this way<sup>35-38</sup>. In preparation of selenoproteins, the thioester was produced by fusing the N-terminal fragment of the protein of interest (POI) with an intein at its C-terminus or directly synthesized by solid-phase peptide synthesis (SPPS). While the production of the C-terminal fragment of the POI that bears an N-terminal Sec was prepared by SPPS. Secmediated NCL was also used to prepare dehydroalanine-containing peptides that have potential applications for protein conjugation<sup>39</sup>. Recently, Sec-mediated NCL has regained people's interest in terms of its ease of deselenization following tris(2-carboxyethyl)phosphine (TCEP) treatment, and thus expands ligation sites at Ala, Phe, Pro, and Ser from its Sec derivative<sup>40-49</sup>.

However, mouse TrxR whose Sec is located in the penultimate position is the only naturally found selenoprotein that has been semisynthesized by NCL<sup>50</sup>. The wide usage of this method is hampered by the challenge of synthesizing long N-terminal Sec peptides<sup>51</sup>. In addition, selenopeptide synthesis itself is difficult, such as the challenge of minimizing racemization during coupling, and the need to prevent  $\beta$ -elimination<sup>34, 50</sup>, though progress has been achieved to minimize these problems<sup>52, 53</sup>.



Limitation of peptide folding or length No limitation of peptide folding or length Figure 5.1. Sec-mediated native chemical ligation. Left: Previous work used solid phase peptide synthesis (SPPS) that may be limited by peptide length and properties; Right: This work introduced a TEV protease cleavage site, and restricted the Sec peptide in a specific fragment and was released after TEV protease cleavage.

# 5.1.7 Expressed Selenoprotein Ligation

Hence, inventing new ways to make sufficient amounts of N-terminal Sec peptide will boost the applicability of Sec-mediated NCL. Inspired by the structural similarity between Sec and Cys<sup>54</sup>, and the high tolerance of most of the 20 amino acids, including Cys in P1' position of tobacco etch virus protease (TEV) recognition site  $(ENLYFQ/X(P1'))^{55, 56}$ , we thus hypothesized that a substrate with Sec in the P1' position may also be efficiently cleaved by TEV protease and thus release an Nterminal Sec peptide for expressed protein ligation (Figure 5.1), hereon referred to as ESL. It has already been reported that all Cys in thioredoxin can be randomly substituted for Sec, with a good yield and high efficiency by overexpressing with a Cys auxotrophic strain<sup>12</sup>.

Selenoproteins usually have additional Cys residues besides the Sec, though Cys is less abundant in eukaryotes. In order to overcome the issue of non-specificity, we isolated the N-terminal Sec peptide that does not contain any Cys (Figure 5.1) but only Sec. The Sec was incorporated into the protein by feeding the bacteria with L-selenocystine during protein expression. A cytosolic maltose binding protein (MBP) that does not contain any cysteines was fused to the N-terminal of the N-terminal Sec fragment to promote protein expression and solubility<sup>9, 57</sup>. For the convenience of removal of MBP later, we also included a six-histidine tag at the N-terminal MBP, and the TEV recognition site (ENLFQC) was introduced at its C-terminus. To simplify the procedure of protein expression for Sec incorporation we optimized a previously reported protocol that uses a BL21(DE3) strain<sup>21</sup>. Here we demonstrated the wide application of ESL by preparing native selenoprotein M and selenoprotein W, as well as non-selenoproteins, and introducing facile chemical conjugations.

### 5.2 Results

# 5.2.1 Incorporation of Sec into MBP-SelM<sup>CT</sup> and TEV Protease Cleavage

To demonstrate the method capabilities, a Trx fold like selenoprotein M (SelM), an endoplasmic reticulum (ER) localized, with a yet unknown function, was selected<sup>58, 59</sup>. For stabilization during expression the N-terminal Sec fragment of SelM (SelM<sup>CT</sup>) is fused with maltose binding protein (MBP). This inhibits its aggregation and increases both yield and solubility. Sec was incorporated using a modified cysteinyl-tRNA misloading protocol (Table 5.2). Sec could substitute Cys up to more

than 90% as MBP-ENLYFQU-SelM<sup>CT</sup> in general after amylose column purification (Figure 5.2). After purification, MBP-ENLYFQU-SelM<sup>CT</sup> dimerized through diselenide bonds that masked the cleavage site of TEV protease. Indeed, there was almost no cleavage in the absence of DTT, but more than 95% of MBP-ENLYFQU-SelM<sup>CT</sup> was successfully cleaved in the presence of 10 mM DTT (Figure 5.3 and 5.4).



Figure 5.2. MBP-SelM<sup>CT</sup> characterization by mass spectrometry. Deconvoluted electrospray ionization mass spectrum of reduced MBP-SelM<sup>CT</sup>. The calculated molecular mass of MBP-SelM<sup>CT</sup> is 53674 Da. The 53933 Da form is due to a 6-phosphogluconolactone modification, most likely on His<sub>6</sub>-tag at the N-terminal of MBP<sup>60</sup>. Inset shows the M/Z spectrum.



Figure 5.3. Cleavage of a Sec-containing TEV recognition sequence in MBP-SelM<sup>CT</sup> by TEV protease in the presence of reductant. (A) Cleavage by TEV protease is fastest when small amino acids such as Gly or Ser are placed after the Gln of the recognition sequence. It is less efficient when Glu or Asp follow the Gln. Here we demonstrate that TEV protease can cleave when Sec is at the end of its recognition sequence ENLYFQU. The recognition site was placed between the MBP tag and POI<sup>CT</sup> to release N-terminal Sec peptide. (B) MBP-SelM<sup>CT</sup> can be efficiently cleaved by TEV protease in the presence of reductant (here DTT). Proteins are identified to the right of a reducing 16% Tris-Glycine SDS-PAGE. MBP-SelM<sup>CT</sup> was incubated with TEV protease at a molar ratio of 5:1 in 50 mM sodium phosphate (pH 7.5), 200 mM NaCl, 2 mM EDTA, and 10 mM DTT.



Figure 5.4. Cleavage of a Sec-containing TEV recognition sequence in MBP-SelM<sup>CT</sup> by TEV protease in the absence of reductant. MBP-SelM<sup>CT</sup> was not cleaved by TEV protease in the absence of reductant. The failure of TEV protease to cleave is most likely due to the formation of intermolecular diselenide bonds in MBP-SelM<sup>CT</sup>, which masked the cleavage site. MBP-SelM<sup>CT</sup> was incubated with TEV protease at a molar ratio of 5:1 in 50 mM sodium phosphate (pH 7.5), 200 mM NaCl, and 2 mM EDTA. Proteins are identified on a 16% Tris-Glycine SDS-PAGE with reducing agents (right of the protein markers) or without (left of the protein markers) reducing agents.

# 5.2.2 MBP-SelM<sup>NT</sup> Thioester Preparation

For the preparation of MBP-SelM<sup>NT</sup> thioester, we used intein technology. Briefly, we fused MBP-SelM<sup>NT</sup> at the N-terminal yeast *Sce* VMA intein that contains a chitin-binding domain (Figure 5.5A). The whole fusion MBP-SelM<sup>NT</sup>-VMA was purified by amylose column, and the crude elutant from the amylose column was pooled onto a chitin column. The MBP-SelM<sup>NT</sup> thioester was collected after being cleaved by sodium 2-mercaptoethanesulfonate, and then detected by mass spectrometry (Figure 5.5B).



Figure 5.5. MBP-SelM<sup>NT</sup> thioester characterization by mass spectrometry. (A) Scheme demonstrating MBP-SelM<sup>NT</sup> thioester production. MBP-SelM<sup>NT</sup>-VMA fusion was first purified by amylose column. It was then bound to a chitin column due to the presence of a chitin-binding domain fused at the C-terminal of the intein. During cleavage of the intein, the thioester form of MBP-SelM<sup>NT</sup> (MBP-SelM<sup>NT</sup>-MES) was formed with the addition of 75 mM MESNA. (B) The thioester form of MBP-SelM<sup>NT</sup> purified by a chitin column was detected by mass spectrometry. The theoretical mass of the MBP-SelM<sup>NT</sup>-MES is 45322 Da. The 42809 Da peak arises from a truncation. Inset shows the M/Z spectrum.



Figure 5.6. SelM U48C purification and characterization by mass spectrometry. (A) The purification of SelM U48C prepared by heterologous expression analyzed by 16% Tris-Glycine SDS-PAGE. Lane 1: cell lysate; lane 2: following purification by amylose affinity chromatography; lane 3: TEV cleavage mixture; lane 4: purified SelM U48C after size exclusion chromatography; lane M, protein molecular mass standards (the molecular mass in kDa is noted on the right). (B) Deconvoluted electrospray ionization mass spectrum of SelM U48C. The theoretical molecular mass of SelM U48C is 14121 Da. Inset shows the M/Z spectrum.

#### 5.2.3 Expressed Selenoprotein Ligation of SelM

The ESL of SelM was carried out by incubating the MBP-SelM<sup>CT</sup> TEV cleavage mixture with the MBP-SelM<sup>NT</sup> thioester (Figure 5.7A). After a day, over 90% ligation was seen (Figure 5.7E). The ligated SelM was further purified to homogeneity and characterized by intact mass spectrometry (Figure 5.7F) as well as tandem MS sequencing (Figure 5.7G and 5.7H). The circular dichroism (CD) spectra of SelM before and after refolding and that of SelM U48C prepared by conventional heterologous expression (Figure 5.6) exhibited a high degree of similarity when superimposed (Figure 5.7I). This indicates that, after ligation, SelM folded spontaneously to adopt its native fold.



Figure 5.7. (A) Principle of expressed selenoprotein ligation. The seleno-fragment of interest is expressed in *E. coli* as a fusion with MBP in defined growth medium enriched with selenocystine. The complementary thioester protein fragment is prepared by established intein technology. Immediately prior to ligation with the thioester-fragment, the seleno-fragment is cleaved from MBP by TEV protease. The selenolate attacks the thioester and the native protein is generated spontaneously through an amide bond formation. "POI" stands for protein of interest. (B)

Downstream applications of Sec-mediated ligation: Sec is readily deselenized into an alanine (anaerobically, blue) or serine (aerobically, pink) by TCEP generating a nonselenoprotein; oxidative elimination generates dehydroalanine, an electrophilic tag for chemical labelling; reactions with thiols and selenolates introduce chemical probes and post-translational modifications via the selenvlsulfide, diselenide or selenoether bonds. Panels C-H) ESL preparation of SelM. (C) The minimal Trx fold of Mus musculus SelM (PDB entry 2A2P). Red: thioester fragment: Green: seleno-fragment: Yellow: sulfur atom; Orange: selenium atom. (D) The design of thioester (SelM<sup>NT</sup>) and seleno (SelM<sup>CT</sup>) fragments. SelM residues 1 to 24 constitute an ER targeting sequence that is cleaved in vivo and thus were omitted here. Sec48 was mutated to Cys to allow Sec incorporation using the cysteinyl-tRNA misloading protocol. (E) SelM ligation monitored by SDS-PAGE. Following the generation of the SelM<sup>C1</sup> fragment, the thioester-fragment MBP-SelM<sup>NT</sup> is added and the reaction is initiated as shown in panel A. Since the TEV protease was not removed from the SelM<sup>CT</sup> mixture, it releases the SelM<sup>NT</sup> fragment as the ligation progresses. The ligation was monitored using the appearance of a 14 kDa band. Lane 1: MBP-SelM<sup>CT</sup>; lane 2: same sample as in lane 1 following cleavage by TEV protease; lane 3: MBP-SelM<sup>NT</sup>-VMA; lane 4: MBP-SelM<sup>NT</sup> thioester following cleavage of MBP-SelM<sup>NT</sup>-VMA with thiols and subsequent purification; lanes 5-7 the ligation monitored on days 0, 1, and 2, respectively. The ligation mixture contained MBP-SelM<sup>NT</sup> thioester and MBP-SelM<sup>CT</sup> TEV cleavage mixture. Lane M: protein molecular mass standards (the molecular mass in kDa is noted on the right). Sel $M^{NT}$  is too small to be detected. (F) Deconvoluted ESI-MS of intact SelM (the calculated molecular mass is 14168 Da). (G) ESI-MS spectrum of the iodoacetamide-alkylated SelM peptide (VETCGGUQLNR) that contains the Sec residue. Iodoacetamide is marked with a purple star. The observed isotopic pattern matches the calculated one. (H) Tandem MS sequencing of the iodoacetamide-alkylated VETCGGUQLNR peptide derived from trypsin digestion of SelM further corroborates the presence of Sec. Fragment ions that contain the alkylated Sec are colored red. (I) Superimposed CD spectra of SelM, before and after refolding, and SelM U48C prepared by heterologous expression.

# 5.2.4 Expressed Selenoprotein Ligation of Selenoprotein W

Besides SelM, we have prepared selenoprotein W (SelW), a cell cycle

regulator, which belongs to a different class of selenoproteins from SelM (Figure 5.8-5.13)<sup>61</sup>.



