Published in final edited form as:

Protein Pept Lett. 2005 November ; 12(8): 757–764.

INCORPORATION OF SELENOCYSTEINE INTO PROTEINS USING PEPTIDE LIGATION

Robert J. Hondal^{*,†}

Robert J. Hondal: Robert.Hondal@uvm.edu

^{*}Department of Biochemistry, Vermont College of Medicine, The University of Vermont, Burlington, VT 05405, USA, Phone (802) 656-8282

Abstract

Expressed protein ligation has become a frequently used technique to insert non-standard amino acids into proteins. The technique has been adapted to insert selenocysteine residues in place of cysteine residue in proteins, taking advantage of the similarity in the chemistries of sulfur and selenium. This replacement can confer unique structural and catalytic properties to enzymes and proteins. The development of this technique also allows for naturally occurring selenoproteins to be produced semisynthetically.

Keywords

Selenocysteine; expressed protein ligation; thioredoxin reductase; semisynthetic

Selenium, the mysterious moon-metal

The Swedish chemist, J.J. Berzelius discovered the 34th element of the periodic table in 1817 and named it after the Greek goddess of the moon, Selene [1]. It is wholly appropriate that he named this element after the moon-goddess, because selenium, like the moon, has remained mysterious to science until the modern era

The biological role of selenium began to be realized when Schwarz discovered a dietary factor that was necessary to prevent liver necrosis in rats [2]. He called this mysterious nutritional requirement, “Factor 3”. He then determined that this dietary factor was the trace element selenium. In the 1970’s, Chinese scientists reported that selenium could prevent a childhood cardiomyopathy known as Keshan disease [3], much like vitamin C can prevent the onset of scurvy. Beck and Levander later showed that both influenza and coxsackie viruses accumulate genetic mutations when these viruses have infected animal hosts that are deficient in selenium or vitamin E [4–5]. Keshan disease is caused when coxsackie virus mutates to a more virulent strain in a host deficient in selenium. In the 1980’s, it was discovered that selenium is the *only trace element that is incorporated directly into proteins as part of the polypeptide chain*. [6]. The incorporation of selenium into proteins, known as selenoproteins, is beginning to explain the important role that selenium plays in human health. A number of important examples can be cited. Selenium is vital to the function of the thyroid gland. The key enzyme in the thyroid gland, which converts thyroxine (T4) to triiodothyronine (T3), contains selenium [7]. Selenium is absolutely essential to human male fertility because an enzyme in the testes containing selenium protects developing sperm from oxidative injury [8]. There has long been anecdotal evidence that selenium offers

[†]This work was supported by grant GM070742-01

protection against some form of cancers [9]. This hypothesis was tested recently in clinical trials and selenium was found to offer significant protection against, lung, colorectal, and prostate cancer [10–11].

Selenium and the split personality of the genetic code

Many of the health benefits of selenium can be explained by its incorporation into proteins as the “21st” amino acid, selenocysteine [12–13]. This amino acid, though rare, is ubiquitous in all of the kingdoms of life being found in eukarya, prokarya, and archaea [14–15]. Examples of selenocysteine-containing proteins are found in plants, fungi, mammals, bacteria, and fish [16]. Like the other standard proteinogenic amino acids, selenocysteine is incorporated through the genetic machinery by a unique codon. The opal stop codon, UGA, was discovered to have a dual function (*i.e.* a “split personality”) [17]. In addition to being a STOP signal for protein synthesis, it is also the signal to the cellular protein machinery to insert the amino acid selenocysteine into a protein chain [18]. The UGA codon can be recoded as a sense codon for selenocysteine using an intricate mechanism that requires a stem-loop structure in the *mRNA* known as a selenocysteine insertion sequence element (SECIS element) [19–20]. Gladyshev has elucidated the mammalian selenoproteome by using the SECIS element as a marker to differentiate which genes contain an in-frame sense codon for selenocysteine from genes that merely contain UGA as a stop codon. Thus, 25 such proteins in the human genome have been identified [21]. Reading frames having a selenocysteine sense codon require this structure to correctly interpret the UGA codon.

