PRODUCTION OF A POLYCLONAL ANTIBODY AGAINST HUMAN SELENOPROTEIN M

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERISTY OF HAWAIʻI IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

BIOLOGY

MAY 2008

By
成玉
Cheng Yu

Thesis Committee:

Marla J. Berry, PhD, Chairperson
Peter R. Hoffman, PhD
Pratibha V. Nerurka, PhD

We certify that we have read this dissertation and that, in our opinion, it is satisfactory in scope and quality as a dissertation for the degree of Master of Science in Biology.

THESIS COMMITTEE

_____

Chairperson

_____

_____

ABSTRACT

Cheng Yu

Production of a Polyclonal Antibody against Human Selenoprotein M

Thesis directed by Professor Marla J. Berry, PhD

Selenoproteins contain selenocysteine (Sec), which is encoded in selenoprotein genes by UGA. The Sec insertion sequence (SECIS) element is essential for the recognition of UGA as a codon for Sec rather than as a codon for a stop signal. Functions for many selenoproteins remain unknown, including the endoplasmic reticulum resident selenoprotein M (SelM). The lack of effective antibody against SelM, however, limits the investigation into its functions. Previous attempts in our laboratory to use polypeptides as antigens failed to produce rabbit polyclonal antibodies against SelM. The full-length protein, therefore, was considered as an alternative antigen for antibody production. In this study, full-length, histidine-tagged (His-tagged) human SelM proteins were expressed in *E. coli*. To assist the translation, Sec was mutated to cysteine (Cys). Protein was then purified using a series of techniques, primarily involving immobilized metal affinity chromatography (IMAC) purification for His-tagged proteins and electroelution purification for in-gel proteins. Samples were then analyzed by mass spectrometry and nano-LC-MS/MS to determine the amino acid sequence. After SelM was verified, the protein was concentrated. For future

studies, these samples will be used for animal injection to produce antibodies against human SelM.

The form and content of this abstract are approved. I recommend its publication.

Signed _____
Faculty member in charge of thesis

CONTENTS

FIGURES

ACRONYMS AND ABBREVIATIONS

6X His-tagged	Polyhistidine-tag
ATP	Adenosine Triphosphate
att	Attachment Site
***att*B**	Attachment Site on *E. Coli* Chromosome
***att*L**	Hybridized Attachment Site L
***att*P**	Attachment Site on λ Chromosome
***att*R**	Hybridized Attachment Site R
bp	Base Pair
BSA	Bovine Serum Albumin
cDNA	Complementary DNA
cGPx	Cellular Glutathione Peroxidase
Cm^R	Chloramphenicol Resistance Gene
CVB3	Coxsackievirus B3
Cys	Cysteine
dNTP	Nucleotide
DRI	Dietary Reference Intake
DTT	Dithiothreitol
E. coli	Escherichia Coli
EDTA	Ethylenediaminetetraacetic Acid
EFsec	Elongation Factor for Selenocysteine
fmol	Femtomoles
GPx	Glutathione Peroxidase

GPx-1	Glutathione Peroxidase (Cellular)
GPx-2	Glutathione Peroxidase (Intestinal)
GPx-3	Glutathione Peroxidase (Plasma)
GPx-4	Glutathione Peroxidase 4 (Phospholipid)
GPx-6	Glutathione Peroxidase 6 (Olfactory)
HRP	Horseradish Peroxidase
hSelP	Human Selenoprotein P
IMAC	Immobilized Metal Affinity Chromatography
IPTG	Isopropyl-b-D-thiogalactopyranoside
kDa	Kilo-Dalton
LB	Lysogeny Broth
LC-MS/MS	Liquid Chromatography-Mass Spectrometry with Peptide Mass Fingerprinting
mRNA	Messenger Ribonucleic Acid
MWCO	Molecular Weight Cut-off
ng	Nanograms
NIa	Nuclear Inclusion A
Ni-NTA	Nickel Nitrilotriacetic Acid
OD	Optical Density
ORF	Open Reading Frame
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PVDF	Polyvinylidene Fluoride
Pfu	Pyrococcus Furiosus

$\boldsymbol{pK_a}$	$-\log_{10} \times (K_a)$, where K_a is the ionization constant.
Rdx	Cyclotrimethylenetrinitramine
SBP2	RNA Binding Protein 2 for Selenoprotein Synthesis
SDS-PAGE	Sodium Dodecyl Sulfate-polyacrylamide Gel Electrophoresis
Sec	Selenocysteine
SECIS	Selenocysteine Insertion Sequence
SOC	Super Optimal Broth with Glucose
SPS1	Selenophosphate Synthetase 1
SPS2	Selenophosphate Synthetase 2
TE	Trishydroxymethylaminomethane-Ethylenediaminetetraacetic Acid
TEV	Tobacco Etch Virus
TPN	Total Parenteral Nutrition
TR	Thioredoxin Reductases
tRNA	Transfer Ribonucleic Acid
trsp	Selenocysteine $tRNA^{[Ser]Sec}$ Gene
TRX	Reduced Thioredoxin
Tris	Trishydroxymethylaminomethane
UTR	Untranslated Regions

INTRODUCTION

Selenium is attained through the diet.

Selenium is attained through the diet. Selenium deficiency is relatively rare in healthy, well-nourished individuals. Patients with severely compromised intestinal function or people undergoing Total Parenteral Nutrition (TPN) are at higher risk of selenium deficiency. Alternatively, foods grown from selenium-deficient soil often contain low levels of this micronutrient, which may lead to lower selenium status in those individuals restricted to this food source. In the U.S., the Dietary Reference Intake (DRI) for adults is 55 μg per day. This recommendation is based on levels of plasma glutathione peroxidase (GPx). Alternatively, selenoprotein P has been suggested as a better indicator of selenium nutritional status [37]. This is due to the role of selenoprotein P as a major transporter of Se in the serum, and in certain organs, it is believed to be the local Se storage and recycling protein that directly maintains Se level. Full expression of selenoprotein P requires more than 66 μg per day [46].

Keshan disease, a disease found in the southwestern province of China for which it is named, is potentially fatal, and it is caused by selenium deficiency. There are data that suggest that Keshan disease also involves infection with coxsackievirus B3 (CVB3), and selenium deficiency converts otherwise benign strains of CVB3 into more virulent strains [3]. Selenium deficiency, along with iodine deficiency, also contributes to Kashin-Beck disease, a permanent and

disabling osteoarticular disease involving growth and joint cartilage that occurs in parts of China and central Africa [8]. Studies have indicated reduced prevalence of these diseases with selenium supplementation, and Keshan disease is wholly preventable with selenium supplementation.

Selenium is incorporated into selenoproteins as Sec.

Selenium is incorporated into selenoproteins as the amino acid, selenocysteine (Sec; Figure 1). This involves a complex, unique translation process as described in more detail below. The result of this process is the insertion of Sec into the active site of selenoproteins. The selenium atom in Sec confers a much higher reactivity than cysteine, as its lower pK_a (5.2) allows it to remain ionized at physiological pH. Most selenoproteins use their higher nucleophilic activity to catalyze redox reactions.

The availability of complete genome databases has allowed investigators to search for selenoproteins in the whole genomes, based on the conserved features contained within gene sequences of selenoproteins. Using this approach, twenty-five human selenoprotein genes have been identified, twenty-four of which exist as selenoprotein orthologues in the mouse genome [29]. In addition, alternative mRNA splicing is likely to raise the number of selenoprotein isoforms with different functions well above twenty-five.

Less than half of the known selenoproteins have been extensively characterized. One of the first selenoproteins to be identified and characterized

was cellular glutathione peroxidase (cGPx, also called GPx-1) [38]. In addition to cGPx, there are other GPx proteins including intestinal GPx (GPx-2), plasma GPx (GPx-3), and phospholipid GPx (GPx-4) [45]. GPx-6 is a selenoprotein in humans expressed in olfactory tissues but exists as a cysteine-containing protein in rodents [20].

Another selenoprotein subfamily is the group of Sec-containing thioredoxin reductases (TR-1, TR-2, and TR-3). These enzymes regenerate reduced thioredoxin (TRX), which is used for maintaining cellular antioxidant systems, activating signaling molecules, reducing ribonucleotides to deoxyribonucleotides for synthesis of DNA, and regulating activity of transcription factors [1]. The different TR enzymes have different localizations including the cytoplasm and nucleus (TR-1), mitochondria (TR-2), and sperm (TR-3) [2, 36]. Both TR-1 and TR-2 are widely distributed throughout tissues, and each of the knockouts in mice is embryoniclethal, emphasizing their essential biological role [9, 25].

A third subgroup of selenoproteins that has been functionally characterized includes the deiodinases. Thyroid hormone metabolism is regulated by three selenoprotein enzymes within the iodothyronine deiodinase family: types 1, 2, and 3 (D1, D2, and D3), which are membrane-anchored enzymes of 29–33 kDa that share substantial sequence homology and catalytic properties [5]. Thyroid hormone action is initiated by the activation of T4 to T3, an outer ring monodeiodination reaction catalyzed by D1 or D2. T4 and T3 are irreversibly inactivated via inner ring monodeiodination catalyzed by D3. Thus, thyroid

hormone metabolism is dependent upon the combined actions of the three deiodinases and is regulated mainly through D2 stability in response to changes in iodine supply, to cold exposure, and to changes in thyroid gland function [19, 40, 41]. All three deiodinases are expressed in a number of fetal and adult tissues. Their tissue and developmental expression patterns suggest that deiodinases may control the concentration of active thyroid hormone available to specific tissues or cell types at certain stages of development [22].

A recent study by the Gladyshev laboratory defined and characterized a new subfamily of selenoproteins, designated Rdx, that contain primary sequence motifs and a thioredoxin-like fold that reflect a possible redox function [12]. This subfamily includes SelW, SelV, SelT and SelH, as well as cysteine-containing proteins of unknown function (Rdx12). SelR has been characterized as a methionine sulfoxidase enzyme that reduces proteins that have been reversibly oxidated on methionine groups, and SelR appears to be widely distributed throughout tissue and cell types [26]. Two other partially characterized selenoproteins include SelM and the distantly related 15-kDa selenoprotein (Sep15). The structures of SelM and Sep15 have recently been solved establishing these proteins as endoplasmic reticulum-resident redox proteins [16, 27]. The function of these two selenoproteins, however, remains unknown.

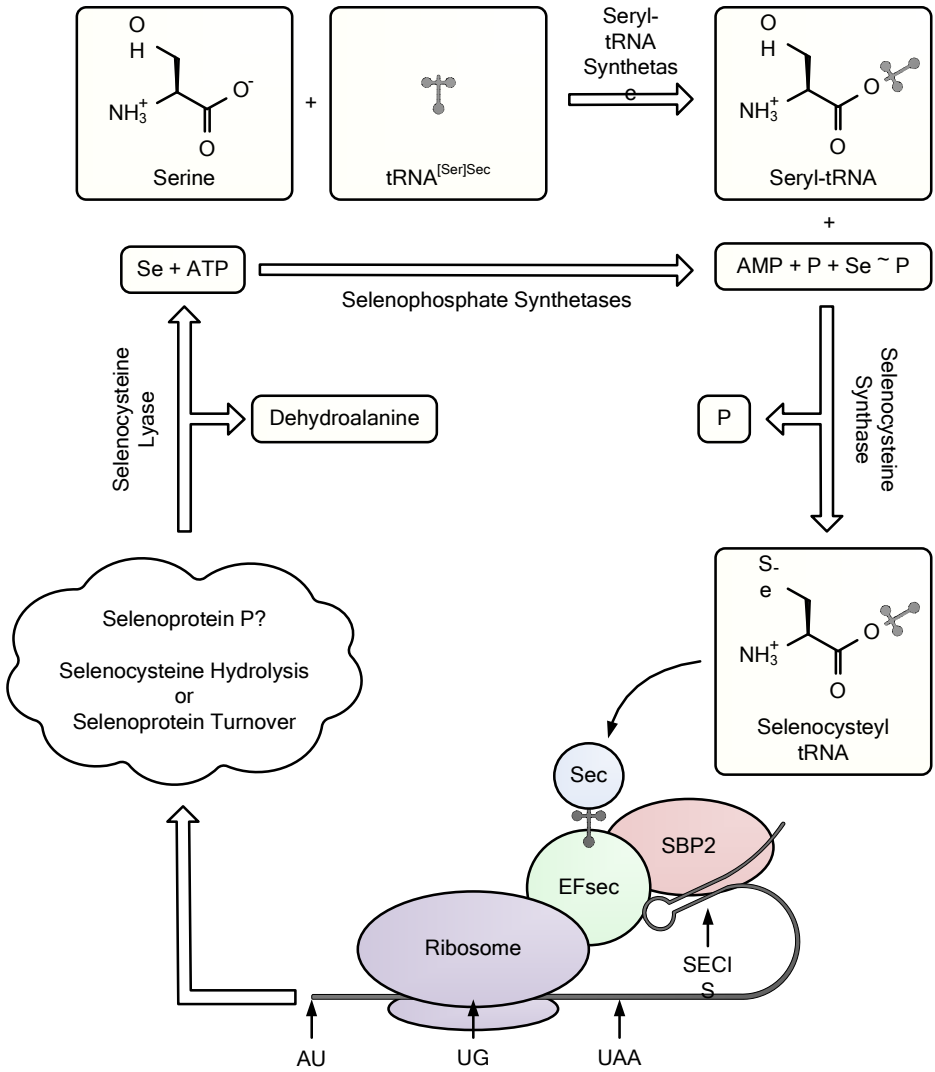

Figure 1. Co-translational insertion of Sec into selenoproteins. Unlike other essential trace elements such as zinc and iron, selenium is incorporated directly into proteins as the amino acid, Sec. This process requires several dedicated, unique elements.

Recoding of UGA codons involves special translational components.

Unlike other essential trace elements such as zinc and iron, selenium is incorporated directly into proteins as an amino acid, Sec. Translation of selenoproteins is similar to generalized protein translation in that it consists of three main steps: initiation, elongation, and termination. The special feature of selenoprotein translation lies in the recoding of the UGA codon from a stop codon to Sec. That is, the translational machinery within the cell typically reads the UGA codon as a termination signal, releasing the nascent polypeptide from the ribosome. During translation of selenoproteins, the machinery is redirected to insert Sec at UGA codons instead of terminating polypeptide synthesis. Most selenoprotein mRNAs, with the exception of selenoprotein P, contain a single UGA codon encoding a single Sec residue per polypeptide chain. The incorporation of Sec requires unique cellular mechanisms and factors to synthesize this amino acid, recognize the site at which it is to be inserted, and carry out the insertion process. The recoding process in eukaryotes involves unique, dedicated factors including a specific secondary structure in the mRNA, a unique tRNA, an RNA binding protein (SBP2), and a specialized elongation factor (EFsec), all of which are discussed in detail below.

In eukaryotic selenoprotein mRNAs, required stem-loop structures are located in the 3' untranslated regions (UTR) of the messages at variable and often considerable distances from the Sec codons (Figure 2). These stretches of RNA are

termed Sec insertion sequences, or SECIS elements, and they serve to recode the entire message. The SECIS element functions for any upstream in-frame UGA codon [4, 23], provided a minimal spacing requirement of ≈60 nucleotides is met [35].

In addition to the *cis*-acting SECIS elements signaling recoding, cotranslational incorporation of Sec requires selenocysteyl-tRNA (Sec-tRNA^{Sec}). The necessity of Sec-tRNA^{Sec} in selenoprotein synthesis was first demonstrated by the study in which homozygous deletion of the Sec-tRNA^{Sec} gene (*trsp*) resulted in embryonic lethality in mice [6]. Subsequent data revealed that conditional deletion of the gene in liver abrogated selenoprotein synthesis in this tissue and selenium supply to tests [7, 30]. Sec-tRNA^{Sec} is synthesized via a metabolic pathway that has been characterized in prokaryotes, particularly in *E. coli*, but is still not fully elucidated in eukaryotes [32, 39, 43]. The synthesis of cytoplasmic selenophosphate from ATP and selenide is catalyzed by the enzyme selenophosphate synthetase. Two forms of selenophosphate synthetase, SPS1 and SPS2, have been described in eukaryotes [21, 34]. In humans, one variant contains a cysteine or Sec residue at position 17 in the active site and catalyses the formation of selenophosphate. The other human variant, although sharing sequence similarity with the previous one, does not have a cysteine or Sec at position 17. Selenophosphate then serves as the selenium donor molecule in the conversion of seryl-tRNA^{Sec} (Ser-tRNA^{Sec}) into Sec-tRNA^{Sec}, a reaction carried out by Sec synthase in eukaryotes [47].

