JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

Convergent Solid-Phase Synthesis of Macromolecular MUC1 Models Truly Mimicking Serum Glycoprotein Biomarkers of Interstitial Lung Diseases

Naoki Ohyabu,[†,§] Kiyoshi Kakiya,[‡] Yasuhiro Yokoi,[§] Hiroshi Hinou,[§,∥] and Shin-Ichiro Nishimura*[,§,∥]

[†]Shionogi & Company, Ltd., 3-1-1 Futaba-cho, Toyonaka-shi, Osaka 561-0825, Japan

[‡]Shionogi & Company, Ltd., 2-1-3 Kuiseterashima, Amagasaki 660-0813, Japan

[§]Graduate School of Life Science and Faculty of Advanced Life Science, Hokkaido University, N21, W11, Kita-ku, Sapporo 001-0021, Japan

[∥]Medicinal Chemistry Pharmaceuticals, Co., Ltd., N21, W12, Kita-ku, Sapporo 001-0022, Japan

Supporting Information

ABSTRACT: Synthetic macromolecular MUC1 glycopeptides have been used to unravel molecular mechanisms in antibody recognition of disease-specific epitopes. We have established a novel synthetic strategy for MUC1 tandem repeats having complex O-glycosylation states at each repeating unit based on convergent solid-phase fragment condensation under microwave irradiation. We have accomplished the synthesis of 77 amino acid MUC1 glycopeptides (MW = 12 759) having three major antigenic O-glycoforms [Tn, core 1 (T), and core 2 structures] at 10 designated positions out of 19 potential O-glycosylation sites. We demonstrate that the macromolecular MUC1 glycopeptide displaying the essential glycopeptidic neoepitope Pro-Asp-Thr(sialyl-T)-Arg-Pro-Ala-Pro at two different tandem repeats is an excellent serum MUC1 model showing ideal stoichiometric binding with anti-KL6/MUC1 antibody in the sandwich ELISA to quantify human serum KL6/MUC1 levels as a critical biomarker of interstitial lung diseases.

MUC1 is an endothelial cell membrane-bound glycoprotein having a variable number of tandem repeating extracellular domains.[1] Each repeat is composed of a highly conserved 20-mer peptide unit, Pro-Pro-Ala-His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala, that contains five Ser/Thr residues with potential to be glycosylated. It is well-documented that MUC1 is highly overexpressed in over 90% of breast cancer and many other cancers such as hepatocellular carcinoma, pancreatic cancer, colon cancer, and ovarian carcinoma.[2,3] CA15-3 is one of the most important serum MUC1 biomarkers indicating the early recurrence of breast cancer and the efficacy of the treatments for metastatic breast cancer.[4]

