Glycosylation (2)

DOI: 10.1002/anie.200804735

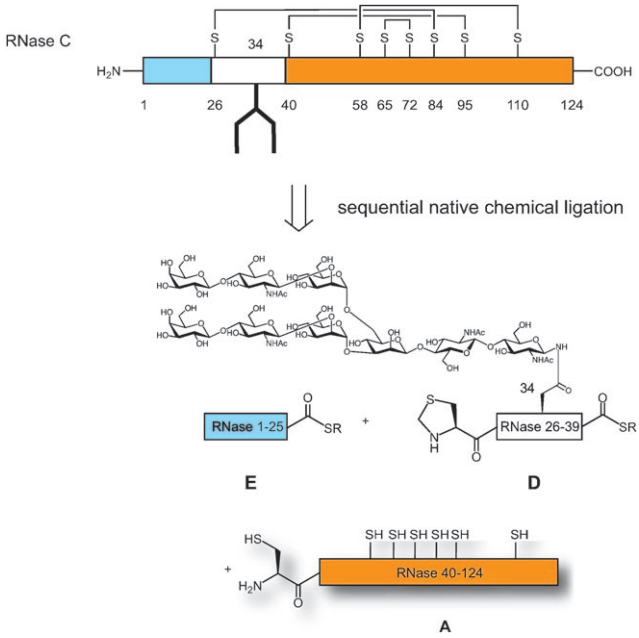
Semisynthesis of a Homogeneous Glycoprotein Enzyme: Ribonuclease C: Part 2**

*Christian Piontek, Daniel Varón Silva, Christian Heinlein, Claudia Pöhner, Stefano Mezzato, Petra Ring, Andreas Martin, Franz Xaver Schmid, and Carlo Unverzagt**

In memory of Bruce Merrifield

The synthesis of homogeneously glycosylated glycoproteins[1–7] and their analogues[8,9] provides a novel approach to study the importance of the carbohydrate moieties for the functions of therapeutic glycoproteins. These efforts were initially stimulated by the need to produce defined glycoforms for such investigations, and this research has gained further momentum by the evidence that individual glycoforms[10] can have distinctly different functions.[11,12] The currently used methodology for the synthesis of N-glycoproteins comprises solid-phase glycopeptide synthesis,[13,14] the use of transglycosidases[15] or enzymatic remodeling,[16] the use of engineered cell lines,[17,18] and in particular the application of chemoselective ligation methods.[19–21] In the preceding Communication[22] we demonstrated that active RNase A can be obtained by expressed chemical ligation using a recombinant RNase segment that was chemically stabilized by transformation of its seven cysteines into mixed disulfides. We have taken this approach further to the synthesis of a single glycoform of bovine RNase carrying a biantennary complex-type nonasaccharide (RNase C[23]) by sequential[21,24] native chemical ligations (Scheme 1).

In our initial strategy to obtain N-glycosylated RNase we envisioned a single ligation site at Cys 40 (Scheme 1) linking a recombinant RNase fragment 40–124 with a synthetic glycopeptide thioester RNase 1–39. A thioester of RNase 1–39 containing an acetylated *N*-acetylglucosamine at Asn 34 was synthesized by Fmoc strategy on a polystyrene resin equipped with a safety-catch linker.[25] After alkylation, thiolysis, and deprotection an RNase 1–39 glycopeptide benzylthioester was obtained, albeit in low yield and purity. These difficulties initiated the development of a safety-catch/Rink amide double-linker PEGA resin that would be useful for the synthesis of glycopeptide thioesters containing complex oligosaccharides.[26,27] Surprisingly, when we attempted the stepwise solid-phase synthesis of the RNase 1–39 glycopeptide with an acetylated complex-type N-glycan nonasaccharide on double-linker PEGA resin, only glycopeptides less than 20 amino acids in length could be obtained as a result of numerous side reactions. Thus the use of an additional ligation site at Cys 26 was considered. This shortens the synthetic glycopeptide thioester to a 14-mer (26–39) containing Cys 26 suitably masked for sequential ligation with an RNase 1–25 thioester **E** (Scheme 1). We envisioned the recently developed sequential native chemical ligation method based on thiazolidine protection of the N-terminal cysteine, a procedure that can be performed in a one-pot fashion.[28,29]