Figure 5.8. MBP-SelW<sup>NT</sup> thioester characterization by mass spectrometry. Deconvoluted mass spectrum of MBP-SelW<sup>NT</sup> thioester after chitin affinity chromatography. The calculated molecular mass of the thioester form of MBP-SelW<sup>NT</sup> (MBP-SelW<sup>NT</sup>-MES) is 43934 Da. Inset shows the M/Z spectrum.



Figure 5.9. MBP-SelW<sup>CT</sup> characterization by mass spectrometry. Deconvoluted electrospray ionization mass spectrum of reduced MBP-SelW<sup>CT</sup>. The calculated molecular mass of MBP-SelW<sup>CT</sup> is 50510 Da. The calculated molecular mass for the protein with Cys is 50463 Da (denoted with \*). The 50767 Da form is due to a 6-phosphogluconolactone modification, most likely on the His6-tag at the N-terminal of MBP<sup>60</sup>. Inset shows the M/Z spectrum.



Figure 5.10. The ligation, purification and characterization of SelW. (A) The refolding of ESL SelW (denoted with \*\*) monitored by Tricine-SDS-PAGE. Lane1: refolding mixture after guanidine hydrochloride was removed; lane 2: pellet of refolding mixture; lane 3: supernatant of refolding mixture; lane M: protein molecular mass standards (the molecular mass in kDa is noted on the right). (B) Additional purification by size exclusion chromatography monitored by Tricine-SDS-PAGE. (C) No SelW was found in the pellet following refolding, but unreacted SelW<sup>CT</sup> corresponds to \* in lane 2 of panel A. Calculated molecular masses: 8166 Da SelW<sup>CT</sup> as a monomer and 16332 as the dimer. (D) SelW was predominantly found in the supernatant following refolding. The calculated molecular mass of SelW is 9614 Da.



Figure 5.11. The isotope pattern corresponding to SelW's Sec-containing peptide VTVLYCGAUGYKPK. (A) Experimental isotopic distribution of double charged ion VTVLYCGAUGYKPK from iodoacetamide alkylated and trypsin digested SelW. (B) Calculated isotope pattern. (C) The peptide alkylated with iodoacetamide also identified as a triply charged ion. (D) Calculated isotope pattern.



Figure 5.12. SelW U13C characterization by mass spectrometry. Deconvoluted electrospray ionization mass spectrum of SelW U13C prepared by heterologous recombinant expression. The theoretical and observed molecular mass is 9567 Da. Inset shows the M/Z spectrum.



Figure 5.13. Characterization of SelW by CD spectroscopy. (A) SelW was characterized by intact mass spectrometry, detected as 9613 Da, theoretically 9614 Da. (B) CD spectra of SelW U13C and ESL SelW were comparable which demonstrated that refolding yields a similar content of secondary elements.

#### 5.2.5 Expressed Selenoprotein Ligation Followed by Deselenization of Trx

Importantly, ESL can generate a wide range of proteins because the Sec can be readily deselenized into either serine or alanine<sup>44, 62</sup>. To demonstrate the conversion and its specificity, we prepared a Sec-containing form of *E. coli* thioredoxin (Trx) using residue 94 as the ligation site (Figure 5.14A-5.14C, Figure 5.15-5.18). Trx A94U was converted to native Trx by treating it with a 200-fold excess of TCEP and DTT under anaerobic conditions to promote deselenization of Sec into alanine (Figure 5.14D, 5.19). Trx has two catalytic Cys residues, which were not affected by this mild treatment. The disulfide reductase activity of Trx formed by deselenization was comparable to that of Trx prepared by heterologous expression (Figure 5.14E, Figure 5.20-5.21).

#### 5.2.6 Site Selective Modification through Sec Conjugation

Subsequently, Trx A94U was used to selectively introduce post-translational modifications. A phosphate group was introduced by reacting the protein with thiophosphate (Figure 5.14F, Figure 5.22). Trx A94U was conjugated to ubiquitin G76C via a selenylsulfide bond, and to ubiquitin G76U via a diselenide bond (Figure 5.14G-H, Figure 5.23). The diselenide bond has a lower redox potential compared to that of the disulfide bond and can therefore be retained, under mild reducing conditions, for downstream applications such as activity assays or for preserving the diselenide linkage while reducing undesired disulfide bonds. Finally, we also demonstrated that it was possible to selectively label Sec over Cys in a mixture of ubiquitin G76C and ubiquitin G76U (Figure 5.14I-J).



Figure 5.14. Expressed selenoprotein ligation can be used to generate nonselenoproteins and to introduce a unique site for labeling. Panels A-D) chemical ligation of *E. coli* Trx with Sec at position 94 and subsequent deselenization to form wild-type Trx. (A) The structure of *E. coli* Trx A94U based on PDB entry 2TRX. Red: thioester fragment; Green: seleno-fragment; Yellow: sulfur atoms; Orange: selenium atom. (B) The design of thioester (Trx<sup>NT</sup>) and seleno (Trx<sup>CT</sup>) fragments. (C) Deconvoluted ESI-MS of intact Trx A94U (the calculated molecular mass is 11754 Da). (D) Deconvoluted ESI-MS of wild-type Trx generated by deselenization of Sec to Ala (the calculated molecular mass 11675 Da; the \* denotes Trx A94S). (E) The enzymatic activity of Trx generated by deselenization and that of Trx prepared by heterologous expression are similar as measured by insulin turbidity assays. (F)

Labelling of Trx A94U with a thiophosphate (the calculated molecular mass is 11864 Da; the peak marked with **\*\*** is a contaminant from the commercial thiophosphate). (G) Conjugation of Trx A94U to ubiquitin G76C via the formation of a selenylsulfide bond (the calculated molecular mass is 20521 Da). (H) Conjugation of Trx A94U to ubiquitin G76U via the formation of a diselenide bond (the calculated molecular mass is 20568 Da). (I-J) Due to Sec's low pK<sub>a</sub> it can be selectively alkylated. (I) A mixture of ubiquitin G76C and ubiquitin G76U prior to alkylation (the calculated molecular mass for ubiquitin G76C is 8769 Da and that of ubiquitin G76U is 8816 Da). (J) The same sample from panel I alkylated with MM(PEG)<sub>24</sub> (the calculated molecular mass for MM(PEG)<sub>24</sub>-alkylated ubiquitin G76U is 10055 Da). Only ubiquitin G76U was alkylated.



Figure 5.15. Trx<sup>NT</sup> thioester characterization by mass spectrometry. Deconvoluted mass spectrum of reduced Trx<sup>NT</sup> thioester. Calculated molecular masses: Trx<sup>NT</sup> thioester is 10100 Da; hydrolyzed Trx<sup>NT</sup> is 9976 Da and if an additional oxygen is lost from hydrolyzed Trx<sup>NT</sup> the mass is 9960 Da (denoted with a \*). Inset shows the M/Z spectrum.


Figure 5.16. MBP-Trx<sup>CT</sup> characterization by mass spectrometry. Deconvoluted electrospray ionization mass spectrum of reduced MBP-Trx<sup>CT</sup>. The calculated molecular mass: the Sec-containing MBP-Trx<sup>CT</sup> at 44140 Da; the Cys-containing MBP-Trx<sup>CT</sup> with Cys at 44093 Da (denoted with \*); and MBP-Trx<sup>CT</sup> with 6-phosphogluconolactone modification at 44398 Da. Inset shows the M/Z spectrum.



Figure 5.17. Trx A94U characterization by mass spectrometry. (A) Steps in the ESL visualized by 16% Tris-tricine SDS-PAGE. Lane 1: MBP-Trx<sup>NT</sup>-VMA purified by amylose affinity chromatography; lane 2: MBP-Trx<sup>NT</sup>-VMA in the presence of TEV protease; lane 3: Trx<sup>NT</sup> thioester purified by chitin affinity chromatography; lane 4: MBP-Trx<sup>CT</sup> in the presence of TEV protease (the cleaved peptide is 1.8 kD and can not be visualized in this gel); lane 5: Ligation mixture of Trx<sup>NT</sup> thioester with MBP-Trx<sup>CT</sup> TEV cleavage mixture; lane 6: Purified Trx A94U; lane M: protein molecular mass standards (the molecular mass in kDa is noted on the right). (B) Purified ESL Trx A94U from IMAC flowthrough was immediately detected by mass spectrometry. The peaks at 11893 Da and 11920 Da were due to Trx A94U bound to MES through a selenylsulfide bond (theoretical mass: 11894 Da); and with MPAA (theoretical mass: 11921 Da), respectively. Inset shows the M/Z spectrum.



Figure 5.18. The isotope pattern corresponding to Trx A94U's Sec-containing peptide VGULSK. (A) Experimental isotopic distribution of VGULSK from iodoacetamide alkylated and trypsin-digested Trx A94U. (B) The theoretical isotope pattern. (C) Sequencing by tandem-MS of iodoacetamide-alkylated VGULSK.



Figure 5.19. Characterization of Trx prepared by deselenization of Trx A94U. (A) The isotopic selenium pattern of VGALSK. (B) The tandem mass sequencing of the peptide VGALSK. Blue:  $Y^+$  ions; Red:  $B^+$  ions.



Figure 5.20. Characterization of Trx prepared by heterologous expression in *E. coli*. The calculated molecular mass of Trx is 11675 Da. Inset shows the M/Z spectrum.



Figure 5.21. Reduction of insulin by Trx prepared by ESL and deselenization and by Trx prepared by recombinant expression. Insulin reduction was visualized using 16% Tricine-SDS-PAGE. Lane 1: no enzyme; lane 2: 8  $\mu$ M recombinant Trx; lane 3: 4  $\mu$ M recombinant Trx; lane 4: 8  $\mu$ M ESL Trx; lane 5: 4  $\mu$ M ESL Trx; lane M: protein molecular mass standards (the molecular mass in kDa is noted on the right).



Figure 5.22. Characterization of the thiophosphate-conjugated Trx A94U. Experimental isotopic distribution of thiophosphate-conjugated VGULSK from trypsin-digested thiophosphate-conjugated Trx A94U.



Figure 5.23. Conjugation of ubiquitin to E. coli Trx. (A) Deconvoluted mass spectrum of ubiquitin G76C (the calculated mass is 8769 Da). (B) Deconvoluted mass spectrum of TNB-activated ubiquitin G76C (the calculated mass of TNB-conjugated ubiquitin G76C is 8966 Da and that of the ubiquitin G76C dimer is 17536 Da). (C) Deconvoluted mass spectrum of TNB-activated ubiquitin G76U. The calculated mass of TNB-conjugated ubiquitin G76U is 9013 Da and that of the ubiquitin G76U dimer is 17630 Da. (D) The conjugation of ubiquitin G76C and Trx A94U monitored by Tricine-SDS-PAGE. Lane 1: TNB-conjugated ubiquitin G76C; lane 2: Trx A94U; lane 3: ligation mixture of Trx A94U and TNB-activated ubiquitin G76C; lane M: protein molecular mass standards (the molecular mass in kDa is noted on the right).

#### 5.2.7 Selective Generation of Dehydroalanine for Site Specific Labeling

Another promising application of ESL is to promote selenium elimination from Sec in order to form dehydroalanine, which serves as an electrophilic site for Michael addition reactions, typically with thiol nucleophiles<sup>63</sup>. Thus, introducing dehydroalanine by Sec deselenization would provide a unique site for protein modifications and a path to introduce conjugates that closely resemble the original biological modifications<sup>63</sup>. We have tested the specificity of the reagent  $\alpha_{,}\alpha'$ -dibromo-adipyl(bis)amide (DBAA) using SelM because it has a Sec and a Cys in close proximity. We found that at pH 6.5, essentially all Sec was converted to DHA, as is apparent from the mass loss of 81 Da (Figure 5.24), corresponding to a selenium atom and two hydrogen atoms compared to that of the unreacted protein. When the sample with DHA was incubated with the alkylating reagent iodoacetamide the protein gained a mass of 57 Da, corresponding to Cys alkylation with one iodoacetimide. Reactions at higher pH were not specific, instead generating a bridge between them (Figure 5.25). Hence, it is possible to generate DHA selectively in proteins that contain cysteines by carrying out the labeling at a pH value lower than the pK<sub>a</sub> of Cys, and closer to the pK<sub>a</sub> of Sec.



Figure 5.24. Chemoselective dehydroalanine formation from Sec. (A) Sec deselenization into dehydroalanine can be promoted through two successive alkylation steps. The reaction can be carried out selectively at low pH, even when cysteines are present, due to Sec's low  $pK_a$  (5.2) and high nucleophilicity.

(B) Dehydroalanine formation in SelM at pH 6.5. The calculated molecular mass of SelM is 14168 Da and that of SelM U48dehydroalnine is 14087 Da. (C) The sample from panel B treated with excess IAM at pH 6.5. The sole Cys was available for alkylation at that pH even though it was close to the dehydroalanine site. The molecular weight of iodoacetamide-alkylated SelM U48dehydroalnine is 14144 Da. (D) Tandem mass of trypsin-digested sample C further confirms the presence of dehydroalanine in position 48. Identifiable daughter ions are labeled. Red color indicates ions containing dehydroalanine.