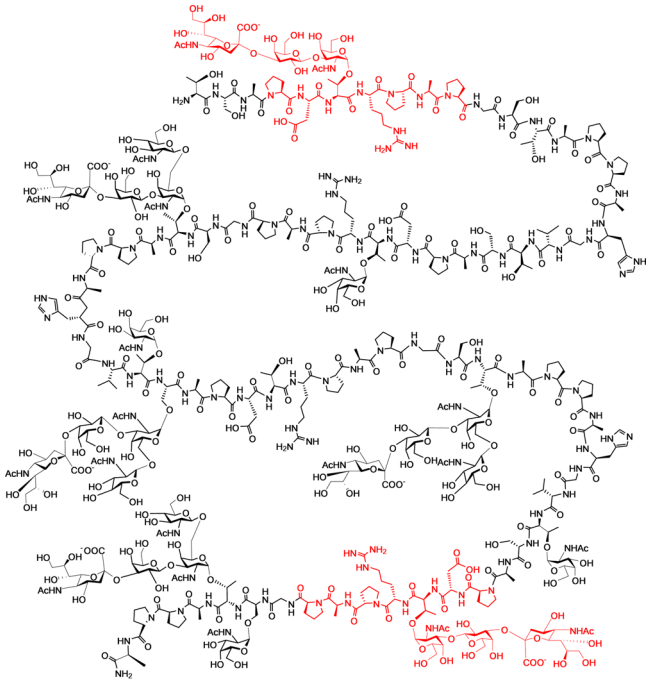
Krebs von den Lungen-6 (KL6/MUC1) is also clinically important serum MUC1 biomarker for diagnosing various lung diseases associated with interstitial pneumonitis.[5] KL6 was approved in 1999 by the Japanese Health Insurance Program as a diagnostic marker for interstitial lung diseases (ILDs), and recent European cohorts also indicated that KL6/MUC1 is currently one of the best and most reliable serum biomarkers available for various ILD management.[6] Anti-KL6/MUC1 monoclonal antibody (anti-KL6 mAb) is a promising probe for monitoring KL6/MUC1 levels in the bloodstream.[7] In 2009, we identified for the first time the minimal and essential epitope structure recognized by anti-KL6 mAb to be the heptapeptide Pro-Asp-Thr-Arg-Pro-Ala-Pro having sialyl-T antigen [Neu5Acα(2→3)Galβ(1→3)GalNAcα1→] at the Thr residue.[8] It is important to note the accumulated evidence that anti-KL6 mAb reacts often with serum MUC1 fragments in patients suffering from various cancers in addition to the ILDs.[9] Recently it was also revealed that common anti-MUC1 mAbs such as DF3 (a probe antibody for CA15-3) and SM3 directing serum MUC1 fragments cannot discriminate any *O*-glycan structures attached at this universal epitope region, notably the broad binding characteristics toward the variously glycosylated PDTR motif.[10] Surprisingly, DF3 and SM3 react even with the KL6 epitope, a critical marker of interstitial pneumonia, and their binding affinities are strongly influenced by the multiple O-glycosylation states at neighboring Ser/Thr residues within the 20-mer MUC1 tandem repeating unit.[11] It is clear that the broad binding of DF3 to the glycans in the epitope region of serum MUC1 leads to inaccurate conclusions in diagnosis.

The serum levels of both CA15-3 and KL6 are commonly represented as the relative concentration (units/mL or milliunits/well) defined by using a calibrator, the standard MUC1 glycoproteins prepared from cultured human cell lines.[12] Importantly, these MUC1 calibrators are heterogeneous mixtures of MUC1 tandem repeats containing variable numbers of epitope regions displaying unidentified *O*-glycans,[13] even though the calibrators are feasible in ordinary diagnosis. It should be emphasized that the present assays using MUC1 calibrators do not provide any absolute concentrations of the serum KL6/MUC1 or CA15-3 levels.

In the present study, we challenged the synthesis of 77 amino acid MUC1 glycopeptide **1** (MW = 12 759) that truly mimicks human serum glycoproteins having multiple and highly complicated *O*-glycans, as shown in Figure 1. It was thought that the monodispersed macromolecular MUC1 tandem repeat

Received: May 14, 2016
Published: June 24, 2016

DOI: 10.1021/jacs.6b04973
J. Am. Chem. Soc. 2016, 138, 8392−8395



1

Figure 1. The 77 amino acid MUC1 tandem repeat **1** as a designated calibrator for anti-KL6/MUC1 mAb. Two glycopeptide regions (shown in red) indicate the essential epitope of anti-KL6/MUC1 mAb.

1 displaying the KL6 epitope at two different regions would be an ideal MUC1 model that would enable quantitation of the serum KL6/MUC1 level as an absolute molar concentration. MUC1 model **1** contains Tn, core 1-, and core 2-type structures at 10 (eight Thr and two Ser residues) out of 19 potential O-glycosylation sites (13 Thr and six Ser residues), in which two Thr residues among four PDTR motifs are modified with sialyl-T antigen, as shown in red in Figure 1. The other eight *O*-glycans (four Tn and four core 2-type tetrasaccharides), notably "non-antigenic structures" for the anti-KL6 mAb, are distributed between the two epitope regions at the N- and C-terminal tandem repeats.