There are differences in the position of the SECIS element in relation to the UGA codon between prokaryotes and eukaryotes as shown in Figure 1. The SECIS element is located in the 3′ untranslated region of the *mRNA* in eukaryotes, whereas in prokaryotes it is proximal to the UGA codon [22]. In addition, there are sequence and structural differences between eukaryotic and prokaryotic SECIS elements. In both eukarya and prokarya, accessory proteins are required to assist in the decoding of the UGA codon. In eukaryotes these accessory factors are SBP2 (SECIS-binding protein), efSEC (an elongation factor specific to selenocysteine), and a Sec-specific tRNA [23]. Homologues of these accessory factors are present in prokaryotes as well. The differences between the Sec-inserting machineries of prokaryotes and eukaryotes makes it difficult to express mammalian selenocysteine-containing proteins in bacteria using recombinant DNA technology. However, two research groups have been able to express mammalian thioredoxin reductase (TR) in *E. coli* at high levels by altering the *mRNA* of the gene for TR so that it contains a bacterial SECIS element [24–25]. This is possible for TR because the selenocysteine residue is the penultimate amino acid in the sequence. In addition, this method does not result in complete read-through of the UGA codon, thus resulting in the formation of some truncated product. This technique is not amenable for other selenoproteins in which the UGA codon is at the 5′ end of the gene. In one example in which a Sec residue was inserted near the 5′ end of the gene for a mutant form of glutathione S-transferase, 85% of the translated product was truncated and the wild-type sequence of the protein had to be altered to accommodate the native sequence of the bacterial SECIS element [26]. For high levels of expression, the accessory genes required for decoding UGA must also be expressed. Another biological approach was taken by Muller and coworkers who used a cysteine auxotroph to produce a selenium-containing variant of thioredoxin [27].

In contrast to the genetic approach by Arner, my group and several others have used expressed protein ligation to construct proteins containing selenocysteine semisynthetically [28–31]. Prior to these efforts, Hilvert had used a total chemical approach to create the first artificial selenoenzyme by activating a serine residue in subtilisin with phenylmethanesulfonyl fluoride and then treating the activated ester with hydrogen selenide

to complete the conversion to selenocysteine. This was possible because subtilisin contains a very reactive serine hydroxyl group that made conversion to selenocysteine more facile [32]. This review focuses on semisynthetic approaches to constructing selenoproteins by using peptide ligation.

Physico-chemical properties of selenocysteine

In comparison to the thiol group of cysteine, selenocysteine is much more acidic and much more nucleophilic than its sulfur cousin. The selenol group of selenocysteine has a pK_a of 5.2 (compare to 8.5 for a typical thiol) and is thus ionized at physiological pH [33]. The redox potential of a diselenide bond (-380 mV) and a mixed selenylsulfide bond (-326 mV) are much lower than that of a typical disulfide bond (-180 mV) [34]. Hence, selenium compounds will quickly oxidize in oxygenated buffer and favor the formation of a diselenide bond. This makes working with synthetic peptides containing selenocysteine problematic. A number of oxidation states are available to selenium, but selenium in a protein is usually found as a selenolate due to its low pK_a . Glutathione peroxidase is postulated to use a selenenic acid (R-Se-OH) intermediate during its cycle of catalysis, similar to other enzymes that use the sulfenic acid form of cysteine [35]. Thioredoxin reductase and probably other selenoenzymes use a selenylsulfide bond during their catalytic cycle [36]. However, diselenide bonds have not been found in proteins to date since only one known protein (Selenoprotein P) contains multiple Sec residues [37]. The low redox potential of a diselenide makes it unlikely to be important biologically, but Moroder and colleagues have taken advantage of the redox properties of selenium to synthesize stable, diselenide analogues of apamin [38]. The covalent radius of selenium is only slightly larger than that of sulfur (1.16 Å vs. 1.03 Å), thus a sulfur-selenium bond has a length of 2.2 Å vs. 2.0 Å for a typical sulfur-sulfur bond [39]. These properties make selenium an acceptable replacement for sulfur in a protein structure.

One feature of selenium that is useful to NMR spectroscopists is that it has a naturally occurring NMR-active nuclide (^{77}Se) which has a natural abundance of 7.5% and a spin of $1/2$. This has made biological ^{77}Se spectroscopy possible and several examples are known in the literature [40–41].

