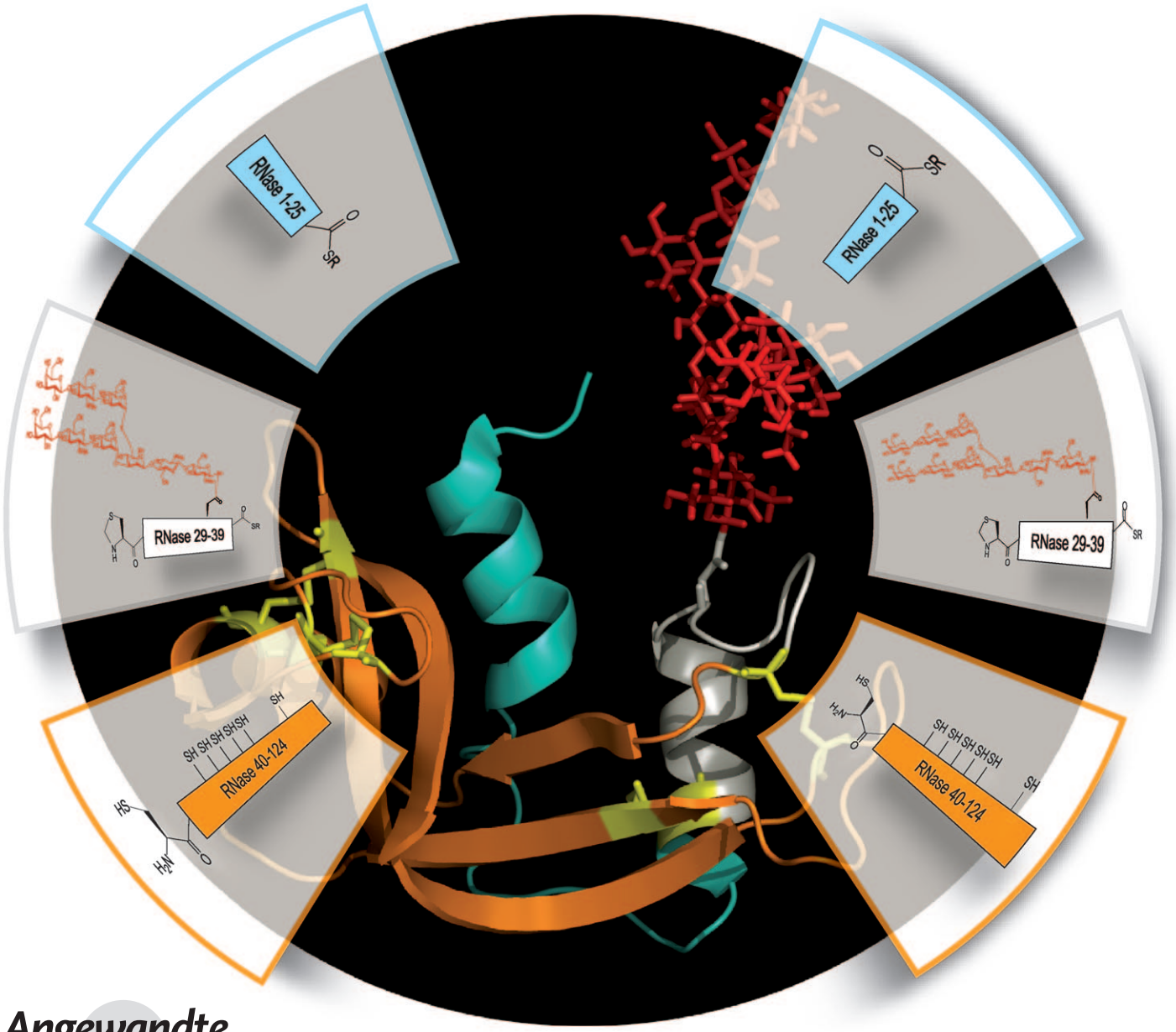
Glycosylation (1)

