Synthesis Design

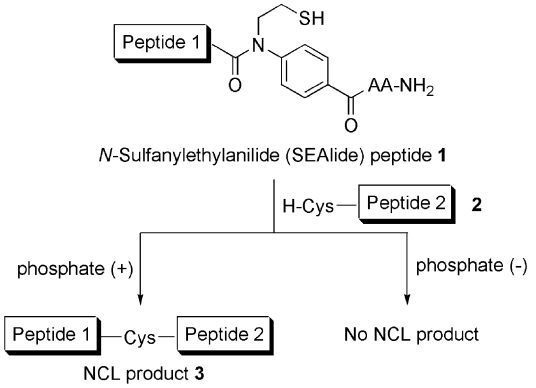
DOI: 10.1002/ange.201303390

Chemical Synthesis of Biologically Active Monoglycosylated GM2-Activator Protein Analogue Using *N*-Sulfanylethylanilide Peptide**

*Kohei Sato, Akira Shigenaga, Keisuke Kitakaze, Ken Sakamoto, Daisuke Tsuji, Kohji Itoh, and Akira Otaka**

Protein therapeutics containing glycoproteins have been attracting increasing attention as potential agents for treating diseases that were once thought to be incurable.[1] In therapeutics development using proteins, a critical issue is how protein molecules are produced. Genetic engineering procedures represent a significant advance in the production of naturally occurring proteins, however, medicinal-chemistry-oriented examination of protein therapeutics incorporating unnatural structural units is far from satisfactory using genetic protocols. One alternative to genetic protocols is chemical synthesis of proteins in which the protein backbone is constructed using ligation chemistry such as native chemical ligation (NCL).[2] Chemically synthesized proteins could be useful in medicinal chemistry.[3] Recent advances in NCL have placed the chemical synthesis of small- and medium-sized proteins within our reach. In particular, combinations of sequential and convergent NCL protocols provide technical improvements for the protein synthesis.[4] A representative sequential NCL protocol uses a kinetically controlled NCL (KCL)[5] developed by Kent and co-workers, in which a peptide chain can be elongated based on the differences between the reactivities of aryl and alkyl thioesters in a one-pot, N-to-C-directed process.[6,7] The KCL protocol has enjoyed great success in providing the large N-segment alkyl thioesters for convergent synthesis. However, KCL reactions using highly reactive alkyl thioesters are inefficient.[5c]

In the course of developing synthetic protocols for peptide thioesters using Fmoc solid-phase peptide synthesis (Fmoc SPPS),[8] we found that the *N*-sulfanylethylanilide (SEAlide) peptides **1**, as crypto-thioesters, can be successfully used for one-pot, N-to-C-directed ligation under kinetically controlled conditions. The high kinetic control is achieved by selecting an appropriate buffer salt (phosphate) in the reaction medium, that is, the SEAlide unit **1** remains intact under NCL conditions with N-terminal cysteinyl peptides **2** in the absence of phosphate salts, whereas it functions as a thioester in the presence of phosphate salts and participates in NCL to yield **3** (Scheme 1).[7,9]



Scheme 1. Reactivity of SEAlide peptide as crypto-thioester is tuned by phosphate salts.

Herein, we address the chemical synthesis of the monoglycosylated GM2-activator protein (GM2AP) analogue **4** using SEAlide peptides. GM2AP is an essential cofactor for the lysosomal degradation of ganglioside GM2 by β-hexosaminidase A (HexA).[10] Functional deficiencies of GM2AP result in a fatal neurological disease, and thus we explored a convergent synthesis platform compatible with various GM2AP analogues to develop a medical treatment. Generally, synthetic strategies based on the incorporation of a sugar chain acceptor residue, such as monoglycosylated asparagine, and subsequent site-selective transfer of a sugar chain on the acceptor site have been used for glycoprotein synthesis.[11,12] GM2AP consists of 162 residues with N-linked glycans at Asn32 (Figure 1). Our route to a monoglycosylated GM2AP analogue relies on replacement of Asn32 by Cys, on which an *N*-acetylglucosamine moiety is incorporated by an S alkylation of Cys32 with iodoacetyl-*N*-acetylglucosamine.[13] The use of the S-alkylation protocol, which provides one additional ligation site, facilitates the SPPS of the peptide fragments. The NCL-mediated construction of naturally occurring GM2AP requires the synthesis of a thioester fragment of at least 67 residues (8–74). Straightforward SPPS does not always guarantee successful synthesis of such large peptides (over 40–50 residues). Our convergent synthetic strategy for the GM2AP analogue **4** is shown in Scheme 2.