Synthesis of Selenocysteine Derivatives

The amino acid selenocysteine can be synthesized from sodium diselenide, which is synthesized *in situ* by reduction of elemental selenium by NaBH_4 . The addition of β -chloroalanine as shown in Scheme 1 to the reaction mixture produces selenocystine as a yellow crystal [42–43]. Selenocystine can then be reduced *in situ* with sodium borohydride and treated with *p*-methoxybenzyl chloride to produce the pMob protected derivative [42]. The α -nitrogen atom can then be protected with an Fmoc group by treating *Se-p*-methoxybenzyl selenocysteine with Fmoc-OSu to produce *N*-9-fluorenylmethoxycarbonyl-*Se*-4-methoxybenzylselenocysteine (Fmoc-*Se*(pMob)-OH). This derivative can be recrystallized with an overall yield of ca. 60% [44].

Early work on the synthesis of peptides containing selenocysteine was done by Walter and coworkers who prepared *Se*-benzyl derivatives of selenocysteine and synthesized selenium analogs of glutathione, oxytocin, and somatostatin. The benzyl group is removed by sodium in liquid ammonia (Na/NH_3 liq.) and is accompanied by side reactions [45–47]. Kenji Soda pioneered the use of the *p*-methoxybenzyl group for selenocysteine [42]. The 4-methoxybenzyl group (pMob) is much more acid labile than the benzyl protecting group. The pMob group can be removed by using a cleavage cocktail of TFA, *m*-cresol, thioanisole, and trimethylsilyl trifluoromethanesulfonate (TMSOTf) [42]. Trimethylsilyl bromide has also been used [29].

van der Donk and colleagues have developed a simpler synthesis of Fmoc-Sec derivatives by converting the allyl ester of Fmoc-serine to the *O*-tosyl derivative. Selenium nucleophiles generated by *in situ* reduction of diselenides can then be used to displace the *O*-tosyl group to produce the protected selenocysteine derivative. As pointed out by van der Donk in a recent review, the simplicity and scalability of this route is advantageous over the more cumbersome route of starting with elemental selenium [48].

Synthesis conditions for peptides containing selenocysteine

The biggest problem with using Fmoc-Sec(pMob)-OH in solid phase peptide synthesis is that it is prone to racemization during coupling. A second problem is that in subsequent elongation steps using a tertiary base such as diisopropylethylamine, the protected amino acid is prone to formation of dehydroalanine via a C α hydrogen atom abstraction mechanism. The dehydroalanine residue thus formed during the synthesis can be attacked by piperidine to form β -piperidine adducts. Besse and Moroder have diminished these problems by using the pentafluorophenyl esters of both Sec and Cys derivatives and eliminating the use of tertiary bases during the elongation step. Their optimized procedure also includes minimizing the time that the resin is exposed to piperidine in the deprotection step [49].

A procedure that has worked well in my laboratory is a protocol developed by Barany who has long recognized that racemization for Cys derivatives is also a problem in solid-phase peptide synthesis [50]. This protocol eliminates any preactivation step and uses a weak base (collidine) in place of DIEA or NMM. Further reduction in racemization is accomplished by using 2-fold less base and a less polar solvent system, CH₂Cl₂-DMF (1:1), during coupling. These coupling conditions result in only 0.5% formation of *D*-peptide. I have used this protocol in my laboratory to synthesize a tripeptide of sequence Cys-Sec-Gly, which corresponds to the C-terminal tripeptide of the selenoenzyme thioredoxin reductase. A HPLC chromatogram of this peptide after deprotection with TFA/water (95:5) is shown in Figure 2. Mass spectral analysis of the peaks in Figure 2 reveal that 95% of the peak area belongs to the protected Sec(pMob) peptide (peak 2) and 5% corresponds to the cyclic deprotected peptide (peak 1). Using the same HPLC conditions as Barany, we could not detect any formation of the *D*-peptide, although it is probably present to a very low extent [51].

Deprotection of Sec residues

Even though the pMob group can be removed from a Sec residue in a peptide by the use of strong Lewis acids, there are some problems associated with this method of removal. For example, *m*-cresol is not very soluble in ether and peptides treated with a cleavage cocktail consisting of TFA, *m*-cresol, thioanisole, and TMSOTf, tend to oil out upon lyophilization, especially small peptides containing Sec. In addition, deselenization happens to an appreciable extent using this cleavage cocktail.