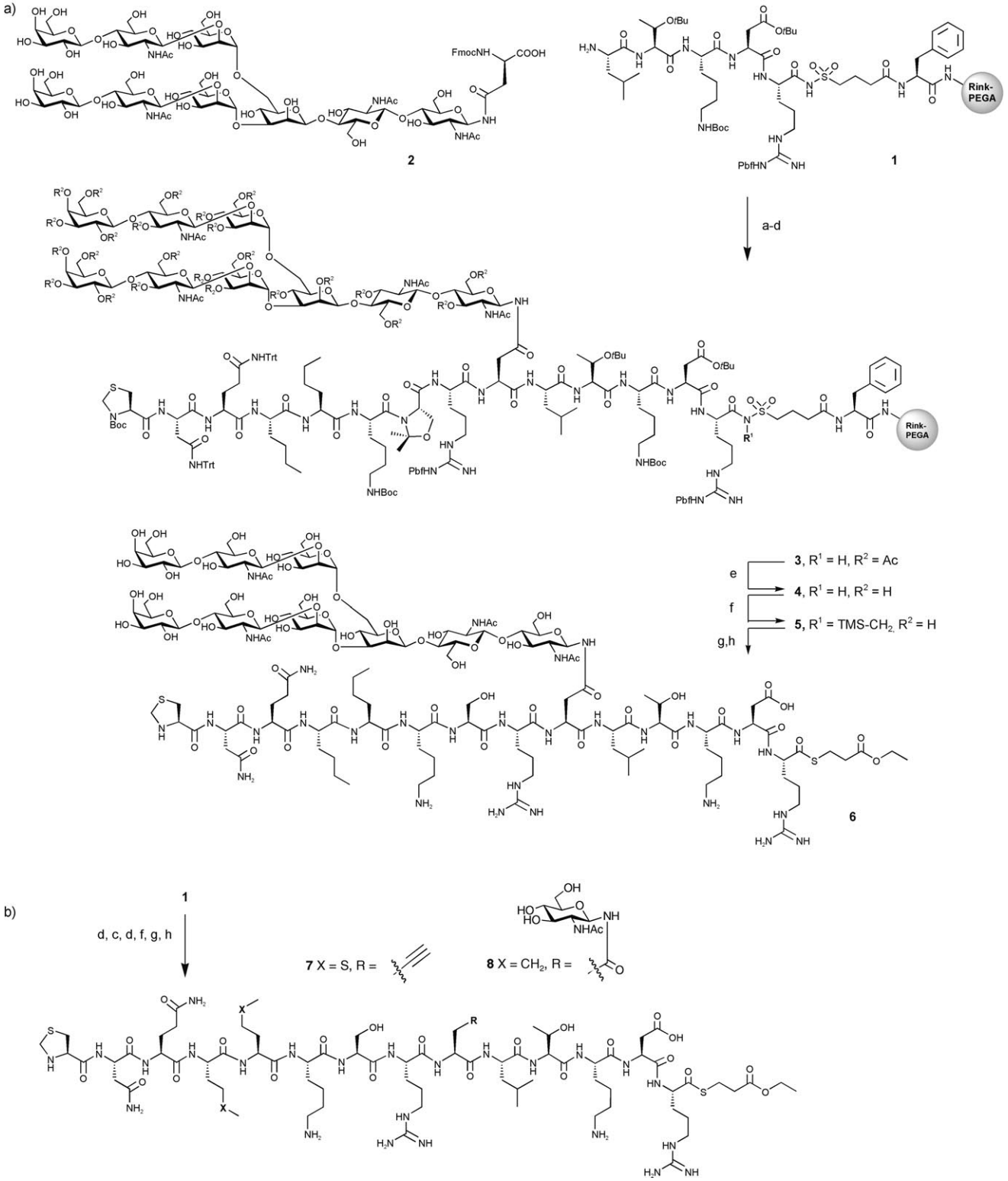
Scheme 1. Retrosynthesis of homogeneously glycosylated RNase C.