DOI: 10.1002/anie.200804734

Semisynthesis of a Homogeneous Glycoprotein Enzyme: Ribonuclease C: Part 1**

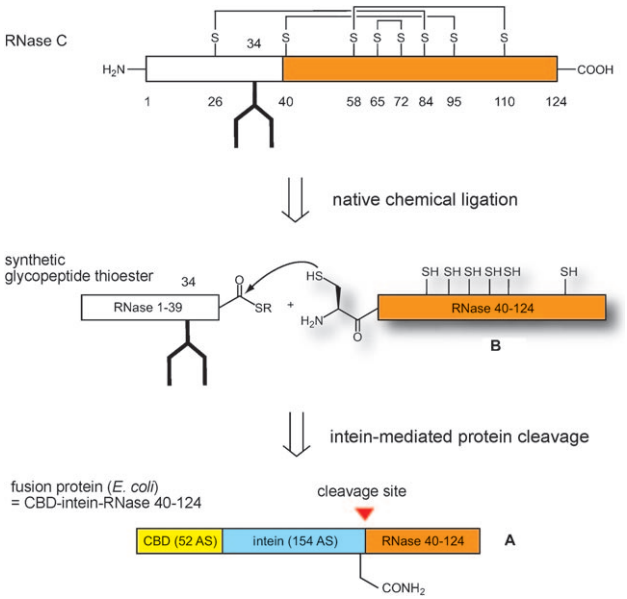
*Christian Piontek, Petra Ring, Olaf Harjes, Christian Heinlein, Stefano Mezzato, Nelson Lombana, Claudia Pöhner, Markus Püttner, Daniel Varón Silva, Andreas Martin, Franz Xaver Schmid, and Carlo Unverzagt**

Dedicated to Professor Chi-Huey Wong on the occasion of his 60th birthday



With the advent of chemoselective coupling methods, the total synthesis of proteins has advanced considerably.[1] Of particular importance is native chemical ligation (NCL),[2] which creates a native peptide bond at the ligation site. NCL is based on the selective coupling of a peptide thioester with a peptide containing an N-terminal cysteine. The ligation of recombinant fragments is termed "expressed protein ligation" (EPL).[3] Up to now recombinant peptide thioesters can be obtained only by the thiolytic cleavage of an intein.[3b] Protein fragments with an N-terminal cysteine are mainly liberated from precursors by specific proteases[3f–i] or BrCN[3j] cleavage. An intein-based method is commercially available; however, only a few examples have been described.[4] The synthesis of homogeneously glycosylated therapeutic glycoproteins is of great interest since the homogeneous natural glycoforms required for structure–activity studies appear to be available mainly by peptide ligation.[5] Pancreatic ribonucleases (RNases) are well-established reference systems in protein synthesis.[6a–c] They occur as an unglycosylated form (Rnase A) and as differently glycosylated forms (RNases B, C, and D). RNase B contains an oligomannosidic and RNase C a complex-type N-glycan.[6] We selected bovine ribonuclease C (RNase C)[6d] as a model for glycoprotein semisynthesis since many therapeutic glycoproteins (e.g. hormones, enzymes, antibodies) also contain several disulfides and complex-type N-glycans. A seminal study has been conducted with the different glycoforms pancreatic RNase B.[6e] In the course of our studies we developed a recombinant method for the production of chemically stabilized protein fragments. We found that the use of mixed disulfides greatly facilitates the isolation of thiol-rich protein fragments with an N-terminal cysteine.

Bovine ribonuclease (124 amino acids) contains one N-glycosylation site at Asn34 and eight cysteines engaged in four disulfide bridges. We initially planned to synthesize RNase by NCL from the fragment 40–124 (**B**) and the glycopeptide thioester 1–39 (**C**; Scheme 1), which should be obtained from a PEGA-double linker resin.[7] For RNase fragment **B** a recombinant approach was envisioned allowing the rapid production of the Cys-fragment **B**. Since inteins are very valuable for the generation of thioesters for EPL we attempted to employ inteins also for fragment **B**. The C-terminal protease activity of inteins is pH-dependant and based on the formation of a succinimide from a C-terminal asparagine.[4b] It was envisioned to express RNase fragment **B** as part of the fusion protein **A** containing an *Ssp* DnaB intein and a chitin-binding domain (CBD), which serves as an affinity tag for purification.



Scheme 1. Retrosynthesis of homogeneously glycosylated RNase C.