Standard NCL condensation of the alkylthioester Fr 1 (**5**), which has an *S*-acetamidomethyl (Acm) protecting group on Cys8, with Fr 2 (**6**), which is S-Acm-protected at Cys68 and the SEAlide moiety,[14] followed by S glycosylation at Cys32

[*] K. Sato, Prof. Dr. A. Shigenaga, K. Kitakaze, K. Sakamoto, Prof. Dr. D. Tsuji, Prof. Dr. K. Itoh, Prof. Dr. A. Otaka
Institute of Health Bioscience and Graduate School of Pharmaceutical Sciences, The University of Tokushima
Shomachi, Tokushima 770-8505 (Japan)
E-mail: aotaka@tokushima-u.ac.jp

[**] This research was supported in part by a Grant-in-Aid for Scientific Research (KAKENHI) and research grants from the Takeda Science and the Uehara Memorial Foundations. K.S. is grateful for a scholarship from the Yoshida Scholarship Foundation.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201303390.

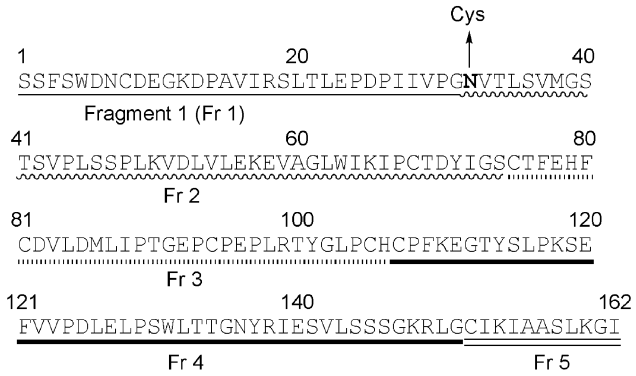


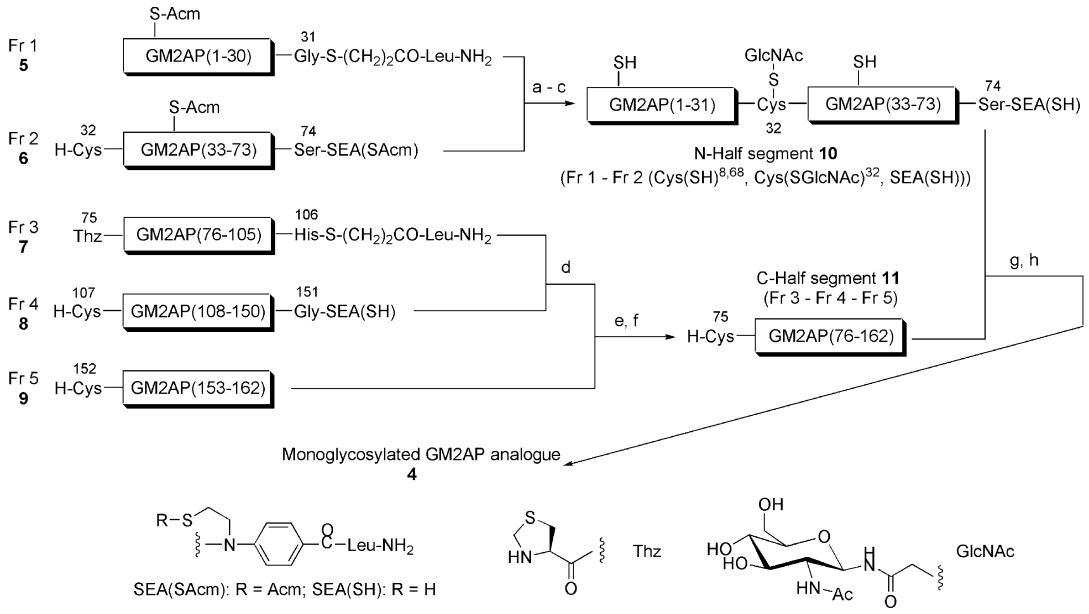
Figure 1. Entire amino acid sequences of GM2AP and the corresponding Asn32-mutated analogue. Five peptide fragments for the convergent synthesis of the mutated analogue.