Figure 5.25. Conjugations of thiol nucleophiles to SelM Sec48dehydroalanine. Deconvoluted mass spectra of (A) Conjugation of SelM Sec48dehydroalanine to 1thio- $\beta$ -D-glucose. Calculated molecular weights are 14087 Da for SelM Sec48dehydroalanine and 14283 Da after 1-thio- $\beta$ -D-glucose incorporation. (B) SelM incubated with DBBA at pH 8.0. A SelM-DBAA adduct formed with both Cys45 and Sec48. Calculated molecular weight of SelM with a DBAA adduct is 14308 Da. (C) Sample from B was reacted with 20 mM idoacetamide. No alkylation was observed confirming that Cys45 and Sec48 are not exposed for modifications. (D) SelM prior to incubation with DBAA is fully alkylated when the procedure at panel C is applied. Calculated molecular weight for SelM with two idoacetamide is 14282 Da.

#### 5.3 Discussion

ESL is as straightforward to implement as Cys-mediated EPL but offers significantly higher efficiency and greater flexibility in selecting the ligation site while also providing a site-specific handle for chemical conjugation. Due to Sec's high reactivity, it should be helpful in improving yield, particularly for large proteins as well as for challenging reactions such as assembling the protein of interest by sequential ligations. It has the significant advantages of high yield and low cost as compared to non-natural amino acid incorporation.

The experiments presented here demonstrate the application of a new chemical biology driven method for production of selenoproteins in sufficient amounts for their biochemical and biophysical characterization. Such information will be invaluable for identifying selenoproteins' enzymatic functions, contribution of Sec, and their roles in sensing and signaling of oxidative stress. Owing to our method, a large fraction of the selenoproteome can finally be prepared for detailed characterization. In addition to the preparation of naturally occurring selenoproteins, ESL also enables the introduction of Sec residues as well as novel selenylsulfide and diselenide bonds in proteins to generate new functions, folding paths or to design novel therapeutics<sup>6</sup>.

Sec's unique chemistry that is distinct from that of the canonical 20 amino acids offers paths for site-specific incorporation of bioconjugations, and even the coupling of two proteins. As we have shown here ESL is compatible with the presence of cysteines and it is possible to adjust the reaction conditions to preferentially label Sec. Overall, Sec's potential for introducing posttranslational modifications has so far not been systematically explored because previously there was no general method to prepare selenoproteins at high yield and with positioning of Sec at any position<sup>64</sup>. As we show here, Sec can be used to reversibly introduce bioconjugates via selenylsulfide

199

and diselenide bonds and irreversibly via selenoether bonds. We have also shown that it is possible to introduce DHA in a selective fashion in multiple proteins. This generates new possibilities for incorporation of bioconjugations with the distinct advantage of introducing PTM that closely resembles those that occur in vivo, e.g. with similar charge, length and conformational degrees of freedom.

In conclusion, our work provides new tools for preparing and modifying proteins. The ESL procedure is versatile, robust, straightforward to implement and requires no synthesis or special expertise with selenium chemistry. Overall, ESL extends the ease and generality of biosynthetic protein preparations, the class of proteins that can be prepared, and offers new avenues to exploit Sec's high chemical reactivity.

# 5.4 Materials and Methods

# 5.4.1 Molecular Cloning

*Homo sapiens* selenoprotein M (SelM) gene (accession number: NP\_536355.1) and *Gallus gallus* selenoprotein W (SelW) gene (accession number: NP\_001159799.1) were codon optimized for expression in *E. coli* using Genscript OptimumGene<sup>TM</sup> and synthesized by GenScript. The Sec in both SelM and SelW was mutated to a Cys. The cDNA for *Homo sapiens* ubiquitin (plasmid #12647) and *E. coli* thioredoxin (Trx) (plasmid #11516) were purchased from Addgene<sup>65, 66</sup>. Primers were synthesized and purified by Sigma-Aldrich, and plasmids were sequenced by GENEWIZ. All molecular biology reagents were obtained from New England Biolabs and mutagenesis was accomplished with their Q5<sup>®</sup> Site-Directed Mutagenesis Kit.

Table 5.1. Primers used for cloning.

Primer	Oligonucleotide Sequence (5'→3')
SelM	ATTTCCATGGGCACCGCTTACCGCCCG
NcoI For	
SelM	GCGAATGGATCCTTACAGGTCGGCATGATC
BamHI	
Rev	
MBP-	GCCGCATATGAAAATCCATCACCATCACCATCACGAAG
POI <sup>NT</sup>	
Ndel For	
SelM <sup>NT</sup>	ATATGCTCTTCCGCAACCACCGCACGTTTCCACACGC
SapI Rev	
SelM <sup>CT</sup>	TGTCAGCTGAACCGCCTG
For	
SelM <sup>CT</sup>	CTGAAAATACAGGTTTTCAGAGCTC
Rev	
SelW	GTTGTTCCATGGGCCCGCTGCGTGTGACCGTG

NcoI For	
SelW	GTTGGTGGATCCTTACGGCAGCGCCGCCTGAATC
BamHI	
Rev	
SelW <sup>NT</sup>	AATAATGCTCTTCCGCACGCACCGCAATACAGCAC
SapI Rev	
SelW <sup>CT</sup>	TGCGGTTACAAGCCGAAG
For	
SelW <sup>CT</sup>	CTGAAAATACAGGTTTTCAGAGC
Rev	
Trx NcoI	GGTGGGGGTCCATGGATAAAATTATTCACCTGACTGACG
For	
Trx	ATTATTGGATCCTTACGCCAGGTTAGCGTCGAGGAACTC
BamHI	
Rev	
Trx step2	GATAAAATTATTCACCTGACTGAC
For	
Trx step2	GGACTGAAAATACAGGTTTTC
Rev	
pMAL-	AATTCGAGCTCTGAAAACC
c5X	
His <sub>6</sub> -POI	
For	
pMAL-	GTGATGGTGATGGTGATG
c5X	
His <sub>6</sub> -POI	
Rev	
Trx <sup>NT</sup>	ATAATAATGCTCTTCCGCAACCCACTTTGGTTGCCGCCAC
SapI Rev	
Trx <sup>CT</sup>	CCTGTCTAAAGGTCAGTTGAAAG
For	
Trx <sup>CT</sup>	CACTGAAAATACAGGTTTTCAGAG
Rev	
Ubiquitin	GGTGGTCCATGGCTCAGATCTTCGTCAAGAC
G76C	
NcoI For	
Ubiquitin	GGTGGTGGATCCTTAGCAACCTCTTAGTCTTAAGAC
G76C	
BamHI	
Rev	

DNA sequences of SelM and SelW (optimized for expression in *E. coli*) SelM:

TCACTGCTGCTGCCGCCGCTGGCTCTGCTGCTGCTGCTGGCTGCCCTGGTT GCTCCGGCTACCGCTGCGACCGCTTACCGCCCGGATTGGAACCGTCTGAGT GGCCTGACCCGTGCGCGTGTGGAAACGTGCGGTGGTTGTCAGCTGAACCG CCTGAAAGAAGTGAAGGCTTTTGTTACCCAAGATATTCCGTTCTATCATAA TCTGGTTATGAAACACCTGCCGGGCGCAGACCCGGAACTGGTCCTGCTGG GTCGTCGCTACGAAGAACTGGAACGTATTCCGCTGAGCGAAATGACCCGC GAAGAAATCAATGCGCTGGTTCAGGAACTGGGTTTTTATCGTAAAGCAGC ACCGGATGCACAAGTCCCGCCGGAATACGTGTGGGCACCGGCTAAGCCGC CGGAAGAAACGTCTGATCATGCCGACCTGTAA SelW:

MBP-SelM U48C: The primers SelM NcoI For and SelM BamHI Rev were used to clone full length SelM with U48C mutation, but excluding the signal sequence (AA 1-24) into pMAL-c5X, in which a hexahistidine tag was inserted between I3 and E4 of cMBP. The translated amino acid sequence is as follows: MKIHHHHHHEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHP DKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFT WDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSA LMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVD LIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPT FKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAV ALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGR QTVDEALKDAQTNSSSENLYFQSMGTAYRPDWNRLSGLTRARVETCGGCQL NRLKEVKAFVTQDIPFYHNLVMKHLPGADPELVLLGRRYEELERIPLSEMTRE EINALVQELGFYRKAAPDAQVPPEYVWAPAKPPEETSDHADL

Black: His<sub>6</sub>-cytosolic MBP; Blue: TEV cleavage site; Orange: SelM sequence (AA 25-145) excluding the signal peptide; Green: U48 was mutated to C48; Gray: Following cleavage with TEV protease, SMG were present before the native AA 25-145.

MBP-SelM N-terminus (MBP-SelM<sup>NT</sup>-VMA): The primers MBP-POI<sup>NT</sup> NdeI For and SelM<sup>NT</sup> SapI Rev were used to clone full length MBP-SelM<sup>NT</sup> from MBP-SelM U48C into pTYB1 vector. The translated amino acid sequence is as follows:

MKIHHHHHHEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHP DKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFT WDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSA LMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVD LIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPT FKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAV ALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGR QTVDEALKDAQTNSSSENLYFQSMGTAYRPDWNRLSGLTRARVETCGGCFA KGTNVLMADGSIECIENIEVGNKVMGKDGRPREVIKLPRGRETMYSVVQKSQ HRAHKSDSSREVPELLKFTCNATHELVVRTPRSVRRLSRTIKGVEYFEVITFEM GQKKAPDGRIVELVKEVSKSYPISEGPERANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPILYENDHFFDYMQKSKFHLTIEGPKVLAYLLGL WIGDGLSDRATFSVDSRDTSLMERVTEYAEKLNLCAEYKDRKEPQVAKTVNL YSKVVRGNGIRNNLNTENPLWDAIVGLGFLKDGVKNIPSFLSTDNIGTRETFL AGLIDSDGYVTDEHGIKATIKTIHTSVRDGLVSLARSLGLVVSVNAEPAKVDM NVTKHKISYAIYMSGGDVLLNVLSKCAGSKKFRPAPAAAFARECRGFYFELQ ELKEDDYYGITLSDDSDHQFLLGSQVVVHACGGLTGLNSGLTTNPGVSAWQV NTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVPALWQLQ

Black: His<sub>6</sub>-cytosolic MBP; Blue: TEV cleavage site; Orange: SelM Nterminus sequence (AA 25-47) excluding the signal peptide; Purple: VMA intein tag (the Sce VMA intein/chitin binding domain); Gray: After cleavage with TEV protease, SMG were present before the native AA 25-145.

MBP-SelM C-terminus (MBP-SelM<sup>CT</sup>): The primers MBP-SelM<sup>CT</sup> was generated using SelM<sup>CT</sup> For and SelM<sup>CT</sup> Rev using the MBP-SelM U48C as a template. The translated amino acid sequence is as follows:

MKIHHHHHHEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHP DKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFT WDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSA LMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVD LIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPT FKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAV ALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGR QTVDEALKDAQTNSSSENLYFQCQLNRLKEVKAFVTQDIPFYHNLVMKHLPG ADPELVLLGRRYEELERIPLSEMTREEINALVQELGFYRKAAPDAQVPPEYVW APAKPPEETSDHADL

Black: His6-cytosolic MBP; Blue: TEV cleavage site; Orange: SelM sequence (AA 48-145) excluding the signal peptide; Green: U48 was mutated to C48.

MBP-SelW U13C: The primers SelW NcoI For and SelW BamHI Rev were used to clone full length SelW with a U13C mutation into pMAL-c5X vector, in which a hexahistidine tag was inserted between I3 and E4 of cMBP. The translated amino acid sequence is as follows:

MKIHHHHHHEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHP DKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFT WDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSA LMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVD LIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPT FKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAV ALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGR QTVDEALKDAQTNSSSENLYFQSMGPLRVTVLYCGACGYKPKYERLRAELEK RFPGALEMRGQGTQEVTGWFEVTVGSRLVHSKKNGDGFVDTDAKLQRIVAA IQAALP

Black: His<sub>6</sub>-cytosolic MBP; Blue: TEV cleavage site; Orange: SelW sequence (AA 2-85); Green: U13 was mutated to C13; Gray: After cleavage with TEV protease, SMG were present before the native AA 2-85.