The third factor required for selenoprotein synthesis is SBP2, an 846 amino acid protein that binds specifically to the SECIS element [10, 17, 33]. Immunodepletion of SBP2 from reticulocyte lysates has been shown to abolish Sec incorporation in vitro, while addition of recombinant SBP2 restored Sec insertion, clearly demonstrating SBP2 protein is required for Sec incorporation in vitro [10]. Importantly, SBP2 contains neither elongation factor homology nor activity. This observation is reconciled by the fact that SBP2 binds to a fourth factor that is involved in selenoprotein synthesis with elongation activity, EFsec [44]. EFsec specifically interacts with aminoacylated Sec-tRNA^{Sec}, but not Ser-tRNA^{Ser} or Ser-tRNA^{Sec} [15, 44]. Co-precipitation studies identified the C-terminal 64 amino acids of EFsec as sufficient for the interaction with SBP2 [48]. Binding of Sec-tRNA^{Sec} by the N-terminal elongation factor domain of EFsec strongly enhances interaction with SBP2 and, in addition, stabilizes the C-terminal domain of EFsec. The latter finding implies that Sec-tRNA^{Sec} binding may affect the conformation of the C-terminal interaction domain of EFsec. An overall schematic illustrating the structure of the SECIS element is shown in Figure 3.

Other protein factors have been identified that may play important roles in selenoprotein synthesis. One such factor is the SECp43 protein that has been shown to bind to Sec-tRNA^{Sec} [13]. Another factor is a protein that was identified by its ability to bind to Sec-tRNA^{Sec} and is recognized by autoantibodies in patients presenting with type-1 autoimmune hepatitis [11, 18]. How these proteins are involved in selenoprotein synthesis remains to be determined, but they may

participate in the biosynthesis of Sec-tRNA^{Sec} from Ser-tRNA^{Sec}, modification of Sec-tRNA^{Sec}, or some other step upstream of the cotranslational incorporation of Sec.

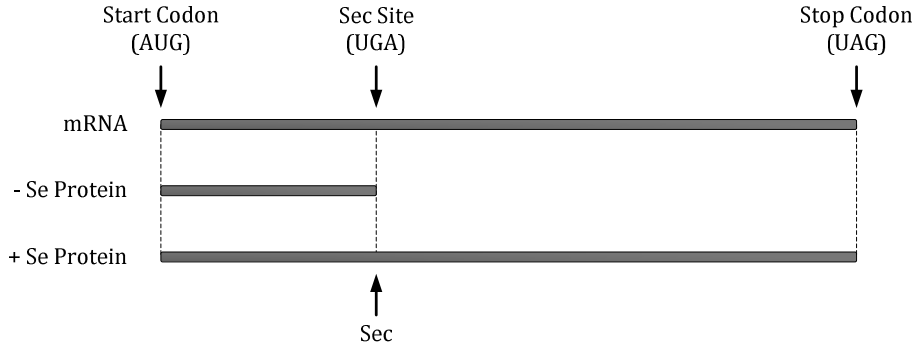

Figure 2. The requirement of dietary selenium on the expression of full-length selenoproteins. Selenoprotein contains a UGA codon, which is used as a stop codon in the absence of Sec-tRNA^{Sec}. When selenium is provided and Sec-tRNA^{Sec} is generated in the cytoplasm, UGA functions as a Sec codon to incorporate Sec into protein translation.

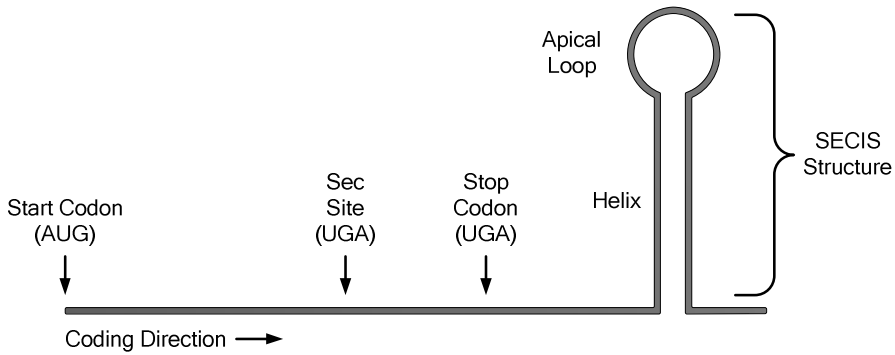

Figure 3. General structure of the SECIS element. The UGA codon is converted from a stop codon to a Sec insertion codon during a process directed in part by the SECIS element located in the 3' UTR. Location of the Sec insertion site varies for different selenoprotein mRNA as does the distances between the UGA codon, stop codon, and SECIS elements. Furthermore, some selenoprotein mRNA contain an additional loop (mini-helix) structure in their SECIS element. The structure of human selenoprotein P (hSelP) mRNA is quite different from those of other selenoproteins in that it contains 10 Sec amino acids and 2 SECIS elements [24].

SelM, a new eukaryotic selenoprotein, has been identified.

The 3-kb human SelM encoding gene has five exons and is located at chromosome position 22q12.2 [20] (Figure 4). Chromosome 22 was the first chromosome sequenced by the Human Genome Project, and it is known to contain at least 545 protein encoding genes. The SelM encoding gene was not initially correctly identified because ATG was processed as the stop signal.

The 3' UTR of the SelM encoding gene lacks a canonical SECIS element. Instead, Sec is incorporated in response to a conserved mRNA structure in which cytidines (C) are present in place of adenosines (A). Adenosine was previously considered invariant. This new structure of SECIS element was found by modified SECIS searching software that accepts a CC sequence in place of AA. The SelM SECIS is more similar to the type II SECIS element, which differs from the type I SECIS element in that it contains an additional mini-helix in the apical loop. Like both previously proposed SECIS element forms, SelM SECIS contains the UGA…GA motif that forms the non-Watson-Crick base-paired quartet and is involved in SBP2 binding. The study of the SelM SECIS element indicates that two factors are essential for the SECIS function. They are the UGA…GA motif in the quartet and the actual three-dimensional structure [28].

SelM cDNA sequences have been characterized, and the selenoprotein has been expressed in various mammalian cell lines. Human SelM is localized and retained in the perinuclear structures, a rare location for selenoproteins. Its N-terminal signal peptide is necessary for protein translocation since the first 23

residues contain an ER-signal sequence. As judged by mRNA levels, SelM expresses in many tissues with the highest levels in the brain and the lowest levels in liver and spleen.

Similar to many other selenoproteins, the SelM sequence contains a thioredoxin reductase-like redox structure, CXXU motif. It indicates that SelM could be a redox active protein [20]. No commercial antibody is effective against SelM, and this limits investigation into its functions.

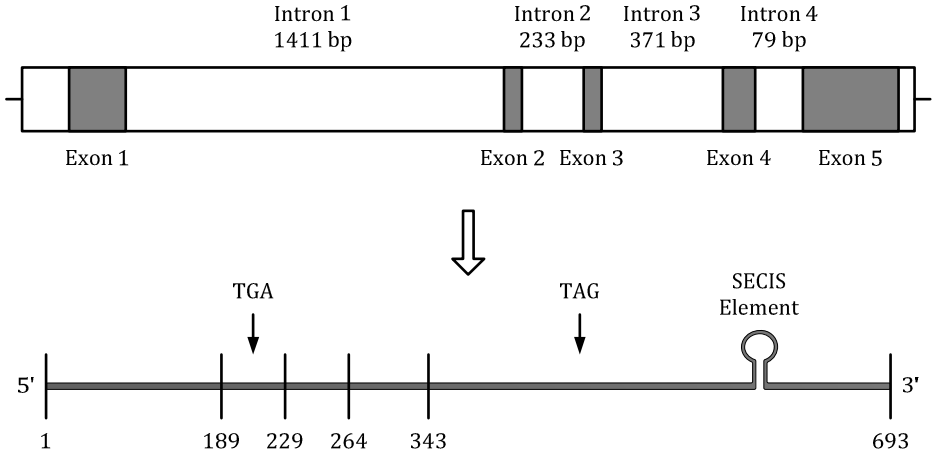

Figure 4. Organization of the human SelM encoding gene. Human SelM has five exons. The SECIS element is located at the 3' UTR region of the gene. TGA, which incorporates Sec into human SelM, is found at the N-terminal of the gene [28].

HYPOTHESIS

The full-length SelM protein serves as a better antigen than SelM peptides for producing antibodies.

Like other proteins, studying the role of selenoproteins in different biological processes requires effective antibody reagents. A limited selection of commercially available antibodies, however, have proven effective in detecting selenoproteins by several laboratories, including our laboratory. Commercially available antibodies against human SelM have been evaluated and none of them have been found to effectively recognize SelM by western blot or immunofluorescence. In fact, previous work in our laboratory was carried out in attempt to produce rabbit polyclonal antibodies recognizing five different selenoproteins for which there are currently no effective commercially available antibody reagents. These included selenoproteins I, K, M, P, and W. Peptides were designed for each selenoprotein based on antigenic indices for each amino acid sequence. These peptides were then injected into rabbits by a commercial vendor (ProSci, Inc.), and affinity purified antibodies were obtained. Of the five antibodies produced in this manner, only one (anti-SelW) was effective at recognizing its cognate protein by western blot, immunohistochemistry, or immunofluorescence. This led to the hypothesis that full-length selenoproteins might serve as better antigens for producing rabbit polyclonal antibodies.

METHODS

Proteins were extracted from human cell lines.

Human cell lines, U937 and Jurkat cells, were cultured until the density was approximately 10^6 cells per mL. Cells were collected into 15 mL conical tubes by centrifuging for 5 min at 2,500 × g and room temperature. The supernatant was decanted and discarded, and then the cell pellets were washed once with PBS and centrifuged for 5 min at 2,500 × g. The supernatant was again decanted and discarded, and then the cell pellet was re-suspended in CelLytic™ MT Reagent (Sigma-Aldrich, Inc.) containing 1 mM of DTT, 1 × protease inhibitor cocktail (Calbiochem®), and 5 mM of EDTA.

The cells were incubated for 15 min on a shaker, and then the cell lysate was collected. The cellular debris was pelleted by centrifuging at 20,000 × g (Microfuge® Microcentrifuge; Beckman Coulter, Inc.), and then the protein-containing supernatant was transferred to chilled tubes and stored at –70°C.

SelM antibodies were tested using western blots.

Western blot was used to analyze effectiveness of polyclonal antibodies for detecting human SelM. Thirty μg of protein was combined with reduced Laemmli buffer, boiled at 95°C for 10 min, and then cooled on ice. The protein sample was loaded into wells of polyacrylamide gradient gels (Bio-Rad Laboratories, Inc.). Electrophoresis was applied, and then the protein was transferred from the gel to PVDF membranes (Bio-Rad Laboratories, Inc.) under constant currency of 0.12 A

for one night. The membranes were blocked for 1 h with 5% BSA and were then probed for 1 h with primary antibodies, including all the commercially available anti-Human SelM antibodies. They were incubated with the appropriate HRP-conjugated secondary antibodies from Jackson ImmunoResearch Laboratories, Inc. with the membranes for 45 min. The membranes were detected using ECL Plus™ (GE Healthcare, General Electric Company), and images were captured using the KODAK Gel Logic 200 and KODAK Molecular Imaging Software (Carestream Health, Inc.) for the autoradiographic films.

SelM cDNA was transferred to an entry clone using BP reaction.

A complete SelM cDNA sequence was transferred from pCMV•SPORT6 vectors to pDONR-211 vectors through the recombination of *att*B sites on the pCMV•SPORT6 and *att*P sites on the pDONR-211.

Two μL (158.4 ng) of the linearized *att*B SPORT6 expression clone was added into a 0.2 mL PCR tube and mixed with 1 μL (150 ng) of pDONR-211 vector (supercoiled) and 5 μL of pH 8.0 TE buffer. The Gateway® BP Clonase™ II Enzyme Mix (Invitrogen™ Corporation) was removed, thawed on ice for 2 s, and then briefly vortexed twice. Two μL of the mix was added to the sample, and then the sample was mixed by vortexing twice briefly. The reactions were incubated at 25°C for 1 h, and then 1 μL of proteinase K was added to each reaction. They were incubated for 10 min at 37°C to stop the reaction, and then the protease in the system was removed.

Fifty femtomoles of linearized *att*B SPORT6 expression clone was required for the BP reaction. The size of the clone was 4,800 bp. To determine the measurable amount of DNA to use for the reaction, the following formula was used to convert femtomoles (fmol) to nanograms (ng).

$$N(\eta, \rho) = (\eta \text{ bp}) \times (\rho \text{ fmol}) \times \left(\frac{660 \text{ fg}}{\text{fmol}}\right) \times \left(\frac{1 \text{ ng}}{10^6 \text{ fg}}\right) \times \left(\frac{1}{\text{bp}}\right)$$

$$= (4800) \times (50) \times (6.6 \times 10^{-4}) \text{ ng}$$

$$= 158.4 \text{ ng}$$

Plasmids were transformed into One Shot® TOP10 *E. coli*.

pDONR-211 vectors with SelM cDNA were transformed into *E. coli* TOP10 strain cells. One μL of the BP reaction product was added into a vial of One Shot® TOP10 (Invitrogen™ Corporation) cells and mixed gently. It was incubated on ice for 30 min, followed by 30 s of heat shock at 42°C without shaking. The tubes were incubated on ice for 2 min, and then 250 mL of room temperature SOC was added. The tube was shaken and incubated horizontally at 37°C for 1 h. The transformation was spread on a pre-warmed LB plate with chloramphenicol for pDONR-211 clones (and later on, ampicillin for pET160 clones). It was incubated overnight at 37°C and then the colonies were picked up for analysis.

Single nucleotide mutation was introduced by mutagenesis inverse PCR.

Single nucleotide mutation from TGA to TGC was introduced in SelM cDNA insertion by mutagenesis inverse PCR. It was carried out with two controls, one

with less template plasmid (10 ng) and one without primers to test the digestion efficiency of *Dpn*I protease. PCR primers were designed to be adjacent but non-overlapping at their 5' ends, with the desired site mutation at the 5' end of one primer. The primers were phosphorylated. The primers were */phosphate 5'/-TGC CAG CTG AAC CGC CTA AAG G-3'* and */phosphate 5'/-TCC CCC GCA GGT CTC TAC CC-3'*.

Pfu Deep Vent$_R$™ DNA polymerase (New England BioLabs®, Inc.) 1.25 U within 0.65 μL was used in the 50-μL PCR system. Fifty ng of template plasmids, 50 pmol (each) phosphorylated primer was added. There were also 1 μL of 10 mM dNTP mix (Invitrogen™ Corporation) and 5 μL of 10 × *Pfu* ThermoPol Reaction Buffer (New England BioLabs®, Inc.). Reaction conditions were 94°C for 4 min, followed by 15 cycles of 94°C for 1 min, annealing for 1 min at 10°C below the calculated melting temp of 4G/C + 2A/T (≈56°C), and extension at 72°C for 2.5 min per kb of the plasmid length. The extension for each cycle took 12 min (4.8 × 2.5). Five μL of product was stored for each sample and run on a gel to test the PCR efficiency. The rest was dissolved in 18 μL of distilled water and then added with 2 μL of 10 × T4 DNA Ligase Reaction Buffer (New England BioLabs®, Inc.). Ten μL was transferred to a new tube and added with 5 U of 0.1 μL T4 DNA ligase. The remaining 10 μL served as a ligation control (no ligase was added). It was incubated at room temperature for 2 h, followed by 20 min at 68°C to inactivate the ligase and stop the reaction. Five μL of the ligation product and 5 μL of the ligation control was digested with *Dpn*I. Five μL of ligation or ligation control was

combined with 12 μL of water, 2 μL of 10 × Buffer 4 (New England BioLabs®, Inc.), and 20 units of 1 μL *Dpn*I nuclease (New England BioLabs®, Inc.). It was incubate at 37°C for 1.5 h, and 1 μL was used to transform the TOP10 competent cells.