and subsequent removal of the Acm protections should afford the monoglycosylated N-half segment **10** of the SEAlide peptide. For the preparation of the C-half segment **11**, one-pot, N-to-C-directed sequential ligation using the SEAlide peptide was used. The first NCL of the Fr 3 (**7**) containing N-terminal 1,3-thiazolidine-4-carbonyl (Thz) and C-terminal alkyl thioester with the N-terminal cysteinyl SEAlide peptide Fr 4 (**8**) in the absence of phosphate salts, followed by a second ligation of Fr 5 (**9**) in the presence of phosphate salts and subsequent opening of the 1,3-thiazolidine ring[15] in the Thz residue, gives the desired C-half segment **11**. Standard ligation between **10** and **11** in the presence of phosphate salts and subsequent folding would provide the desired product.

The requisite peptide fragments, Fr 1 (**5**) and Fr 3 (**7**), and Fr 2 (**6**), Fr 4 (**8**), and Fr 5 (**9**), were prepared by Boc and Fmoc SPPS, respectively. To achieve the selective glycosylation, Cys8, Cyc68, and the SEAlide moiety in Fr 1 and Fr 2, except for Cys32, were protected by Acm groups. Peptides corresponding to **6** and **8** were synthesized on Fmoc-Ser(*t*Bu)-incorporated-*N*-acetoamidomethylsulfanylethylaniline[16] and Fmoc-Gly-incorporated-*N*-triphenylmethylsulfanylethylaniline-linked resin, respectively, using Fmoc SPPS. In situ neutralization protocols were used in the Boc SPPS.[17]

The progress of the reactions for the synthesis of **10** are summarized in the Supporting Information (SI-Figures 1 and 2). NCL for the synthesis of **10** from **5** and **6** in 6 M guanidine (Gn)·HCl-0.2 M 3-[4-(2-hydroxyethyl)piperazin-1-yl]propane-1-sulfonic acid (HEPPS) in the presence of 3% (v/v) thiophenol, proceeded smoothly to afford the Acm-protected 74-residue peptide **12** (Fr 1–Fr 2 (Cys(SAcm)^{8,68}, Cys(SH)^{32}, SEA(SAcm))) (SI-Figure 1). Glycosylation of the Cys32 peptide **12** with iodoacetyl-*N*-acetylglucosamine in 6 M Gn·HCl-0.2 M HEPPS yielded the monoglycosylated *S*-Acm peptide **13** (Fr 1–Fr 2 (Cys(SAcm)^{8,68}, Cys(SGlcNAc)^{32}, SEA(SAcm))). in 90% yield upon isolation. The Acm groups on **13** were removed with AgOTf in the presence of anisole in TFA at 4 °C for 24 h,[18] followed by incubation in the presence of 3% (w/v) dithiothreitol (DTT), and subsequent HPLC purification to afford the desired **10** in 42% yield upon isolation (SI-Figure 2).

Monitoring of the reactions for the synthesis of the C-half segment **11** is summarized in Figure 2. The first ligation of **7** with **8** in 6 M Gn·HCl-0.1 M HEPPS, 50 mM tris(2-carboxyethyl)phosphine (TCEP),[19] and 100 mM (4-carboxymethyl)-thiophenol (MPAA),[20] in the absence of phosphate salts, proceeded under kinetic conditions to yield the desired ligated peptide (Fr 3–Fr 4; Figure 2B). A chemoselective reaction between the thioester group in **7** and the N-terminal cysteinyl residue in **8** was achieved because the anilide-type SEAlide moiety remained intact in the absence of phosphate salts. A solution of **9** in 0.4 M phosphate buffer was added to



Scheme 2. Convergent synthetic strategy for the preparation of monoglycosylated GM2AP analogue. a) NCL. b) Glycosylation using alkylation protocol. c) Removal of Acm groups. d) NCL in the absence of phosphate salts. e) NCL in the presence of phosphate salts. f) Opening of thiazolidine ring. g) NCL in the presence of phosphate salts. h) Folding.