MBP-SelW N-terminus (MBP-SelW<sup>NT</sup>-VMA): The primers MBP-POI<sup>NT</sup> NdeI For and MBP-SelW<sup>NT</sup> SapI Rev were used to clone the MBP-SelW<sup>NT</sup> from MBP-SelW U13C template into pTYB1 vector. The translated amino acid sequence is as follows:

MKIHHHHHHEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHP DKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFT WDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSA LMFNLOEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVD LIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPT FKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAV ALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGR **QTVDEALKDAQTNSSSENLYFQ**SMGPLRVTVLYCGACFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLPRGRETMYSVVQKSQHRAHKSDSSREVP ELLKFTCNATHELVVRTPRSVRRLSRTIKGVEYFEVITFEMGQKKAPDGRIVEL VKEVSKSYPISEGPERANELVESYRKASNKAYFEWTIEARDLSLLGSHVRKAT YQTYAPILYENDHFFDYMQKSKFHLTIEGPKVLAYLLGLWIGDGLSDRATFSV DSRDTSLMERVTEYAEKLNLCAEYKDRKEPQVAKTVNLYSKVVRGNGIRNN LNTENPLWDAIVGLGFLKDGVKNIPSFLSTDNIGTRETFLAGLIDSDGYVTDEH **GIKATIKTIHTSVRDGLVSLARSLGLVVSVNAEPAKVDMNVTKHKISYAIYMS** GGDVLLNVLSKCAGSKKFRPAPAAAFARECRGFYFELQELKEDDYYGITLSD DSDHQFLLGSQVVVHACGGLTGLNSGLTTNPGVSAWQVNTAYTAGQLVTYN GKTYKCLQPHTSLAGWEPSNVPALWQLQ

Black: His<sub>6</sub>-cytosolic MBP; Blue: TEV cleavage site; Orange: SelW Nterminus sequence (AA 2-12); Purple: VMA Intein tag (the Sce VMA intein/chitin binding domain); Gray: After cleavage with TEV protease, SMG were present before the native AA 2-12.

MBP-SelW C-terminus (MBP-SelW<sup>CT</sup>): The primers SelW<sup>CT</sup> For and SelW<sup>CT</sup> Rev were used to generate MBP-SelW<sup>CT</sup> using the MBP-SelW U13C plasmid as a template. The translated amino acid sequence is as follows:

MKIHHHHHHEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHP DKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFT WDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSA LMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVD LIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPT FKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAV ALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGR QTVDEALKDAQTNSSSENLYFQCGYKPKYERLRAELEKRFPGALEMRGQGTQ EVTGWFEVTVGSRLVHSKKNGDGFVDTDAKLQRIVAAIQAALP

Black: His<sub>6</sub>-cytosolic MBP; Blue: TEV cleavage site; Orange: SelW Cterminus AA 13-85; Green: U13 mutated to C13.

His<sub>6</sub>-Trx: The primers Trx NcoI For and Trx BamHI Rev were used to clone *E. coli* Trx into pMAL-c5X vector, in which a hexahistidine tag was inserted between I3 and E4 of cMBP. A second step of site mutagenesis was used to delete the first Met using the primers Trx step2 For and Trx step2 Rev (abbreviated as MBP-Trx). Finally, the primers pMAL-c5X His<sub>6</sub>-POI For and pMAL-c5X His<sub>6</sub>-POI Rev were used to delete cMBP. The translated amino acid sequence is as follows:

# MKIHHHHHHNSSSENLYFQSDKIIHLTDDSFDTDVLKADGAILVDFWA EWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKN GEVAATKVGALSKGQLKEFLDANLA

Black: His6 tag; Blue: TEV cleavage site; Orange: E. coli Trx (AA 2-109).

MBP-Trx N-terminus (MBP-Trx<sup>NT</sup>-VMA): The primers MBP-POI<sup>NT</sup> NdeI For and Trx<sup>NT</sup> SapI Rev were used to clone MBP-Trx<sup>NT</sup> from MBP-Trx into pTYB1 vector. The translated amino acid sequence is as follows:

MKIHHHHHHEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHP DKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFT WDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSA LMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVD LIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPT FKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAV ALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGR **QTVDEALKDAQTNSSSENLYFQSDKIIHLTDDSFDTDVLKADGAILVDFWAE** WCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNG **EVAATKVG**CFAKGTNVLMADGSIECIENIEVGNKVMGKDGRPREVIKLPRGR ETMYSVVQKSQHRAHKSDSSREVPELLKFTCNATHELVVRTPRSVRRLSRTIK **GVEYFEVITFEMGQKKAPDGRIVELVKEVSKSYPISEGPERANELVESYRKAS** NKAYFEWTIEARDLSLLGSHVRKATYQTYAPILYENDHFFDYMQKSKFHLTIE **GPKVLAYLLGLWIGDGLSDRATFSVDSRDTSLMERVTEYAEKLNLCAEYKDR KEPQVAKTVNLYSKVVRGNGIRNNLNTENPLWDAIVGLGFLKDGVKNIPSFLS** TDNIGTRETFLAGLIDSDGYVTDEHGIKATIKTIHTSVRDGLVSLARSLGLVVS VNAEPAKVDMNVTKHKISYAIYMSGGDVLLNVLSKCAGSKKFRPAPAAAFA

# RECRGFYFELQELKEDDYYGITLSDDSDHQFLLGSQVVVHACGGLTGLNSGL TTNPGVSAWQVNTAYTAGQLVTYNGKTYKCLQPHTSLAGWEPSNVPALWQ LQ

Black: His<sub>6</sub>-cytosolic MBP; Blue: TEV cleavage site; Orange: *E. coli* Trx AA 2-93; Purple: VMA Intein tag (the Sce VMA intein/chitin binding domain).

MBP-Trx C-terminus (MBP-Trx<sup>CT</sup>): The primers Trx<sup>CT</sup> For and Trx<sup>CT</sup> Rev were used for mutagenesis of MBP-Trx. The translated amino acid sequence is as follows:

MKIHHHHHHEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHP DKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFT WDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSA LMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVD LIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPT FKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAV ALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGR QTVDEALKDAQTNSSSENLYFQCLSKGQLKEFLDANLA

Black: His6-cytosolic MBP; Blue: TEV cleavage site; Orange: *E. coli* Trx (AA 94-109). Green: A94 mutated to C94.

MBP-ubiquitin G76C: The primers ubiquitin G76C NcoI For and ubiquitin G76C BamHI Rev were used to clone ubiquitin G76C into pMAL-c5X vector, in which a hexahistidine tag was inserted between I3 and E4 of cMBP. The translated amino acid sequence is as follows:

MKIHHHHHHEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHP DKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFT WDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSA LMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVD LIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPT FKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAV ALKSYEEELVKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGR QTVDEALKDAQTNSSSENLYFQSMAQIFVKTLTGKTITLEVEPSDTIENVKAKI QDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGC

Black: His<sub>6</sub>-cytosolic MBP; Blue: TEV cleavage site; Orange: ubiquitin AA 2-76; Green: G76 mutated to C76; Gray: After cleavage with TEV protease, SMA were present before the native AA 2-76.

#### 5.4.2 Protein Expression and Purification

The pMHTDelta238 plasmid expressing tobacco etch virus (TEV) protease fused with the cytoplasmic maltose binding protein (cMBP)<sup>67</sup> was purchased from the Protein Structure Initiative: Biology Materials Repository<sup>68</sup>. Chromatography media was supplied by GE Healthcare Life Sciences and New England Biolabs. Reagents were purchased from Acros Organics, Gold Biotechnology, Sigma-Aldrich, New England Biolabs, and Chem-Impex International.

Expression and purification of SelM U48C: For protein expression of SelM U48C mutant, the plasmids were transformed into E. coli strain BL21(DE3). Cells were grown in low salt LB (10 g tryptone, 5 g yeast extract, 5 g NaCl), supplemented with 0.2% glucose at 37 °C, with good aeration and relevant antibiotic selection (100  $\mu$ g/mL ampicillin). When the optical density (OD) at 600 nm reached 0.5, the temperature was lowered to 18 °C, and the cells were allowed to shake at the lower temperature for an additional 1 h. Protein expression was induced with 0.5 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG). The cells were harvested after 18-20 h, and the cell paste (7 g/L) was resuspended in 50 mM sodium phosphate (pH 7.5) and 200 mM NaCl (amylose buffer), supplemented with 0.5 mM benzamidine, 1 mM phenylmethanesulfonyl fluoride (PMSF), and 2 mM ethylenediaminetetraacetic acid (EDTA). Cells were lysed using a high-pressure homogenizer (Emulsi-Flex-C5, Avestin) on ice, and all subsequent procedures were conducted at 4 °C. Cell debris was removed by centrifugation at 20000 g for 1 h. The supernatant was loaded onto an amylose column, and the column was washed with the amylose buffer. The SelM U48C fusion protein was eluted using amylose buffer containing 20 mM maltose. Cleavage of the fusion partner cMBP was conducted by incubating a 1:20 molar ratio of TEV protease to MBP-SelM U48C at 4 °C overnight. Following cleavage, the

protein was dialyzed to remove EDTA and loaded onto a 5 mL histidine affinity column (Histrap FF, GE Healthcare) to remove MBP and TEV. The flowthrough was pooled, concentrated and loaded onto a size exclusion column (HiPrep 16/60 Sephacryl S-100 HR column, GE Healthcare), as a polishing step to further purify SelM U48C to homogeneity. Protein purity, as determined with a 16% Tris-glycine sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) gel, was greater than 95%.

Expression and purification of MBP-SelM<sup>NT</sup> thioester: For protein expression of MBP-SelM<sup>NT</sup>-VMA, the plasmid was transformed into *E. coli* strain BL21(DE3). Cells were grown in low salt LB, supplemented with 0.2% glucose at 37 °C, with good aeration and the relevant antibiotic selection (100  $\mu$ g/mL ampicillin). When the optical density (OD) at 600 nm reached 0.5, the temperature was lowered to 18 °C, and the cells were allowed to shake at the lower temperature for an additional half hour. Protein expression was induced with 0.5 mM IPTG. The cells were harvested after 18-20 h. Cell paste (13 g/L) was resuspended in amylose buffer supplemented with 0.5 mM benzamidine, 1 mM PMSF, and 2 mM EDTA. Cells were lysed using a highpressure homogenizer on ice, and all subsequent procedures were conducted at 4 °C. Cell debris was removed by centrifugation at 20000 g for 1 h. The supernatant was loaded on an amylose column, and the column was washed first with 15 column volumes (CV) of the amylose buffer followed by 2 CV with 25 mM 2-(Nmorpholino)ethanesulfonic acid (MES) (pH 6.5), 200 mM NaCl, and 2 mM EDTA (wash buffer). The MBP-SelM<sup>NT</sup>-VMA fusion protein was eluted using wash buffer containing 20 mM maltose. The MBP-SelM<sup>NT</sup> thioester was generated either by on column or in solution cleavage by 2-mercaptoethanesulfonic acid sodium salt

(MESNA). For on column cleavage, the elution from the amylose column was loaded onto a 40 mL chitin column, then incubated with 2 CV 25 mM MES (pH 6.5), 200 mM NaCl, 2 mM EDTA, and 75 mM MESNA (thioester cleavage buffer) after the column was drained completely. For off column cleavage, 75 mM MESNA was added to the above MBP-SelM<sup>NT</sup> amylose elution, and left to react at room temperature (RT) for 12 h. The cleavage mixture was then loaded onto a 40 mL chitin column, and MBP-SelM<sup>NT</sup> was collected in the flowthrough when the column was washed with 2 CV of wash buffer 2. The process was repeated twice to obtain maximum yield. The formation of MBP-SelM<sup>NT</sup> thioester was confirmed by mass spectrometry.

Expression and purification of MBP-SelM<sup>CT</sup>: For protein expression of MBP-SelM<sup>CT</sup>, the expression vector was transformed into *E. coli* strain BL21(DE3). The procedure for Sec incorporation was modified from previous protocols<sup>21</sup>. 1 mL of an overnight culture in LB broth was spun at 2500 g for 5 min, then resuspended in modified MDAG media from Studier<sup>69</sup> as listed in Table 5.2 step 1. Cells were grown at 37 °C with good aeration. When the OD at 600 nm reached 1.5 (about 10-14 h), the growth medium was supplemented as detailed in Table 5.2 step 2, and continuously shaken for an additional 10 min. The temperature was then lowered to 18 °C, and the cells were allowed to recover for 10 min. Protein expression was induced with 0.5 mM IPTG, and the cells were grown for 20-24 h at 18 °C. Subsequent purification by amylose affinity chromatography was the same as described for SelM U48C. The yield of MBP-SelM<sup>CT</sup> ranged from 60 to 80 mg per L. The ratio of Sec incorporation into MBP-SelM<sup>CT</sup> was evaluated by mass spectrometry after its reduction by 100 molar excess DTT at 37 °C for 20 min. MBP-SelM<sup>CT</sup> at 25 °C for 16 h in amylose

214

elution buffer supplemented with 10 mM DTT. The cleavage was monitored by 16% Tris-Glycine SDS-PAGE.

Expression and purification of SelW U13C and SelW fragments: The procedures for expression and purification of SelW U13C and SelW fragments are the same as for the expression and purification of SelM U48C, and SelM fragments, respectively.