The commercially available pTWIN1 vector[8] was chosen as the expression vector. The DNA gene fragment of RNase 40-124 **B** was synthesized by PCR from five synthetic oligonucleotides, which were optimized in codon usage for expression in *E. coli* (see Scheme S1 in the Supporting Information). The amplicon was cloned into the vector and transfected into *E. coli* K12 (BER2566). Overexpression of the fusion protein was induced by isopropylthiogalactoside (IPTG, 3 h, 36 °C), resulting in the formation of insoluble inclusion bodies (Figure 1a) in which the intein domain was inactive.[4c] Since the inclusion bodies could be purified readily[9] (Figure 1b), the use of a chitin column for affinity purification was omitted. In order to cleave the target peptide **B** after refolding of the CBD-intein domain, the fusion protein **A** was first dissolved in 6 M guanidine hydrochloride (GdmCl; pH 8, 5 mM TCEP (tris(2-carboethoxy)phosphine)[10]) under argon followed by slow dilution with buffers (pH 7.4) containing additives promoting refolding. Rapid dilution or dialysis[4d] led to precipitation. Subsequently the refolding solution was kept at pH 6.9 for 24 h. The highest cleavage was found (Figure 1c) when 0.5 M L-arginine[9] and 5 mM TCEP were added to the refolding buffer, with significant cleavage occurring only after dilution below 2 M GdmCl. The desired RNase 40-124 **B** (9 kDa) showed only a weak band on the SDS gel and could neither be isolated after reduction nor detected by HPLC-MS analysis. Fragment **B** contains seven cysteines, which are susceptible to oxidation and complicate isolation. As a solution for the difficulties, a selective protection of the free thiols was envisioned, for example, as mixed disulfides[11] or *S*-sulfonates.[12] It was examined if, after the formation of mixed disulfides on the fusion protein, the refolding and the cleavage of the intein is still functional. Thus, the reduced fusion protein was reacted with glutathione (GSSG, pH 9.3)[9] and refolded by slow

[*] Dr. C. Piontek, P. Ring, O. Harjes, C. Heinlein, Dr. S. Mezzato, N. Lombana, C. Pöhner, Dr. M. Püttner, Dr. D. Varón Silva, Dr. A. Martin, Prof. F. X. Schmid, Prof. C. Unverzagt
Bioorganische Chemie, Gebäude NW1
Universität Bayreuth, 95440 Bayreuth (Germany)
Fax: (+49) 921-555-365
E-mail: carlo.unverzagt@uni-bayreuth.de

[**] This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Deutschen Chemischen Industrie.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200804734.