Fragment deletion was carried out by mutagenesis inverse PCR.

Fragment deletion including 5' UTR was accomplished in SelM cDNA by mutagenesis inverse PCR. It was also carried out with two controls, one with less template plasmid (10 ng) and one without primers to test the digestion efficiency of *Dpn*I protease. PCR primers were designed to be nonadjacent at their 5' ends because of the deletion, and the construct was left with a right open reading frame for the cloned gene, SelM. Primers were phosphorylated at their 5' ends. The following sets of oligonucleotides were used: */phosphate 5'/-ATG AGC CTC CTG TTG CCT CCG-3'* and */phosphate 5/'-TCA GCC TGC TTT TTT GTA CAA AGT TGG CA-3'*. Polymerase chain reaction was applied at the same condition as above. The following day, the T4 ligation was carried out to close the vector, *Dpn*I digestion was performed to eliminate the template vector, and the amplified vector was transformed into TOP10 competent cells.

SelM cDNA was transferred to an expression clone using LR reaction.

SelM-encoding cDNA was transferred from pDONR-211 vectors to pET160 vectors through the recombination of *att*L sites on pDONR-211 and *att*R sites on pET160.

One μL (50-150 ng) of the entry clone was added into a 0.2 mL PCR tube and mixed with 1 μL (150 ng) of pET160 vector and 6 μL of pH 8.0 TE buffer. The Gateway® LR Clonase™ II Enzyme Mix (Invitrogen™ Corporation) was removed, thawed on ice for 2 min, and then briefly vortexed twice. Two μL of the mix was added to the sample and mixed by pipetting. The reactions were incubated at 25°C for 1 h, and then 1 μL of proteinase K was added to each reaction, followed by incubation for 10 min at 37°C to stop the reaction and digest the enzyme. One μL was transformed into One Shot® TOP10 Competent Cells (Invitrogen™ Corporation), as described above, for characterizing analysis and plasmid storage. After sequencing and confirming the vector construct, a fresh transformation was applied to transform vectors that were extracted from the TOP10 cells into the expression host DE3 competent cells.

Small quantities of bacteria were cultured and collected for pilot expression.

Small quantities of SelM were expressed, and cells were collected at six time points. Ten mL of LB containing ampicillin was inoculated with 500 μL of the DE3 transformants. Bacteria were grown for two h at 37°C with shaking until the OD600 was about 0.5-0.8 (mid-log). The culture was split into two 5 mL cultures. IPTG was added to one of the cultures at a final concentration of 1 mM. This yielded two bacteria cultures, one with IPTG induction, and one without induction. An aliquot of 500 μL was removed from each culture. The cultures were then

centrifuged at a maximum speed in a microcentrifuge for 30 s (Microfuge® Microcentrifuge; Beckman Coulter, Inc.), and the supernatant was aspirated. The cell pellets were frozen at −20°C, and the cultures continued to be incubated at 37°C with shaking. Samples were taken for each culture every hour for 6 h. For each time point, 500 µL were removed from the induced and non-induced cultures, and the freeze pellets were collected.

Protein samples from the pilot expression were extracted, and whole cell lysates were collected. The cell pellets from the pilot expression were thawed and re-suspended in 50 µL of a CelLytic™ B lysis buffer (Sigma-Aldrich, Inc.). Fifteen µL of each sample was transferred to a fresh tube for the analysis of the total bacteria lysates.

Polyacrylamide gels were detected with Lumio™ Fluorescent Reagent.

Lumio™ fluorescence was used for the in-gel detections. Five µL of 4 × Lumio™ Gel Sample Buffer was added to each 15 µL lysate sample. The Lumio™ Green Detection Reagent was thawed and mixed well by pipetting up and down. Each sample was added with 0.2 µL of the Lumio™ Green Detection Reagent. The samples were mixed well and incubated at 70°C for 10 min. The samples were allowed to cool for 1-2 min and centrifuged briefly at high speed in a microcentrifuge. The Lumio™ In-Gel Detection Enhancer was thawed and mixed well by pipetting up and down. Two µL of Lumio™ In-Gel Detection Enhancer were added to each sample. The samples were mixed well by pipetting up and down and

incubated at room temperature for 5 min. For each sample, 5-20 μL was loaded on protein polyacrylamide gel and electrophoresis was performed. The gel was placed on a UV 302 nm transilluminator (FOTO/Prep® UV Transilluminator; Fotodyne®, Incorporated), and the ethidium bromide filter was selected on the camera to take the fluorescence image.

Bacteria were grown in large scale, and soluble proteins were extracted.

Large-scale bacteria were grown and induced with IPTG for 4 h, and then soluble proteins were extracted. One L of LB containing ampicillin was inoculated with 50 mL of the fresh overnight-cultured DE3 transformants. It was grown for two h at 37°C with shaking until OD600 was about 0.5-0.8 (mid-log). IPTG was added to a final concentration of 1 mM. The culture was incubated at 37°C with shaking for 4 h. The cell culture was spun in 50 mL conical tubes, and the bacterial cells were collected by centrifuging at 5,000 × g for 10 min. Medium was carefully removed from the cell pellets, and the cell pellets were frozen at –80°C. Ten mL of the cell lysis buffer, CelLytic™ B (Sigma-Aldrich, Inc.), was added for each pellet. It was mixed well, and the cells were completely re-suspended. The extraction suspension was incubated with shaking at room temperature for 10-15 min. The extract was centrifuged at 16,000 × g for 10 min to pellet the insoluble material. The supernatant containing the soluble protein fraction was transferred to new tubes.

Poly His-tagged SelM was purified with Ni-NTA columns.

Soluble fractions of bacteria proteins were purified with Ni-NTA (Qiagen, Inc.) columns. The cleared lysate was mixed with a 50% Ni-NTA slurry at a ratio of 1:4. It was incubated at 4°C with shaking for 1 h. The lysate Ni-NTA mixture was loaded and passed through a column. The first column flow-through was collected. The mixture was washed twice with 10 mM of imidazole wash buffer at the ratio of 4:5. Wash fractions were collected for protein polyacrylamide gel analysis. The protein mixture was eluted at least 4 times with 250 mM of imidazole elution buffer at the ratio of 1:10. The elution was collected in separate tubes for analysis and protein concentration determination.

Protein buffer solutions were exchanged with TEV buffer using dialysis.

Dialysis was used to change protein-buffering environments to TEV working conditions. TEV protease buffer (50 mM Tris-HCl, 0.5 mM EDTA, pH 8.0) was prepared in large volume as dialysate. The dialysis cassette (Thermo Scientific Slide-A-Lyzer® Dialysis Cassettes; Pierce, Thermo Fisher Scientific, Inc.) was pre-hybridized for 30 s in deionized water. The sample was injected into the cassette with a needle, the air was removed, and dialysis was initiated by placing the cassette in dialysate (300 times of the volume of the sample volume) for 2 h at room temperature. The dialysis buffer was changed once, and dialysis continued at 4°C overnight. The sample was transfer carefully with a needle into new tubes.

SelM N-terminal tags were cut off using TEV protease.

TEV protease (AcTEV™ Protease, Invitrogen™ Corporation) was used to cleave the N-terminal tag from SelM. Reactions were carried out in buffer containing protein samples with 1% 0.1 M DTT and 0.65-2.6% AcTEV™ Protease (5-40 units). Two controls and five samples with increasing concentrations of TEV protease (2-40 U of TEV) were included in the experiment. The first control was an elution fraction after 6X His-tagged protein purification. The second control was a cutting product without dialysis. These reactions were incubated at room temperature overnight. Small aliquots were removed for analysis, and the remaining samples were stored at –20°C.

In-gel SelM was extracted and purified by electroelution.

After cleavage of N-terminal tags, the remaining SelM protein products were extracted and purified from protein gels by electroelution. Commercially available dialyzers (D-Tube™ Dialyzer; Novagen®; EMD Chemicals, Inc.) were hydrated by adding 3 mL of deionized water and incubating upright for 5 min, and then the water was removed. The gel slices were transferred into the dialyzer chamber along with 2 mL of protein running buffer. The dialyzer was placed in the traditional nucleic acid agarose electrophoresis tank containing the protein-running buffer. The two membranes of dialyzer chamber were placed perpendicular to the electric field. The electric current was applied at 100 V for 2 to 3 h until the protein eluted from the gel slice. The polarity of electric current was reversed for 2 min to remove

the proteins that stuck in the membrane pores. The eluted protein solution was transferred to clean tubes with care to avoid the transferring of gel slices.

Polyacrylamide gels were detected using silver staining.

An established Silver staining technique (SilverXpress® Silver Staining Kit, Invitrogen™ Corporation) was used to detect protein samples in polyacrylamide gels. The gel was fixed with 200 mL of fixing solution (50% methanol, 10% acetic acid) for 10 min with shaking. The gel was incubated in two changes of sensitizing solution for 10 min, rinsed twice with ultra-pure water for 5 min, and then incubated with 200 mL of staining solution for 15 min with shaking in a fuming hood. Each gel was rinsed twice with ultra-pure water for 5 min. The gel was incubated with 200 mL of developing solution for 5 min until the desired darkness of banks appeared. Stopping solution was added directly to the developing solution, and it was incubated for 10 min with shaking until the desired staining intensity was reached, followed by two rinses with ultra-pure water. Images were captured using the KODAK Gel Logic 200 and KODAK Molecular Imaging Software (Carestream Health, Inc.).

SelM was concentrated with protein concentrator.

The protein extracted from protein gels using electroelution was concentrated with protein concentrators (5 kDa Spin Concentrator, Agilent Technologies). The concentrator was hydrated by running through 1 mL of PBS at 3500 × g (Damon

IEC HN-SII Centrifuge; GMI, Inc.). Electroeluted protein solutions were spun at 3500 × g until the desired volume was achieved. The sample was transferred from the concentrator into new microcentrifuge tubes. Small amounts were removed for purity and concentration analysis, and the remaining solution was stored –20°C.

Bradford assay was used to test the protein concentration.

Bradford assay was used to determine the protein concentration. Thirty mL of 1 × Bradford Assay Reagent (Bio-Rad Laboratories, Inc.) was added to BSA standard solutions or protein samples in triplicate to a 96-well plate. Samples were read on a Spectra Max 340 (Molecular Devices) at 595 nm between 5 min and 1 h after the addition of the Bradford Assay Reagent.

RESULTS

Molecular cloning produced complete SelM expression vector.

Gateway® Technology was used for the molecular cloning of vector construction. Based on the bacteriophage λ site-specific recombination system, vectors of the Gateway® system share the same recombinant sites for transferring DNA segment among different vectors. Thus, this system was chosen due to its versatility (Figure 5).

Recombination reactions occur at the specific attachment (*att*) sites of the interacting DNA molecules. The reaction is conservative and requires no DNA synthesis. After the recombination reactions complete, the inserted gene remains the same with no net gain or loss of nucleotides. The initial *att* sites are *att*B on the *E. coli* chromosome and *att*P on the λ chromosome. Upon the integration, *att*L and *att*R sites are generated. The strand exchange of the recombination reaction occurs between homologous 15 bp core regions, which are common on all of the *att* sites. The surrounding sequences, however, are needed because they contain the binding sites for the recombination proteins [31].

The attachment sites were *att*B on SPORT6 vectors, *att*P on pDONR-211 vectors, and *att*R on the pET160 vectors. After recombination reactions between *att*B on SPORT6 vectors and *att*P on pDONR-211 vectors, hybrid sequences were generated. They were *att*L sites of the derived pDONR-211 vectors.

Generally, the recombination occurs between plasmid DNA of any topology: supercoiled, linear, or relaxed. As long as the *att* sites are accessible, recombination may also occur between other types of DNA sequences such as PCR products, cDNA clones, or restriction fragments. The efficiency of these different reactions may vary.

Supercoiled plasmid DNA was used for both of the recombination reactions. The *att*B sites on SPORT6 vectors reacted with *att*P on pDONR-211 vectors (BP reaction, Figure 6), and *att*L on pDONR-211 vectors reacted with *att*R on pET160 vectors (LR reaction, Figure 7). The lower efficiency of supercoiled plasmid was adequate for the successful transformation. An enzyme mix consisting of λ and *E. coli*-encoded recombination proteins was used for the recombination reactions of SelM constructs. λ bacteriophages used this same enzyme combination to integrate into the *E. coli* chromosome. Because the recombination was conservative, the SelM construct was built in an entry clone pDONR-211 and then moved into an expression clone pET160. The insertion was kept the same with no ORF shifting.

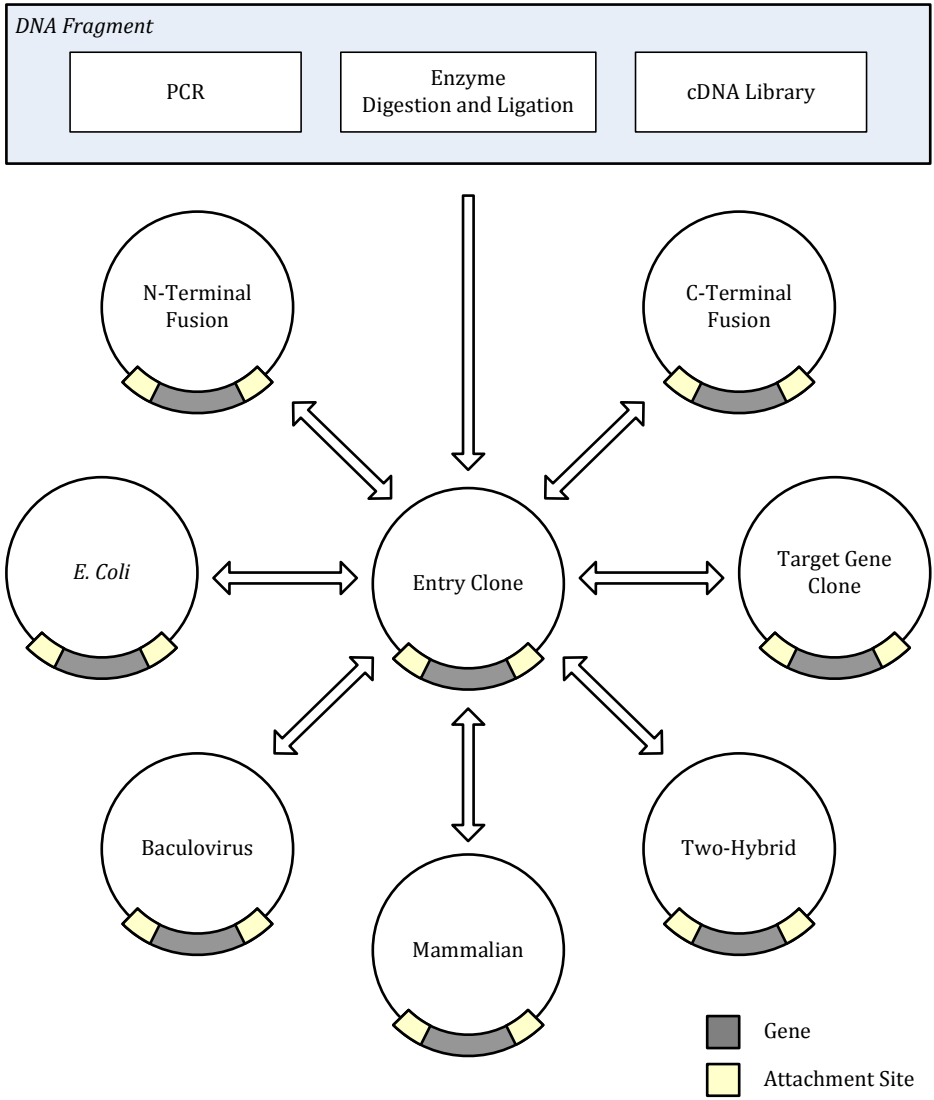

Figure 5. Gateway® system. Vectors of the Gateway® system share the same recombinant sites for transferring DNA segments among different vectors. This system provides a rapid and efficient route to multiple expression and functional analysis options.