Oxidative deprotection using I₂ has been a strategy used by van der Donk's group to remove the pMob group [52]. However, great care must be used with this procedure because the selenyliodide adduct that forms tends to deselenate. In our hands, excess I₂ leads to rapid deselenization. This is because carbon-selenium bonds are weak and a stable selenyliodide anion can be formed. This seems somewhat surprising since the enzyme triiodothyronine deiodinase presumably attacks an iodine atom of an aryl iodide [53]. Presumably the enzyme does not deselenate, but rather a reaction coordinate is available to the enzyme that is not available in solution. Thus conditions for oxidative deprotection should be optimized for each individual Sec(pMob) containing peptide.

My group has found that at least in the synthesis of tripeptide Cys-Sec-Gly, removal of the pMob group can be accomplished with mild conditions if the protecting group for the thiol is a *S*(*t*-butyl) group. The presence of a nearby disulfide apparently allows the protected Se group to form a bridged selenonium intermediate that decomposes in the presence of scavengers to yield an bridged selenylsulfide bond. This oxidized tripeptide can then be used directly to make a semisynthetic thioredoxin reductase [54]. This unique deprotection mechanism is shown in Figure 3.

Incorporation of Sec residues into proteins using native chemical ligation and intein-mediated peptide ligation

Several groups have found that selenocysteine can be incorporated into a polypeptide chain by using either the technique of native chemical ligation as developed by Kent and coworkers [55], or by using a semisynthetic approach through the use of inteins [56–57]. Both of these methods make use of the fact that selenium has similar properties to sulfur and can easily attack an α -thioester bond required for such ligation techniques. In fact, using a model chromogenic thioester shown in Scheme 2, it was shown that the rate of attack on the thioester is 10^3 higher at pH 5 and 10 fold higher at pH 8 (Table 1) [28].

Both native chemical ligation and intein-mediated peptide ligation use a N-terminal cysteine (or selenocysteine) residue to attack a C-terminal α -thioester. The difference in the techniques is the way in which the α -thioester is generated. For the intein-mediated approach, the intein is fused to the C-terminus of the protein of interest. Inteins are naturally occurring proteins that catalyze protein splicing reactions and do so by generating thioester intermediates [58–59]. If an intein is fused to the C-terminus of a target protein, the peptide bond at the junction site will exist in equilibrium between the amide and thioester forms. The thioester formed can be cleaved with a small molecule thiol to generate a thioester-tagged protein. A synthetic peptide containing selenocysteine (or cysteine) at the N-terminus can then be used to attack the thioester and ligate the two peptides together. This semisynthetic approach to constructing a protein containing selenocysteine is shown in Figure 4.

The advantage to the semisynthetic approach outlined in Figure 4 is that the protein can essentially be divided into two modules, a non-selenium containing module that can be produced as a recombinant intein-fusion protein in *E. coli*, and a selenium containing module that is produced as a synthetic peptide. The recombinant fusion-protein can be expressed at a high level, and can yield many milligrams of thioester-tagged protein per milliliter of culture. Although there are some difficulties associated with synthesizing a selenocysteine-containing peptide as discussed above, the overall yield of semisynthetic selenoprotein produced in this manner is easily in the ten's of milligrams. This strategy can be used to replace a naturally occurring cysteine residue in the protein to yield a mutant with unique properties, or naturally occurring selenoproteins can be made. Both types of examples are discussed below.

There are a limited number of examples of this technique in the literature thus far. A semisynthetic form of azurin, a blue copper protein, has been produced by expressed protein ligation (EPL) by replacing Cys112 of the wild-type protein with selenocysteine [30–31]. The copper atom coordinates with Cys112 and two histidine residues in the active site, enabling the copper atom to cycle between Cu^{1+} and Cu^{2+} during biological electron transport [60]. Mutation of the Cys residue to one of the other 19 standard amino acids greatly altered the properties of the enzyme. The creation of a semisynthetic variant containing a Sec residue, having similar chemistry to Cys, allowed for unique spectroscopic studies to be carried out [30–31].