To synthesize the glycopeptide thioester RNase 26–39 (**6**) (Scheme 2 a) the double-linker PEGA resin was used and the RNase 35–39 sequence (**1**) was attached as described previously.[26] The Fmoc-Asn building block **2** containing a biantennary complex-type nonasaccharide was obtained from egg yolk[30] via the disialylated undecasaccharide Fmoc-Asn precursor followed by acid hydrolysis.[31] The glycosyl amino acid **2** (0.8 equiv) was coupled to the resin **1**, and the remaining free amino groups were capped by acetic anhydride/acetic acid/pyridine (1:1:1.5) without acetylation of the safety-catch linker.[26] Under these conditions all OH groups

[*] Dr. C. Piontek, Dr. D. Varón Silva, C. Heinlein, C. Pöhner, Dr. S. Mezzato, P. Ring, 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. We are grateful to Prof. Holger Dobbek for providing the nitrogen tent.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200804735.

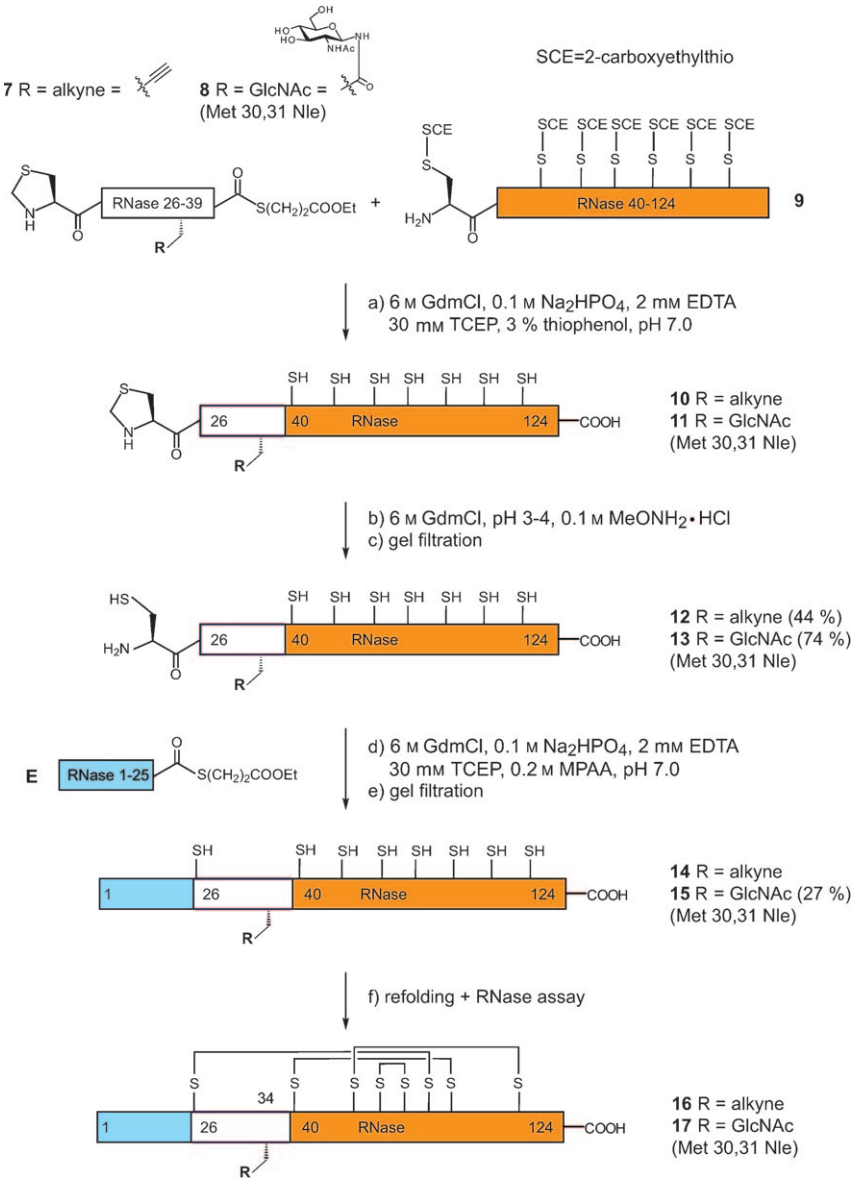


Scheme 2. Synthesis of RNase 26–39 thioester: a) Fmoc-Asn-(nonasaccharide)-OH (**2**) (R^2 = OH), PyBOP, DIPEA, DMSO/NMP; b) Ac_2O/AcOH/pyridine (1:1:1.5) (R^2 = Ac); c) 20% piperidine/NMP; d) manual solid-phase peptide elongation using Fmoc-amino acid (4 equiv), HCTU (4 equiv), Cl-HOBt (4 equiv), DIPEA (9 equiv), NMP; e) 10% hydrazine in THF/MeOH (1:1); f) $TMSCHN_2$, *n*-hexane, CH_2Cl_2; g) ethyl-3-mercaptopropionate, DMF; h) TFA, ethyl-3-mercaptopropionate, TES, H_2O. PyBOP = benzotriazolyl-1-oxytripyrrolidinophosphonium hexafluorophosphate, DIPEA = diisopropylethylamine, DMSO = dimethylsulfoxide, NMP = *N*-methylpyrrolidone, HCTU = *N*-[(1*H*-6-chlorobenzotriazol-1-yl)(dimethylamino)methylene]-*N*-methylmethanaminium hexafluorophosphate *N*-oxide, HOBt = 1-hydroxybenzotriazole, TES = triethylsilane, Boc = *tert*-butoxycarbonyl, Trt = triphenylmethyl, Pbf = 2,2,4,6,7-pentamethyl-2,3-dihydro-1-benzofuran-5-sulfonyl.