We hypothesized that solid-phase condensation of the key peptide segments and glycopeptide-bound resins may facilitate the convergent synthesis of macromolecular and highly complicated glycopeptides as models for the serum MUC1 fragments shedding from endothelial cell surfaces. As shown in Figure 2, it was considered that compound **1** could be synthesized using the three non-glycosylated peptide segments A (**2**, blue), B (**3**, purple), and C (**4**, green) along with Ser and Thr residues as potential O-glycosylation sites (red). In this approach, the Thr and Ser residues must be placed definitely between two of the above three peptide segments. The merit of this method is clear because the combined use of various Fmoc-Ser/Thr derivatives having the designated *O*-glycans[8,11] and non-glycosylated Fmoc-Ser/Thr provides flexibility of the molecular design at all of the potential O-glycosylation sites and the reactions can be performed independently from the fragment condensation. Four cycles of this solid-phase synthesis under microwave irradiation[14] and subsequent treatment with 2,3-(*O*)-sialyltransferase[11,15,16] and CMP-Neu5Ac affords **1**.

To achieve the total synthesis of macromolecular MUC1 glycopeptide **1**, it is important to note that epimerization at the activated C-terminal residue of segments B and C during the

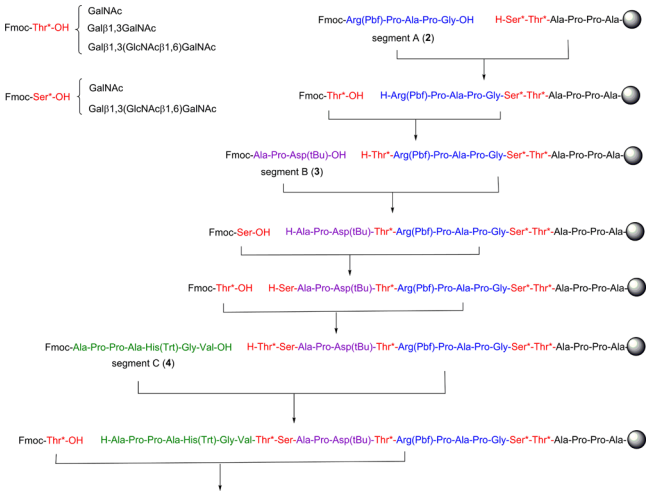


Figure 2. Synthetic strategy for macromolecular MUC1 glycopeptides. A typical protocol based on convergent solid-phase synthesis using the designated segments A (**2**), B (**3**), and C (**4**) and naked/glycosylated Fmoc-Ser/Thr derivatives is shown, representing the synthesis of the first 25 amino acids, one cycle of the 77-mer MUC1 glycopeptide **1**.

condensation remains a critical problem in general procedures for solid-phase peptide synthesis.[17] To overcome this risk in our concept outlined in Figure 2, we aimed to assess the epimerization of various activation methods[18] and identify an optimal protocol suited for the present synthesis by testing them in model coupling reactions. First, segment B [**3**, Fmoc-Ala-Pro-Asp(OtBu)-OH] (3 molar equiv) was allowed to react with H-Thr(tBu)-Arg(Pbf)-Pro-Ala-linkamide PEGA resin under microwave irradiation at 40 or 50 °C for 10 min, and the product peptide resin was treated with cleavage cocktails. The resulting product, the heptapeptide (H-Ala-Pro-Asp-Thr-Arg-Pro-Ala-NH_2), was characterized by common reversed-phase HPLC (Figures S5−S7).

As summarized in Table S1, it was demonstrated that the reaction promoted by means of *N*,*N*-diisopropylcarbodiimide (DIC) with 3,4-dihydro-3-hydroxy-4-oxobenzotriazine (HOOBt) or Cl-HOOBt in CH_2Cl_2 (entries 6 and 9) gave the best result (>99% yield and <0.3% epimerization) among all of the conditions tested. On the contrary, reactions conducted in aprotic solvents (DMF/DMSO) afforded the products with lower stereoselectivity (entries 1−3, 10, and 11). Since the coupling reactions performed by a protocol using DIC/HOOBt/CH_2Cl_2 did not give any byproducts, coupling reactions accelerated under microwave irradiation appeared to suppress both racemization and side reactions due to Lossen rearrangement.[19,20] However, the mechanism of the highly efficient and stereoselective reactions achieved in the cases of HOOBt and Cl-HOOBt remains to be elucidated, while HOBt and HOAt have also been widely used in general peptide synthesis as reagents showing a racemization-suppressing effect.