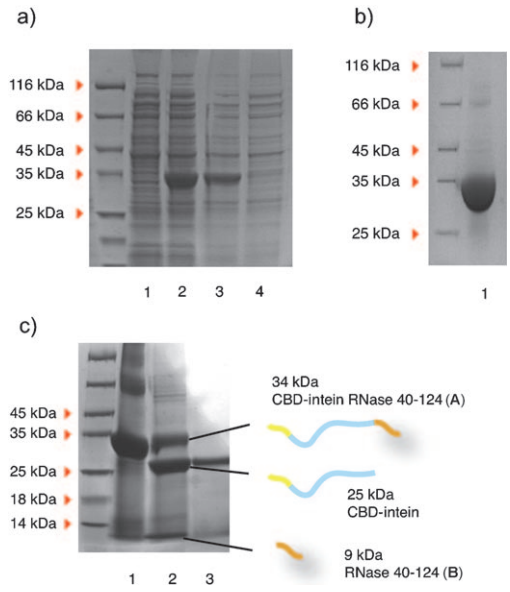
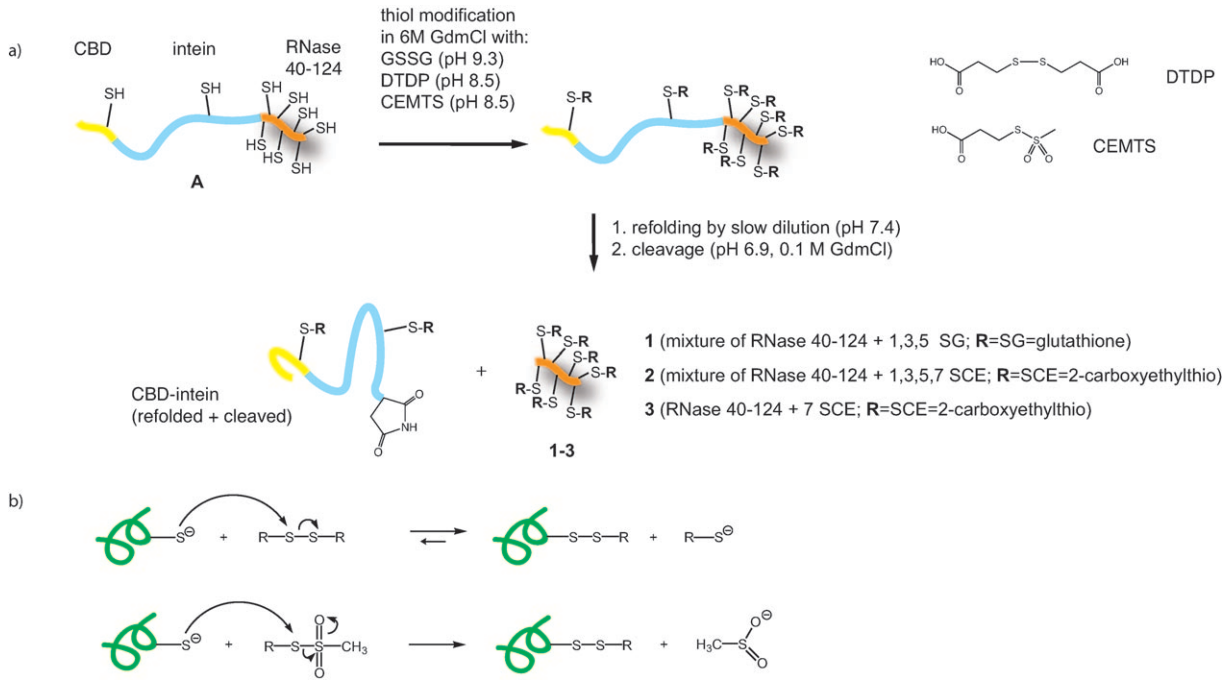


Figure 1. SDS-PAGE. a) Lane 1: before induction, lane 2: after induction, lane 3: insoluble proteins after lysis, lane 4: soluble proteins after lysis; b) lane 1: purified inclusion bodies; c) refolding and cleavage after dilution to: 2 M GdmCl (lane 1), 0.5 M GdmCl (lane 2), 0.1 M GdmCl (lane 3).