A second example of using Sec residues in place of Cys in EPL was the replacement of Cys38 with Sec in bovine pancreatic trypsin inhibitor (BPTI). The selenium-containing mutant of BPTI inhibited trypsin with nearly identical values of K_i [29]. Hilvert has further extended ligations with selenium to include selenohomocysteine to create a selenium variant of bovine pancreatic polypeptide (bPP) [61]. The selenohomocysteine residue incorporated into the peptide was then alkylated to create a selenomethionine residue. This approach would allow site-specific incorporation of selenomethionine residues which may be useful for spectroscopic and structural studies.

Further work on creating semisynthetic selenoenzymes was done with the model protein ribonuclease A (RNase A) [28,44]. RNase A contains 8 cysteine residues that are oxidized to form 4 disulfide bonds. Potentially, each cysteine residue could be replaced with selenocysteine using EPL. We have created three such mutants of RNase A in which both Cys110 and Cys 95 have been replaced with Sec. A double mutant C110U-C95U has also been constructed. Cys110 participates in a disulfide bond with Cys58, while Cys95 is linked with Cys40. Both disulfide bonds are crucial for activity. Sec mutants thus constructed are made by reducing all of the disulfide bonds and then refolding the enzyme. The refolding process becomes a crucial test to demonstrate the formation of the proper Se-S cross-links in the mutant enzyme (see Figure 5). RNase A molecules that do not have the correct S-S (or S-Se in this case) connectivities will not have catalytic activity. Table 2 summarizes the activities of RNase A and its selenium-containing mutants.

My group has recently created a semisynthetic thioredoxin reductase (TR). TR is amenable to EPL because the Sec residue is contained in a C-terminal tripeptide with sequence Cys-Sec-Gly. Thus, 487 amino acids can be biosynthesized as a thioester-tagged polypeptide in *E. coli* with the remaining three amino acids made synthetically [54]. The semisynthetic TR as shown in Figure 6, can be made at a level of 3 mg/L of culture. This semisynthetic approach will allow mechanistic details of the enzyme to be studied in great detail. As shown in Figure 7, the semisynthetic TR has very high activity with its natural substrate thioredoxin. The semisynthetic enzyme has a k_{cat} value of 1500 min^{-1} as compared to a value of 2500 min^{-1} for rat TR1 prepared from rat liver [62]. The use of semisynthesis to create thioredoxin reductase also demonstrates that the Sec residue need not be at the site of ligation, but can be distal from the ligation site as long as a N-terminal cysteine is present. In comparison to the method of Arner, who used genetic manipulation to produce a recombinant thioredoxin reductase, this method does not require exotic bacterial strains or overexpression of accessory proteins to generate TR. The technology is simple, however there are some barriers to overcome in synthesizing peptides containing selenocysteine. The semisynthetic system will allow for incorporation of non-natural moieties that will allow us to explore structure-function relationships that would not be possible with routine recombinant technologies

Acknowledgments

The author would like to acknowledge R.T. Raines for allowing him to pursue semisynthesis as a strategy to make TR. The author is grateful to J.S. Madalenoitia for reading this manuscript and to K.M. Harris for her devotion to peptide synthesis.

References

1. Berzelius JJ. Afhandl Fys Kemi Mineralogi. 1818; 6:42.
2. Schwarz K, Foltz TM. J Am Chem Soc. 1957; 79:3292–3293.
3. Yang GL, Lin ZH, Li SG, Guo BI, Yin YS. J Trace Elem Electrolytes Health Dis. 1988; 2:157–163. [PubMed: 2908794]
4. Beck MA, Shi Q, Morris VC, Levander OA. Nat Med. 1995; 1:433–436. [PubMed: 7585090]

