Selenocysteine in Native Chemical Ligation and Expressed Protein Ligation

Robert J. Hondal,[†] Bradley L. Nilsson,[‡] and Ronald T. Raines*[,†,‡]

Department of Biochemistry and Department of Chemistry, University of Wisconsin-Madison Madison, Wisconsin 53706

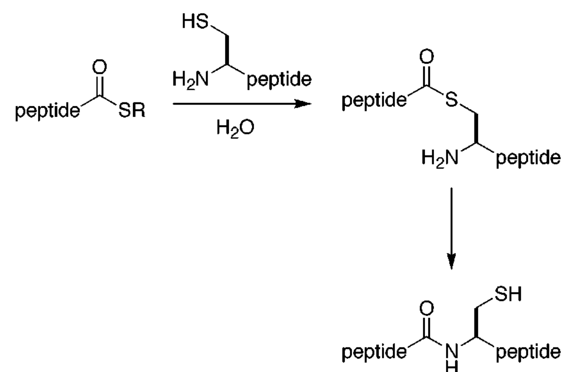
Received December 14, 2000

L-Selenocysteine (Sec or U) has been called the "21st amino acid".[1] Like the twenty common amino acids, selenocysteine is inserted during the translation of mRNA and has its own $tRNA^{Sec}$ and codon, UGA. This codon also serves as the *opal* stop codon. Decoding a UGA codon as one for selenocysteine requires a special structure in the 3′ untranslated region of the mRNA called a *se*lenocysteine *i*nsertion *s*equence (SECIS) element. Because eukaryotic and prokaryotic cells use a different SECIS element to decode UGA as selenocysteine, the production of eukaryotic selenocysteine-containing proteins in prokaryotes is problematic.[2] Here, we describe a general semisynthetic route to proteins containing selenocysteine.[3,4]

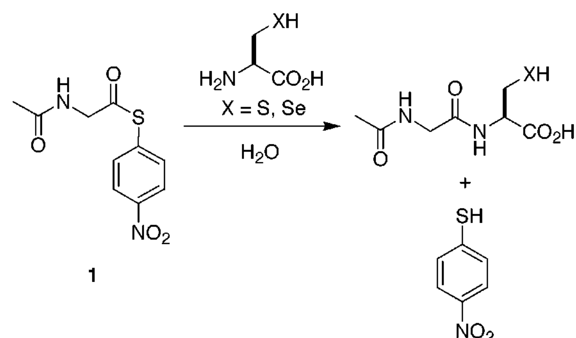
In "native chemical ligation", the thiolate of an N-terminal cysteine residue in one peptide attacks a C-terminal thioester in another peptide to produce, ultimately, an amide bond between the two peptides (Scheme 1).[5] "Expressed protein ligation" is an extension in which the C-terminal thioester is produced by using recombinant DNA (rDNA) technology.[6] We reasoned that selenocysteine, like cysteine, could effect both native chemical ligation and expressed protein ligation, and thereby provide a means to incorporate selenocysteine into proteins.

We used $AcGlySCH_2C(O)NHCH_3$ as a model thioester to test the feasibility of using selenocysteine in native chemical ligation.[7] Reaction with cystine ($(CysOH)_2$) in the presence of the reducing agent tris-(2-carboxyethyl)phosphine (TCEP) produced AcGly-CysOH, as well as some $(AcGlyCysOH)_2$. When selenocystine ($(SecOH)_2$) was used in the same reaction, the product was $(AcGlySecOH)_2$.[8]

Scheme 1



Scheme 2



A selenolate (RSe^-) is more nucleophilic than is its analogous thiolate (RS^-).[9] Moreover, the pK_a of a selenol (RSeH) is lower than that of its analogous thiol (RSH).[9a,10] These properties suggested to us that native chemical ligation with selenocysteine could be more rapid than with cysteine, especially at low pH. To test this hypothesis, we used the chromogenic thioester AcGly-SC_6H_4-*p*-NO_2 (**1**; Scheme 2) to determine the rate of native chemical ligation as a function of pH.[11] The resulting pH-rate profile is shown in Figure 1. Reaction with selenocysteine is 10^3-fold faster than with cysteine at pH 5.0. Thus, native chemical ligation with selenocysteine can be chemoselective.[12]

Having demonstrated the effectiveness of selenocysteine in native chemical ligation, we next set out to explore its utility in expressed protein ligation. As a model protein, we chose ribonuclease A (RNase A; EC 3.1.27.5; Figure 2), which has been the object of much seminal work in protein chemistry.[14] RNase A has 8 cysteine residues that form 4 disulfide bonds in the native

* To whom correspondence should be addressed. Telephone: (608) 262-8588. Fax: (608) 262-3453. E-mail: raines@biochem.wisc.edu.
† Department of Biochemistry.
‡ Department of Chemistry.