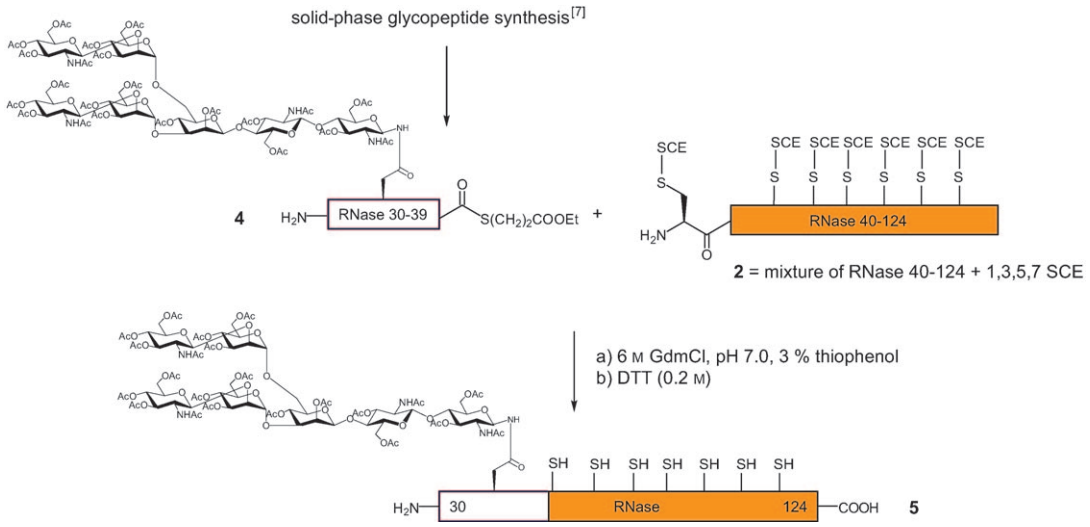
dilution without intermediate purification (Scheme 2a). After cleavage of the intein (2 d, 50% cleavage efficiency according to SDS-PAGE) the reaction mixture was concentrated and purified by RP-HPLC methods (yield, 2.1 mg of **1** per liter of culture). However, HPLC-MS analysis revealed that a mixture of one, three, and five modifications with glutathione was obtained. Also the use of the small disulfide reagent 3,3′-dithiodipropionic acid (DTDP, 50 mM) gave only a mixture of fragment **B** with one to seven disulfides (**2**), and cleavage of the fusion protein reached only 50%. The introduction of mixed disulfides by disulfide exchange is an equilibrium reaction and does not go to completion. In comparison, the reaction with thiosulfonates is much faster and nearly irreversible since no thiols are liberated (Scheme 2b).[13] After derivatization with 2-carboxyethylmethanethiosulfonate (CEMTS, 50 mM, Scheme S2) a cleavage efficiency of 80–90% was observed (Scheme 4) with all seven cysteines converted into mixed disulfides (Figure 2c) in the purified fragment **3** (Schemes S2 and S4 in the Supporting Information). In addition, fragment **3** provided better yields than **2** (3.5 mg of **3** per liter of culture vs. 2.8 mg of **2**). Fragment **3** proved to be resistant to oxidation and can be stored in dry state over several months. In an alternative approach, **3** was obtained after disulfide formation of the partially refolded fusion protein (1M GdmCl).

To assess the reactivity of the RNase fragments **1**–**3** in native chemical ligations, an analytical reaction of compound **2** was conducted with the glycopeptide thioester RNase 30–39 (**4**)[7] (1.5 equiv) in 6 M GdmCl in the presence of 3% of thiophenol (Scheme 3). The mixed disulfides of **2** were reduced in situ; however, after prolonged reaction times (2 d) the reaction mixture became turbid. After reduction with 0.2 M dithiothreitol (DTT) the reaction mixture cleared up, and the ligation product RNase 30–124 (**5**) could be detected by HPLC-MS methods (Scheme S5 in the Supporting Information).

The reaction conditions for the ligation to full-length RNase[14] were optimized with the recombinant thioester Met-RNase 1–39 (**6**) Scheme 4, obtained from a synthetic gene using the intein method and the pTWIN1 vector[8]) and



Scheme 2. a) Conversion of SH groups of the fusion protein **A** into mixed disulfides with subsequent refolding and cleavage of the intein. b) Mechanisms for the formation of mixed disulfides.



Scheme 3. Ligation of the protected RNase 40–124 fragment **2** with the synthetic glycopeptide thioester **4**.