were acetylated concomitantly, preventing esterification during subsequent elongations. Attachment of the terminal amino acids was greatly facilitated after incorporation of a Lys-Ser pseudoproline dipeptide.[32] Additionally, norleucine residues were incorporated instead of methionines 30 and 31, since the known susceptibility of this RNase sequence to sulfoxide formation could not be suppressed efficiently during manual synthesis. Norleucine substitutions are known not to affect the folding or the activity of most proteins.[33,34] As a further improvement, the generation of a glycopeptide thioester 26–39 with a deprotected N-glycan was attempted in order to simplify the subsequent native chemical ligations and prevent side reactions. Thus on-resin cleavage of the acetyl groups[35] was tested by brief treatment of resin **3** with hydrazine hydrate in water, THF, or methanol. Complete deprotection was achieved only when an optimized solvent mixture (MeOH/THF 1:1) was used, indicating that good swelling of the resin (MeOH) and solubility of the protected peptide (THF) are required simultaneously. This set the stage for the alkylation of the safety-catch linker with TMS-diazomethane[36] in an optimized solvent mixture (hexane/$CHCl_3$ 1:1), followed by thiolysis and global deprotection in the presence of ethyl-3-mercaptopropionate. Despite accompanying methylation of the sugar moieties (10–20%), pure thioester **6** was obtained in a final yield of 18% after HPLC purification (Scheme S3 in the Supporting Information).

Thioester 1–25 (**E**) was assembled on double-linker PEGA resin employing a Ser-Ser pseudoproline dipeptide at positions 15,16 and 21,22; the previously described conditions were used for thioester formation and deprotection (20% yield after HPLC).