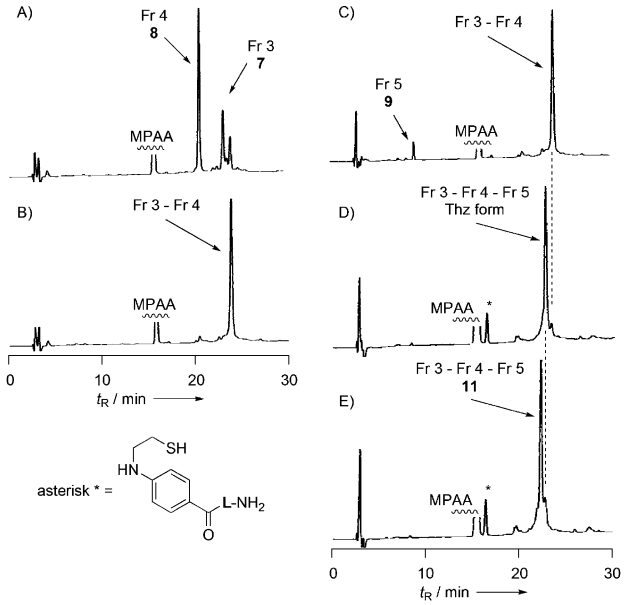


Figure 2. HPLC monitoring of reactions for the synthesis of the C-half segment **11**. A) First NCL, kinetically controlled conditions ($t<1$ min): components **7** (2.0 mM) and **8** (2.0 mM) were ligated in Gn·HCl [6 M, HEPPS (0.1 M), TCEP (50 mM), MPAA (100 mM), pH 7.0, 37 °C]. B) First NCL ($t=3$ h). C) Second NCL ($t<1$ min): To the above reaction mixture was added a solution of the segment **9** (1.1 equiv) in Gn·HCl [6 M, Na phosphate (0.4 M)]. [Final concentrations: Gn·HCl (6 M), Na phosphate (0.32 M), HEPPS (20 mM), MPAA (20 mM), TCEP (10 mM), segments **7** and **8** (0.4 mM each)]. D) Second NCL ($t=24$ h). E) Opening of 1,3-thiazolidine ring ($t=3$ h): to the above reaction mixture, 0.2 M NH_2OH·HCl was added. HPLC conditions: Cosmosil 5C18 AR-II column (4.6×250 mm) with a linear gradient of MeCN/0.1% aq. TFA (20:50–50:50 over 30 min) at a flow rate 1.0 mL min^{-1}, detection at 220 nm.

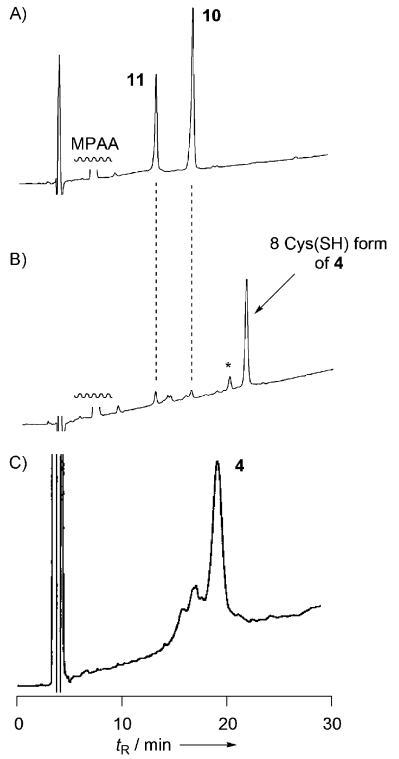


Figure 3. HPLC monitoring of NCL of the N-half segment **10** with **11** (A and B) and subsequent folding reaction (C). A) NCL ($t=0$): components **10** (0.5 mM) and **11** (0.5 mM) were ligated in Gn·HCl [6 M, Na phosphate (0.5 M), TCEP (50 mM), MPAA (50 mM), pH 6.0, 37 °C]. B) NCL ($t=24$ h). C) HPLC-purified reduced **4** [8 Cys(SH) form] was folded in Gn·HCl [1 M, Na phosphate (16 mM), Tris·HCl (42 mM), reduced form glutathione (1.7 mM), oxidized form glutathione (0.17 mM), Tween 20 (0.0025%, v/v), pH 8.0, 0.1 mg mL^{-1} protein]. HPLC conditions: Cosmosil Protein-R column (4.6×250 mm) with a linear gradient of MeCN/0.1% aq. TFA (35:65 to 55:45 over 30 min) at a flow rate of 1.0 mL min^{-1}, detection at 220 nm. *=Ser74-epimerized product (observed mass was identical to that of desired product).

the reaction mixture to initiate the second NCL of the SEAlide moiety with the cysteinyl residue of **9** (Figure 2C). This reaction also proceeded efficiently to yield the desired Thz peptide (Fr 3–Fr 4–Fr 5; Figure 2D), which then underwent opening of the 1,3-thiazolidine ring upon addition of NH_2OH·HCl, to give **11** in 47% yield after HPLC purification (Figure 2E). Kinetically controlled one-pot, N-to-C sequential ligation was therefore successful in assembling three peptide fragments (**7**, **8**, and **9**).