Expression and purification of Trx<sup>CT</sup> and Trx<sup>NT</sup> thioester: The expression and purification of MBP-Trx<sup>CT</sup> and MBP-Trx<sup>NT</sup>-VMA are the same as those described for MBP-SelM<sup>CT</sup> and MBP-SelM<sup>NT</sup>-VMA, respectively. Following purification MBP-Trx<sup>NT</sup>-VMA was cleaved with TEV protease at a 10:1 molar ratio. Cleavage was carried out at 4 °C overnight. The TEV cleavage mixture was loaded onto a chitin column and washed with 2 CV of 25 mM MES (pH 6.5), 200 mM NaCl, 75 mM MESNA, and 2 mM EDTA (elution buffer). The column was then capped to allow on-column cleavage to proceed at 25 °C for 12 h. The Trx<sup>NT</sup> thioester was eluted from the column with 2 CV of elution buffer, and the on-column cleavage and elution process were repeated two times for maximum yield. The molecular mass of Trx<sup>NT</sup> thioester was confirmed by mass spectrometry.

Expression and purification of Trx: For protein expression of *E. coli* Trx, the His<sub>6</sub>-Trx expression vector was transformed into *E. coli* strain BL21(DE3). Cells were grown in regular LB at 37 °C with good aeration and the relevant antibiotic selection (100  $\mu$ g/mL ampicillin). When the optical density (OD) at 600 nm reached 0.5, the temperature was lowered to 18 °C, and the cells were allowed to shake at the lower temperature for an additional 0.5 h. Protein expression was induced with 0.5 mM IPTG. The cell pellet was resuspended in 50 mM sodium phosphate (pH 7.5), 200 mM

NaCl, and 10 mM imidazole (IMAC equilibration buffer), supplemented with 1 mM PMSF, and 1 mM benzamidine. Cells were disrupted by homogenization and the cell debris was removed by centrifugation at 20000 g for 1 h. The supernatant was pooled and loaded onto a 10 mL IMAC column (HisTrap FF) that was pre-equilibrated with 5 CV IMAC equilibration buffer. His<sub>6</sub>-Trx was eluted with a linear gradient (0 mM - 1 M imidazole). The fractions containing His<sub>6</sub>-Trx were combined, concentrated and loaded onto a HiPrep 26/10 desalting column and buffer exchanged to 50 mM sodium phosphate (pH 7.5), 200 mM NaCl, and 10 mM imidazole. TEV was added at a TEV to His<sub>6</sub>-Trx molar ratio of 1:5 for overnight cleavage at RT. A second IMAC column was used to remove TEV and uncleaved His<sub>6</sub>-Trx. The flowthrough that contained Trx was collected, concentrated, and further purified on Hiprep 16/60 Sephacryl S-100 HR. The purity of Trx after size exclusion as estimated by Tris-Glycine SDS-PAGE was greater than 95%.

Expression and purification of ubiquitin G76C: The procedures for expression and purification of ubiquitin G76C are the same as for expression and purification of SelM U48C. The procedure of expression for ubiquitin G76U was the same as for the seleno-fragments.

Component	Volume (mL)
Step 1	
$H_2O$ (autoclaved)	926
1 M MgSO <sub>4</sub>	2
1000 X metals*	0.2

Table 5.2. Defined growth medium.

Vitamin mix*	1
Ampicillin 100 mg/mL	1
40 % glucose (W/V)	5
50 X M*	20
Step 2	
40 % glucose (W/V)	5
25 % aspartate (W/V)	10
17 amino acids (10 mg/mL) *	20
25 mg/mL methionine	4
Selenocystine*	0.1 g

\* The recipes to prepare the 1000 X metals, Vitamin mix, and 50 X M (salts mixture) and 17 amino acids are as described by Studier<sup>69</sup>.

\*\* Excess selenocystine can lead to decreased yield due to toxicity. In our experience, 0.1 g L-selenocystine per L growth medium resulted in incorporation of over 90% Sec into target proteins. If incorporation levels are low then L-selenocystine concentration should be increased to 0.15 g per L growth medium.

### 5.4.3 Expressed Selenoprotein Ligation

Expressed selenoprotein ligation and purification of SelM: MBP-SelM<sup>CT</sup> TEV cleavage mixture in 50 mM sodium phosphate (pH 7.5), 200 mM NaCl, and 2 mM EDTA was mixed with MBP-SelM<sup>NT</sup> thioester in 25 mM MES (pH 6.5), 200 mM NaCl, 2 mM EDTA, at a 4:1 molar ratio of SelM<sup>NT</sup> thioester to SelM<sup>CT</sup>. The reaction mixture was supplemented with 100 mM MESNA, 50 mM 4-mercaptophenylacetic acid (MPAA), 5 mM DTT, and the pH was adjusted to 7.0. The ESL progress was

monitored by 16% Tris-Glycine SDS-PAGE and by mass spectrometry. Subsequently SelM was purified to homogeneity using the same protocol as that of SelM U48C. The yield was 10 mg from 6 L LB media to express MBP-SelM<sup>NT</sup>-VMA and 2 L MDAG media to express MBP-SelM<sup>CT</sup>.

Expressed selenoprotein ligation and purification of SelW: MBP-SelW<sup>NT</sup> thioester in 25 mM MES (pH 6.5), 200 mM NaCl, and 2 mM EDTA, was mixed with MBP-SelW<sup>CT</sup> TEV cleavage mixture at a 6:1 molar ratio. The solution was supplemented with 100 mM MESNA, 50 mM MPAA, 5 mM DTT and the pH was adjusted to 7.0. The ligation mixture was allowed to shake gently at RT for two days. The mixture was centrifuged at 20000 g for 30 min at 4 °C. The pellet containing ligated SelW was solubilized in 20 mM Tris-HCl (pH 7.0), 50 mM NaCl, 2 mM EDTA, 10% glycerol, and 6 M guanidine hydrochloride. Refolding was carried out by dialysis to gradually remove the denaturant by 1 M increments. Following refolding, the solution was centrifuged again at 20000 g for 30 min at 4 °C. SelW in the supernatant was further purified using Superdex 75 10/300 GL size exclusion column. The yield is 2 mg from 4 L LB media to express MBP-SelW<sup>NT</sup>-VMA and 2 L MDAG media to express MBP-SelW<sup>CT</sup>.

Expressed selenoprotein ligation and purification of Trx A94U: The expressed selenoprotein ligation and purification of Trx A94U were similar to SelM, except that a 6:1 molar ratio of Trx<sup>NT</sup>-MES thioester to MBP-Trx<sup>CT</sup> TEV cleavage mixture was used. The ligation was monitored by running 16% Tricine-SDS-PAGE. The remaining purification steps of Trx A94U were the same as for the purification of Trx described above. Successful ligation and protein intact mass were verified by mass spectrometry.

The yield is 4 mg total TrxA94U from 6 L LB media to express MBP-Trx<sup>NT</sup>-VMA and 2 L MDAG media to express MBP-Trx<sup>CT</sup>.

## 5.4.4 Deselenization of Trx A94U to Trx

 $100 \mu$ M Trx A94U in purged and degassed buffer composed of 50 mM sodium phosphate (pH 7.5), 200 mM NaCl, and 2 mM EDTA, was treated with 20 mM TCEP and 5 mM DTT (also purged and degassed) at 25 °C for 12 h. The ligation mixture was then concentrated and loaded onto a Superdex 75 10/300 GL size exclusion column to remove excess DTT and TCEP. Deselenization was confirmed by mass spectrometry.

## 5.4.5 Enzymatic Activity of Trx

Insulin turbidity assays were conducted to compare the enzymatic activity of Trx prepared by ESL and deselenization with that of Trx prepared by recombinant expression<sup>70</sup>. Briefly, the reaction was carried out in 100 mM potassium phosphate (pH 7.0), 2 mM EDTA, 0.18 mM insulin, with either 4 or 8  $\mu$ M ESL Trx (after deselenization); 4 or 8  $\mu$ M recombinant Trx; or without Trx. The reaction was initiated by the addition of 1 mM DTT, and the reaction progress was monitored by recording the change of absorbance at 650 nm with respect to time for 20 min at 25 °C. After that time, 12 mM N-ethylmaleimide (NEM) was added to quench the reaction and the reduction of insulin was verified by running 16% Tricine-SDS-PAGE.

### 5.4.6 Conjugation of Thioglucose and Thiophosphate to Trx A94U

20 μM Trx A94U in 50 mM sodium phosphate (pH 7.5), 200 mM NaCl, and 2 mM EDTA, was reduced with 10 mM DTT for 2 h at RT. The buffer was subsequently exchanged to 100 mM HEPES (pH 8.2) using a Nap-5 desalting column.

The reduced Trx A94U was immediately reacted with 200 mM of either sodium thiophosphate or 1-thio- $\beta$ -D-glucose, and the reaction mixture was incubated at RT for 3 h. The ligated protein mass was confirmed by mass spectrometry. When the same procedure was carried for Trx expressed by heterologous expression as a negative control, no labeling was detected by ESI-MS.

### 5.4.7 Conjugation of Ubiquitin G76C or Ubiquitin G76U to Trx A94U

The TEV cleavage mixture of either 0.5 mM ubiquitin G76C or ubiquitin G76U in 50 mM sodium phosphate (pH 7.5), 200 mM NaCl, 2 mM EDTA, and 10 mM DTT was treated with 100 mM 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) at RT for 1 h. The protein was then purified on a Superdex 75 10/300 GL size exclusion column. Subsequently 50  $\mu$ M 2-nitro-5-thiobenzoic acid (TNB) activated ubiquitin was added to 5  $\mu$ M reduced Trx A94U in 100 mM HEPES (pH 8.2). The mixture was incubated for 3 h at RT. The conjunction of ubiquitin G76C or ubiquitin G76U to Trx A94U was monitored by 16% Tricine-SDS-PAGE as well as by mass spectrometry. When the same procedure was carried for Trx expressed by heterologous expression as a negative control, no labeling was detected by ESI-MS.

## 5.4.8 Selective Alkylation of Ubiquitin G76U

Ubiquitin G76C and ubiquitin G76U were mixed at equal concentration (final ubiquitin concentration was 10  $\mu$ M) in 50 mM sodium phosphate (pH 7.5), 200 mM NaCl, and 2 mM EDTA. The mixture was reduced with 10 mM DTT for 2 h at RT. The ubiquitin buffer was exchanged using a Nap-5 desalting column into 200 mM sodium phosphate (pH 5.0) and 20 mM DTT. The alkylating reagent MM(PEG)24

(ThermoFisher Scientific) was added to a final concentration of 0.5 mM, and incubated at RT for 10 min followed by analysis by mass spectrometry.

## 5.4.9 Selective Generation of Dehydroalanine

 $\alpha, \alpha'$ -di-bromo-adipyl(bis)amide (DBAA) was synthesized according to the procedure described by Davis and colleagues<sup>9</sup> with the following modifications: thionyl chloride was removed by distillation and not by reduced pressure, and the reaction time was extended to 1 h. DBAA was subsequently purified by preparative Waters C18 Xbridge BEH C4 column (300 Å, 3.5 µm, 2.1 mm x 50 mm), using an acetonitrile gradient from 5-90%, with 0.1% formic acid.

To generate dehydroalanine, 5  $\mu$ M SelM in 50 mM sodium phosphate (pH 7.5), 200 mM NaCl, and 2 mM EDTA was reduced with 10 mM DTT at room temperature for 2 h, then the protein was buffer exchanged to 200 mM sodium phosphate and 2 mM DTT at a final pH of 5.0, 6.0 or 6.5. DBAA was added from a stock of 0.5 M in dimethylformamide (DMF) to a final concentration of 5 mM. The mixture was incubated at RT for 30 min, followed by a second incubation at 37 °C for 2 h. Elimination of selenium to dehydroalanine was confirmed by mass spectrometry.

For conjugation reactions, either 300 mM 1-thio- $\beta$ -D-glucose sodium or  $\beta$ mercaptoethanol was added to the sample described above and the pH was adjusted to 8.0. The sample was incubated at RT overnight.

## 5.4.10 Mass Spectrometry

Electrospray Ionization Mass Spectrometry Mass spectra of intact proteins were obtained using a QTOF Ultima (Waters) mass spectrometer, operating under positive electrospray ionization (+ESI) mode, connected to an LC-20AD (Shimadzu) liquid chromatography unit. Protein samples were separated from small molecules by reverse phase chromatography on a Waters Xbridge BEH C4 column (300 Å, 3.5  $\mu$ m, 2.1 mm x 50 mm), using an acetonitrile gradient from 30-71.4%, with 0.1% formic acid. Each analysis was 25 min under constant flow rate of 0.2 mL/min at RT. Data were acquired from m/z 350 to 2500, at a rate of 1 sec/scan. Alternatively, spectra were acquired by Xevo G2-S QTOF on a Waters ACQUITY UPLC Protein BEH C4 reverse-phase column (300 Å, 1.7  $\mu$ m, 2.1 mm x 150 mm). An acetonitrile gradient from 5%-95% was used with 0.1% formic acid, over a run time of 5 min and constant flow rate of 0.5 mL/min at RT. Spectra were acquired from m/z 350 to 2000, at a rate of 1 sec/scan. The spectra were deconvoluted using maximum entropy in MassLynx.