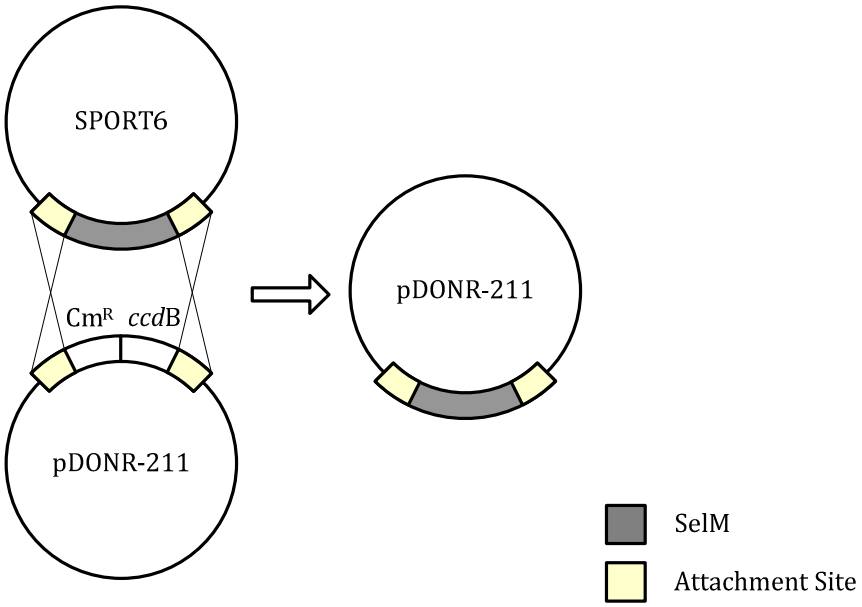

Figure 6. Schematic of BP reaction showing the recombination reaction between *att*B and *att*P sites. The original human SelM cDNA in pCMV•SPORT6 plasmid was purchased from the Invitrogen™ Corporation. After the sequence was confirmed, BP reaction was performed to transfer the complete insertion into the Gateway® entry clone, pDONR-211 plasmid. Recombination reactions occurred between the *att*B sites on pCMV•SPORT6 vectors and *att*P sites on pDONR-211 vectors.

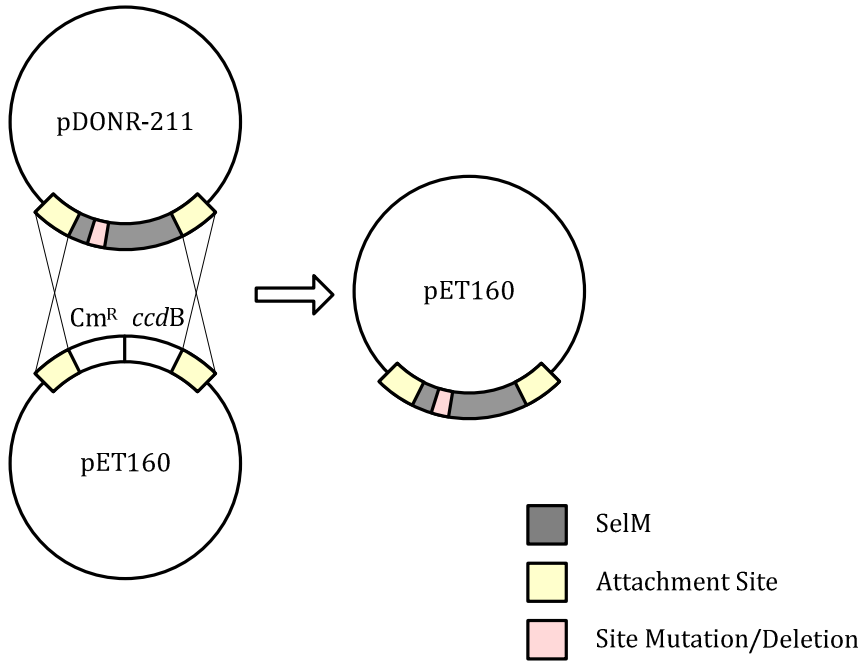

Figure 7. Schematic of LR reaction showing the recombination reaction between *att*L and *att*R sites. Desired mutations were performed on the SelM cDNA in the entry clone pDONR-211, including a mutation from Sec to Cys and the 5' UTR deletion. The 5' UTR deletion allowed for a correct ORF and insertion of the minimum number of extra nucleotides upstream of the SelM start codon. The construct was then transferred into the expression vector system by another round of recombination reactions at the *att*L sites of the entry clone, pDONR-211, and the *att*R sites of the expression clone, pET160.

Mutagenesis for construct building was applied by inverse PCR.

Mutagenesis by inverse PCR was developed from traditional site-directed mutagenesis. Site-directed mutagenesis has been a great revolution for the study of protein structures and functions. It enables well-controlled and systematic production of mutant proteins. The first revolutionary application of site-directed mutagenesis involved the use of a mutated oligonucleotide primer to synthesize a target single-stranded DNA template. The second application used single-stranded, uracil-containing DNA molecules isolated from *ung*-dut-*E. coli* strains. The parent strands were removed after introduction into wild-type *E. coli*. This application increased the mutation efficiency. A more recent application of site-directed mutagenesis used a two-PCR reaction system. Two sets of mutated oligonucleotide primers were used for PCR. The product mixture was transformed into *E. coli* after denaturization and renaturation. Compared to the previous techniques, mutagenesis by inverse PCR (Figure 8) bypasses the series of sub-cloning steps involving single-stranded M13 phage clones. Inverse PCR mutagenesis has other advantages including a convenient closing of the vector after the PCR reaction by simply adding T4 ligase into the system. The ligase links the 5' phosphorylated ends and recovers the circle plasmids [14].

Pfu DNA polymerase was used for inverse PCR. The main difference between *Pfu* and other enzymes is *Pfu*'s superior thermostability and "proofreading" properties when compared to other thermostable polymerases. Unlike Taq DNA

polymerase, *Pfu* DNA polymerase possesses 3' to 5' exonuclease proofreading activity, which allows it to work its way along the DNA from the 5' end to the 3' end and correct nucleotide-misincorporation errors. This allows for *Pfu* DNA polymerase-generated PCR fragments to contain fewer errors than Taq-generated PCR fragments. The 3' to 5' exonuclease also digest single-stranded DNA at 3', making them shorter and less specific. To avoid this, the enzyme was added last, and the PCR machine was pre-heated to 94°C. The results showed that the inverse PCR reactions gave rise to self-annealing of primers. With a small mismatch in the middle, the others were annealed to the target sequence of plasmid containing SelM. Then, the proof-reading *Pfu* DNA was used to amplify the entire plasmid and thus incorporate the primer as the new (mutant) sequence. Only 15 PCR cycles were performed on a relatively large amount of plasmid template (50 ng) to minimize the chance of expanding PCR sequence errors, followed by ligation of the linear PCR plasmids into circular plasmids using T4 ligase. Finally, the plasmids were digested with *Dpn*I, which is a 4-cutter that cuts only methylated DNA, in order to remove the template plasmid DNA.

Mutagenesis PCR primers were designed adjacent at their 5' ends but non-overlapping (site mutation). They contained the desired mutation at the 5' end of one primer, and they were both phosphorylated at the 5' ends (Figure 9). Based on the notion that secondary structures and other factors may affect the performance of primers, several pairs of primers were tested simultaneously. These were

designed to vary from each other by complementing slightly different template sequences, while still retaining the same mutation site.

The goal of first round of PCR was to generate Cys-mutated SelM. Sec is encoded by UGA, which functions as a stop codon under normal circumstances. The rationale for this was that Cys incorporation would be more efficient during translation in bacteria, and presumably, this would not affect antibody production in inoculated rabbits. To ensure the continuance of translation, Sec was mutated to Cys by one nucleotide modification of the gene sequence from TGA to TGC. The products were analyzed by agarose gel electrophoresis, which included a negative control with no primer added and included samples produced from different primer and different template concentrations (Figure 10). Two out of three primer sets generated right-side PCR products. Both controls showed negative results as expected.

The goal of the second round of inverse PCR was to introduce a deletion in SelM. The deletion started at the end of the *att* site, which was located on the plasmid before the gene insertion. The deletion extended to the start codon of the protein encoding sequence of the insertion. Other than the deletion, the primers were also designed to generate in-frame protein product. Because the tag sequence of the expression system was with an ORF plus one extra nucleotide, insertion was kept in its own ORF plus two extra nucleotides at the 5' end. These PCR products were analyzed by agarose gel electrophoresis (Figure 11). Two out of three pairs of

primers produced the expected amplification product. A difference in efficiency was observed with the other two sets of primers.

High efficiencies for both steps of mutagenesis were observed. After transformation, approximately half of the single clones were confirmed with a desired mutation site (data not shown). One final entry clone construct (Figure 12) was picked to perform an LR reaction with the expression clone: pET160 empty vector.

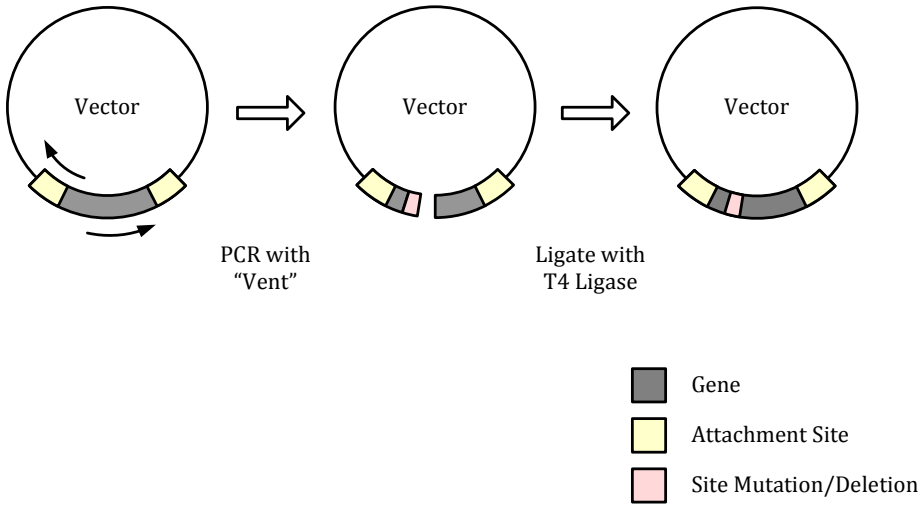

Figure 8. Flowchart for mutagenesis by inverse PCR. First, primers were designed based on the sequence information of both the cDNA to be inserted and the plasmid. The primers needed to contain a desirable mutation site or be able to introduce a deletion into the product plasmids by PCR. Primers were made with 5' phosphorylated ends, which allowed subsequent circularization of the plasmid using T4 ligase.

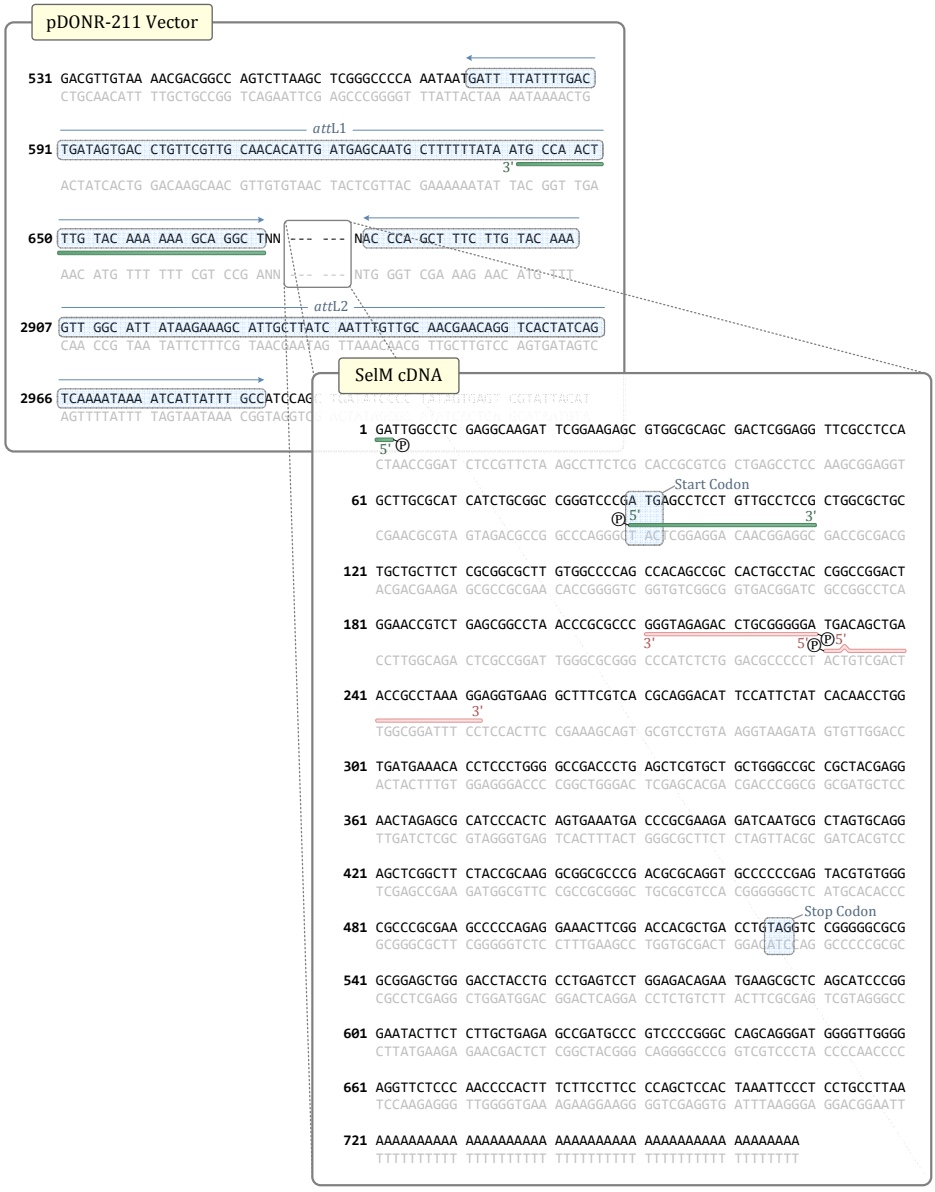

Figure 9. Primer design for mutagenesis by inverse PCR. Two rounds of PCR reaction were applied to generate Cys-mutated SelM with a 5' UTR deletion for more efficient expression in bacteria. Two primers were designed with either TGA to TGC mutation sites or 5' UTR deletion.

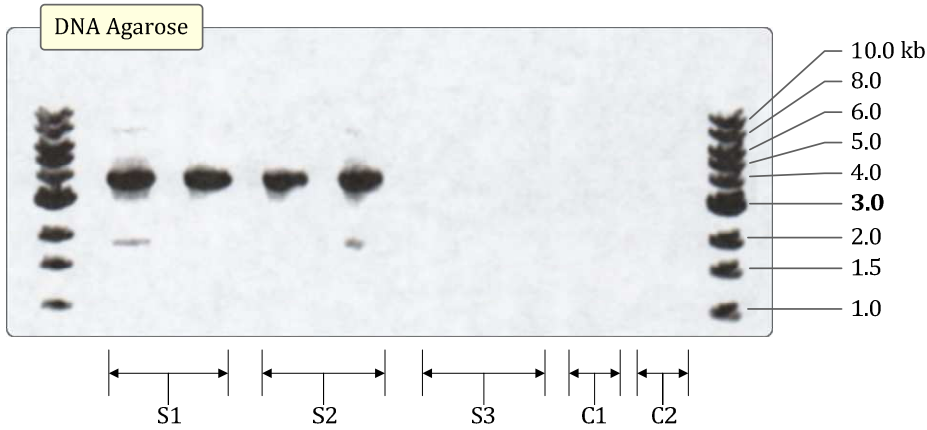

Figure 10. Analysis of inverse PCR products using agarose gel electrophoresis. As a first step, the Sec- to Cys-codon mutation was carried out with three different primer sets, two different template concentrations, and a negative control with no primers. The PCR products from the three primer sets are shown in the figure as S1, S2, and S3. The products from the lower template at 10 ng/reaction (C1) and negative control (C2). PCR reaction conditions followed the general rules of inverse PCR: an initial denaturation for 4 min at 94°C, followed by 15 cycles of 1-min denaturation at 94°C, 1-min annealing at 50°C, and 12-min extension at 72°C. Amplification was completed by a final extension step at 72°C for 10 min. PCR products were detected for two of the designed primer sets, and they were approximately the size of the pDONR-211 plasmid. Both C1 and C2 showed negative results as expected.