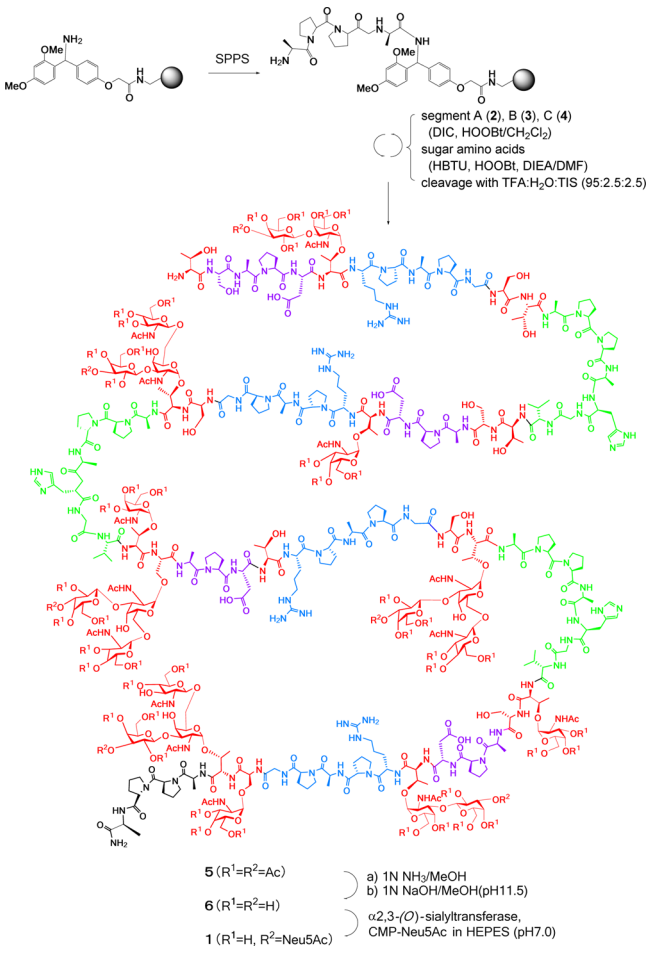
Next, the feasibility of these optimized conditions in the reaction with glycopeptide-bound resins was proved by two model reactions: (a) reactions between segment B (**3**) and H-Thr(R)-Arg(Pbf)-Pro-Ala-linkamide PEGA resin bearing three glycoforms represented as "R" [GalNAc (Tn), core 1 (T), and core 2 trisaccharide] at the N-terminal Thr residue and (b) reactions between segment C [**4**, Fmoc-Ala-Pro-Pro-Ala-His-(Trt)-Gly-Val-OH] and various glycopeptide bound-resins [H-Thr(R)-Ser(tBu)-Ala-link amide resin]. Analysis of the glycopeptides revealed that the reaction using DIC/HOOBt/CH_2Cl_2

DOI: 10.1021/jacs.6b04973
J. Am. Chem. Soc. 2016, 138, 8392−8395

is a promising protocol without any significant isomerization at the linkages formed between all segments and glycopeptide-bound resins considered in the convergent synthesis of 77-mer MUC1 glycopeptide **1** (Figures S8–S15).