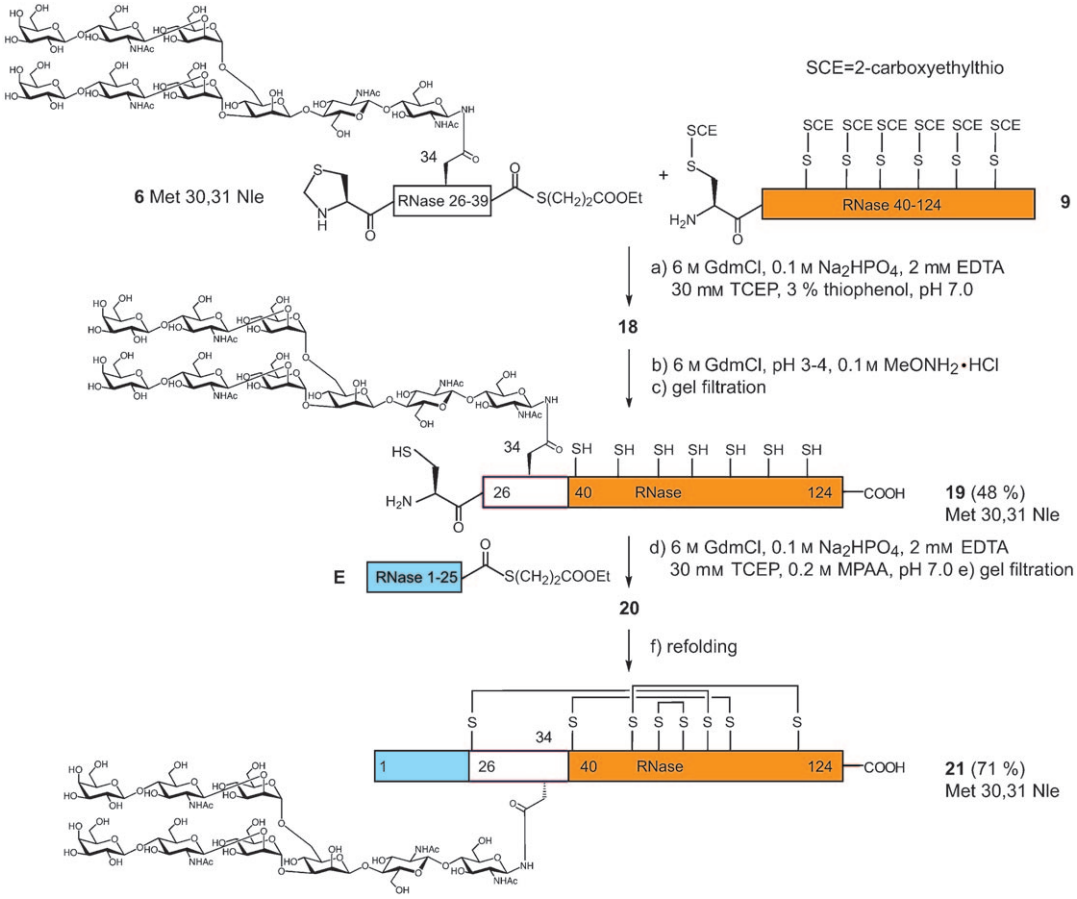
With the three segments of RNase C in hand, we carried out initial ligation experiments as a one-pot procedure in order to avoid the isolation of thiol-rich intermediates, which are sensitive to oxidation. The disulfide-protected fragment **B** (**9**) and thioester **6** were ligated in high efficiency in the presence of tris(2-carboxyethyl)phosphine (TCEP)[37] and thiophenol followed by deprotection of the N-terminal cysteine moiety by incubation with methoxyamine[28] (0.2 M, pH 3–4, 4 h). However, the subsequent ligation with thioester **C** gave only a small amount of full-length RNase; instead, the C-terminal *N*-methoxyamide[38] of **E** was formed in a pH-dependent manner. Despite the addition of further portions of thioester **E** at lower pH (6.5), only incomplete ligation (≈20%) was observed. Therefore we had to isolate the deprotected glycopeptide RNase 26–124 containing eight thiols. Owing to the sensitivity of this intermediate to oxidation, this procedure was optimized employing the more readily available alkyne-[8] and GlcNAc-containing thioesters **7** and **8**, respectively (Scheme 2b and Scheme S3 in the Supporting Information). Compounds **7** (14% yield) and **8** (18% yield) were obtained conveniently from double-linker (PEGA) resin.[26] Ligations of the thioesters **7** and **8** with **9** were conducted in 6 M guanidinium chloride in the presence of 30 mM TCEP and 3% thiophenol in glass vials and under a nitrogen atmosphere (< 10 ppm O_2) (Scheme 3). After two days high conversions were observed according to HPLC-MS analysis, and the thiazolidines of the ligation products **10** (Scheme S4) and **11** (Scheme S5 in the Supporting Information) were opened by addition of 0.1 M methoxy-