5. Nelson HK, Shi Q, Van Dael P, Schiffrin EJ, Blum S, Barclay D, Levander OA, Beck MA. *FASEB J*. 2001; 15:1846–1848. [PubMed: 11481250]
6. Zinoni F, Bikmann A, Stadtman TC, Bock A. *Proc Natl Acad Sci USA*. 1986; 83:4650–4654. [PubMed: 2941757]
7. Behne D, Dyriakopoulos A, Meinhold H, Kohrle J. *Biochem Biophys Res Commun*. 1990; 173:1143–1149. [PubMed: 2268318]
8. Ursini F, Heim S, Kiess M, Maiorino M, Roveri A, Wissing J, Flohe L. *Science*. 1999; 285:1393–1395. [PubMed: 10464096]
9. Willett WC, Stampfer MJ. *Science*. 1988; 259:573–574.
10. Clark LC, Combs GF Jr, Turnbull BW, Slate EH, Chalker DK, Chow J, Davis LS, Glover RA, Graham GF, Gross EG, Krongrad A, Leshner JL Jr, Park HK, Sanders BB Jr, Smith CL, Taylor JR. *JAMA*. 1996; 276:1957–1963. [PubMed: 8971064]
11. Duffield-Lillico AJ, Slate EH, Reid ME, Turnbull BW, Wilkins PA, Combs GF Jr, Park HK, Gross EG, Graham GF, Stratton MS, Marshall JR, Clark LC. *J Natl Cancer Inst*. 2003; 95:1477–81. [PubMed: 14519754]
12. Atkins JF, Gesteland RF. *Nature*. 2000; 407:463–465. [PubMed: 11028985]
13. Bock A, Forchhammer K, Heider J, Leinfelder W, Sawers G, Veprék B, Zinoni F. *Mol Microbiol*. 1991; 5:515–520. [PubMed: 1828528]
14. Gladyshev VN, Kryukov GV. *Biofactors*. 2001; 14:87–92. [PubMed: 11568444]
15. Rother M, Resch A, Wilting R, Bock A. *Biofactors*. 2001; 14:75–83. [PubMed: 11568443]
16. Whanger PD. *J Am Coll Nutr*. 2002; 21:223–32. [PubMed: 12074249]
17. Hatfield D, Diamond A. *Trends Genet*. 1993; 9:15–16.
18. Chambers I, Frampton J, Goldfarb P, Affara N, McBain W, Harrison PR. *EMBO J*. 1986; 5:1221–1227. [PubMed: 3015592]
19. Heider J, Baron C, Bock A. *EMBO J*. 1992; 11:3759–3766. [PubMed: 1396569]
20. Tujebajeva RM, Copeland PR, Xu XM, Carlson BA, Harney JW, Driscoll DM, Hatfield DL, Berry MJ. *EMBO Rep*. 2000; 1:158–163. [PubMed: 11265756]
21. Kryukov GV, Castellano S, Novoselov SV, Lobanov AV, Zehetab O, Guigo R, Gladyshev VN. *Science*. 2003; 300:1439–43. [PubMed: 12775843]
22. Krol A. *Biochimie*. 2002; 84:765–74. [PubMed: 12457564]
23. Mehta A, Rebsch CM, Kinzy SA, Fletcher JE, Copeland PR. *J Biol Chem*. 2004; 279:37852–9. [PubMed: 15229221]
24. Arner ESJ, Sarioglu H, Lottspeich F, Holmgren A, Bock A. *J Mol Biol*. 1999; 292:1003–1016. [PubMed: 10512699]
25. Bar-Noy S, Gorlatov SN, Stadtman TC. *Free Rad Biol Med*. 2001; 30:51–61. [PubMed: 11134895]
26. Jiang Z, Arner ES, Mu Y, Johansson L, Shi J, Zhao S, Liu S, Wang R, Zhang T, Yan G, Liu J, Shen J, Luo G. *Biochem Biophys Res Commun*. 2004; 321:94–101. [PubMed: 15358220]
27. Muller S, Senn H, Gsell B, Vetter W, Baron C, Bock A. *Biochemistry*. 1994; 33:3404–12. [PubMed: 8136378]
28. Hondal RJ, Nilsson BL, Raines RT. *J Am Chem Soc*. 2001; 123:5140–5141. [PubMed: 11457362]
29. Quaderer R, Sewing A, Hilvert D. *Helv Chim Acta*. 2001; 84:1191–1206.
30. Berry SM, Giesselman MD, Nilges MJ, van der Donk WA, Lu Y. *J Am Chem Soc*. 2002; 124:2084–2085. [PubMed: 11878940]
31. Ralle M, Berry SM, Nilges MJ, Giesselman MD, van der Donk WA, Lu Y, Blackburn NJ. *J Am Chem Soc*. 2004; 126:7244–7256. [PubMed: 15186162]
32. Wu ZP, Hilvert D. *J Am Chem Soc*. 1989; 111:4513–4514.
33. Huber RE, Criddle RS. *Arch Biochem Biophys*. 1967; 122:164–173. [PubMed: 6076213]
34. Besse D, Siedler F, Diercks T, Kessler H, Moroder L. *Angew Chem Int Ed Engl*. 1997; 36:883–885.
35. Wendel A, Pilz W, Ladenstein R, Sawatzki G, Weser U. *Biochim Biophys Acta*. 1975; 377:211–215. [PubMed: 1122287]