Convergent assembly of **10** and **11** was accomplished by NCL in the presence of phosphate salts. The reaction of **10** with **11** in 6 M Gd·HCl-0.5 M Na phosphate, 50 mM TCEP, 50 mM MPAA, proceeded at a reasonable reaction rate and almost went to completion within 24 hours to yield the desired ligated 162-residue protein (reduced form of **4**). After HPLC purification, folding in the presence of reduced and oxidized forms of glutathione was performed to yield the folded monoglycosylated 162-residue GM2AP analogue **4**. The SEAlide-mediated ligation protocol is also undoubtedly of use in convergent assembly, considering the successful coupling of 74- and 88-residue peptides. HPLC analysis of the coupling reaction indicated that a material with a molecular weight identical to that of the desired ligation product was obtained (Figure 3B, denoted by asterisk). From our previous investigation of SEAlide-mediated coupling, ligation at the Ser-SEAlide site was observed to accompany partial epimerization (9–10%).[16] Epimerization of a Ser thioester has also been reported to occur under normal NCL conditions.[21]

Mass analysis of the synthetic material revealed that the product has the expected molecular weight of the folded monoglycosylated GM2AP analogue **4** (calcd 17830.4; found 17830.3). Circular dichroism (CD) analysis indicated that the synthetic material has a spectrum similar to that reported in the literature (see the Supporting Information).[22] In the hydrolysis-assisted conversion of GM2 into GM3 in the presence of Hex A, the synthetic sample was more active than an *E. coli*-expressed carbohydrate-free sample (Figure 4).[23]

In conclusion, SEAlide peptides are versatile reagents in chemical protein synthesis. In this study, SEAlide peptides were successfully used in the preparation of the N-half and C-half segments required for the convergent strategies for chemical synthesis of proteins. The use of SEAlide peptides allows one-pot, sequential NCL reactions to proceed with high kinetic control. Successful convergent coupling indicated that the SEAlide peptides function as useful synthetic units in convergent synthesis. Establishment of a synthetic platform for obtaining the Cys32 GM2AP analogue enables the

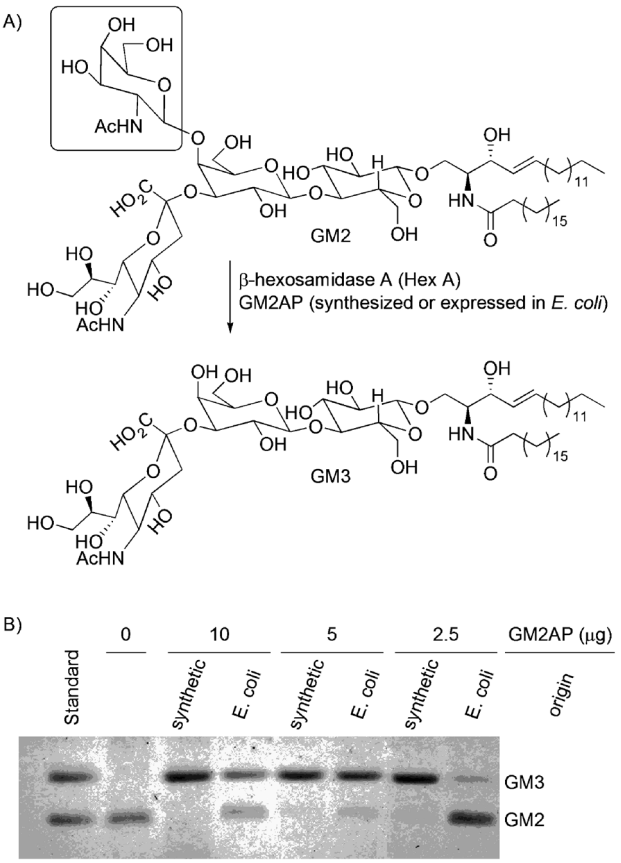


Figure 4. Hydrolysis of GM2 to GM3 with β-hexosamidase A (Hex A) in the presence of GM2APs. A) Conversion of GM2 into GM3. B) Comparable hydrolysis of GM2 with Hex A in the presence of varying amounts of synthesized or expressed GM2AP. Hydrolysis was monitored by TLC of the reaction aliquots (solvent: $CHCl_3$/MeOH 0.2% (w/v) $CaCl_2$ = 55:45:10, (v/v); detect: orcinol reagent, 120°C, 5 min).

preparation of a wide range of GM2AP protein analogues, thus leading to medicinal-chemistry-based evaluation of glycoproteins.