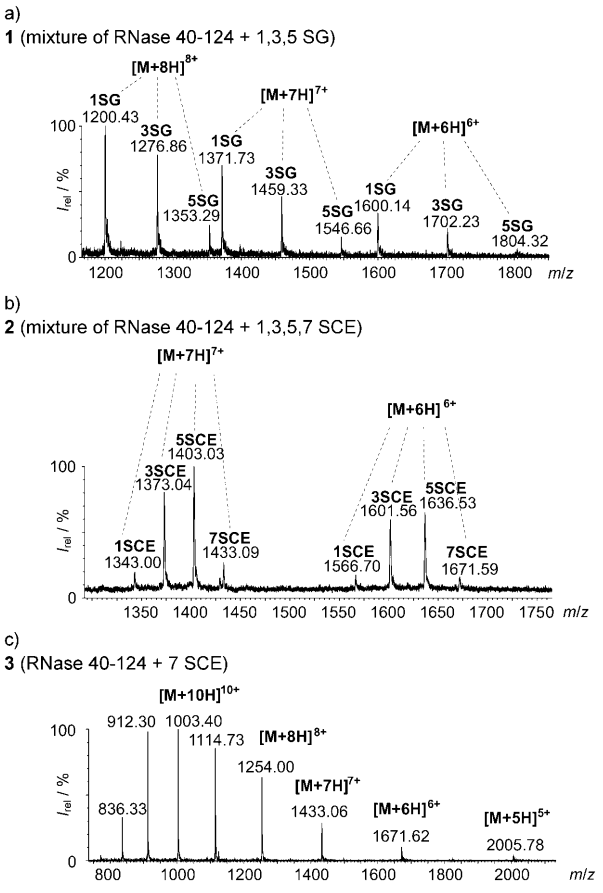


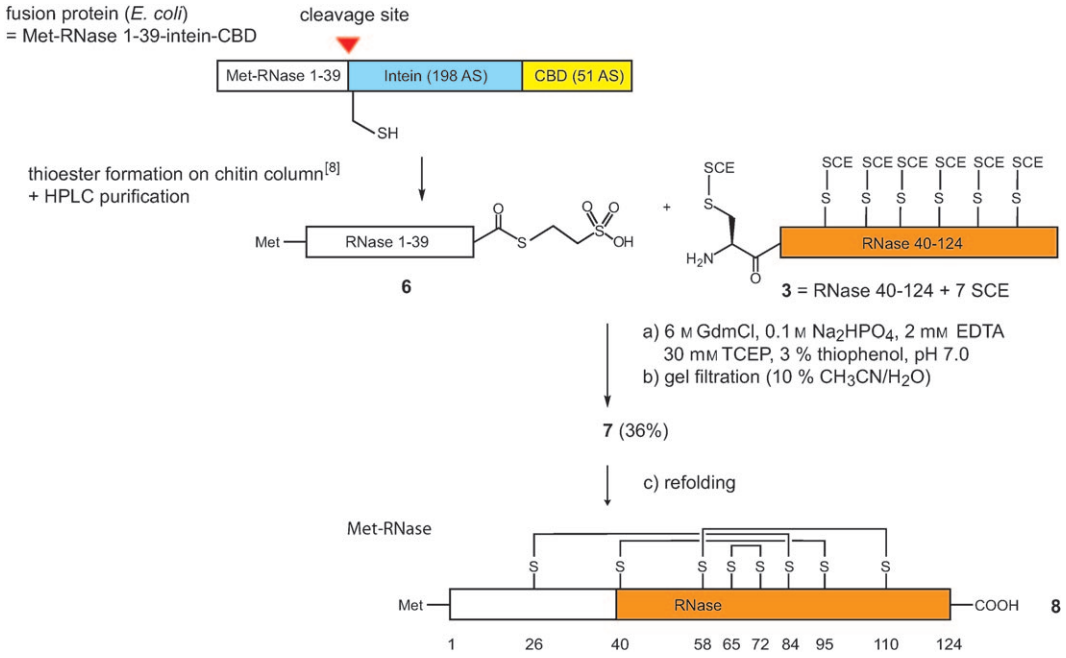
Figure 2. ESI-TOF mass spectra of protected RNase 40–124 fragments **1**–**3**.

fragment **3** carrying seven disulfides. To insure the complete liberation of the seven thiols of RNase 40–124 (**3**) the reducing agent TCEP (30 mM)[10] was added to the ligation buffer. Although the reaction mixture was slightly turbid after 2 d, the ligation product Met-RNase 1–124 (**7**) was observed by HPLC-MS analysis without pretreatment (Scheme S6 in the Supporting Information). Reoxidation by oxygen was efficiently avoided by conducting the reaction under an inert atmosphere in a nitrogen tent (< 10 ppm O_2). Even after prolonged reaction times (> 2 d) no turbidity was observed under these conditions. Product **7** was isolated after gel filtration in 36% yield. Subsequently the synthetic protein was dissolved in 6 M GdmCl, reduced with 0.2 M glutathione (GSH), and refolded by rapid dilution with a buffer containing 0.3 mM oxidized glutathione (GSSG 0.09 equiv). An aliquot from the refolding mixture was used in an RNase assay[15] based on the hydrolysis of cyclic cytidine 2′,3′-monophosphate (cCMP) indicating the RNase activity of the refolded Met-RNase **8** (Scheme S7 in the Supporting Information). Thus the stable and disulfide-modified RNase fragments are well suited for native chemical ligation under reducing conditions. Besides application in the total synthesis of bovine ribonuclease C (see the following Communication),[16] this concept was successfully applied to other cysteine-rich protein domains.[17]