36. Zhong L, Arner ES, Holmgren A. *Proc Natl Acad Sci USA*. 2000; 97:5854–9. [PubMed: 10801974]
37. Hill KE, Lloyd RS, Yang JG, Read R, Burk RF. *J Biol Chem*. 1991; 266:10050–10053. [PubMed: 2037562]
38. Fiori, Pegoraro S, Rudolph-Bohner S, Cramer J, Moroder L. *Biopolymers*. 2000; 53:550–564. [PubMed: 10766951]
39. Pauling, L. *The Nature of the Chemical Bond*. 3. Cornell University Press; Ithaca, NY: 1960. p. 246–260.
40. Gettins P, Wardlaw S. *J Biol Chem*. 1991; 266:3422–3426. [PubMed: 1995605]
41. Duddeck H. *Prog NMR Spectroscopy*. 1995; 27:1–323.
42. Koide T, Itoh H, Otaka A, Yasui H, Masataka K, Esaki N, Soda K, Fuji N. *Chem Pharm Bull*. 1993; 41:502–506. [PubMed: 8477500]
43. Tanaka H, Soda K. *Meth Enzymol*. 1987; 143:240–243. [PubMed: 2958675]
44. Hondal RJ, Raines RT. *Meth Enzymol*. 2002; 347:70–83. [PubMed: 11898440]
45. Theodorou D, Shuwartz IL, Walter R. *Biochemistry*. 1967; 6:3927–3932. [PubMed: 6076637]
46. Hartrodt B, Neuber K, Bierwolf B, Blech W, Jakubke HD. *Tetrahedron Lett*. 1980; 21:2393–2398.
47. Walter R, Chan WY. *J Am Chem Soc*. 1967; 89:3892–3893. [PubMed: 6068786]
48. Gieselman MD, Zhu Y, Zhou H, Galonic D, van der Donk WA. *ChemBioChem*. 2002; 3:709–716. [PubMed: 12203969]
49. Besse D, Moroder L. *J Pep Sci*. 1997; 3:442–453.
50. Han Y, Albericia F, Barany G. *J Org Chem*. 1997; 62:4307–4312. [PubMed: 11671751]
51. Harris KM, Hondal RJ. 2004 submitted to *Pep. Sci*.
52. Gieselman MD, Xie L, van der Donk WA. *Org Lett*. 2001; 3:1331–1334. [PubMed: 11348227]
53. Jacob C, Giles GI, Giles NM, Sies H. *Angew Chem Int Ed Engl*. 2003; 42:4742–4758. [PubMed: 14562341]
54. Eckenroth B, Harris KM, Gladyshev VN, Raines RT, Hondal RJ. 2004 submitted to *Biochemistry*.
55. Dawson PE, Muir TW, Clark-Lewis I, Kent SB. *Science*. 1994; 266:776–9. [PubMed: 7973629]
56. Hackeng TM, Griffin JH, Dawson PE. *Proc Natl Acad Sci U S A*. 1999; 96:10068–73. [PubMed: 10468563]
57. Evans TC Jr, Benner J, Xu MQ. *Protein Sci*. 1998; 7:2256–64. [PubMed: 9827992]
58. Gray HB, Winkler JR. *Annu Rev Biochem*. 1996; 65:537–61. [PubMed: 8811189]
59. Shao Y, Kent SBH. *Chem Biol*. 1997; 4:187–194. [PubMed: 9115411]
60. Cooper AA, Stevens TH. *Trends Biochem Sci*. 1995; 20:351–356. [PubMed: 7482702]
61. Roelfes G, Hilvert D. *Angew Chem Int Ed Engl*. 2003; 42:2275–2277. [PubMed: 12772161]
62. Zhong L, Holmgren A. *J Biol Chem*. 2000; 275:18121–18128. [PubMed: 10849437]

Figure 1.