Scheme 3. Sequential native chemical ligation of RNase variants containing propargylglycine or GlcNAc-β-Asn at Asn34. Intermediates **12** and **13** were isolated and ligated to the full-length RNases and then refolded.

amine at pH 3–4. After complete deprotection (LC-MS) the Cys-peptide **12** (Scheme S4 in the Supporting Information) was isolated by gel filtration under acidic conditions giving a yield of 44%. Additional reduction of **13** with TCEP at pH 6 prior to gel filtration led to a yield of 74% (Scheme S5 in the Supporting Information). Both Cys-peptides were subsequently ligated with excess thioester **E** using 0.2 M *p*-mercaptophenyl acetic acid (MPAA) as a strongly activating thiol[39] under reducing conditions strictly avoiding the entry of oxygen into the reaction vial. Under these stringent conditions the ligations were complete after one day, and the full-length RNases **14** (Scheme S4) and **15** (Scheme S5 in the Supporting Information) were isolated by gel filtration. Refolding was initiated by 60-fold dilution of the RNases **14** and **15** in 6 M guanidinium chloride + 0.2 M glutathione (GSH) with a buffer containing 0.3 mM GSSG. The RNase activity of **16** and **17** was confirmed by a standard assay carried out with an aliquot of the refolding solutions using the ring-opening reaction of cyclic cytidine 2′,3′-monophosphate (cCMP)[40] (Scheme S7 in the Supporting Information).

Finally the optimized reaction conditions and isolation procedures were applied to the sequential ligation of full-length RNase C containing a complex-type nonasaccharide (Scheme 4). The ligation of glycopeptide **6** with **9** (3% thiophenol, TCEP) gave RNase **18**, which was deprotected by addition of methoxyamine to the reaction mixture (pH 3–4). Subsequent reduction with TCEP and gel filtration in 25% acetonitrile yielded the glycosylated RNase 26–124 (**19**) with a free N-terminal cysteine (48%, Schemes S1 and S6 in the Supporting Information). Ligation with the 1–25 thioester (**E**) proceeded smoothly using 0.2 M MPAA as an activator under strictly oxygen-free conditions within one day and furnished RNase 1–124 (**20**) (Scheme S6 in the Supporting Information). Owing to the high solubility of MPAA and its folding-enhancing properties,[41] the reaction mixture was submitted directly to refolding by dilution with a 0.3 mM GSSG solution. In this case MPAA replaces the GSH typically added as a reducing agent. Refolding was allowed to proceed for four days, and the protein was subsequently isolated by gel filtration in 71% yield (Schemes S1 and S7 in the Supporting Information). The refolded synthetic RNase C **21** was assayed[40] for enzymatic activity[42] (Scheme S7 in the Supporting Information). It showed a relative activity of 56% when compared with a sample of commercial RNase A of equal concentration. Additionally, a CD spectrum of **21** was recorded, which shows that the synthetic glycoprotein displays ordered structure similar to native RNase A (Scheme S8 in the Supporting Information). This native-like structure and the high enzymatic activity of the final product **21** demonstrate that the chemical synthesis of a glycoprotein