We have developed a novel method to isolate and stabilize thiol-rich protein fragments with an N-terminal cysteine from intein fusion proteins. The use of thiosulfonates allowed complete conversion of cysteine thiols to disulfides. The disulfides provide several advantages: the solubility of the protein is enhanced, the cysteine moieties are protected from oxidation, and rapid reduction under the conditions of native chemical ligation is possible without side reaction. Many proteins and glycoproteins with therapeutic potential (growth factors, antibodies) contain disulfide bridges and should be accessible by ligating disulfide-protected fragments.

Received: September 28, 2008
Published online: January 28, 2009

Keywords: glycoproteins · glycosylation · native chemical ligation · protecting groups · ribonucleases



Scheme 4. Ligation of the fully protected RNase 40–124 fragment **3** with the Met-RNase 1–39 thioester (**6**) followed by refolding.

[1] T. Kimmerlin, D. Seebach, *J. Pept. Res.* **2005**, *65*, 229.

[2] a) P. E. Dawson, T. W. Muir, I. C. Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776; b) P. E. Dawson, S. B. H. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923; c) S. B. H. Kent, *J. Pept. Sci.* **2003**, *9*, 574.

[3] a) T. W. Muir, D. Sondhi, P. A. Cole, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6705; b) C. J. Noren, J. M. Wang, F. B. Perler, *Angew. Chem.* **2000**, *112*, 458; *Angew. Chem. Int. Ed.* **2000**, *39*, 450; c) review: T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249; d) Z. Machova, A. G. Beck-Sickinger, *Methods Mol. Biol.* **2005**, *298*, 105; e) I. E. Gentle, D. P. De Souza, M. Baca, *Bioconjugate Chem.* **2004**, *15*, 658; f) T. J. Tolbert, C.-H. Wong, *Angew. Chem.* **2002**, *114*, 2275; *Angew. Chem. Int. Ed.* **2002**, *41*, 2171; g) D. A. Erlanson, M. Chytil, G. L. Verdine, *Chem. Biol.* **1996**, *3*, 981; h) D. Liu, R. Xu, K. Dutta, D. Cowburn, *FEBS Lett.* **2008**, *582*, 1163; i) M. P. Malakhov, M. R. Mattern, O. A. Malakhova, M. Drinker, S. D. Weeks, T. R. Butt, *J. Struct. Funct. Genomics* **2004**, *5*, 75; j) D. Macmillan, L. Arham, *J. Am. Chem. Soc.* **2004**, *126*, 9530; k) L. Brunsveld, J. Kuhlmann, K. Alexandrov, A. Wittinghofer, R. S. Goody, H. Waldmann, *Angew. Chem.* **2006**, *118*, 6774; *Angew. Chem. Int. Ed.* **2006**, *45*, 6622; l) T. Kurpiers, H. D. Mootz, *Angew. Chem.* **2007**, *119*, 5327; *Angew. Chem. Int. Ed.* **2007**, *46*, 5234.

[4] a) T. C. Evans, J. Brenner, M. X. Xu, *J. Biol. Chem.* **1999**, *274*, 3923; b) D. W. Wood, V. Derbyshire, W. Wu, M. Chartrain, M. Belfort, G. Belfort, *Biotechnol. Prog.* **2000**, *16*, 1055; c) C. Guo, Z. Li, Y. Shi, M. Xu, J. G. Wise, W. E. Trommer, J. Yuan, *Protein Expression Purif.* **2004**, *37*, 361; d) C. P. R. Hackenberger, M. M. Chen, B. Imperiali, *Bioorg. Med. Chem.* **2006**, *14*, 5043.