The convergent solid-phase synthesis of 77 amino acid MUC1 glycopeptide **6** was performed under the optimized conditions (DIC/HOOBt/CH_2Cl_2) with microwave irradiation using peptide segments **2**–**4** and Fmoc-Ser/Thr derivatives having a sugar moiety (Scheme 1). The consecutive condensation

Scheme 1. Synthetic Scheme for MUC1 Glycopeptide 1



reactions of the peptide segments/sugar amino acids with glycopeptide-bound resins proceeded smoothly, and subsequent de-O-acetylation of intermediate **5** gave asialo-MUC1 (**6**) in 4.2% overall yield from the first solid-supported coupling reaction. Enzymatic sialylation of precursor **6** (2.2 mg) was performed efficiently and gave **1** in 83% isolated yield (2.0 mg) with high chiral purity[21] (Figures 3, S16, and S17).

MUC1 models **1** and **6** were assessed using a commercially available enzyme immunoassay protocol for measuring serum KL6 levels.[6] The binding profiles of compounds **1** and **6** with anti-KL6 mAb were monitored using anti-KL6 mAb conjugated with horseradish peroxidase (HRP) compared to that of the KL6/MUC1 glycoproteins recommended as a calibration standard in the assay kit. As shown in Figure 4A, it was clearly demonstrated that compound **1** exhibits stoichiometrically an ideal binding profile with the antibody compared to that of the calibrator KL6/MUC1. As anticipated, asialo-MUC1 glycopeptide **6** did not interact with anti-KL6 mAb. The saturated binding

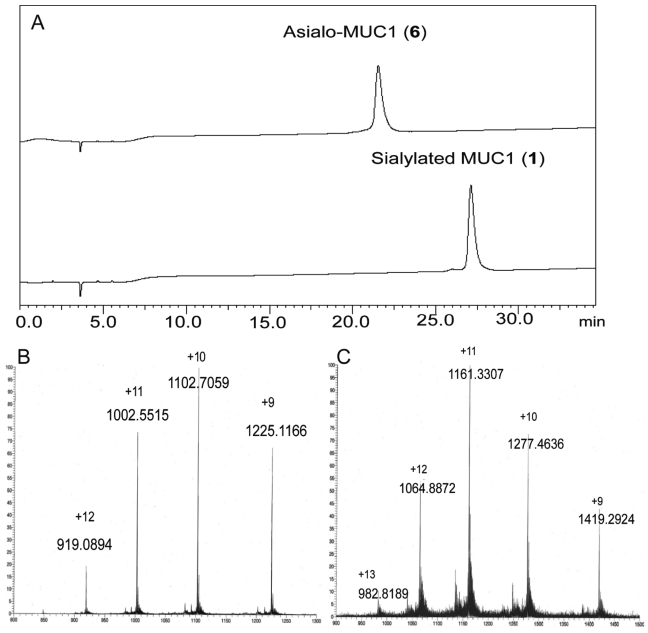


Figure 3. HPLC (A) and ESI-MS of the 77 amino acid MUC1 tandem repeats **6** (B) and **1** (C).

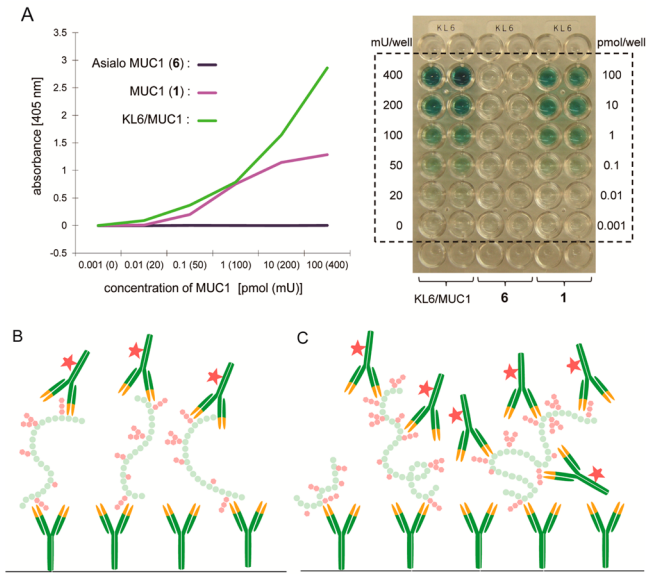


Figure 4. Binding characteristics of anti-KL6 mAb with the macromolecular MUC1s **1** and **6** and the calibration standard KL6/MUC1. (A) Sandwich ELISA obtained using the EITEST KL6 kit, showing the binding profiles of anti-KL6 mAb with **1** and **6** and the calibrator KL6/MUC1. (B) Representative binding mode showing the stoichiometric binding of anti-KL6 mAb with **1**. (C) Highly complicated binding features of anti-KL6 mAb with the calibrator KL6/MUC1 in the kit.