Scheme 4. Optimized sequential native chemical ligations and subsequent refolding of RNase C Met30,31Nle, which contains a biantennary N-glycan at Asn34.

with the complexity of RNase was indeed successful. This synthesis of RNase C Met30,31Nle (**21**) containing a native biantennary complex type N-glycan at Asn34 represents the first synthesis of an N-glycoprotein enzyme with over 100 amino acids by sequential native chemical ligation.

This semisynthesis of the glycoprotein enzyme RNase C by sequential native chemical ligation combines as key features the use of a recombinant RNase fragment, which is chemically stabilized by mixed disulfides, with a synthetic glycopeptide thioester containing only one thiazolidine protecting group. With the thiol-rich peptides, strict exclusion of oxygen was essential for efficient ligations. Based on the improved conditions developed in the course of this project the semisynthesis of thiol-rich N-glycoproteins and derivatives with therapeutic and biochemical interest are within reach.

Received: September 28, 2008
Published online: January 29, 2009

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

[1] B. G. Davis, *Chem. Rev.* **2002**, *102*, 579.
[2] Z. Zhang, J. Gildersleeve, Y. Y. Yang, R. Xu, J. A. Loo, S. Uryu, C. H. Wong, P. G. Schultz, *Science* **2004**, *303*, 371.
[3] Z. Guo, N. Shao, *Med. Res. Rev.* **2005**, *25*, 655.
[4] M. R. Pratt, C. R. Bertozzi, *Chem. Soc. Rev.* **2005**, *34*, 58.
[5] T. Buskas, S. Ingale, G. J. Boons, *Glycobiology* **2006**, *16*, 113R.
[6] C. S. Bennett, C. H. Wong, *Chem. Soc. Rev.* **2007**, *36*, 1227.
[7] H. Hojo, Y. Nakahara, *Biopolymers* **2007**, *88*, 308.
[8] B. G. Davis, *Science* **2004**, *303*, 480.
[9] Y. Ito, S. Hagihara, I. Matsuo, K. Totani, *Curr. Opin. Struct. Biol.* **2005**, *15*, 481.
[10] P. M. Rudd, H. C. Joao, E. Coghill, P. Fiten, M. R. Saunders, G. Opdenakker, R. A. Dwek, *Biochemistry* **1994**, *33*, 17.
[11] R. M. Anthony, F. Nimmerjahn, D. J. Ashline, V. N. Reinhold, J. C. Paulson, J. V. Ravetch, *Science* **2008**, *320*, 373.
[12] R. Sackstein, J. S. Merzaban, D. W. Cain, N. M. Dagia, J. A. Spencer, C. P. Lin, R. Wohlgemuth, *Nat. Med.* **2008**, *14*, 181.
[13] N. Yamamoto, Y. Tanabe, R. Okamoto, P. E. Dawson, Y. Kajihara, *J. Am. Chem. Soc.* **2008**, *130*, 501.
[14] Q. Wan, S. J. Danishefsky, *Angew. Chem.* **2007**, *119*, 9408; *Angew. Chem. Int. Ed.* **2007**, *46*, 9248.
[15] L. X. Wang, *Carbohydr. Res.* **2008**, *343*, 1509.
[16] L. J. Thomas, K. Panneerselvam, D. T. Beattie, M. D. Picard, B. Xu, C. W. Rittershaus, H. C. Marsh, Jr., R. A. Hammond, J. Qian, T. Stevenson, D. Zopf, R. J. Bayer, *Glycobiology* **2004**, *14*, 883.