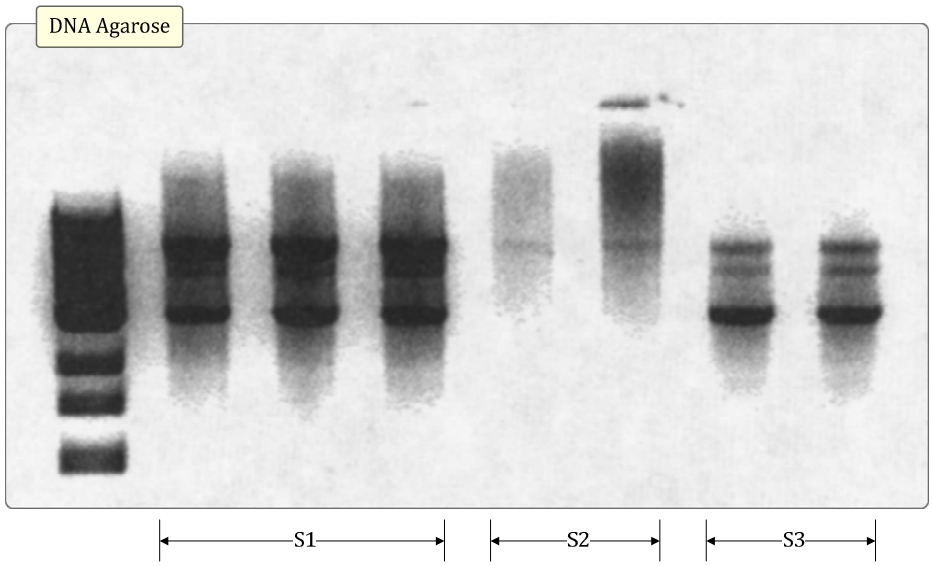

Figure 11. Analysis of inverse PCR generating products with 5' UTR deletion using agarose gel electrophoresis. The PCR reactions were carried out with three different primer sets for 5' UTR deletion, which are displayed in the figure as S1, S2, and S3. The reaction conditions were the same as the first round of mutagenesis by inverse PCR described above. The 80 bp deletion did not significantly alter the extension efficiency during the same extension time. PCR products of S1 and S3 were detected, with lower efficiently for the S3 reaction compared to S1.

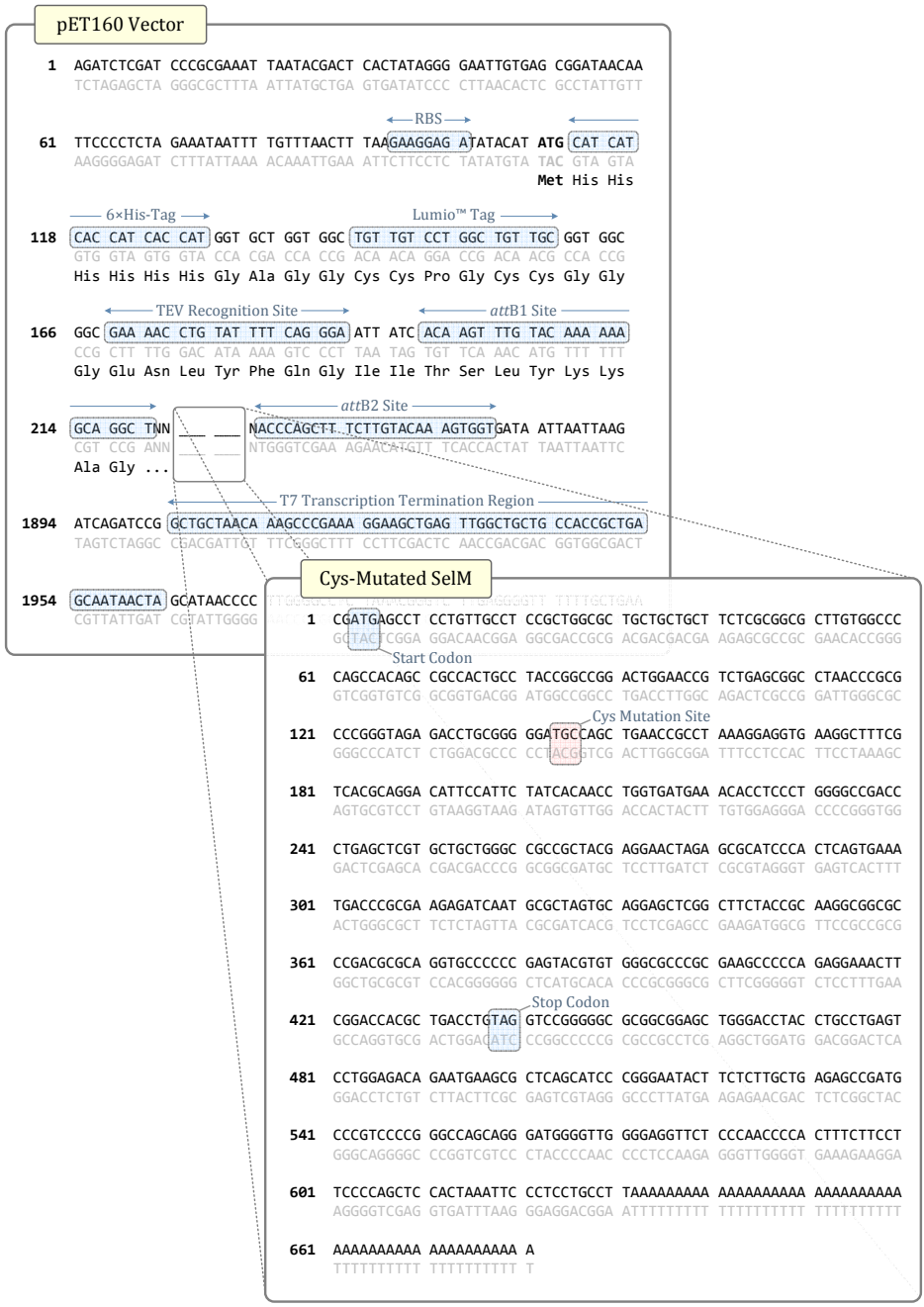

Figure 12. Cloning strategy for inserting SelM in the pET160 expression system.

SelM was successfully expressed in the pET160 system using the expression host, DE3.

The plasmid, pET160, was chosen as an expression vector for several reasons. First, the presence of a short tag upstream of the SelM insertion site contained three desirable features: the Lumio™ site, the tobacco etch viral (TEV) site, and the 6X His-tagged. The Lumio™ site is a six amino acid sequence that is specifically recognized by a biarsenical labeling reagent, which allows in-gel detection via fluorescence. The TEV site is a seven amino acid sequence that shares a general form of E-X_{aa}-X_{aa}-Y-X_{aa}-Q-(G/S). It allows TEV protease to recognize and cleave between Q and G or Q and S. The 6X His-tagged component provides an affinity tag that facilitates binding to Ni-NTA and is used to purify tagged protein with Ni-NTA columns. Generally, these tags do not affect secretion, compartmentalization, or folding of the fusion proteins within the cell. pET160 also contains a T7 promoter, a lac operon, a ribosome binding site, and an ATG translation starting site. Cm^R and *ccd*B are located between *att* sites, and their expression leads to host cell death. Therefore, they act as a selective marker for isolating bacteria containing constructs with genes of interest against empty vector for the pET160 expression system (Figure 13).

Cloning Host

The *E. coli* TOP10 strand was used as the host for cloning. After the construct was built in the entry clone pDONR-211 vector, it was transferred to the

expression clone pET160 (Figure 13). The expression plasmid was transformed into One Shot® TOP10 Competent Cells for the final analysis. The analysis verified successful mutations and correct ORF introduced by oligonucleotides, and it ensured the LR reaction did not introduce any ORF shifting.

The choice of cloning host was based on the features of both pET system and the *E. coli* TOP10 strand. A powerful feature of the pET system is the ability to clone target genes under conditions of extremely low transcriptional activity. Gene expressions of the pET system are under the control of T7 promoters. In the absence of a source of T7 RNA polymerase, the background expression would be minimal because the host RNA polymerases do not initiate from T7 promoters. *E. coli* TOP10 does not generate T7 RNA polymerase, which is the essential factor for the transcription of the pET expression system. For maintaining or analyzing the plasmid construct, it is a poor choice to use an expression system containing the base expression of a T7 RNA polymerase gene. In that case, it can be difficult to grow bacteria, and plasmids tend to become unstable even at low levels of foreign gene expression.

TOP10 *E. coli* cells were convenient hosts for the initial cloning of target SelM and for maintaining the plasmids. The results demonstrated high transformation efficiency and a good yield of plasmid extraction.

Expression Host

Lysogens of bacteriophage BL21(DE3) were used as the expression host. For protein production, recombinant pET plasmid was transformed into the host *E. coli* strain, DE3. A small-scale pilot expression study was carried out in the expression host DE3 with controls and time points. The approach was to determine if successful expression of the protein could be achieved before moving to a large-scale expression system in BL21(DE3).

The choice of cloning host was also based on the features of both the pET system and the BL21(DE3). DE3 contains a chromosomal copy of the gene for T7 RNA polymerase, which facilitates the foreign gene expression in the pET expression systems. DE3 is a λ derivative that has the immunity region for phage 21 and carries a DNA fragment containing the *lac*I gene, the *lac*UV5 promoter, and the gene for T7 RNA polymerase [42]. This fragment is inserted into the *int* gene, preventing DE3 from integrating into or excising from the chromosome without a helper phage. Once a DE3 lysogen is formed, the only promoter known to direct the transcription of the T7 RNA polymerase gene is the *lac*UV5 promoter, which is inducible by isopropyl-b-D-thiogalactopyranoside (IPTG). Addition of IPTG to a growing culture of the lysogen induces T7 RNA polymerase, which in turn transcribes the target DNA in the plasmid.

Pilot expression was carried out that included a control group that excluded IPTG induction. Six time points were taken for both experimental group and control group. Whole cell lysate was collected and run on a polyacrylamide gel.

The biarsenical labeling reagent mixed with the sample showed the fluorescent bands of appropriate size. An increase of SelM production was observed in the experimental group, which reached a plateau after 3 h of IPTG induction. No SelM was produced in the control group without IPTG induction (Figure 14).

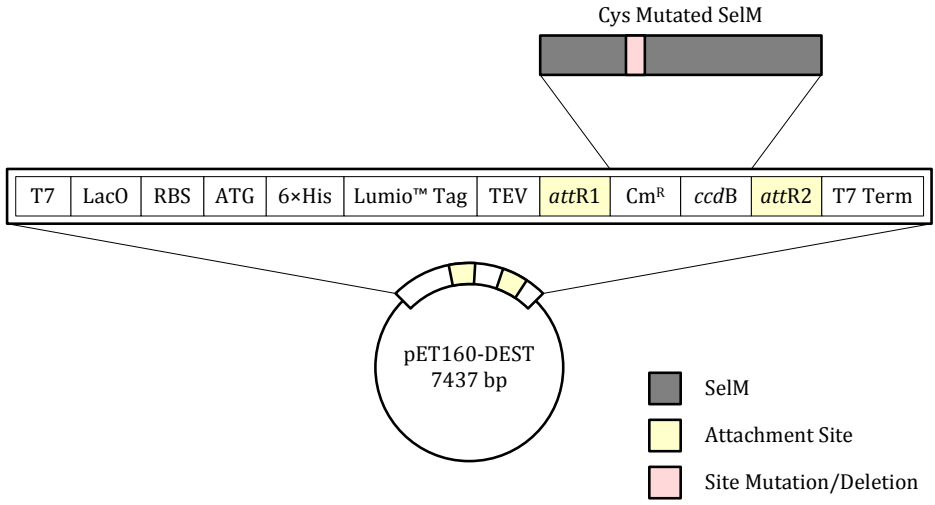

Figure 13. Components of the fusion protein tag included with the pET160 system. The tag consisted of three main components: the Lumio™ site, the tobacco etch viral (TEV) site, and the 6X His-tagged site. The Lumio™ site is a six amino acid sequence that is specifically recognized by a biarsenical labeling reagent, which allows in-gel visualization using fluorescence. The TEV site is a seven amino acid sequence that shares a general form of E-X_{aa}-X_{aa}-Y-X_{aa}-Q-(G/S). It allows TEV protease to recognize and cleave between Q and G or Q and S. The 6X His-tagged component is an affinity tag that facilitates binding to Ni-NTA and is used to purify tagged protein with Ni-NTA columns. Generally, these tags do not affect secretion, compartmentalization, or folding of the fusion protein within the cell.

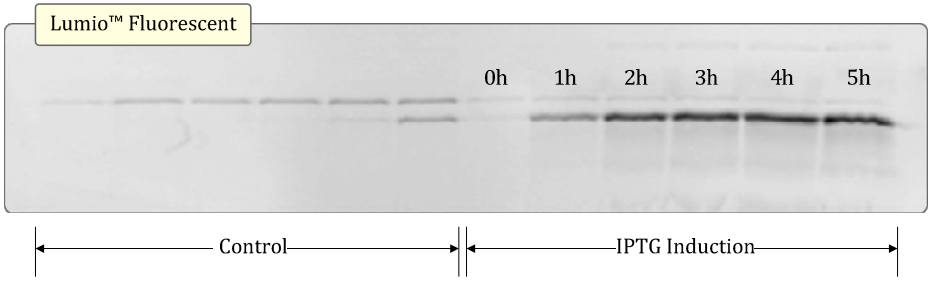

Figure 14. Pilot experiment involving expression of SelM. Fresh transformants were cultured overnight and used to express the SelM. After 2 h of recovering, 1 mM of IPTG was added to the experimental group. Bacteria culture was collected, and proteins were extracted at every hour time point for 6 h in both groups. Fifteen μL of whole cell lysate was mixed with Lumio™ Detection Reagent and run on 10-14.5% polyacrylamide gradient gel. The results were detected with a UV 302 nm transilluminator. SelM protein was expressed at high levels in the IPTG-induced group but not in the control group. SelM expression increased with increasing culture time after IPTG induction, with a plateau point reached approximately 3 h after the IPTG incubation.

Lumio™ reagents were used for in-gel protein detection.

In addition to traditional Coomassie blue and silver staining, Lumio™ fluorescent detection was also used for most of the in-gel detections of SelM. The tetracysteine Lumio™ tag is a six amino acid tag Cys-Cys-Pro-Gly-Cys-Cys. A small-sized tag (585 Da) is less likely to interfere with the structure or biological activity of the protein of interest. When it was fused to the SelM gene, the tag allowed the expressed fusion protein to be specifically recognized by a biarsenical labeling reagent. This procedure was a highly specific and sensitive detection of the tagged SelM (Figure 15).

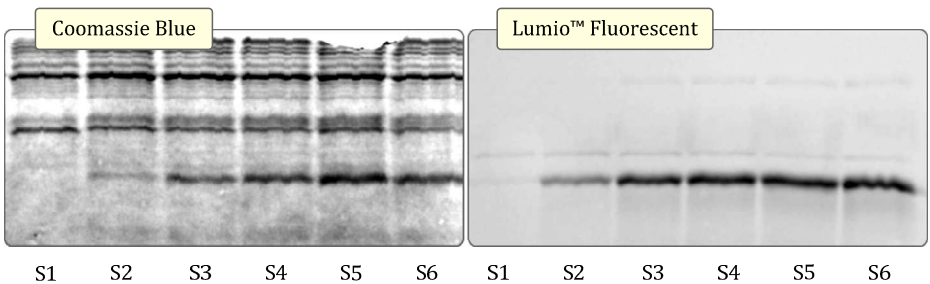

Figure 15. Comparison of Coomassie blue staining and Lumio™ fluorescent detection. The same polyacrylamide gel was stained with a Coomassie blue staining reagent (left) and displayed under a UV 302 nm transilluminator (right). The protein bands were shown in fluorescence by the biarsenical labeling reagent added to the protein samples. The bands for SelM were visible by both approaches. The Coomassie blue stain theoretically displayed all the proteins presented in the samples. The protein of interest could be spotted based on its size. In order to confirm the existence and expression situation of SelM, however, Lumio™ fluorescent detection proved to be a more powerful, sensitive, and specific technique.

