Probing Protein Folding

Incorporation of Selenomethionine into Proteins through Selenohomocysteine-Mediated Ligation**

*Gerard Roelfes and Donald Hilvert**

The chemical synthesis of moderate-sized proteins has benefited enormously from the development of chemoselective methods for the ligation of peptide fragments.[1] In this regard, the reaction of peptide thioesters with peptides containing an N-terminal cysteine has proved particularly powerful.[2] The thiol side chain of cysteine attacks the thioester and subsequent acyl transfer from sulfur to the amine yields a native peptide bond. Homocysteine,[3] sulfur-containing auxiliary groups,[4] and selenocysteine[5] have also been exploited to mediate ligations to thioesters, thus broadening the scope of the method. Herein we show that selenohomocysteine can also facilitate efficient peptide ligation (Scheme 1 a). Subsequent methylation of the resulting selenol affords selenomethionine at the ligation site, which is a useful probe for NMR spectroscopic[6] and X-ray crystallographic[7] investigations of biomolecules.

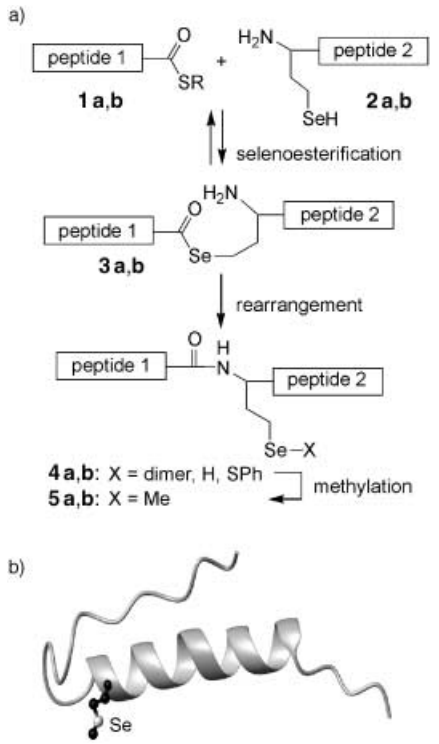
The strategy employed is illustrated by the model ligation of LYRAG-SEt with selenohomocystine. This reaction was carried out at a peptide concentration of 1 mM in Tris-HCl buffer (100 mM, pH 8.5; Tris = tris(hydroxymethyl)aminomethane) in the presence of guanidinium chloride (GdmCl; 6 M) and thiophenol (5 % v/v). Thiophenol helps to activate the thioester and also reduces the diselenide to the corresponding selenol. At pH 7–8, more than 90 % conversion was observed after 4 h, and the reaction was essentially complete within 20 h. Comparable results were obtained with homocysteine in control experiments.[5] The ligation product was isolated as a variable mixture of diselenide, a mixed selenosulfide (with thiophenol), and a species with the mass of the free selenol that was slowly converted into the diselenide upon exposure to air. The combined yield was 98 %. The reaction was somewhat slower in a phosphate buffer but also complete within 24 h. The addition of tris(2-carboxyethyl)phosphane hydrochloride (TCEP) to reduce the diselenide did not increase the rate of reaction, but did lead to the formation

[*] Prof. Dr. D. Hilvert, Dr. G. Roelfes
Laboratorium für Organische Chemie
Swiss Federal Institute of Technology (ETH)
ETH-Hönggerberg, 8093 Zürich (Switzerland)
Fax: (+41) 1-632-1486
E-mail: Hilvert@org.chem.ethz.ch

[**] We thank Brigitte Brandenberg and Bernhard Jaun for assistance with the NMR spectra, Richard Thomas for the analytical ultracentrifugation analysis, and the ETH Zürich and Novartis Pharma for financial support. G.R. wishes to thank the Netherlands Organisation for Scientific Research (NWO) for a Talent Fellowship.

Supporting Information for this article (including detailed experimental procedures, representative HPLC traces, a GdmCl titration curve, and analytical ultracentrifugation and CD data) is available on the WWW under http://www.angewandte.org or from the author.