[5] a) D. Macmillan, C. R. Bertozzi, *Tetrahedron* **2000**, *56*, 9515; b) T. J. W. Tolbert, C.-H. Wong, *J. Am. Chem. Soc.* **2000**, *122*, 5421; c) review: B. G. Davis, *Chem. Rev.* **2002**, *102*, 579; d) C. P. R. Hackenberger, C. T. Friel, S. E. Radford, B. Imperiali, *J. Am. Chem. Soc.* **2005**, *127*, 12882; e) S. Ficht, R. J. Payne, A. Brik, C.-H. Wong, *Angew. Chem.* **2007**, *119*, 6079; *Angew. Chem. Int. Ed.* **2007**, *46*, 5975; f) Q. Wan, S. J. Danishefsky, *Angew. Chem.* **2007**, *119*, 9408; *Angew. Chem. Int. Ed.* **2007**, *46*, 9248; g) review: H. Hojo, Y. Nakahara, *Biopolymers* **2007**, *88*, 308; h) N. Yamamoto, Y. Tanabe, R. Okamoto, P. E. Dawson, Y. Kajihara, *J. Am. Chem. Soc.* **2008**, *130*, 501.

[6] a) D. Y. Jackson, J. Burnier, C. Quan, M. Stanley, J. Tom, J. A. Wells, *Science* **1994**, *266*, 243; b) K. Witte, P. Sears, R. Martin, C.-H. Wong, *J. Am. Chem. Soc.* **1997**, *119*, 2114; c) R. T. Raines, *Chem. Rev.* **1998**, *98*, 1045; d) J. W. Baynes, F. Wold, *J. Biol. Chem.* **1976**, *251*, 6016; e) P. M. Rudd, H. C. Joao, E. Coghill, P. Fiten, M. R. Saunders, G. Opdenakker, R. A. Dwek, *Biochemistry* **1994**, *33*, 17.

[7] S. Mezzato, M. Schaffrath, C. Unverzagt, *Angew. Chem.* **2005**, *117*, 1677; *Angew. Chem. Int. Ed.* **2005**, *44*, 1650.

[8] a) T. C. Evans, Jr., J. Benner, M. Q. Xu, *J. Biol. Chem.* **1999**, *274*, 18359; b) IMPACT-TWIN Instruction Manual, Version 1.2, New England Biolabs Inc.

[9] R. Rudolph, G. Böhm, H. Lilie, R. Jaenicke in *Protein Function* (Ed.: T. E. Creighton), Oxford University Press, Oxford, **1997**, p. 57.

[10] J. A. Burns, J. C. Butler, J. Moran, G. M. Whitesides, *J. Org. Chem.* **1991**, *56*, 2648.

[11] a) R. Wynn, F. M. Richards, *Methods Enzymol.* **1995**, *251*, 351; b) B. G. Davis, R. C. Lloyd, J. B. Jones, *J. Org. Chem.* **1998**, *63*, 9614; c) G. DeSantis, X. Shang, J. B. Jones, *Biochemistry* **1999**, *38*, 13391.

[12] W. W. Chan, *Biochemistry* **1968**, *7*, 4247.

[13] J. W. Baynes, F. Wold, *J. Biol. Chem.* **1976**, *251*, 6016.

[14] other examples of thioester ligations yielding full length bovine RNase A: a) T. C. Evans, Jr, J. Benner, M. Q. Xu, *Protein Sci.* **1998**, *7*, 2256; b) G. S. Beligere, P. E. Dawson, *J. Am. Chem. Soc.* **1999**, *121*, 6332; c) R. J. Hondal, B. L. Nilsson, R. T. Raines, *J. Am. Chem. Soc.* **2001**, *123*, 5140; d) U. Arnold, M. P. Hinderaker, R. T. Raines, *ScientificWorldJournal.* **2002**, *2*, 1823; e) B. L. Nilsson, R. J. Hondal, M. B. Soellner, R. T. Raines, *J. Am. Chem. Soc.* **2003**, *125*, 5268; f) D. J. Boerema, V. A. Tereshko, S. B. H. Kent, *Biopolymers* **2008**, *90*, 278.

[15] E. M. Crook, A. P. Mathias, B. R. Rabin, *Biochem. J.* **1960**, *74*, 234.

[16] C. Piontek, D. Varón Silva, C. Heinlein, C. Pöhner, S. Mezzato, P. Ring, A. Martin, F. X. Schmid, C. Unverzagt, *Angew. Chem.* **2009**, *121*, 1974; *Angew. Chem. Int. Ed.* **2009**, *48*, 1941.

[17] N. Lombana, S. Siebenhaar, Universität Bayreuth, unpublished results.