[17] S. Wildt, T. U. Gerngross, *Nat. Rev. Microbiol.* **2005**, *3*, 119.
[18] C. Ferrara, P. Brunker, T. Suter, S. Moser, U. Puntener, P. Umana, *Biotechnol. Bioeng.* **2006**, *93*, 851.
[19] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776.
[20] T. Kimmerlin, D. Seebach, *J. Pept. Res.* **2005**, *65*, 229.
[21] T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249.
[22] C. Piontek, P. Ring, O. Harjes, C. Heinlein, S. Mezzato, N. Lombana, C. Pöhner, M. Püttner, D. Varón Silva, A. Martin, F. X. Schmid, C. Unverzagt, *Angew. Chem.* **2009**, *121*, 1968; *Angew. Chem. Int. Ed.*, **2009**, *48*, 1936.
[23] J. W. Baynes, F. Wold, *J. Biol. Chem.* **1976**, *251*, 6016.
[24] P. E. Dawson, S. B. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923.
[25] Y. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman, C. R. Bertozzi, *J. Am. Chem. Soc.* **1999**, *121*, 11684.
[26] S. Mezzato, M. Schaffrath, C. Unverzagt, *Angew. Chem.* **2005**, *117*, 1677; *Angew. Chem. Int. Ed.* **2005**, *44*, 1650.
[27] E. Meinjohanns, M. Meldal, H. Paulsen, R. A. Dwek, K. Bock, *J. Chem. Soc. Perkin Trans. 1* **1998**, 549.
[28] D. Bang, S. B. Kent, *Angew. Chem.* **2004**, *116*, 2588; *Angew. Chem. Int. Ed.* **2004**, *43*, 2534.
[29] J. D. Boerema, A. V. Tereshko, S. B. H. Kent, *Biopolymers* **2008**, *90*, 278.
[30] A. Seko, M. Koketsu, M. Nishizono, Y. Enoki, H. R. Ibrahim, L. R. Juneja, M. Kim, T. Yamamoto, *Biochim. Biophys. Acta Gen. Subj.* **1997**, *1335*, 23.
[31] Y. Kajihara, Y. Suzuki, K. Sasaki, L. R. Juneja, *Methods Enzymol.* **2003**, *362*, 44.
[32] T. Wöhr, F. Wahl, A. Nefzi, B. Rohwedder, T. Sato, X. Sun, M. Mutter, *J. Am. Chem. Soc.* **1996**, *118*, 9218.
[33] R. Rocchi, A. Scatturin, L. Moroder, F. Marchiori, A. M. Tamburro, E. Scoffone, *J. Am. Chem. Soc.* **1969**, *91*, 492.
[34] C. B. Anfinsen, L. G. Corley, *J. Biol. Chem.* **1969**, *244*, 5149.
[35] E. Bardají, J. L. Torres, P. Clapes, F. Albericio, G. Barany, R. E. Rodriguez, M. P. Sacristan, G. Valencia, *J. Chem. Soc. Perkin Trans. 1* **1991**, 1755.
[36] R. Ingenito, E. Bianchi, D. Fattori, A. Pessi, *J. Am. Chem. Soc.* **1999**, *121*, 11369.
[37] J. A. Burns, J. C. Butler, J. Moran, G. M. Whitesides, *J. Org. Chem.* **1991**, *56*, 2648.
[38] E. C. Johnson, E. Malito, Y. Shen, D. Rich, W. J. Tang, S. B. Kent, *J. Am. Chem. Soc.* **2007**, *129*, 11480 (see Scheme S2 in the Supporting Information).
[39] E. C. Johnson, S. B. Kent, *J. Am. Chem. Soc.* **2006**, *128*, 6640.
[40] E. M. Crook, A. P. Mathias, B. R. Rabin, *Biochem. J.* **1960**, *74*, 234.
[41] J. D. Gough, E. J. Barrett, Y. Silva, W. J. Lees, *J. Biotechnol.* **2006**, *125*, 39.
[42] R. S. Hodges, R. B. Merrifield, *J. Biol. Chem.* **1975**, *250*, 1231.