 DOI: 10.1002/ange.200250830



Scheme 1. a) Schematic representation of a selenohomocysteine-ligation–methylation strategy. Model reaction: peptide 1 = LYRAG, peptide 2 = -CO_2H; seleno-bPP: peptide 1 = bPP(1-16), APLEPEYPGDNATPEQ-SEt, peptide 2 = bPP(18-36), -AQYAAELRRYINMLTRPRY-NH_2; b) MOLMOL[17] representation of the proposed structure of monomeric seleno-bPP based on the NMR structure of bPP from Ref. [12], which shows the position of the selenomethionine within the typical PP fold.

of more side products.[8] The diselenide was methylated under reducing conditions with methyl 4-nitrobenzenesulfonate, which has been successfully employed in the conversion of homocysteine into methionine.[3] Complete conversion was observed within 30 minutes in the presence of an excess of the methylating agent (2.5–10 equivalents based on the monomeric peptide) and TCEP in a 1:1 mixture of phosphate (200 mM, pH 6.0) and acetonitrile. The pure selenomethionine-containing hexapeptide was obtained from the reaction in 66% yield. Larger excesses of the methylating agent, higher pH values, or longer reaction times resulted in a lower yield and gave rise to numerous side products. That side reactions can be minimized by performing the reaction at a relatively low pH value, because of the low pK_a values of selenols, is an advantage over the methylation of homocysteine.

The utility of the selenohomocysteine ligation strategy was demonstrated by the synthesis of seleno-bPP, a variant of bovine pancreatic polypeptide (bPP) that contains selenomethionine at position 17 as a spectroscopic probe for folding studies. bPP is a member of the neuropeptide Y (NPY) family,[9] involved in the regulation of food uptake.[10] It is 36 amino acids long, has methionine groups at positions 17 and 30, and has a C-terminal carboxamide group. At high concentrations bPP is homodimeric[11] and each monomer adopts the so-called pancreatic-polypeptide (PP) fold (Scheme 1b), which consists of a polyproline helix docked onto the hydrophobic face of an α helix.[12]

Seleno-bPP was synthesized as shown in Scheme 1a. The bPP(1-16) fragment activated as a C-terminal thioester **1b** and bPP(17-36) containing an N-terminal selenohomocysteine **2b** were assembled on a safety-catch resin and Rink-amide resin, respectively, using standard Fmoc chemistry with HBTU/HOBt activation (Fmoc = 9-fluorenylmethoxycarbonyl, HBTU = *O*-(benzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate, HOBt = 1-hydroxy-1*H*-benzotriazole). Compound **2b** was stored as a diselenide and was regenerated in situ by reduction with excess thiophenol in the standard ligation buffer (Tris-HCl (100 mM), GdmCl (6M), pH 8.5). As expected for ligations involving glutamine thioesters,[13] the reaction with **1b** was considerably slower than the model ligation with LYRAG-SEt. However, full conversion could be achieved when a fourfold excess of the thioester was employed and the reaction was allowed to proceed for at least five days. This procedure yielded the expected ligation product as the diselenide and a mixed selenosulfide (with thiophenol) in a combined yield of 59% after purification. Both compounds were methylated independently with a tenfold excess of methyl 4-nitrobenzenesulfonate under the conditions described above to give the selenomethionine-containing bPP derivative in 78% yield. The identity of the final product was confirmed by electrospray-ionization mass spectroscopy (ESI-MS; found: m/z 4272.2, calcd for average isotope distribution: m/z 4273.6).

Seleno-bPP is homodimeric as determined by conventional sedimentation-equilibrium analytical ultracentrifugation,[14] and its circular dichroism (CD) spectrum is identical to that reported for native bPP.[15] It undergoes reversible unfolding in the presence of guanidinium chloride.[14] The denaturation was monitored by CD, which showed a broad transition region. Fitting the data to a two-state model gave $\Delta G(H_2O) = 2.67 \pm 0.08$ kcal mol^{-1}, $m = 705$ cal mol^{-1}M^{-1}, and $[GdmCl]_{1/2} = 3.8$ M.