(1) (a) Bock, A.; Forchhammer, J.; Heider, W.; Leinfelder, G.; Veprek, B.; Zinoni, F. *Mol. Microbiol*. **1991**, *5*, 515−520. For other reviews, see: (b) Odom, J. D. *Structure Bonding* **1983**, *54*, 1−26. (c) Low, S. C.; Berry, M. J. *Trends Biochem. Sci.* **1996**, *21*, 203−208. (d) Stadtman, T. C. *Annu. Rev. Biochem.* **1996**, *65*, 83−100.

(2) Arner, E. S. Sarioglu, H.; Lottspeich, F.; Holmgren, A.; Bock, A. *J. Mol. Biol.* **1999**, *292*, 1003−1016.

(3) For other means to incorporate selenocysteine into semisynthetic and synthetic proteins, see: (a) Wu, Z. P.; Hilvert, D. *J. Am. Chem. Soc.* **1989**, *111*, 4513−4514. (b) Fiori, S.; Pegoraro, S.; Rudolph-Bohner, S.; Cramer, J.; Moroder, L. *Biopolymers* **2000**, *53*, 550−564.

(4) For a means to produce selenopeptide libraries on the surface of M13 bacteriophage, see: Sandman, K. E.; Benner, J. S.; Noren, C. J. *J. Am. Chem. Soc.* **2000**, *122*, 960−961.

(5) (a) Wieland, T.; Bokelmann, E.; Bauer, L.; Lang, H. U.; Lau, H. *Liebigs Ann. Chem.* **1953**, *583*, 129−149. (b) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776−779. For a recent review, see: (c) Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923−960.

(6) (a) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705−6710. For reviews, see: (b) Cotton, G. J.; Muir, T. W. *Chem. Biol.* **1999**, *6*, R247−R256. (c) Evans, T. C., Jr.; Xu, M.-Q. *Biopolymers* **2000**, *51*, 333−342.

(7) $AcGlySCH_2C(O)NHCH_3$ was synthesized as described by Nilsson B. L.; Kiessling, L. L.; Raines, R. T. *Org. Lett.* **2000**, *2*, 1939−1941. We find *N*-methylmercaptoacetamide to be a superb thiol for effecting both native chemical ligation and expressed protein ligation.

(8) L-Selenocystine was synthesized from elemental selenium and β-chloro-L-alanine as described by Tanaka, H.; Soda, K. *Methods Enzymol.* **1987**, *143*, 240−243. Ligation reactions were performed at (23 ± 2) °C under Ar(g) in 0.10 M Tris-HCl buffer (pH 8.0) containing $AcGlySCH_2C(O)NHCH_3$ (25 mM), TCEP (25 mM), and cystine or selenocystine (12.5 mM).

(9) (a) Huber, R. E.; Criddle, R. S. *Arch. Biochem. Biophys.* **1967**, *122*, 164−173. (b) Pearson, R. G.; Sobel, H.; Songstad, J. *J. Am. Chem. Soc.* **1968**, *90*, 319−326. (c) Pleasants, J. C.; Guo, W.; Rabenstein, D. L. *J. Am. Chem. Soc.* **1989**, *111*, 6553−6558. (d) Singh, R.; Whitesides, G. M. *J. Org. Chem.* **1991**, *56*, 6931−6933. (e) Gorlatov, S. N.; Stadtman, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8520−8525. (f) Lee, S.-R.; Bary-Noy, S.; Kwon, J.; Levine, R. L.; Stadtman, T. C.; Rhee, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2521−2526. (g) Zhong, L.; Holmgren, A. *J. Biol. Chem.* **2000**, *275*, 18121−18128.

(10) Arnold, A. P.; Tan, K. S.; Rabenstein, D. L. *Inorg. Chem.* **1986**, *25*, 2433−2437.

(11) Ligation assays were performed at (23 ± 2) °C under Ar(g) in 0.10 M buffer containing thioester **1** (10 μM), TCEP (0.10 mM), and cystine (20 μM-0.50 mM) or selenocystine (1.0−20 μM). Rate = $\partial[HSC_6H_4\text{-}p\text{-}NO_2]/\partial t = k_{obs}[\text{CysOH or SecOH}] - k_{H_2O}$ was determined by using $\epsilon = 11,230\ M^{-1}\ cm^{-1}$ for *p*-nitrothiophenolate at 410 nm.