Poly His-tagged SelM was roughly purified using Ni-NTA columns.

Immobilized metal affinity chromatography (IMAC) was used to purify 6X His-tagged SelM. Purification of His-tagged proteins by IMAC is a widely used method for easily producing large quantities of first-step purified protein in a one-step process. The two most commonly used metal ions in IMAC resins are Co^{2+} and Ni^{2+}. Ni-NTA resin was used for native SelM protein purification. Ni-NTA resin can be used to purify His-tagged proteins in both native and denaturing conditions. The supernatant of the whole cell lysate, which includes soluble proteins, was incubated with Ni-NTA slurry for an hour. The mixture was passed through the column, followed by two steps of washing. Then, the binding proteins were eluted off the Ni-NTA beads using a high concentration of imidazole solution. The majority of the soluble proteins were left in the first flow-through solution. Binding proteins were collected in the elution fractions (Figure 16).

Aliquots of each step were collected for analysis. Thirty μL of each sample were removed and run on 10-14.5% polyacrylamide gradient gels. The gel was stained with Coomassie blue (Figure 17). Most proteins remained in the first flow-through but not the human SelM, which was exclusively found in the elution fractions. The same samples were analyzed with Lumio™ fluorescent detection reagent (Figure 18). Results showed that the major 17-kDa bands on the Coomassie blue-stained gel were also present in the Lumio™-detected gels. The experiment was repeated with longer incubation with Ni-NTA beads. Elution fractions of poly His-tagged protein purification were run on a protein gel and

detected with Lumio™ Fluorescent Reagent (Figure 19). Like previous experiments, highly expressed SelM proteins were detected. The viscosity of slurry mixture greatly increased, however, so the column purification became a time-consuming procedure. Additional force was needed to pump it through.

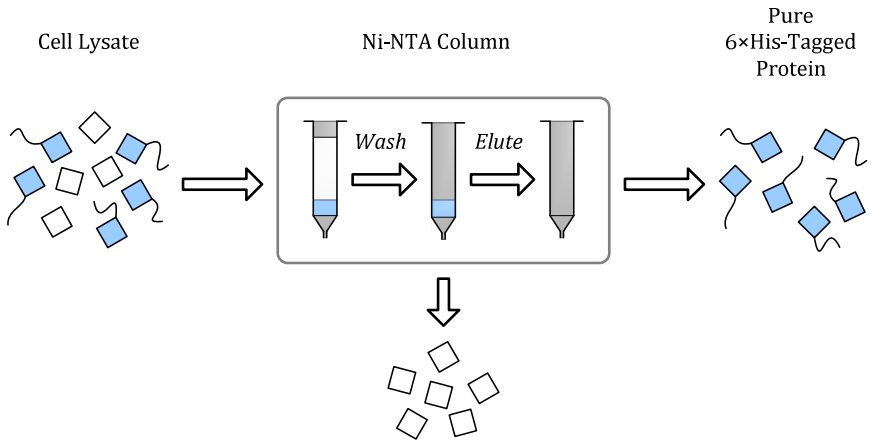

Figure 16. Poly His-tagged protein purification strategy. The supernatant of the whole cell lysate is incubated with Ni-NTA slurry for an hour or overnight. The mixture goes through the column and is then washed twice. The bound proteins are eluted off the beads with a high concentration of imidazole. The majority of the soluble proteins from the protein extraction should be found in the first flow-through solution. Poly His-tagged proteins may then be collected in the elution fractions.

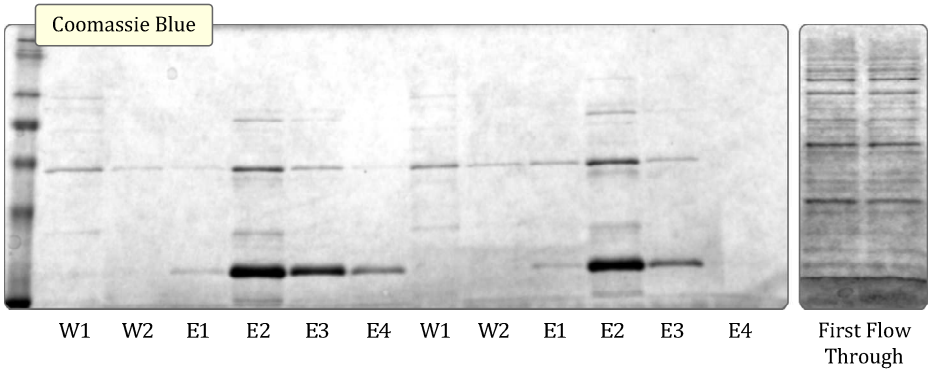

Figure 17. Results from poly His-tagged protein purification as detected by Coomassie blue staining reagent. Samples included the first flow-though, wash buffers, and elution fractions. Aliquots of each sample were run on a piece of 10-14.5% polyacrylamide gradient gel, which was then stained with Coomassie blue. The majority of soluble proteins were left in the solution and detected in the first flow-through. Seventeen-kDa SelM, however, was found exclusively in elution fractions. Other than SelM, several contamination proteins were also detected in the elution.

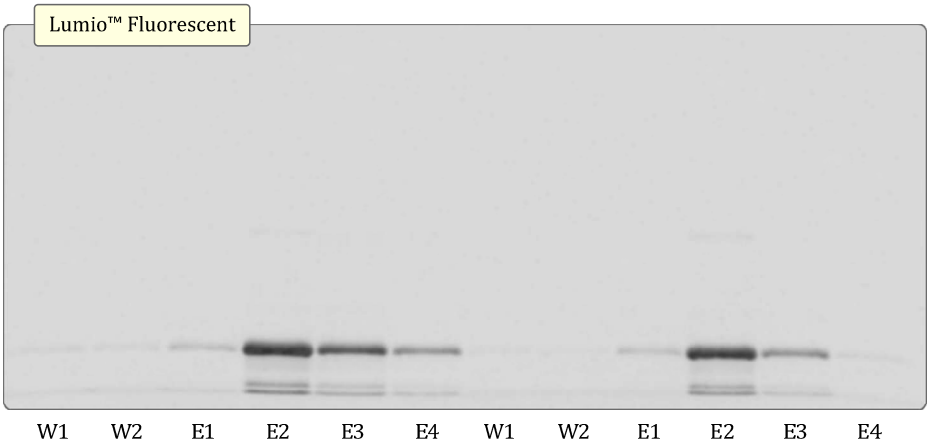

Figure 18. Results from poly His-tagged protein purification as detected by Lumio™ fluorescent reagent. Samples included the first flow-though, wash buffers, and elution fractions, which were detected with biarsenical labeling reagent added in the samples prior to being loaded into the gel. The image was captured using UV 302 nm transillumination. Seventeen-kDa SelM was found exclusively in the elution fractions. The biarsenical labeling reagent detected Lumio™-tagged SelM with high specificity.

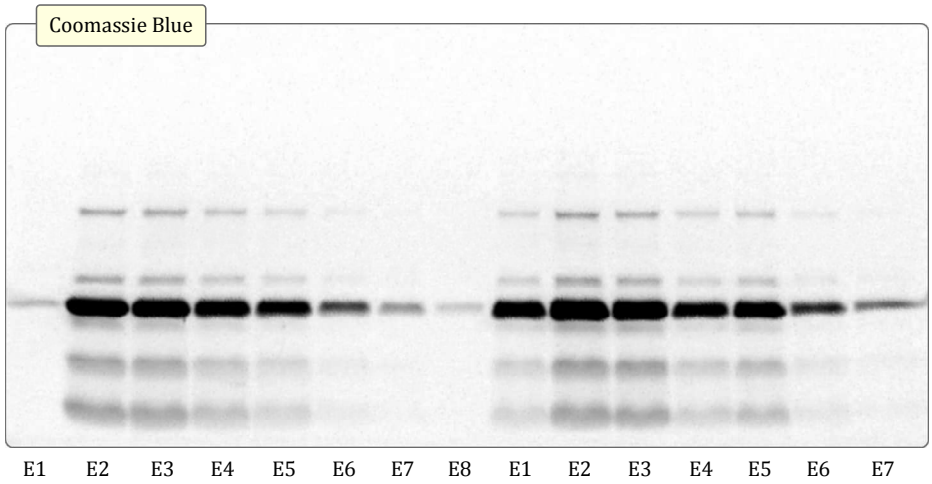

Figure 19. Poly His-tagged protein purification with long Ni-NTA incubation. Lumio™-tagged SelM was detected with the biarsenical-labeling reagent. Elution fractions were collected after the poly His-tagged protein purification sample after an overnight Ni-NTA slurry incubation. High yield of SelM was still observed. Fifteen μL of each sample was mixed with Lumio™ detection reagent and run on a piece of 15% polyacrylamide gel. It was detected using UV 302 nm transilluminator.

N-terminal tags were partially removed by the TEV protease.

AcTEV™ Protease was used to cleave the N-terminal poly peptide tag of SelM. TEV protease is the common name for the 27-kDa catalytic domain of the Nuclear Inclusion A (NIa) protein encoded by the tobacco etch virus (TEV). TEV protease is used for cleaving fusion proteins because its sequence specificity is far more stringent than that of factor Xa, thrombin, or enterokinase. It is also relatively easy to overproduce and purify large quantities of the enzyme. TEV protease recognizes a linear epitope of the general form E-X_{aa}-X_{aa}-Y-X_{aa}-Q-(G/S), with cleavage occurring between Q and G or Q and S. The most commonly used sequence is Glu-Asn-Leu-Tyr-Phe-Gln-Gly, which is the amino acid sequence introduced in the N-terminal introduced by the pET160 expression system. The most efficient substrate is thought to be Glu-Asn-Leu-Tyr-Phe-Gln-Ser.

Results showed that TEV protease cleaved the target site of the tagged SelM and partially removed the fusion protein tag from SelM (Figure 20). The reaction was found to work at a strict buffering condition and required high protease working concentration. Control 1 was the protein sample without dialysis. Control 2 was the protein sample without dialysis and incubated overnight with 2 × TEV. Sample 1 was the protein sample after buffer dialysis and incubated with 2 × TEV. Sample 2 was similar to sample 1 but incubated overnight with 4 × TEV. Sample 3 was similar to sample 1 but incubated overnight with 1 × TEV. Sample 4 was similar to sample 1 but incubated overnight with 0.5 × TEV. Sample 5 was

similar to sample 1 but incubated overnight with 0.2 × TEV. Thirty μL of each sample were taken and run on 10-14.5% polyacrylamide gradient gels. Tagged SelM proteins were cleaved TEV protease. The protease, however, required a stringent buffering environment and high working concentration. The cleavage efficiency was low. This might have been due to the fact that some fusion proteins are intrinsically poor substrates for TEV protease. For example, steric occlusion might form when the protease cleavage site is too close to the ordered structure in the target protein or when the fusion protein exists in the form of soluble aggregates. This problem may be partially overcome by using a large amount of TEV protease (Figure 20).

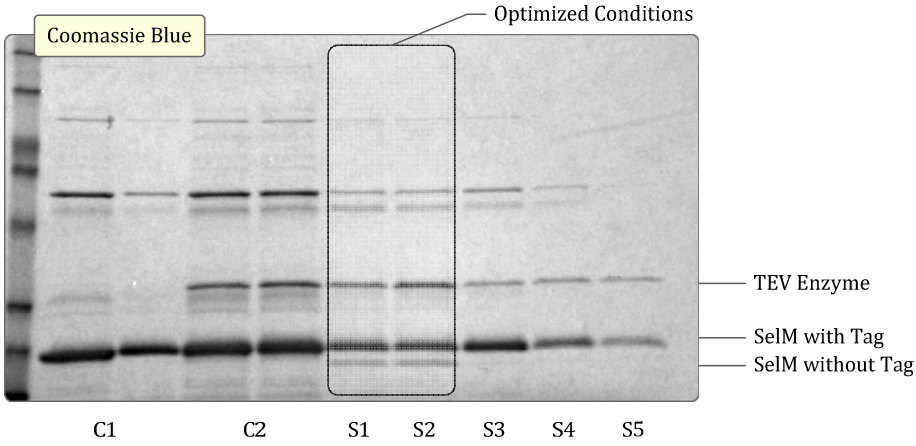

Figure 20. Successful, but inefficient, application of TEV protease cleavage. TEV protease cleaved the target site and removed the fusion protein tag at a condition of strict buffering environment and high protease working concentration. Control 1 represented the protein sample without dialysis. Control 2 represented the protein sample without dialysis and incubated overnight with 2 × TEV. Sample 1 represented the protein sample after buffer dialysis and incubated with 2 × TEV. Sample 2 was similar to sample 1 but incubated overnight with 4 × TEV. Sample 3 was similar to sample 1 but incubated overnight with 1 × TEV. Sample 4 was similar to sample 1 but incubated overnight with 0.5 × TEV. Sample 5 was similar to sample 1 but incubated overnight with 0.2 × TEV. Thirty μL of each sample were taken and run on 10-14.5% polyacrylamide gradient gels, and the gel was detected with Coomassie blue.

Further purification of SelM was not achieved by an additional round of poly His-tagged protein purification.

After overnight incubation with TEV protease, the resulting product was a mixture containing SelM, SelM with tag, TEV protease, two major contamination proteins, and possibly other proteins existing in trace amounts. Except for SelM and TEV protease, all other components came from elution fractions of poly His-tagged protein purification. They were, therefore, Ni-NTA binding proteins. Because the AcTEV™ protease itself is a His-tagged enzyme, this enzyme should be removed using poly His-tagged column purification. Thus, only SelM should have passed through the Ni-NTA column as an unbound protein, allowing SelM purification by another round of poly His-tagged protein column purification. In other words, in theory all of the proteins other than SelM should have bound the beads. The results showed, however, that this was not the case. The mixture of TEV protease and sample was run on a Ni-NTA column, washed twice, and bound proteins eluted with imidazole. Polyacrylamide electrophoresis results included the first flow-through, wash buffers, and elution fractions (Figure 21). Untagged SelM did not separate as expected. Other than untagged SelM, tagged SelM and some other contamination proteins were all found in the first flow-through. TEV protease, however, successfully remained in the elution fractions, but both tagged and untagged SelM were also found in the elution. Although the majority of SelM was collected in the first flow-through, SelM and SelM with an N-terminal tag did

not separate from each other at any stage. Applying another round of Ni-NTA column purification therefore was not an ideal way to purify the SelM protein.

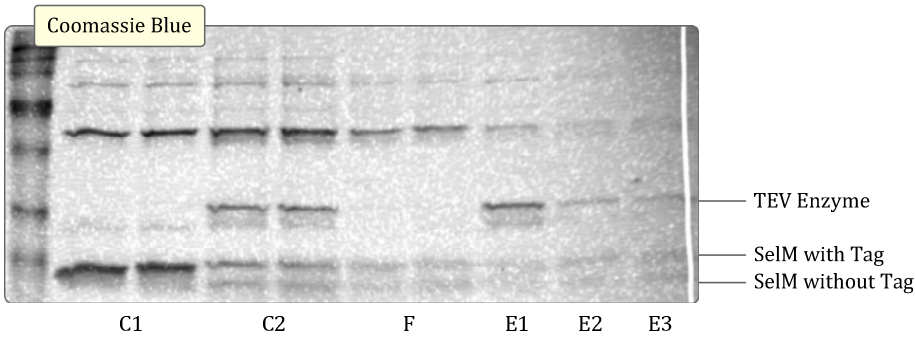

Figure 21. Analysis of purification of the untagged SelM from poly His-tagged proteins using polyacrylamide electrophoresis. Samples cleaved with TEV were incubated overnight with Ni-NTA, the mixture was washed, and then the bounded proteins were eluded with imidazole. The samples run on the gel included the first flow-through, wash buffers, and elution fractions. The gel was detected with Coomassie blue. Untagged SelM did not separate as expected. Other than untagged SelM, tagged SelM and some other contamination proteins were all found in the first flow-through. TEV protease, however, successfully remained in the elution fractions, but both tagged and untagged SelM was found with it.