The unfolding of seleno-bPP was monitored as a function of GdmCl concentration by ^{77}Se NMR spectroscopy, which revealed equilibrium details that were not observable by CD spectroscopy (Figure 1). Relatively high protein concentrations (10.5–12.5 mM) were required as a result of the low natural abundance of ^{77}Se (7.58%). Nevertheless, in the absence of denaturant, the ^{77}Se NMR spectrum of the protein exhibits a single signal at $\delta = 46$ ppm with a linewidth of 101 Hz. In GdmCl (6M) a cluster of peaks in the vicinity of the signal for free selenomethionine in solution ($\delta = 76.8$ ppm under the same conditions) is observed. Apparently, the denatured state in this case consists of several discrete conformations that are stable on the NMR time scale. At intermediate GdmCl concentrations covering the transition region of the CD unfolding curve, two very broad peaks are evident, which indicate the presence of multiple species that interconvert only slowly. As the GdmCl concentration is increased, the intensity of the signal at $\delta \approx 70$ ppm increases relative to that at higher field and can thus be confidently assigned to the denatured protein. This signal is also subject to a modest downfield shift, which can be ascribed to the presence of GdmCl, as it parallels the $\delta = 4.5$ ppm downfield

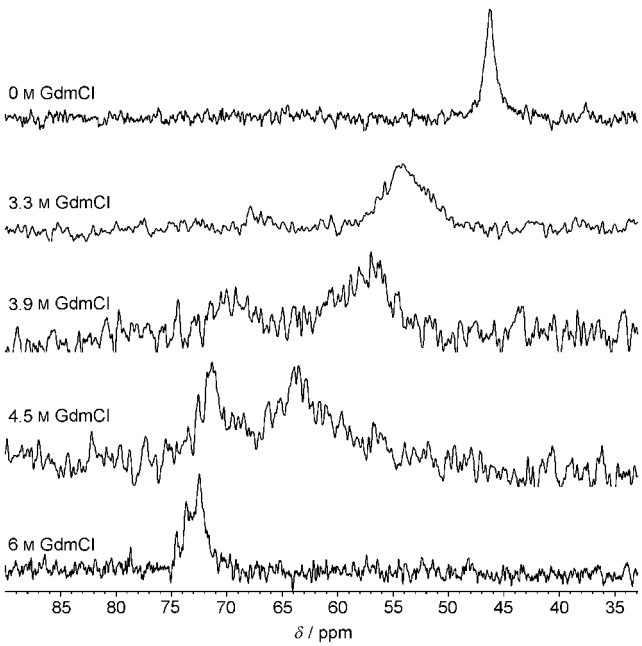


Figure 1. ^{77}Se NMR spectra of seleno-bPP at different GdmCl concentrations. Samples were measured in a solution of acetate in D_2O (150 mM, pH 4.2; uncorrected for the D effect) at 25 °C. Spectra were recorded at 95 MHz and the chemical shifts were referenced externally against Me_2Se (60% v/v) in $CHCl_3$. Chemical shifts were not corrected for the medium effect of GdmCl.

shift seen for free selenomethionine as the GdmCl concentration is increased from 0 to 6 M. The higher-field signal in the partially denatured samples is less readily assigned. However, based on the observed line broadening and chemical shifts it presumably represents an intermediate state that exhibits considerable dynamic behavior and is possibly in rapid equilibrium with the native protein.

A possible explanation for these observations is shown in Figure 2. The unfolding of the native dimeric protein (A) proceeds via an intermediate that further unfolds to give the denatured state (D). The intermediate could either be a monomeric species (B), or perhaps more likely, a dimer (C) in which the polyproline helices are no longer backfolded on the α helix, analogous to the structure of NPY.[16] A dynamic mixture of several intermediates is also possible.

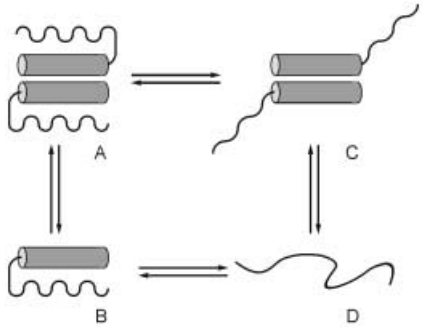


Figure 2. Schematic representation of the proposed unfolding mechanism of seleno-bPP.