Received: April 22, 2013
Revised: May 15, 2013
Published online: June 13, 2013

Keywords: glycoproteins · glycosylation · peptides · solid-phase synthesis · synthesis design

[1] a) B. Leader, Q. J. Baca, D. E. Golan, *Nat. Rev. Drug Discovery* **2008**, *7*, 21 – 39; b) P. J. Carter, *Exp. Cell Res.* **2011**, *317*, 1261 – 1269.
[2] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776 – 779; b) P. E. Dawson, S. B. H. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923 – 960; c) S. B. H. Kent, *Chem. Soc. Rev.* **2009**, *38*, 338 – 351; d) S. B. H. Kent, Y. Sohma, S. Liu, D. Bang, B. Pentelute, K. Mandal, *J. Pept. Sci.* **2012**, *18*, 428 – 436.
[3] a) C. S. Zhang, L. H. Lai, *Biochem. Soc. Trans.* **2011**, *39*, 1382 – 1386; b) M. Zakrzewska, A. Szlachcic, J. Otlewski, *Curr. Pharm. Biotechnol.* **2011**, *12*, 1792 – 1798.
[4] J. Y. Lee, D. Bang, *Pept. Sci.* **2010**, *94*, 441 – 447.
[5] a) D. Bang, B. L. Pentelute, S. B. H. Kent, *Angew. Chem.* **2006**, *118*, 4089 – 4092; *Angew. Chem. Int. Ed.* **2006**, *45*, 3985 – 3988; b) T. Durek, V. Y. Torbeev, S. B. H. Kent, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 4846 – 4851; c) D. Bang, J. Lee, Y. Kwon, B. L. Pentelute, *Bioconjugate Chem.* **2011**, *22*, 1645 – 1649.
[6] a) N. Ollivier, J. Vicogne, A. Vallin, H. Drobecq, R. Desmet, O. El Mahdi, B. Leclercq, G. Goormachtigh, V. Fafeur, O. Melnyk, *Angew. Chem.* **2012**, *124*, 213 – 217; *Angew. Chem. Int. Ed.* **2012**, *51*, 209 – 213; b) R. Yang, W. Hou, X. Zhang, C.-F. Liu, *Org. Lett.* **2012**, *14*, 374 – 377; c) L. Raibaut, N. Ollivier, O. Melnyk, *Chem. Soc. Rev.* **2012**, *41*, 7001 – 7015.
[7] a) K. Sato, A. Shigenaga, K. Tsuji, S. Tsuda, Y. Sumikawa, K. Sakamoto, A. Otaka, *ChemBioChem* **2011**, *12*, 1840 – 1844; b) A. Otaka, K. Sato, H. Ding, A. Shigenaga, *Chem. Rec.* **2012**, *12*, 479 – 490.
[8] S. Tsuda, A. Shigenaga, K. Bando, A. Otaka, *Org. Lett.* **2009**, *11*, 823 – 826.
[9] Other N-substituted type N-to-S acyl transfer devises can be converted into the corresponding thioesters under acidic conditions. See Ref. [6] and a) C. Ozawa, H. Katayama, H. Hojo, Y. Nakahara, *Org. Lett.* **2008**, *10*, 3531 – 3533; b) K. S. A. Kumar, S. N. Bavikar, L. Spasser, T. Moyal, S. Ohayon, A. Brik, *Angew. Chem.* **2011**, *123*, 6261 – 6265; *Angew. Chem. Int. Ed.* **2011**, *50*, 6137 – 6141; c) S. N. Bavikar, L. Spasser, M. Haj-Yahya, S. V. Karthikeyan, T. Moyal, K. S. Ajish Kumar, A. Brik, *Angew. Chem.* **2012**, *124*, 782 – 787; *Angew. Chem. Int. Ed.* **2012**, *51*, 758 – 763.
[10] For a review of GM2AP, see: T. Kolter, K. Sandhoff, *Annu. Rev. Cell Dev. Biol.* **2005**, *21*, 81 – 103.