Tandem Mass Spectrometry Analysis and sequencing of peptides was carried out using a Q Exactive Orbitrap interfaced with Ultimate 3000 LC system. Data acquisition by Q Exactive Orbitrap was as follows: 10  $\mu$ l of trypsin-digested protein was loaded on an Ace UltraCore super C18 reverse-phase column (300 Å, 2.5  $\mu$ m, 75 mm × 2.1 mm) via an autosampler. An acetonitrile gradient from 5%-95% was used with 0.1% formic acid, over a run time of 45 min and constant flow rate of 0.2 mL/min at RT. MS data were acquired using a data-dependent top10 method dynamically choosing the most abundant precursor ions from the survey scan for HCD fragmentation using a stepped normalized collision energy of 28, 30, 35 eV. Survey scans were acquired at a resolution of 70,000 at m/z 200 on the Q Exactive. Theoretical patterns of isotopic patterns of peptides were calculated using UCSF MS-ISOTOPE<sup>71</sup> or enviPat Web 2.1<sup>72</sup>.

# REFERENCES

- Labunskyy, V.M., Hatfield, D.L. & Gladyshev, V.N. Selenoproteins: molecular pathways and physiological roles. *Physiol Rev* 94, 739-777 (2014).
- Bellinger, F.P., Raman, A.V., Reeves, M.A. & Berry, M.J. Regulation and function of selenoproteins in human disease. *Biochemical Journal* 422, 11-22 (2009).
- 3. Yoshizawa, S. & Bock, A. The many levels of control on bacterial selenoprotein synthesis. *Biochim Biophys Acta* **1790**, 1404-1414 (2009).
- Kryukov, G. V., Castellano, S., Novoselov, S. V., Lobanov, A. V., Zehtab, O., Guigo, R. & Gladyshev, V. N. Characterization of mammalian selenoproteomes. *Science* 300, 1439-1443 (2003).
- Liu, J. & Rozovsky, S. Membrane-bound selenoproteins. *Antioxidants & redox* signaling 23, 795-813 (2015).
- Metanis, N. & Hilvert, D. Natural and synthetic selenoproteins. *Curr Opin Chem Biol* 22, 27-34 (2014).
- 7. Arner, E.S., Sarioglu, H., Lottspeich, F., Holmgren, A. & Bock, A. 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**, 1003-1016 (1999).

- Liu, J. & Rozovsky, S. Contribution of selenocysteine to the peroxidase activity of selenoprotein S. *Biochemistry* 52, 5514-5516 (2013).
- 9. Liu, J., Li, F. & Rozovsky, S. The intrinsically disordered membrane protein selenoprotein S is a reductase in vitro. *Biochemistry* **52**, 3051-3061 (2013).
- Kim, H.Y. & Gladyshev, V.N. Methionine sulfoxide reduction in mammals: characterization of methionine-R-sulfoxide reductases. *Mol Biol Cell* 15, 1055-1064 (2004).
- Rackham, O., Shearwood, A. M. J., Thyer, R., McNamara, E., Davies, S. M. K., Callus, B. A., Miranda-Vizuete, A., Berners-Price, S. J., Cheng, Q., Arner, E. S. J. & Filipovska, A. Substrate and inhibitor specificities differ between human cytosolic and mitochondrial thioredoxin reductases: Implications for development of specific inhibitors. *Free Radic Biol Med* **50**, 689-699 (2011).
- Muller, S., Senn, H., Gsell, B., Vetter, W., Baron, C. & Bock, A. The formation of diselenide bridges in proteins by incorporation of selenocysteine residues: biosynthesis and characterization of (Se)2-thioredoxin. *Biochemistry* 33, 3404-3412 (1994).
- Denk, D. & Bock, A. L-cysteine biosynthesis in Escherichia coli: nucleotide sequence and expression of the serine acetyltransferase (cysE) gene from the wild-type and a cysteine-excreting mutant. *J Gen Microbiol* 133, 515-525 (1987).

- Nakamori, S., Kobayashi, S.I., Kobayashi, C. & Takagi, H. Overproduction of L-cysteine and L-cystine by Escherichia coli strains with a genetically altered serine acetyltransferase. *Appl Environ Microbiol* 64, 1607-1611 (1998).
- Jiang, Z. H., Arner, E. S. J., Mu, Y., Johansson, L., Shi, J. M., Zhao, S. Q., Liu,
   S. J., Wang, R. Y., Zhang, T. Z., Yan, G. L., Liu, J. Q., Shen, J. C. & Luo, G.
   M. Expression of selenocysteine-containing glutathione S-transferase in
   Escherichia coli. *Biochem Biophys Res Commun* 321, 94-9101 (2004).
- Yu, Y., Song, J., Guo, X., Wang, S., Yang, X., Chen, L. & Wei, J. Y. Characterization and structural analysis of human selenium-dependent glutathione peroxidase 4 mutant expressed in Escherichia coli. *Free Radic Biol Med* 71, 332-338 (2014).
- Song, J., Yu, Y., Xing, R., Guo, X., Liu, D. L., Wei, J. Y. & Song, H. W. Unglycosylated recombinant human glutathione peroxidase 3 mutant from Escherichia coli is active as a monomer. *Scientific Reports* 4, 6698 (2014).
- Guo, X., Song, J., Yu, Y. & Wei, J. Can recombinant human glutathione peroxidase 1 with high activity be efficiently produced in Escherichia coli ? *Antioxidants & redox signaling* 20, 1524-1530 (2014).
- Guo, X., Yu, Y., Liu, X. X., Zhang, Y. L., Guan, T. C., Xie, G. Q. & Wei, J. Y. Heterologous expression and characterization of human cellular glutathione peroxidase mutants. *IUBMB life* (2014).

- Strub, M. P., Hoh, F., Sanchez, J. F., Strub, J. M., Bock, A., Aumelas, A. & Dumas, C. Selenomethionine and selenocysteine double labeling strategy for crystallographic phasing. *Structure* 11, 1359-1367 (2003).
- 21. Salgado, P.S., Taylor, J.D., Cota, E. & Matthews, S.J. Extending the usability of the phasing power of diselenide bonds: SeCys SAD phasing of CsgC using a non-auxotrophic strain. *Acta Crystallogr D Biol Crystallogr* **67**, 8-13 (2011).
- 22. Sanchez, R. & Sali, A. Comparative protein structure modeling Introduction and practical examples with modeller. *Protein Structure Prediction: Methods and Protocols* 143, 97-129 (2000).
- Dikiy, A., Novoselov, S. V., Fomenko, D. E., Sengupta, A., Carlson, B. A., Cerny, R. L., Ginalski, K., Grishin, N. V., Hatfield, D. L. & Gladyshev, V. N. SelT, SelW, SelH, and Rdx12: genomics and molecular insights into the functions of selenoproteins of a novel thioredoxin-like family. *Biochemistry* 46, 6871-6882 (2007).
- 24. Arner, E.S. Selenoproteins-What unique properties can arise with selenocysteine in place of cysteine? *Exp Cell Res* **316**, 1296-1303 (2010).
- Hondal, R.J., Marino, S.M. & Gladyshev, V.N. Selenocysteine in thiol/disulfide-like exchange reactions. *Antioxidants & redox signaling* 18, 1675-1689 (2013).
- Kryukov, G.V., Kumar, R.A., Koc, A., Sun, Z.H. & Gladyshev, V.N.
   Selenoprotein R is a zinc-containing stereo-specific methionine sulfoxide

reductase. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 4245-4250 (2002).

- Schweizer, U., Schlicker, C., Braun, D., Kohrle, J. & Steegborn, C. Crystal structure of mammalian selenocysteine-dependent iodothyronine deiodinase suggests a peroxiredoxin-like catalytic mechanism. *Proc Natl Acad Sci U S A* 111, 10526-10531 (2014).
- Aldag, C., Brocker, M. J., Hohn, M. J., Prat, L., Hammond, G., Plummer, A. and Soll, D. Rewiring Translation for Elongation Factor Tu-Dependent Selenocysteine Incorporation. *Angewandte Chemie-International Edition* 52, 1441-1445 (2013).
- Thyer, R., Robotham, S.A., Brodbelt, J.S. & Ellington, A.D. Evolving tRNA(Sec) for Efficient Canonical Incorporation of Selenocysteine. *Journal of the American Chemical Society* 137, 46-49 (2015).
- Rakauskaite, R., Urbanaviciute, G., Ruksenaite, A., Liutkeviciute, Z., Juskenas, R., Masevicius, V. & Klimasauskas, S. Biosynthetic selenoproteins with genetically-encoded photocaged selenocysteines. *Chemical communications* 51, 8245-8248 (2015).
- Dawson, P., Muir, T., Clark-Lewis, I. & Kent, S. Synthesis of proteins by native chemical ligation. *Science* 266, 776-779 (1994).
- 32. Hondal, R.J., Nilsson, B.L. & Raines, R.T. Selenocysteine in native chemical ligation and expressed protein ligation. *J Am Chem Soc* 123, 5140-5141 (2001).

- Hondal, R.J. & Raines, R.T. Semisynthesis of proteins containing selenocysteine. *Methods Enzymol* 347, 70-83 (2002).
- Hondal, R.J. Incorporation of selenocysteine into proteins using peptide ligation. *Protein Pept Lett* 12, 757-764 (2005).
- 35. Metanis, N., Keinan, E. & Dawson, P.E. Synthetic seleno-glutaredoxin 3 analogues are highly reducing oxidoreductases with enhanced catalytic efficiency. *J Am Chem Soc* **128**, 16684-16691 (2006).
- Casi, G., Roelfes, G. & Hilvert, D. Selenoglutaredoxin as a glutathione peroxidase mimic. *Chembiochem* 9, 1623-1631 (2008).
- Berry, S.M., Gieselman, M.D., Nilges, M.J., van Der Donk, W.A. & Lu, Y. An engineered azurin variant containing a selenocysteine copper ligand. *J Am Chem Soc* 124, 2084-2085 (2002).
- Gieselman, M.D., Xie, L. & van Der Donk, W.A. Synthesis of a selenocysteine-containing peptide by native chemical ligation. *Org Lett* 3, 1331-1334 (2001).
- Levengood, M.R. & van der Donk, W.A. Dehydroalanine-containing peptides: preparation from phenylselenocysteine and utility in convergent ligation strategies. *Nat Protoc* 1, 3001-3010 (2006).
- 40. Malins, L.R. & Payne, R.J. Synthesis and utility of β-selenol-phenylalanine for native chemical ligation-deselenization chemistry. *Org Lett* 14, 3142-3145 (2012).

- Townsend, S. D., Tan, Z. P., Dong, S. W., Shang, S. Y., Brailsford, J. A. & Danishefsky, S. J. Advances in proline ligation. *J Am Chem Soc* 134, 3912-3916 (2012).
- Metanis, N., Keinan, E. & Dawson, P.E. Traceless Ligation of Cysteine Peptides Using Selective Deselenization. *Angewandte Chemie-International Edition* 49, 7049-7053 (2010).
- Dawson, P.E. Native Chemical Ligation Combined with Desulfurization and Deselenization: A General Strategy for Chemical Protein Synthesis. *Israel Journal of Chemistry* 51, 862-867 (2011).
- 44. Dery, S., Reddy, P. S., Dery, L., Mousa, R., Dardashti, R. N. & Metanis, N. Insights into the deselenization of selenocysteine into alanine and serine. *Chemical Science* 6, 6207-6212 (2015).
- Malins, L.R., Mitchell, N.J., McGowan, S. & Payne, R.J. Oxidative Deselenization of Selenocysteine: Applications for Programmed Ligation at Serine. *Angew Chem Int Ed Engl* 54, 12716-12721 (2015).
- 46. Malins, L.R. & Payne, R.J. Synthesis and utility of beta-selenol-phenylalanine for native chemical ligation-deselenization chemistry. *Organic letters* 14, 3142-3145 (2012).
- 47. Metanis, N., Keinan, E. & Dawson, P.E. Traceless ligation of cysteine peptides using selective deselenization. *Angewandte Chemie* **49**, 7049-7053 (2010).