The UGA codon in the *mRNA* can be decoded as a sense codon if a special stem-loop structure is present in the *mRNA*. In prokaryotes such as *E. coli*, this SECIS element is located proximal to the UGA codon, while in eukaryotes, it is located in the 3' untranslated region. There are structural differences as well, making it difficult to express a eukaryotic selenoprotein in a prokaryote such as *E. coli*.

Figure 2. HPLC trace of tripeptide CUG synthesized with the optimized conditions as reported by Barany [50]. Separation was achieved on a Waters C-18 column using a linear gradient of 0.1% TFA in CH_3CN and 0.1% aqueous TFA from 0% CH_3CN to 60% CH_3CN in 60 minutes. Only the *L*-peptide could be detected (peak 2). Some cyclized product is present (peak 1).

A

B

C

Figure 3.

(A) Mechanism of deprotection of Sec(pMob) by TMSOTf. (B) Oxidative deprotection of Sec(pMob) peptides by I_2 . (C) Possible mechanism of deprotection of a Sec(pMob) group by neighboring group participation. The Se atom can attack the adjacent disulfide resulting in the formation of a selenonium ion. Removal of the pMob group is then facilitated by the presence of scavengers in the cleavage cocktail.

Figure 4. Mechanism of intein-mediated peptide ligation with peptides containing N-terminal Sec residues. The fusion protein between the protein of interest and the intein exists in equilibrium between amide and thioester forms (**A–B**). Thiol mediated cleavage occurs when a small molecule thiol is added at high concentration to produce a thioester-tagged protein (**C**). Addition of a peptide containing a N-terminal Sec residue allows for attack on the thioester to produce a selenoester (**D**) which rapidly rearranges to the amide form (**E**).

Figure 5. Ribbon drawing of RNase A showing positions of disulfide bonds and mixed selenylsulfide bonds formed with semisynthetic RNase A that incorporates selenocysteine. RNase A contains 4 disulfide bonds that occur between Cys40-Cys95, Cys84-Cys26, Cys110-Cys58, and Cys72-Cys65. Cys110 and Cys95 have been replaced with selenocysteine to create unique variants containing selenium.

Figure 6. Strategy for creating a semisynthetic thioredoxin reductase. The wild-type enzyme has a C-terminal extension consisting of the tripeptide Cys-Sec-Gly (**A**). A semisynthetic enzyme can be made by producing a thioester-tagged enzyme (**B**) missing this tripeptide and then ligating this tripeptide *in vitro* to create the semisynthetic enzyme (**C**). Such a strategy divides the protein into two distinct modules, one with selenium and one without.

Figure 7. Michaelis-Menten plot of semisynthetic and mutant thioredoxin reductases. The selenium-containing semisynthetic enzyme has very high activity compared to the mutant enzyme containing cysteine.

Scheme 1.

Synthesis of the amino acid selenocystine from elemental selenium.

Scheme 2.

Rate of attack on a model thioester with cysteine or selenocysteine (X=Se,S). Cystine or selenocystine were first reduced with an equimolar amount of triscarboxyethylphosphine to produce a free thiol or selenol, respectively. The rate was then measured as a function of pH.

Table 1

Se/S selectivity for the reaction in Scheme 2.

pH	$k_{\text{obs}}(\text{Se}) \times 10^6 \text{ sec}^{-1}$	$k_{\text{obs}}(\text{S}) \times 10^6 \text{ sec}^{-1}$	$k_{\text{obs}}(\text{Se})/\text{S}$
5.0	2480.0	2.5	992
5.5	6570	15	438
6.0	9300	38	245
6.5	9000	39	231
7.0	9900	107	92
7.5	10000	320	31.5
8.0	9500	665	14.3

Table 2

Properties of RNase A and Se-variants

Enzyme	$a k_{\text{cat}}/K_m \text{ s}^{-1} \cdot \text{M}^{-1}$	$b \text{ m/z}$
wild-type	$1.13 \pm .06 \times 10^7$	13,820.5
C110U	$1.14 \pm .12 \times 10^7$	13,865.6
C95U	$1.01 \pm .08 \times 10^7$	not determined
C95U-C110U	$1.05 \pm .15 \times 10^7$	13,917.2

^aFluorescein-dArUdAdA-TAMARA is used as the substrate in this assay in 100 mM MES buffer, pH 6.0, 100 mM NaCl. The change in fluorescence is monitored over time.

^bDetermined by MALDI-TOF MS.