[11] For recent reviews of glycoprotein synthesis, see: a) D. P. Gamblin, E. M. Scanlan, B. G. Davis, *Chem. Rev.* **2009**, *109*, 131 – 163; b) C. Unverzagt, Y. Kajihara, *Chem. Soc. Rev.* **2013**, *42*, 4408 – 4420.
[12] For recent achievements on glycoprotein synthesis, see: a) I. Sakamoto, K. Tezuka, K. Fukae, K. Ishii, K. Taduru, M. Maeda, M. Ouchi, K. Yoshida, Y. Nambu, J. Igarashi, N. Hayashi, T. Tsuji, Y. Kajihara, *J. Am. Chem. Soc.* **2012**, *134*, 5428 – 5431; b) M. Murakami, R. Okamoto, M. Izumi, Y. Kajihara, *Angew. Chem.* **2012**, *124*, 3627 – 3632; *Angew. Chem. Int. Ed.* **2012**, *51*, 3567 – 3572; c) H. Hojo, H. Tanaka, M. Hagiwara, Y. Asahina, A. Ueki, H. Katayama, Y. Nakahara, A. Yoneshige, J. Matsuda, Y. Ito, Y. Nakahara, *J. Org. Chem.* **2012**, *77*, 9437; d) O. Boutureira, G. J. L. Bernardes, M. Fernandez-Gonzalez, D. C. Anthony, B. G. Davis, *Angew. Chem.* **2012**, *124*, 1461 – 1465; *Angew. Chem. Int. Ed.* **2012**, *51*, 1432 – 1436; e) V. Ullmann, M. Rädisch, I. Boos, J. Freund, C. Pöhner, S. Schwarzinger, C. Unverzagt, *Angew. Chem.* **2012**, *124*, 11734 – 11738; *Angew. Chem. Int. Ed.* **2012**, *51*, 11566 – 11570; f) P. Wang, B. Aussedat, Y. Vohra, S. J. Danishefsky, *Angew. Chem.* **2012**, *124*, 11739 – 11743; *Angew. Chem. Int. Ed.* **2012**, *51*, 11571 – 11575.
[13] T. Ackrill, D. W. Anderson, D. Macmillan, *Pept. Sci.* **2010**, *94*, 495 – 503.
[14] For strategy using S protection of the N-to-S acyl transfer devices and application to protein syntheses, see Ref. [6a, 9b], and [9c].
[15] M. Villain, J. Vizzavona, K. Rose, *Chem. Biol.* **2001**, *8*, 673 – 679.
[16] K. Sakamoto, K. Sato, A. Shiganaga, K. Tsuji, S. Tsuda, H. Hibino, Y. Nishiuchi, A. Otaka, *J. Org. Chem.* **2012**, *77*, 6948 – 6958.
[17] M. Schnölzer, P. Alewood, A. Jones, D. Alewood, S. B. H. Kent, *Int. J. Pep. Protein Res.* **1992**, *40*, 180 – 193.
[18] N. Fujii, A. Otaka, T. Watanabe, A. Okamachi, H. Tamamura, H. Yajima, Y. Inagaki, M. Nomizu, K. Asano, *J. Chem. Soc. Chem. Commun.* **1989**, 283 – 284.
[19] H. Rohde, J. Schmalisch, Z. Harpaz, F. Diezmann, O. Seitz, *ChemBioChem* **2011**, *12*, 1396 – 1400.

[20] E. C. B. Johnson, S. B. H. Kent, *J. Am. Chem. Soc.* **2006**, *128*, 6640 – 6646.
[21] T. Kawakami, S. Aimoto, *Tetrahedron* **2009**, *65*, 3871 – 3877.
[22] D. Ravasi, M. Masserini, G. Vecchio, Y.-T. Li, S.-C. Li, *Neurochem. Res.* **2002**, *27*, 785 – 792.
[23] The *E. coli*-expressed recombinant carbohydrate-free GM2AP protein was already reported to be in every respect at least as effective as the purified native glycosylated GM2AP. See: H. Klima, A. Klein, G. van Echten, G. Schwarzmann, K. Suzuki, K. Sandhoff, *Biochem. J.* **1992**, *292*, 571 – 576.