(12) The template-directed ligation of a phosphoroselenol and iodoribose is approximately 4-fold faster than that of an analogous phosphorothiol at pH 7.0 (Xu, Y.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, 9040−9041). We observe a similar difference in Se vs S reactivity at high pH (Figure 1).

(13) Danehy, J. P.; Parameswaran, K. N. *J. Chem. Eng. Data* **1968**, *13*, 386−389.

(14) Raines, R. T. *Chem. Rev.* **1998**, *98*, 1045−1065.

10.1021/ja005885t CCC: $20.00 © 2001 American Chemical Society
Published on Web 05/04/2001

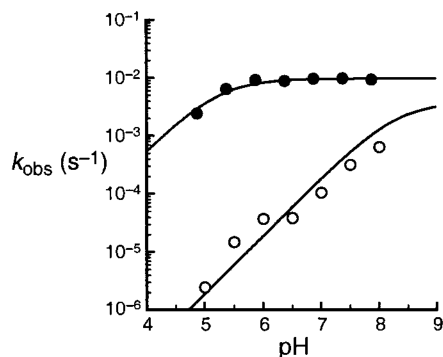


Figure 1. pH-Rate profile for the reaction in Scheme 2 with CysOH (○) and SecOH (●).[11] Data were fitted to the equation: $k_{obs} = k[\mathbf{1}]/(1 + [H^+]/K_a)$ with $pK_a^{Cys} = 8.30$[13] and $pK_a^{Sec} = 5.24$,[8a] yielding $k = (3.7 \times 10^2)$ M^{-1} s^{-1} for CysOH and $k = (9.5 \times 10^2)$ M^{-1} s^{-1} for SecOH.

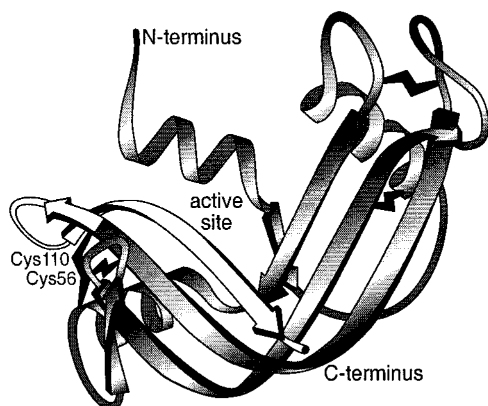


Figure 2. Ribbon diagram of the three-dimensional structure of ribonuclease A.[14] C-Terminal residues 110−124 (white) can derive from a synthetic peptide via expressed protein ligation.

enzyme. Of these 8 cysteine residues, Cys110 is closest to the C-terminus.

To determine whether selenocysteine can effect expressed protein ligation, we used rDNA technology to prepare a fragment corresponding to residues 1−109 with a C-terminal thioester.[15] We used standard solid-phase methods to synthesize a peptide corresponding to residues 110−124 with a cysteine or a selenocysteine residue[16] at position 110. We then ligated the thioester fragment and a peptide fragment, and folded and purified the ligation product. Semisynthetic wild-type RNase A has *m/z* 13 819, and C110U RNase A has *m/z* 13 865. The $\Delta m/z$ of 46 is in gratifying agreement with the $\Delta M_r = 47$ expected for the replacement of sulfur with selenium.

The gain of function provides another indication of a successful ligation. The fragment corresponding to residues 1−109 has no detectable ribonucleolytic activity.[17,18] In contrast, semisynthetic wild-type RNase A and C110U RNase A have equivalent ribonucleolytic activities [$k_{cat}/K_m = (1.13 \pm 0.06) \times 10^7$ M^{-1} s^{-1} and $(1.1 \pm 0.1) \times 10^7$ M^{-1} s^{-1}, respectively],[19] indicating that C110U RNase A is not only intact but also folded properly (Figure 2). We conclude that the isomorphous replacement of sulfur with selenium can be effected with expressed protein ligation.

Incorporation of selenocysteine residues into proteins could have much utility. For example, proteins containing selenosulfide (Se−S) or diselenide (Se−Se) bonds could have high conformational stability in a reducing environment (such as the cytosol), as the reduction potential of selenosulfide and diselenide bonds is less than that of the corresponding disulfide bond.[19] In addition, incorporation of selenocysteine could be used to reveal structure−function relationships with ^{77}Se NMR spectroscopy[20] or determine the phase of protein crystals.[21] Finally, chemoselective alkylation of the side chain of a selenocysteine residue[9] or reduction of the C−Se bond to create alanine[22] would expand the scope of protein semisynthesis.