- Reddy, P.S., Dery, S. & Metanis, N. Chemical Synthesis of Proteins with Non-Strategically Placed Cysteines Using Selenazolidine and Selective Deselenization. *Angewandte Chemie* 55, 992-995 (2016).
- Malins, L.R. & Payne, R.J. Recent extensions to native chemical ligation for the chemical synthesis of peptides and proteins. *Current opinion in chemical biology* 22, 70-78 (2014).
- 50. Hondal, R.J. Using chemical approaches to study selenoproteins focus on thioredoxin reductases. *Biochimica et biophysica acta* **1790**, 1501 (2009).
- Dawson, P.E. & Kent, S.B. Synthesis of native proteins by chemical ligation.
   *Annu Rev Biochem* 69, 923-960 (2000).
- 52. Harris, K.M., Flemer, S. & Hondal, R.J. Studies on deprotection of cysteine and selenocysteine side-chain protecting groups. *J Pept Sci* **13**, 81-93 (2007).
- 53. Flemer, S., Jr. Selenol protecting groups in organic chemistry: special emphasis on selenocysteine Se-protection in solid phase peptide synthesis.
   *Molecules* 16, 3232-3251 (2011).
- 54. Arner, E.S.J. Selenoproteins-What unique properties can arise with selenocysteine in place of cysteine? *Exp Cell Res* **316**, 1296-1303 (2010).
- Kapust, R.B., Tozser, J., Copeland, T.D. & Waugh, D.S. The P1' specificity of tobacco etch virus protease. *Biochem Biophys Res Commun* 294, 949-955 (2002).

- 56. Tolbert, T.J. & Wong, C.H. New methods for proteomic research: preparation of proteins with N-terminal cysteines for labeling and conjugation. *Angew Chem Int Ed Engl* **41**, 2171-2174 (2002).
- Liu, J., Srinivasan, P., Pham, D.N. & Rozovsky, S. Expression and purification of the membrane enzyme selenoprotein K. *Protein Expression and Purification* 86, 27-34 (2012).
- Pitts, M. W., Reeves, M. A., Hashimoto, A. C., Ogawa, A., Kremer, P., Seale,
   L. A. & Berry, M. J. Deletion of selenoprotein M leads to obesity without
   cognitive deficits. *The Journal of biological chemistry* 288, 26121-26134
   (2013).
- Shchedrina, V.A., Zhang, Y., Labunskyy, V.M., Hatfield, D.L. & Gladyshev,
   V.N. Structure-function relations, physiological roles, and evolution of
   mammalian ER-resident selenoproteins. *Antioxidants & redox signaling* 12,
   839-849 (2010).
- Geoghegan, K. F., Dixon, H. B. F., Rosner, P. J., Hoth, L. R., Lanzetti, A. J., Borzilleri, K. A., Marr, E. S., Pezzullo, L. H., Martin, L. B., LeMotte, P. K., McColl, A. S., Kamath, A. V. & Stroh, J. G. Spontaneous alpha-N-6phosphogluconoylation of a "His tag" in Escherichia coli: the cause of extra mass of 258 or 178 Da in fusion proteins. *Anal Biochem* 267, 169-184 (1999).
- Aachmann, F.L., Fomenko, D.E., Soragni, A., Gladyshev, V.N. & Dikiy, A.
   Solution structure of selenoprotein W and NMR analysis of its interaction with 14-3-3 proteins. *J. Biol. Chem.* 282, 37036-37044 (2007).

- Malins, L.R., Mitchell, N.J., McGowan, S. & Payne, R.J. Oxidative deselenization of selenocysteine: Applications for programmed ligation at serine. *Angew. Chem.* 54, 12716-12721 (2015).
- Chalker, J. M., Gunnoo, S. B., Boutureira, O., Gerstberger, S. C., Fernandez-Gonzalez, M., Bernardes, G. J. L., Griffin, L., Hailu, H., Schofield, C. J. & Davis, B. G. Methods for converting cysteine to dehydroalanine on peptides and proteins. *Chemical Science* 2, 1666-1676 (2011).
- Johansson, L., Gafvelin, G. & Arnér, E.S.J. Selenocysteine in proteins properties and biotechnological use. *Biochim. Biophys. Acta* 1726, 1-13 (2005).
- Bennett, M. J., Huey-Tubman, K. E., Herr, A. B., West, A. P., Ross, S. A. & Bjorkman, P. J. A linear lattice model for polyglutamine in CAG-expansion diseases. *Proc Natl Acad Sci U S A* 99, 11634-11639 (2002).
- Brzovic, P.S., Lissounov, A., Christensen, D.E., Hoyt, D.W. & Klevit, R.E. A UbcH5/ubiquitin noncovalent complex is required for processive BRCA1directed ubiquitination. *Molecular cell* 21, 873-880 (2006).
- Blommel, P.G. & Fox, B.G. A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expression and Purification* 55, 53-68 (2007).
- Cormier, C. Y., Mohr, S. E., Zuo, D. M., Hu, Y. H., Rolfs, A., Kramer, J., Taycher, E., Kelley, F., Fiacco, M., Turnbull, G. & LaBaer, J. Protein Structure Initiative Material Repository: an open shared public resource of structural

genomics plasmids for the biological community. *Nucleic Acids Research* **38**, D743-D749 (2010).

- 69. Studier, F.W. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* **41**, 207-234 (2005).
- Holmgren, A. Thioredoxin catalyzes the reduction of insulin disulfides by dithiothreitol and dihydrolipoamide. *The Journal of biological chemistry* 254, 9627-9632 (1979).
- 71. Baker, P.R., Clauster, K.R. http://prospector.ucsf.edu.
- Loos, M., Gerber, C., Corona, F., Hollender, J. & Singer, H. Accelerated isotope fine structure calculation using pruned transition trees. *Analytical chemistry* 87, 5738-5744 (2015).

## Chapter 6

## **OUTLOOK AND FUTURE PERSPECTIVES**

With the many findings described in Chapter 2-5, we here briefly highlight the major functions of SelS and SelK, as well as the method using ESL to prepare selenoprotein and its applications. We then propose the new venue for the future study of them.

### 6.1 The Interaction between SelS and p97

The majority of selenoproteins were discovered to be enzymes, and SelS was further characterized as a reductase and contained a selenylsulfide bond. A selenylsulfide bond usually exists in thioredoxin fold like selenoproteins possessing a CXXU, or UXXC redox motif. However, SelS' selenylsulfide bond has 13 residues between them (abbreviated  $CX_{13}U$ ) rendering it more flexible<sup>1</sup>. Our finding established that U188 is the attacking residue, and the C174 is the resolving residue in a thiol/selenol-disulfide exchange reaction. SelS could directly interact with its protein partner through a selenylsulfide linkage or indirectly through different conformational changes upon reduction or oxidation. Therefore, previous pull-down assays did not identify SelS' specific substrates, possible due to the interaction being transient<sup>2</sup>. However, binding partners are known and future research should investigate SelS' Sec contribution to the interaction with these partners.

Among the strongly interacting partners is the ATPase p97 (Section 1.5), a general contributor of energy for the disassembly of protein machineries and for

pulling misfolded proteins out of the ER lumen<sup>3, 4</sup>. It would be interesting to investigate how SelS interacts with p97, and how p97 is recruited to the ERAD complex. It will help better understand the regulatory role SelS plays in ERAD.

In the future, we will set out to probe the interactions between SelS and p97 *in vitro*. Previous NMR studies using p97's N domain have shown that SelS VCPinteracting motif (VIM) in residues 78-88 (with the consensus sequence RX<sub>5</sub>AAX<sub>2</sub>R) bind a hydrophobic cleft in p97's N domain<sup>5</sup>. However, binding of SelS also requires residues 174-189, which include the active site Sec at the 188<sup>th</sup> position<sup>6</sup>. To explore the interaction between SelS and p97, we would generate a series of SelS mutants, including SelS wild type, SelS U188C, SelS C174S, and SelS U188S that will be used to probe the interaction with p97. Reduced or oxidized SelS will be incubated with p97, and their interaction can be investigated through size exclusion chromatography or blue native polyacrylamide gel electrophoresis (BN-PAGE). Once a strong interaction is found, we will clone, express and purify different domains of p97 to characterize the specific region responsible for this interaction. Förster resonance energy transfer (FRET) or isothermal titration calorimetry (ITC) will be used to probe the interactions between them.

The interaction of SelS with p97 could change the conformation of the intrinsically disordered region, which may influence SelS' reductase activity. On the other hand, many p97 binding partners have been observed to modify its ATPase activity by influencing its conformation<sup>7-9</sup>; by analogy, interaction of SelS with p97 might also influence p97 oxidation state or conformation, which may affect its ATPase activity.

## 6.2 Exploration of Selenoprotein K Auto-proteolytic Cleavage

During our studies, we found that SelK forms an intermolecular diselenide bond, and it is the first example of a human selenoprotein to do so<sup>10</sup>. SelK's redox potential of -257 mV suggests that it might be biologically relevant *in vivo* because the value is within the redox potential of the cytoplasm<sup>11</sup>. Indeed, it can be reduced by thioredoxin reductase<sup>10</sup>. The diselenide bond could be transient *in vivo*, since the cytosol is rich with reductases and small molecular thiols. A diselenide bond could be formed upon an increased level of reactive oxygen species, and thus could lead to its conformational change.



Figure 6.1. SelK auto-proteolytic cleavage may be inhibited by its binding partners.

In addition, we found SelK has auto-proteolytic activity. The cleavage occurs between S<sub>55</sub> and S<sub>56</sub>, and it generates a Sec-containing peptide and a membrane embedded Sec-less protein. This cleavage is dependent on the presence of Sec, and SelK phosphorylation inhibits it (unpublished results). A recent study reported that SelK is required for the palmitoylation of inositol-1, 4, 5-triphosphate receptor (IP3R), through the interaction with protein acyltransferase, DHHC6<sup>12, 13</sup>. It was proposed that SelK promotes the palmitoylation process and serves as a cofactor. We hypothesize that SelK may be stabilized through interactions with its binding partners. Future studies will investigate whether the auto-proteolytic cleavage is a regulatory function between SelK and its binding partners, and whether interaction of SelK with DHHC6 will stabilize it (Figure 6.1). In addition, we will test the contribution of Sec to DHHC6's enzymatic activity.

# 6.3 Future Expansions of Expressed Selenoprotein Ligation for Protein Engineering

As described in chapter 5, Sec-mediated protein ligation is a powerful tool to prepare selenoproteins, as well as non-selenoproteins through selective deselenization. Prior to our report, the Sec peptide could only be prepared by SPPS. The synthesis of a long peptide can be limited by its size, solubility, and cost<sup>14</sup>. Even more so for the chemical synthesis of Sec peptide which, due to the specialized chemistry, is considerably more involved than that of conventional peptides<sup>15</sup>. Our study established that the Sec containing fragments can be made by heterologous protein expression and can be directly used for native chemical ligation. In this thesis, we focused on one-step ligation to prepare selenoproteins. For proteins that have additional cysteines on both sides of the Sec containing fragment, a sequential ligation is required. A fragment that

has either an N-terminal Sec or a C-terminal thioester must be masked during sequential ligation, to prevent self-ligation, or non-specific ligation<sup>16, 17</sup>. It is possible to do sequential ligations using N-terminal to C-terminal or C-terminal to N-terminal strategies (Figure 6.2).

Preparing the selenoproteins by ESL could also be used to enrich the selenoproteins with <sup>77</sup>Se, the only NMR active isotope of selenium<sup>18</sup>. This can be achieved by employing <sup>77</sup>Se-selenocystine with the ESL method or by randomly substituting all sulfur to <sup>77</sup>Se by adding <sup>77</sup>Se-selenite to a defined growth medium<sup>19</sup>. <sup>77</sup>Se NMR has a wide range between +2400 ppm and -900 ppm<sup>20-22</sup>, and is very sensitive to the chemical environment. These properties allow selenium NMR to probe the local environment of seleno-amino acids or the bonding pattern between selenium and other atoms<sup>23</sup>.

In addition, we have shown the utilization of the special chemistry of Sec, for site selective conjugations. Sec has a  $pK_a$  5.2; that is 3.1 pH units lower than cysteine's  $pK_a$  (8.3). In the same protein, it is possible to site selectively label Sec based on the considerable concentration differences of selenolate versus thiolate and their different nucleophilicity at lower pH environment.



Figure 6.2. Sequential ligation for preparation of selenoprotein with multiple cysteines. The challenge here is how to prepare a thioester for the second ligation step. A) A proposed sequential ligation from C-terminal to N-terminal. B) A proposed sequential ligation from N-terminal to C-terminal. T refers to the TEV protease cleavage site and X refers to factor Xa protease's cleavage site. In both cases, the Sec fragment contains a N-terminal split intein providing a thioester precursor. The N-terminal split intein fragment will only generate a thioester when incubated with the other complementary C-terminal split intein fragment.

Another expansion of ESL relies on the observation that selenoproteins can potentially also be generated with protein trans-splicing. Protein splicing is a posttranslational modification to generate active proteins in a controlled fashion, a process by which intein cleaves itself out of a larger peptide fragment while ligating the two flanking N and C exteins into a single folded protein<sup>24</sup>. A trans-splicing intein is to split the intein into N-terminal and C-terminal splicing domains. The N and C intein is fused to the N and C exteins respectively, and only the association of the full intein activates itself and induces protein splicing and joins the N and C exteins (Figure 6.3). Because Sec is more reactive and has lower pK<sub>a</sub>, substitution of the cysteine or serine in the intein into Sec might improve its reaction reactivity, and thus could change its substrate specificity.