SelM was purified by electroelution.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was used as the analytical tool to assess protein purification because of its excellent ability to resolve individual components of complex mixtures. SDS-PAGE could be used not only for evaluating purity but also as an active step in the purification process (Figure 22). Protein purification by gel electrophoresis has been used in various applications such as antigen preparation for antibody generation and isolation of proteins for identification by N-terminal sequencing and mass spectrometry.

The first step in purifying proteins from polyacrylamide gels was to visualize the electrophoresed SelM in the gel (Figure 23). For in-gel detection, the protein ladder was used to estimate the 17 kDa SelM, which was excised from the gel. The electroelution was performed as described in the Methods section. In brief, gel pieces containing SelM were placed in an electroelution chamber. The proteins were eluted from the gel matrix into a buffer solution using an electrical field. The buffer was in the chamber surrounded by dialysis membranes. The membrane was selected with an appropriate molecular weight cut-off, which is 6-8 kDa for SelM. Results showed that SelM was purified using this method (Figure 24). Control 1 was a protein sample before TEV cleavage. Control 2 was a protein sample after overnight TEV cleavage. Elution 1 to 4 were the elution fractions of electroelution. Because silver stain is considered very sensitive (it is able to detect as low as 0.8 ng

protein in protein gel), the clear background of the silver staining image meant the SelM protein was highly purified.

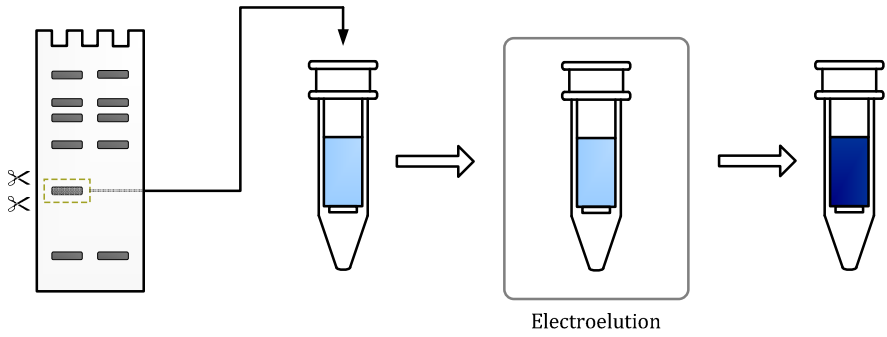

Figure 22. Flowchart for electroelution. Gel pieces containing protein were placed in an electroelution chamber. The proteins were eluted from the gel matrix into a buffer solution using an electrical field, and they were captured against a dialysis membrane. The membrane was selected with an appropriate molecular weight cut-off, which is 6-8 kDa for SelM. The buffer solution, which contained SelM, was collected in the dialyzer chamber.

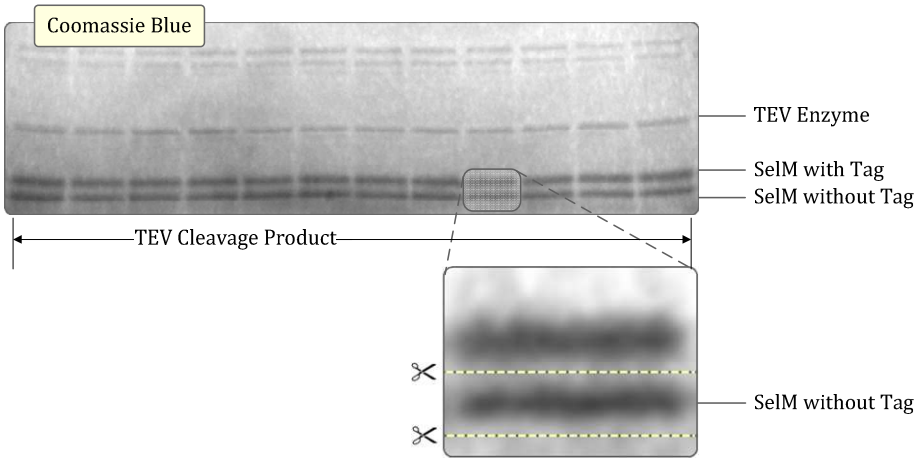

Figure 23. Analysis of electroeluted proteins as detected by Coomassie blue. After condition optimization for TEV cutting, samples were run on the gels. Gel stripes were used for electroelution. Approximately 40 μL of sample were taken for each well. Electrophoresis was applied on polyacrylamide gradient gels. This sample gel was stained with Coomassie blue.

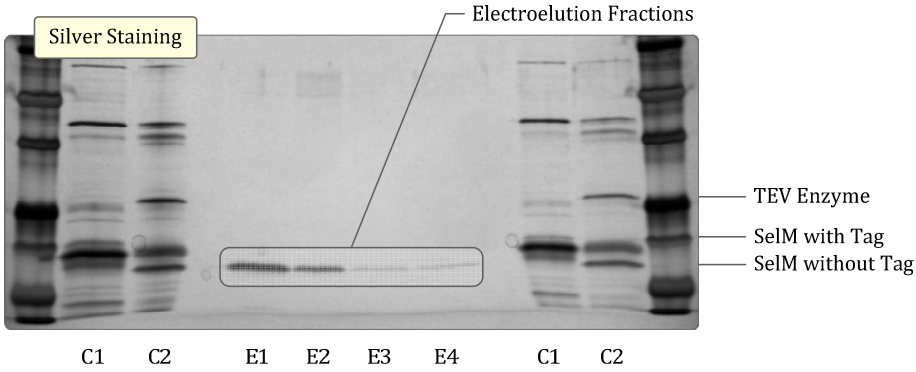

Figure 24. Analysis of electroeluted proteins by silver staining. Electric current forced the protein gel slices to release the non-fixed proteins from the gel. The proteins were then dissolved into the surrounding buffer. SelM was the only protein detected in the electroelution chambers. Thirty μL of each sample were run on a piece of 10-14.5% polyacrylamide gradient gel. Silver staining was applied to display the bands.

Protein was sequenced and verified as SelM.

Protein amino acids were sequenced and SelM was verified by Protein Identification Service (Midwest Bio Services, LLC), as identified by tandem mass spectrometry and the nano-LC-MS/MS technique. The peptide sequences resulting from these analyses were compared to known amino acid sequences on several databases including NCBInr, species, Bos Taurus, masses, and monoisotopic (Figure 25). The Xcorr score showed how well the observed peptide fragment mass spectrum matched the one theoretically predicted based on the sequence (Figure 25-A). The Xcorr was expected to be higher than 1.5-1.7 for peptides charged +1, higher than 2.0-2.5 for peptides charged +2, and higher than 2.5-3.5 for peptides charged +3. The probability score for each peptide showed the probability that the peptide was a random match (Figure 25-B). Therefore, the smaller values were better. According to these two most essential parameters, the probability score and the Xcorr score, selenoprotein M was the only match.

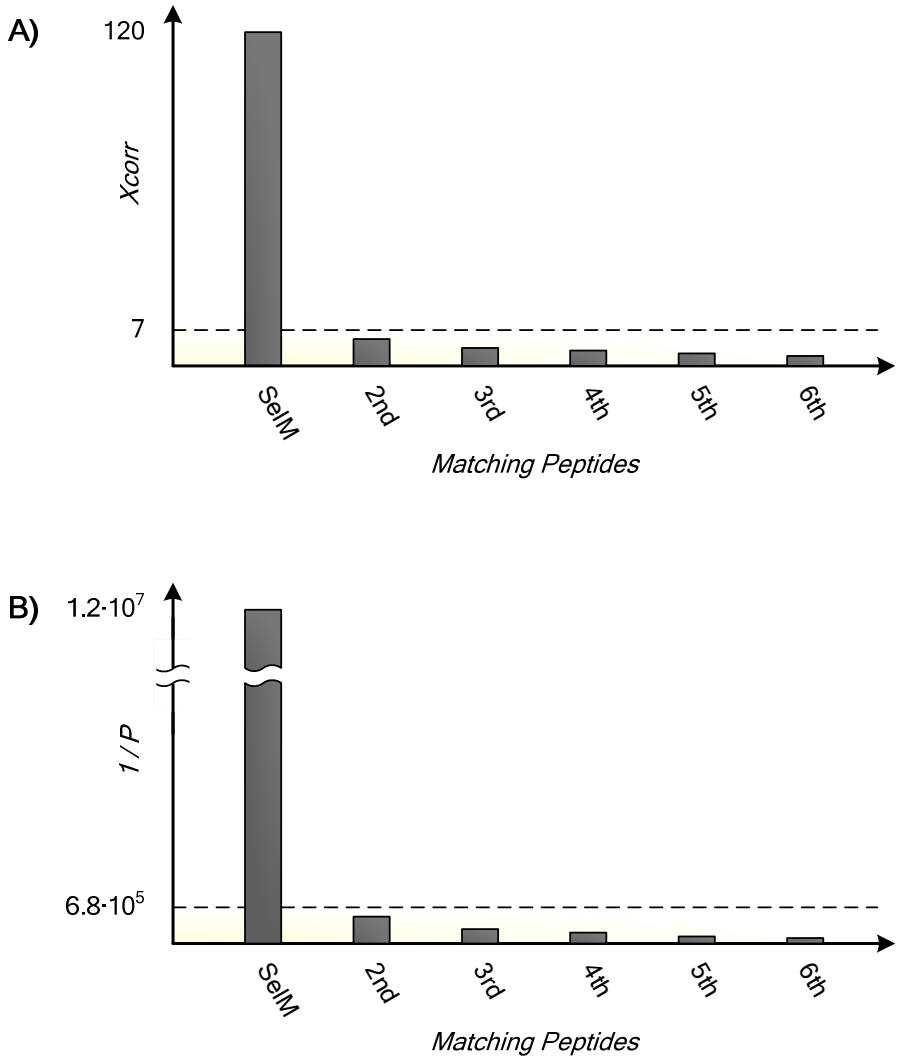

Figure 25. Protein identification. A) The Xcorr score showed how well the observed peptide fragment mass spectrum matched the one theoretically predicted based on the sequence. SelM was the only one that crossed the threshold and verified as the protein. B) The probability score for the protein for each peptide showed the probability that the peptide was a random match. Therefore, the larger 1/P values were better.

Elution fractions were concentrated using protein concentrators.

A proper concentrator was selected based on the molecular weight cut-off (MWCO). Generally, the pore size of the concentrator membrane should be two to three times smaller than the molecular weight of the protein (i.e., if the protein is 30 kDa, 10 kDA MWCO should be used). Because human SelM is a 14-kDa protein, 5 kDa MWCO concentrators were used. Concentrator membrane was conditioned and rinsed by PBS buffer before the addition of the protein sample. A relatively small amount of sample was spun in the concentrator at ≈3500 × g for 15-20 min.

Concentrated protein sample and the flow-through buffer were collected and run on a polyacrylamide gel to test the efficiency of the selected protein concentrator. Gel was stained with Coomassie blue (Figure 26). Results showed that for both samples, this process greatly increased the protein concentrations. For the electroelution product in particular, the concentration was increased from Coomassie blue non-detectable to detectable with a strong signal (Figures 27, 28). No SelM was detected in the flow-through (data not shown). Since the precipitation ability of the SelM was unknown, the concentration was processed in incremental steps in order to avoid over-concentration and precipitation. The main "contamination" protein detected in the samples after poly His-tagged protein purification and after TEV protease cleavage was the polymer of human SelM (Figure 28).

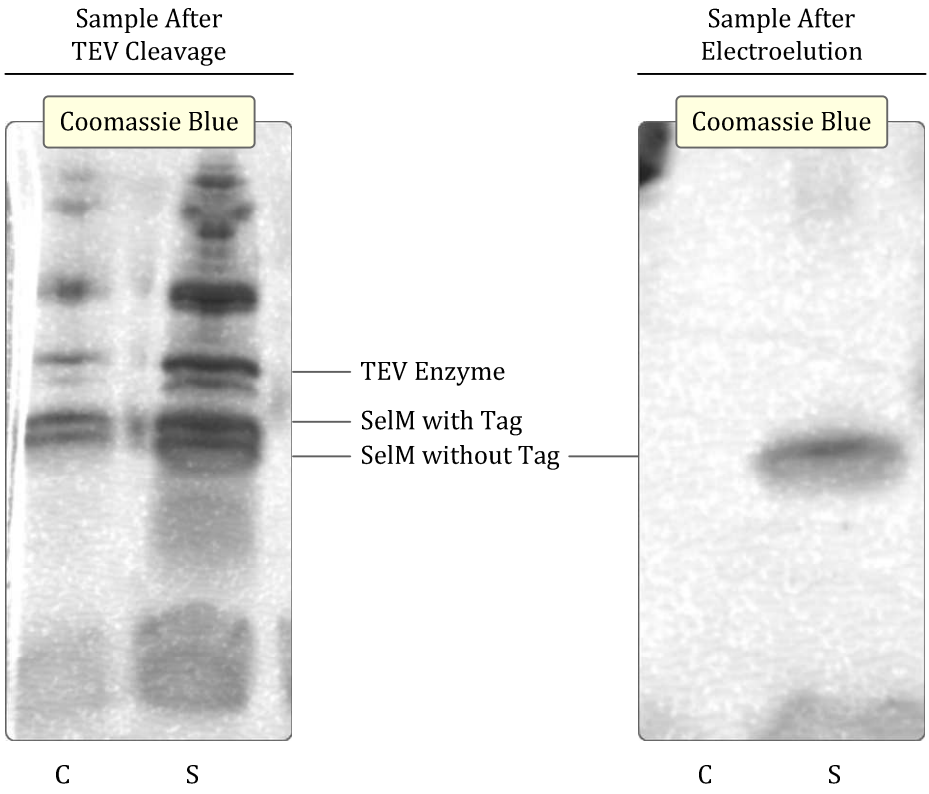

Figure 26. Protein concentration using 5-kDa MWCO concentrators. The sample concentrations were significantly increased by protein concentrators. The sample was reduced to 1/10 of its original volume in the protein concentrator. The concentration of electroelution product was increased from only sliver staining detectable (C right) to Coomassie blue detectable (S right).

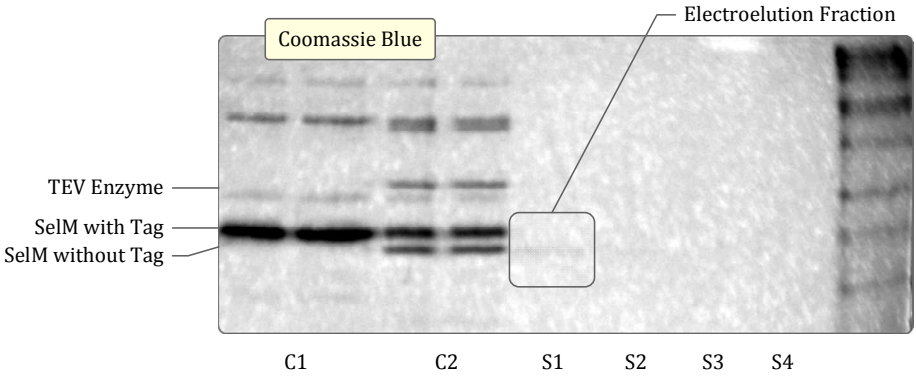

Figure 27. Electroelution fractions on Coomassie blue staining gel. C1 was the SelM protein sample before the TEV protease cleavage. C2 was the SelM protein sample after the TEV protease cleavage. S1, S2, S3, and S4 were the electroelutions. Thirty μL of each sample was used to run on a piece of polyacrylamide gradient gel, and the gel was stained with Coomassie blue. The SelM bands in the electroelution fraction were only barely visible.