Many natural proteins contain selenocysteine residues. For example, mammalian thioredoxin reductase uses a selenosulfide bond in its catalytic mechanism.[9e] Thiol:disulfide oxidoreductases have a CXXC motif in their active site.[23] Likewise, selenoprotein W has a CXXU motif, and selenoprotein P has a UXXC motif as well as 16 other cysteine and 9 other selenocysteine residues.[24] The role of selenocysteine in the function of selenoprotein W or selenoprotein P is unknown. The methods described herein make these and other selenocysteine-containing proteins readily amenable to detailed structure−function analyses.

Acknowledgment. We thank H. J. Reich and C. Park for advice, and T. C. Evans, Jr. for plasmid pTXB1-RNase. R.J.H. was supported by Postdoctoral Fellowship GM20180 (NIH). This work was supported by Grant GM44783 (NIH).

Note Added in Proof. Subsequent to the submission of this manuscript, van der Donk and co-workers reported the synthesis of a selenocysteine-containing peptide by native chemical ligation. Gieselman, M. D.; Xie, L.; van der Donk, W. A. *Org. Lett.* **2001**, *3*, 1331−1334.

Supporting Information Available: Procedures for all syntheses and analyses reported herein (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA005885T

(15) This fragment was produced with plasmid pTXB1-RNase, which directs the production in *Escherichia coli* of the *Mxe* intein fused to residues 1−109 of RNase A plus an *N*-terminal methionine. Evans, T. C., Jr.; Benner, J.; Xu, M.-Q. *Protein Sci.* **1998**, *7*, 2256−2264.

(16) FmocSec(PMB)OH was synthesized essentially as described by Koide, T.; Itoh, H.; Otaka, A.; Yasui, H.; Kuroda, M.; Esaki, N.; Soda, K.; Fujii, N. *Chem. Pharm. Bull.* **1993**, *41*, 502−506.

(17) This fragment lacks His119 and the Cys58−Cys110 disulfide bond. For their roles, see: (a) Thompson, J. E.; Raines, R. T. *J. Am. Chem. Soc.* **1994**, *116*, 5467−5468. (b) Klink, T. A.; Woycechowsky, K. J.; Taylor, K. M.; Raines, R. T. *Eur. J. Biochem.* **2000**, *267*, 566−572.

(18) Catalysis was assessed by using the fluorogenic substrate 6-carboxyfluorescein∼dArU(dA)$_2$∼6-carboxytetramethylrhodamine as described by Kelemen, B. R.; Klink, T. A.; Behlke, M. A.; Eubanks, S. R.; Leland, P. A.; Raines, R. T. *Nucleic Acids Res.* **1999**, *18*, 3696−3701.

(19) (a) Besse, D.; Siedler, F.; Diercks, T.; Kessler, H.; Moroder, L. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 883−885. (b) Pegoraro, S.; Fiori, S.; Rudolph-Bohner, S.; Watanabe, T. X.; Moroder, L. *J. Mol. Biol.* **1998**, *284*, 779−792.

(20) For an example, see: House, K. L.; Dunlap, R. B.; Odom, J. D.; Wu, Z. P.; Hilvert, D. *J. Am. Chem. Soc.* **1992**, *114*, 8573−8579.

(21) For examples with selenomethionine, see: (a) Hendrickson, W. A.; Horton, J. R.; LeMaster, D. M. *EMBO J.* **1990**, *9*, 1665−1672. (b) Doublie, S. *Methods Enzymol.* **1997**, *276*, 523−530. (c) Shao, W.; Fernandez, E.; Wilken, J.; Thompson, D. A.; Siani, M. A.; West, J.; Lolis, E.; Schweitzer, B. I. *FEBS Lett.* **1998**, *441*, 77−82.

(22) Possible routes involve photochemical, tin-hydride, or hydrogenolytic reduction. For examples, see: (a) Pandey, G.; Soma Sekhar, B. B. V.; Bhalerao, U. T. *J. Am. Chem. Soc.* **1990**, *112*, 5650−5651. (b) Light, J.; Breslow, R. *Org. Synth.* **1995**, *72*, 199−208. (c) Yan L. Z.; Dawson P. E. *J. Am. Chem. Soc.* **2001**, *123*, 526−533.

(23) Chivers, P. T.; Prehoda, K. E.; Raines, R. T. *Biochemistry* **1997**, *36*, 4061−4066.

(24) Burk, R. F.; Hill, K. E. *Bioessays* **1999**, *21*, 231−237.