In summary, future work will focus on extending ESL to include Sec-mediated protein trans-splicing, multi-step expressed protein ligation, or a combination of them for preparation of complex targets for biochemical characterization. We will also develop enrichment with <sup>77</sup>Se isotope and applications for site-selective chemical conjugations.



Figure 6.3. Sec-mediated protein trans-splicing. Both the intein's first residue and the first residue of the C-extein (POI<sup>CT</sup>) contain conserved residues such as Cys, Ser, or Thr. We hypothesized that their substitution into Sec may enhance protein trans-splicing because selenolate is a stronger nucleophile than thiolate, and also because the selenoester is more reactive than thioester.

### REFERENCES

- 1. Liu, J., Li, F. & Rozovsky, S. The intrinsically disordered membrane protein selenoprotein S is a reductase in vitro. *Biochemistry* **52**, 3051-3061 (2013).
- Turanov, A. A., Shchedrina, V. A., Everley, R. A., Lobanov, A. V., Yim, S. H., Marino, S. M., Gygi, S. P., Hatfield, D. L. & Gladyshev, V. N. Selenoprotein S is involved in maintenance and transport of multiprotein complexes. *Biochem J* 462, 555-565 (2014).
- Haenzelmann, P. & Schindelin, H. The structural and functional basis of the p97/Valosin-containing protein (VCP)-interacting motif (VIM). *Journal of Biological Chemistry* 286, 38679-38690 (2011).
- Meyer, H. & Weihl, C.C. The VCP/p97 system at a glance: connecting cellular function to disease pathogenesis. *Journal of Cell Science* 127, 3877-3883 (2014).
- Christensen, L. C., Jensen, N. W., Vala, A., Kamarauskaite, J., Johansson, L., Winther, J. R., Hofmann, K., Teilum, K. & Ellgaard, L. The human selenoprotein VCP-interacting membrane protein (VIMP) is non-globular and harbors a reductase function in an intrinsically disordered region. *Journal of Biological Chemistry* 287, 26388-26399 (2012).

- Lee, J. H., Kwon, J. H., Jeon, Y. H., Ko, K. Y., Lee, S. R. & Kim, I. Y. Pro178 and Pro183 of selenoprotein S are essential residues for interaction with p97(VCP) during endoplasmic reticulum-associated degradation. *The Journal* of biological chemistry 289, 13758-13768 (2014).
- Trusch, F., Matena, A., Vuk, M., Koerver, L., Knaevelsrud, H., Freemont, P. S., Meyer, H. & Bayer, P. The N-terminal Region of the Ubiquitin Regulatory X (UBX) Domain-containing Protein 1 (UBXD1) Modulates Interdomain Communication within the Valosin-containing Protein p97. *The Journal of biological chemistry* 290, 29414-29427 (2015).
- Zhang, X., Gui, L., Zhang, X., Bulfer, S. L., Sanghez, V., Wong, D. E., Lee, Y., Lehmann, L., Lee, J. S., Shih, P. Y., Lin, H. J., Iacovino, M., Weihl, C. C., Arkin, M. R., Wang, Y. & Chou, T. F. Altered cofactor regulation with disease-associated p97/VCP mutations. *Proc Natl Acad Sci U S A* 112, E1705-1714 (2015).
- 9. Niwa, H., Ewens, C. A., Tsang, C., Yeung, H. O., Zhang, X. and Freemont, P.
  S. The role of the N-domain in the ATPase activity of the mammalian AAA
  ATPase p97/VCP. *The Journal of biological chemistry* 287, 8561-8570 (2012).
- Liu, J., Zhang, Z. & Rozovsky, S. Selenoprotein K form an intermolecular diselenide bond with unusually high redox potential. *FEBS Lett* 588, 3311-3321 (2014).
- Go, Y.M. & Jones, D.P. Redox compartmentalization in eukaryotic cells. Biochim Biophys Acta 1780, 1273-1290 (2008).

- 12. Fredericks, G.J. & Hoffmann, P.R. Selenoprotein K and protein palmitoylation. *Antioxidants & redox signaling* **23**, 854-862 (2015).
- Fredericks, G. J., Hoffmann, F. W., Rose, A. H., Osterheld, H. J., Hess, F. M., Mercier, F. & Hoffmann, P. R. Stable expression and function of the inositol 1,4,5-triphosphate receptor requires palmitoylation by a DHHC6/selenoprotein K complex. *Proc Natl Acad Sci U S A* **111**, 16478-16483 (2014).
- Dawson, P.E. & Kent, S.B. Synthesis of native proteins by chemical ligation.
   *Annu Rev Biochem* 69, 923-960 (2000).
- Muttenthaler, M. & Alewood, P.F. Selenopeptide chemistry. *J Pept Sci* 14, 1223-1239 (2008).
- Quaderer, R. & Hilvert, D. Selenocysteine-mediated backbone cyclization of unprotected peptides followed by alkylation, oxidative elimination or reduction of the selenol. *Chem Commun (Camb)*, 2620-2621 (2002).
- Raibaut, L., Ollivier, N. & Melnyk, O. Sequential native peptide ligation strategies for total chemical protein synthesis. *Chemical Society Reviews* 41, 7001-7015 (2012).
- Liu, J., and Rozovsky, S. 77Se NMR of selenoproteins. In "Selenium: Its molecular biology and role in human health" (Accepted). (2016).
- Schaefer, S. A., Dong, M., Rubenstein, R. P., Wilkie, W. A., Bahnson, B. J., Thorpe, C. and Rozovsky, S. (77)Se enrichment of proteins expands the biological NMR toolbox. *Journal of molecular biology* 425, 222-231 (2013).

- Rozovsky, S. in Biochalcogen Chemistry: The Biological Chemistry of Sulfur, Selenium, and Tellurium, Vol. 1152 127-142 (American Chemical Society, 2013).
- Duddeck, H. in Annual Reports on NMR Spectroscopy, Vol. Volume 52 105-166 (Academic Press, 2004).
- 22. Duddeck, H. Selenium-77 nuclear magnetic resonance spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy* **27**, 1-323 (1995).
- Struppe, J., Zhang, Y. & Rozovsky, S. (77)Se chemical shift tensor of L-selenocystine: experimental NMR measurements and quantum chemical investigations of structural effects. *The journal of physical chemistry*. *B* 119, 3643-3650 (2015).
- 24. Paulus, H. Protein splicing and related forms of protein autoprocessing. *Annual review of biochemistry* **69**, 447-496 (2000).

# Appendix

## COPYRIGHT

## Chapter 1

Parts of the text and figures are reprinted/adapted with permission from:

Jun Liu, and Sharon Rozovsky. Membrane-bound selenoproteins. (2015) Antioxidants

& redox signaling 2015, 23, 795–813.

Copyright (2015), Mary Ann Liebert, Inc.

## Chapter 2

Parts of the text and figures are reprinted/adapted with permission from:

Jun Liu, Prabhavathi Srinivasan, Diane Pham, and Sharon Rozovsky. Expression and purification of the membrane enzyme selenoprotein K. (2012) Protein Expression and Purification 2012, 86, 27-34.

Copyright (2012), Elsevier Inc.

## Chapter 3

Parts of the text and figures are reprinted/adapted with permission from:

Jun Liu, Fei Li, and Sharon Rozovsky. The intrinsically disordered membrane protein selenoprotein S is a reductase in vitro. (2013) Biochemistry 2013, 52, 3051-3061. Copyright (2013), American Chemical Society.

Jun Liu, and Sharon Rozovsky. Contribution of selenocysteine to the peroxidase activity of selenoprotein S. (2013) Biochemistry 2013, 52, 5514-5516.

Copyright (2013), American Chemical Society.

# Chapter 4

Parts of the text and figures are reprinted/adapted with permission from:

Jun Liu, Zhengqi Zhang, and Sharon Rozovsky. Selenoprotein K form an intermolecular diselenide bond with unusually high redox potential. (2014) FEBS Letters 2014, 588, 3311–3321.

Copyright (2014), John Wiley and Sons

# Chapter 5

This material was submitted for peer review.





Title:	Membrane-Bound Selenoproteins
Author:	Jun Liu, Sharon Rozovsky
Publication:	Antioxidants & Redox Signaling
Publisher:	Mary Ann Liebert, Inc.
Date:	Oct 1, 2015
Copyright © 201	15, Mary Ann Liebert, Inc.

#### **Permissions Request**

Mary Ann Liebert, Inc. publishers does not require authors of the content being used to obtain a license for their personal reuse of full article, charts/graphs/tables or text excerpt.







Title:	Expression and purification of the membrane enzyme selenoprotein K
Author:	Jun Liu,Prabhavathi Srinivasan,Diane N. Pham,Sharon Rozovsky
Publication:	Protein Expression and Purification
Publisher:	Elsevier
Date:	November 2012
Copyright © 201	2 Elsevier Inc. All rights reserved.

#### **Order Completed**

Thank you for your order.

This Agreement between Jun Liu ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

#### Get the printable license.

License Number	3895690038979
License date	Jun 24, 2016
Licensed Content Publisher	Elsevier
Licensed Content Publication	Protein Expression and Purification
Licensed Content Title	Expression and purification of the membrane enzyme selenoprotein K
Licensed Content Author	Jun Liu,Prabhavathi Srinivasan,Diane N. Pham,Sharon Rozovsky
Licensed Content Date	November 2012
Licensed Content Volume	86
Licensed Content Issue	1
Licensed Content Pages	8
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	ENZYMATIC CHARACTERIZATION OF SELENOPROTEIN K AND S AND HARNESSING SELENOCYSTEINE FOR PROTEIN ENGINEERING
Expected completion date	Jul 2016
Estimated size (number of pages)	250
Elsevier VAT number	GB 494 6272 12
Requestor Location	Jun Liu Brown laboratory 136, the green University of Delaware
	NEWARK, DE 19716 United States Attn: Jun Liu
Total	

Copyright © 2016 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement. Terms and Conditions. Comments? We would like to hear from you. E-mail us at <a href="mailto:customercare@copyright.com">customercare@copyright.com</a>



Title:	The Intrinsically Disordered
	Membrane Protein Selenoprotein
	S Is a Reductase in Vitro
Author:	Jun Liu, Fei Li, Sharon Rozovsky
Publication	Biochemistry
Publisher:	American Chemical Society
Date:	May 1, 2013
Copyright © 20	13, American Chemical Society

#### PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional
  uses are granted (such as derivative works or other editions). For any other uses, please
  submit a new request.

BACK CLOSE WINDOW

Copyright © 2016 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement. Terms and Conditions. Comments? We would like to hear from you. E-mail us at customercare@copyright.com



UDIICATIONS I. Most Cited. Most Read.	Title:	Contribution of Selenocysteine to the Peroxidase Activity of Selenoprotein S
	Author:	Jun Liu, Sharon Rozovsky
	Publication:	Biochemistry
	Publisher:	American Chemical Society
	Date:	Aug 1, 2013
	Copyright © 201	13, American Chemical Society

#### PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional
  uses are granted (such as derivative works or other editions). For any other uses, please
  submit a new request.

BACK

CLOSE WINDOW

Copyright © 2016 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement. Terms and Conditions. Comments? We would like to hear from you. E-mail us at customercare@copyright.com



WILEY	Title:	Selenoprotein K form an intermolecular diselenide bond with unusually high redox potential
	Author:	Jun Liu,Zhengqi Zhang,Sharon Rozovsky
	Publication:	FEBS Letters
	Publisher:	John Wiley and Sons
	Date:	Aug 10, 2014
	FEBS Letters 58 of European Bio	8 (2014) 1873-3468 © 2015 Federation chemical Societies

#### **Order Completed**

Thank you for your order.

This Agreement between Jun Liu ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

Your confirmation email will contain your order number for future reference.

#### Get the printable license.

License Number	3895710546623
License date	Jun 24, 2016
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	FEBS Letters
Licensed Content Title	Selenoprotein K form an intermolecular diselenide bond with unusually high redox potential
Licensed Content Author	Jun Liu,Zhengqi Zhang,Sharon Rozovsky
Licensed Content Date	Aug 10, 2014
Licensed Content Pages	11
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Title of your thesis / dissertation	ENZYMATIC CHARACTERIZATION OF SELENOPROTEIN K AND S AND HARNESSING SELENOCYSTEINE FOR PROTEIN ENGINEERING
Expected completion date	Jul 2016
Expected size (number of pages)	250
Requestor Location	Jun Liu Brown laboratory 136, the green University of Delaware
	NEWARK, DE 19716 United States Attn: Jun Liu
Publisher Tax ID	EU826007151
Billing Type	Invoice
Billing address	Jun Llu Brown laboratory 136, the green University of Delaware
	NEWARK, DE 19716 United States Attn: Jun Liu
Total	0.00 USD



Copyright © 2016 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement, Terms and Conditions. Comments? We would like to hear from you. E-mail us at customercare@copyright.com