In summary, selenohomocysteine-mediated ligation of peptides, followed by methylation of the resulting selenol, provides ready access to selenomethionine-containing peptides. The advantage of this strategy over the commonly used biosynthetic approach, with methionine auxotrophic strains supplemented with selenomethionine, is that it allows the site-selective introduction of a single selenomethionine into proteins in the presence of other methionines. As shown with seleno-bPP, this residue can provide unique information about the structure and dynamics of protein molecules.

Received: December 23, 2002 [Z50830]

Keywords: bioorganic chemistry · ligation · peptides · protein folding · selenium

[1] a) P. E. Dawson, S. B. H. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923–960; b) J. P. Tam, J. Xu, K. D. Eom, *Biopolymers* **2001**, *60*, 194–205; c) R. M. Hofman, T. W. Muir, *Curr. Opin. Biotechnol.* **2002**, *13*, 297–303.

[2] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776–779; b) T. Wieland, E. Bokelmann, L. Bauer, H. U. Lang, H. Lau, *Justus Liebigs Ann. Chem.* **1953**, *583*, 129–149.

[3] J. P. Tam, Q. Yu, *Biopolymers* **1998**, *46*, 319–327.

[4] a) D. W. Low, M. G. Hill, M. R. Carrasco, S. B. H. Kent, P. Botti, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6554–6559; b) J. Offer, C. N. C. Boddy, P. E. Dawson, *J. Am. Chem. Soc.* **2002**, *124*, 4642–4646.

[5] a) R. Quaderer, A. Sewing, D. Hilvert, *Helv. Chim. Acta* **2001**, *84*, 1197–1206; b) R. J. Hondal, B. L. Nilsson, R. T. Raines, *J. Am. Chem. Soc.* **2001**, *123*, 5140–5141; c) M. D. Gieselman, L. Xie, W. A. van der Donk, *Org. Lett.* **2001**, *3*, 1331–1334.

[6] a) N. P. Luthra, J. O. Odom in *The Chemistry of Organic Selenium and Tellurium Compounds, Vol. 1* (Eds.: S. Patai, Z. Rappaport), Wiley, Weinheim, **1986**, pp. 189–241; b) J. O. Boles, W. H. Tolleson, J. C. Schmidt, R. B. Dunlap, J. D. Odom, *J. Biol. Chem.* **1992**, *267*, 22217–22223.

[7] W. A. Hendrickson, J. R. Horton, D. M. LeMaster, *EMBO J.* **1990**, *9*, 1665–1672.

[8] The main side product (up to 18% yield) was the deselenated ligation product. Similar observations have been made in other selenium-mediated ligations.[5c]

[9] A. Bettio, A. G. Beck-Sickinger, *Biopolymers* **2001**, *60*, 420–437.

[10] G. Katsuura, A. Asakawa, A. Inui, *Peptides* **2002**, *23*, 323–329.

[11] M. Lerch, V. Gafner, R. Bader, B. Christen, G. Folkers, O. Zerbe, *J. Mol. Biol.* **2002**, *322*, 1117–1133.

[12] X. Li, M. J. Sutcliffe, T. W. Schwartz, C. M. Dobson, *Biochemistry* **1992**, *31*, 1245–1253.

[13] T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068–10073.

[14] See Supporting Information.

[15] M. E. Noelken, P. J. Chang, J. R. Kimmel, *Biochemistry* **1980**, *19*, 1838–1843.

[16] a) R. Bader, A. Bettio, A. G. Beck-Sickinger, O. Zerbe, *J. Mol. Biol.* **2001**, *305*, 307–329; b) A. Bettio, M. C. Dinger, A. G. Beck-Sickinger, *Protein Sci.* **2002**, *11*, 1834–1844.

[17] R. Koradi, M. Billeter, K. Wüthrich, *J. Mol. Graphics* **1996**, *14*, 51–55.