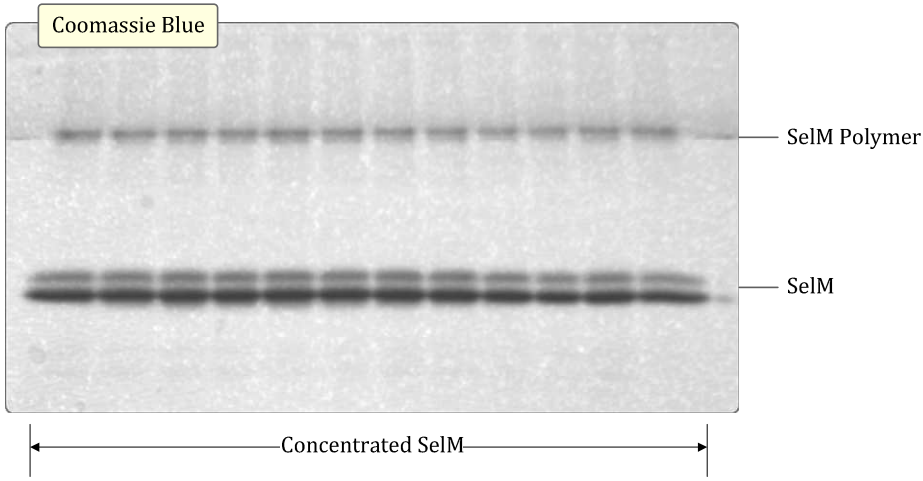

Figure 28. Concentrated protein on Coomassie blue staining gel. SelM protein concentration increased dramatically. The main "contamination" protein was the polymer of human SelM. It was detected in the samples after poly His-tagged protein purification and TEV protease cleavage.

CONCLUSION

Full-length human SelM was successfully expressed in bacteria in soluble fraction.

Gateway® technology was proved effective and convenient for DNA transferring among different vectors. Recombination reactions occurred at the specific attachment sites between the interacting plasmids. The sequences of *att* sites in different plasmids might be slightly different, but they all share the common 15-bp core region. The difference of surrounding sequences did not affect the specificity of the recombination reaction. Sub-cloning steps dealing with restriction sites on different vectors were no longer needed.

Mutagenesis by inverse PCR precisely synthesized large DNA molecules. Some primer sets that were designed according to the same rules, however, did not work. This might have been caused by the secondary structure of specific regions that were complemented by the primers. This explains why mutagenesis with a two-PCR system is not as effective as expected. Ensuring two sets of connected, non-overlapping primers work at the same time could be problematic. The primers need to include desirable site mutation, and there could be fewer combinations than people expect.

His-tagged SelM was efficiently purified using IMAC but inefficiently cleaved by TEV protease.

IMAC has proven to be effective for rough purification of poly His-tagged target proteins against the majority of contamination proteins. An appropriate concentration of imidazole in wash solution needed to be determined before attempting a large-scale purification. Within a certain range, higher concentrations of imidazole yielded purer elution fractions. Protein loss was detected when the concentration was increased to 30 mM.

TEV protease efficiency is poor. Approximately 50% of the substrates were cleaved after condition optimization. This might have been caused by the fact that some fusion proteins are intrinsically poor substrates for TEV protease, which in turn is caused by steric occlusion when the protease cleavage site is too close to the ordered structure in the target protein or when the fusion protein exists in the form of soluble aggregates. This problem can sometimes be mitigated by using a large amount of TEV protease.

IMAC failed to purify untagged SelM; electroelution successfully purified untagged SelM but reduced yield.

The N-terminal tag of the fusion protein was partially cleaved by TEV protease. Untagged SelM was not separated by applying another round of poly His-tagged protein purification. This might have been caused by protein molecules interacting with each other at native condition.

The use of denatured protein gels was considered as an active step for protein extraction and purification. Electroelution successfully separated the target proteins from other components of the samples. Silver detection, which was considered as a highly sensitive detection method, was used to detect the sample purity. A highly purified but diluted target protein solution was abstained.

Electroelution generated diluted protein, so protein concentrators were necessary, which further reduced yield.

Electroelution reduced the yield and dramatically reduced the target protein concentration by increasing the sample volume. An additional concentrating step, therefore, was necessary to recover the samples into concentrated solution. Appropriate protein concentrators were selected, and it was found that concentrators with an MWCO of about one-third the size of the target protein were effective.

REFERENCES

1. Allan, C. B., Lacourciere, G. M. & Stadtman, T. C. (1999). Responsiveness of selenoproteins to dietary selenium. *Annu Rev Nutr* **19**, 1-16.

2. Arner, E. S. & Holmgren, A. (2000). Physiological functions of thioredoxin and thioredoxin reductase. *Eur J Biochem* **267**, 6102-9.

3. Beck, M. A., Kolbeck, P. C., Rohr, L. H., Shi, Q., Morris, V. C. & Levander, O. A. (1994). Benign human enterovirus becomes virulent in selenium-deficient mice. *J Med Virol* **43**, 166-70.

4. Berry, M. J., Banu, L., Harney, J. W. & Larsen, P. R. (1993). Functional characterization of the eukaryotic SECIS elements which direct selenocysteine insertion at UGA codons. *EMBO J* **12**, 3315-22.

5. Bianco, A. C., Salvatore, D., Gereben, B., Berry, M. J. & Larsen, P. R. (2002). Biochemistry, cellular and molecular biology, and physiological roles of the iodothyronine selenodeiodinases. *Endocr Rev* **23**, 38-89.

6. Bosl, M. R., Takaku, K., Oshima, M., Nishimura, S. & Taketo, M. M. (1997). Early embryonic lethality caused by targeted disruption of the mouse selenocysteine tRNA gene (Trsp). *Proc Natl Acad Sci U S A* **94**, 5531-4.

7. Carlson, B. A., Novoselov, S. V., Kumaraswamy, E., Lee, B. J., Anver, M. R., Gladyshev, V. N. & Hatfield, D. L. (2004). Specific excision of the selenocysteine tRNA[Ser]Sec (Trsp) gene in mouse liver demonstrates an essential role of selenoproteins in liver function. *J Biol Chem* **279**, 8011-7.

8. Chasseur, C., Suetens, C., Nolard, N., Begaux, F. & Haubruge, E. (1997). Fungal contamination in barley and Kashin-Beck disease in Tibet. *Lancet* **350**, 1074.

9. Conrad, M., Jakupoglu, C., Moreno, S. G., Lippl, S., Banjac, A., Schneider, M., Beck, H., Hatzopoulos, A. K., Just, U., Sinowatz, F., Schmahl, W., Chien, K. R., Wurst, W., Bornkamm, G. W. & Brielmeier, M. (2004). Essential role for mitochondrial thioredoxin reductase in hematopoiesis, heart development, and heart function. *Mol Cell Biol* **24**, 9414-23.

10. Copeland, P. R., Fletcher, J. E., Carlson, B. A., Hatfield, D. L. & Driscoll, D. M. (2000). A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. *EMBO J* **19**, 306-14.

11. Costa, M., Rodriguez-Sanchez, J. L., Czaja, A. J. & Gelpi, C. (2000). Isolation and characterization of cDNA encoding the antigenic protein of the human tRNP(Ser)Sec complex recognized by autoantibodies from patients withtype-1 autoimmune hepatitis. *Clin Exp Immunol* **121**, 364-74.

12. 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. (2007). SelT, SelW, SelH, and Rdx12: genomics and molecular insights into the functions of selenoproteins of a novel thioredoxin-like family. *Biochemistry* **46**, 6871-82.

13. Ding, F. & Grabowski, P. J. (1999). Identification of a protein component of a mammalian tRNA(Sec) complex implicated in the decoding of UGA as selenocysteine. *RNA* **5**, 1561-9.

14. Dominy, C. N. & Andrews, D. W. (2003). Site-directed mutagenesis by inverse PCR. *Methods Mol Biol* **235**, 209-23.

15. Fagegaltier, D., Hubert, N., Yamada, K., Mizutani, T., Carbon, P. & Krol, A. (2000). Characterization of mSelB, a novel mammalian elongation factor for selenoprotein translation. *EMBO J* **19**, 4796-805.

16. Ferguson, A. D., Labunskyy, V. M., Fomenko, D. E., Arac, D., Chelliah, Y., Amezcua, C. A., Rizo, J., Gladyshev, V. N. & Deisenhofer, J. (2006). NMR structures of the selenoproteins Sep15 and SelM reveal redox activity of a new thioredoxin-like family. *J Biol Chem* **281**, 3536-43.

17. Fletcher, J. E., Copeland, P. R., Driscoll, D. M. & Krol, A. (2001). The selenocysteine incorporation machinery: interactions between the SECIS RNA and the SECIS-binding protein SBP2. *RNA* **7**, 1442-53.

18. Gelpi, C., Sontheimer, E. J. & Rodriguez-Sanchez, J. L. (1992). Autoantibodies against a serine tRNA-protein complex implicated in cotranslational selenocysteine insertion. *Proc Natl Acad Sci U S A* **89**, 9739-43.

19. Gereben, B., Goncalves, C., Harney, J. W., Larsen, P. R. & Bianco, A. C. (2000). Selective proteolysis of human type 2 deiodinase: a novel

ubiquitin-proteasomal mediated mechanism for regulation of hormone activation. *Mol Endocrinol* **14**, 1697-708.

20. Gromer, S., Eubel, J. K., Lee, B. L. & Jacob, J. (2005). Human selenoproteins at a glance. *Cell Mol Life Sci* **62**, 2414-37.

21. Guimaraes, M. J., Peterson, D., Vicari, A., Cocks, B. G., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Ferrick, D. A., Kastelein, R. A., Bazan, J. F. & Zlotnik, A. (1996). Identification of a novel selD homolog from eukaryotes, bacteria, and archaea: is there an autoregulatory mechanism in selenocysteine metabolism? *Proc Natl Acad Sci U S A* **93**, 15086-91.

22. Hernandez, A. & St Germain, D. L. (2003). Thyroid hormone deiodinases: physiology and clinical disorders. *Curr Opin Pediatr* **15**, 416-20.

23. Hill, K. E., Lloyd, R. S. & Burk, R. F. (1993). Conserved nucleotide sequences in the open reading frame and 3' untranslated region of selenoprotein P mRNA. *Proc Natl Acad Sci U S A* **90**, 537-41.

24. Hoffmann, P. R. & Berry, M. J. (2005). Selenoprotein synthesis: a unique translational mechanism used by a diverse family of proteins. *Thyroid* **15**, 769-75.

25. Jakupoglu, C., Przemeck, G. K., Schneider, M., Moreno, S. G., Mayr, N., Hatzopoulos, A. K., de Angelis, M. H., Wurst, W., Bornkamm, G. W., Brielmeier, M. & Conrad, M. (2005). Cytoplasmic thioredoxin reductase is essential for embryogenesis but dispensable for cardiac development. *Mol Cell Biol* **25**, 1980-8.

26. Kim, H. Y. & Gladyshev, V. N. (2004). Methionine sulfoxide reduction in mammals: characterization of methionine-R-sulfoxide reductases. *Mol Biol Cell* **15**, 1055-64.

27. Korotkov, K. V., Kumaraswamy, E., Zhou, Y., Hatfield, D. L. & Gladyshev, V. N. (2001). Association between the 15-kDa selenoprotein and UDP-glucose:glycoprotein glucosyltransferase in the endoplasmic reticulum of mammalian cells. *J Biol Chem* **276**, 15330-6.

28. Korotkov, K. V., Novoselov, S. V., Hatfield, D. L. & Gladyshev, V. N. (2002). Mammalian selenoprotein in which selenocysteine (Sec) incorporation is supported by a new form of Sec insertion sequence element. *Mol Cell Biol* **22**, 1402-11.

29. Kryukov, G. V., Castellano, S., Novoselov, S. V., Lobanov, A. V., Zehtab, O., Guigo, R. & Gladyshev, V. N. (2003). Characterization of mammalian selenoproteomes. *Science* **300**, 1439-43.

30. Kumaraswamy, E., Carlson, B. A., Morgan, F., Miyoshi, K., Robinson, G. W., Su, D., Wang, S., Southon, E., Tessarollo, L., Lee, B. J., Gladyshev, V. N., Hennighausen, L. & Hatfield, D. L. (2003). Selective removal of the selenocysteine tRNA [Ser]Sec gene (Trsp) in mouse mammary epithelium. *Mol Cell Biol* **23**, 1477-88.

31. Landy, A. (1989). Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annu Rev Biochem* **58**, 913-49.

32. Leinfelder, W., Forchhammer, K., Zinoni, F., Sawers, G., Mandrand-Berthelot, M. A. & Bock, A. (1988). Escherichia coli genes whose products are involved in selenium metabolism. *J Bacteriol* **170**, 540-6.

33. Lesoon, A., Mehta, A., Singh, R., Chisolm, G. M. & Driscoll, D. M. (1997). An RNA-binding protein recognizes a mammalian selenocysteine insertion sequence element required for cotranslational incorporation of selenocysteine. *Mol Cell Biol* **17**, 1977-85.

34. Low, S. C., Harney, J. W. & Berry, M. J. (1995). Cloning and functional characterization of human selenophosphate synthetase, an essential component of selenoprotein synthesis. *J Biol Chem* **270**, 21659-64.

35. Martin, G. W., 3rd, Harney, J. W. & Berry, M. J. (1996). Selenocysteine incorporation in eukaryotes: insights into mechanism and efficiency from sequence, structure, and spacing proximity studies of the type 1 deiodinase SECIS element. *RNA* **2**, 171-82.

36. Mustacich, D. & Powis, G. (2000). Thioredoxin reductase. *Biochem J* **346 Pt 1**, 1-8.

37. Papp, L. V., Lu, J., Holmgren, A. & Khanna, K. K. (2007). From selenium to selenoproteins: synthesis, identity, and their role in human health. *Antioxid Redox Signal* **9**, 775-806.

38. Rotruck, J. T., Pope, A. L., Ganther, H. E., Swanson, A. B., Hafeman, D. G. & Hoekstra, W. G. (1973). Selenium: biochemical role as a component of glutathione peroxidase. *Science* **179**, 588-90.

39. Stadtman, T. C. (1996). Selenocysteine. *Annu Rev Biochem* **65**, 83-100.

40. Steinsapir, J., Harney, J. & Larsen, P. R. (1998). Type 2 iodothyronine deiodinase in rat pituitary tumor cells is inactivated in proteasomes. *J Clin Invest* **102**, 1895-9.

41. Steinsapir, J., Harney, J. & Larsen, P. R. (1998). Type 2 iodothyronine deiodinase in rat pituitary tumor cells is inactivated in proteasomes. *J Clin Invest* **102**, 1895-9.

42. Studier, F. W. & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**, 113-30.

43. Sunde, R. A. & Evenson, J. K. (1987). Serine incorporation into the selenocysteine moiety of glutathione peroxidase. *J Biol Chem* **262**, 933-7.

44. Tujebajeva, R. M., Copeland, P. R., Xu, X. M., Carlson, B. A., Harney, J. W., Driscoll, D. M., Hatfield, D. L. & Berry, M. J. (2000). Decoding apparatus for eukaryotic selenocysteine insertion. *EMBO Rep* **1**, 158-63.

45. Ursini, F., Maiorino, M., Brigelius-Flohe, R., Aumann, K. D., Roveri, A., Schomburg, D. & Flohe, L. (1995). Diversity of glutathione peroxidases. *Methods Enzymol* **252**, 38-53.

46. Xia, Y., Hill, K. E., Byrne, D. W., Xu, J. & Burk, R. F. (2005). Effectiveness of selenium supplements in a low-selenium area of China. *Am J Clin Nutr* **81**, 829-34.

47. Xu, X. M., Carlson, B. A., Mix, H., Zhang, Y., Saira, K., Glass, R. S., Berry, M. J., Gladyshev, V. N. & Hatfield, D. L. (2007). Biosynthesis of selenocysteine on its tRNA in eukaryotes. *PLoS Biol* **5**, e4.

48. Zavacki, A. M., Mansell, J. B., Chung, M., Klimovitsky, B., Harney, J. W. & Berry, M. J. (2003). Coupled tRNA(Sec)-dependent assembly of the selenocysteine decoding apparatus. *Mol Cell* **11**, 773-81.