curve observed in the MUC1 model **1** indicated that two KL6 epitope regions within this 77-mer glycopeptide can be consumed mostly for the interaction with two antibody molecules in the sandwich assay (Figure 4B). In contrast, the KL6/MUC1, heterogeneous polydispersed MUC1 fragments from cultured human cells, appeared to show exponentially enhanced binding with antibodies, suggesting highly complicated mechanisms in the binding with HRP-labeled anti-KL6 mAbs (Figure 4C). This result clearly shows that monodispersed macromolecular MUC1 model **1** (MW = 12 759) for the first

DOI: 10.1021/jacs.6b04973
J. Am. Chem. Soc. 2016, 138, 8392–8395

time enables quantitation of the human serum KL6/MUC1 level as the molar concentration (e.g., pmol/well or nmol/mL).

Herein we have described a synthetic strategy for macromolecular mucin glycoprotein models with highly complicated O-glycosylation states based on convergent solid-phase fragment condensation of the designated peptide segments and sugar amino acid derivatives. Synthetic macromolecular MUC1 glycopeptides truly mimicking human serum MUC1 fragments elicited molecular mechanisms in the interaction between anti-KL6 mAb and its essential epitope motif for ILDs. The merit of this strategy is clear because the repertories of synthetic glycoproteins with highly complex glycans can be expanded when combined complementarily with the powerful synthetic methods directing macromolecular peptides and glycopeptides.[22] In view of the accumulated evidence that epimerization of Asp and Ser residues is often detected in proteins implicated in some neurodegenerative diseases,[23] it is of emerging importance to pay particular attention to the effects of the chiral purity of synthetic peptides/glycopeptides on the focused biological activities to be tested. The present results may motivate us to attempt the synthesis of various mucin domains to obtain insights into the significance of post-translational glycosylations[10,11,24] and the discovery of disease-relevant glycopeptidic epitopes of antibodies/vaccines against cancers, neurodegenerative diseases, and inflammatory diseases.[25]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b04973.

Experimental details and characterization data (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*shin@sci.hokudai.ac.jp

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was partly supported by a grant (Innovation COE Program for Future Drug Discovery and Medical Care) from the Ministry of Education, Culture, Science, and Technology of Japan (MEXT) and JSPS (KAKENHI Grant 25220206).

■ REFERENCES

(1) Hattrup, C. L.; Gendler, S. J. *Annu. Rev. Physiol.* **2008**, *70*, 431−457.
(2) Kufe, D. W. *Nat. Rev. Cancer* **2009**, *9*, 874−884.
(3) Bafna, S.; Kaur, S.; Batra, S. *Oncogene* **2010**, *29*, 2893−2904.
(4) Kumpulainen, E. J.; Keskikuru, R. J.; Johansson, R. T. *Breast Cancer Res. Treat.* **2002**, *76*, 95−102.
(5) Ishikawa, N.; Hattori, N.; Yokoyama, A.; Kohno, N. *Respir. Investig.* **2012**, *50*, 3−13.
(6) Horimasu, Y.; Hattori, N.; Ishikawa, N.; Kawase, S.; Tanaka, S.; Yoshioka, K.; Yokoyama, A.; Kohno, N.; Bonella, F.; Guzman, J.; Ohshimo, J. S.; Costabel, U. *Respir. Med.* **2012**, *106*, 1756−1764.
(7) Kohno, N.; Akiyama, M.; Kyoizumi, S.; Hakoda, M.; Yamakido, K. M. *Jpn. J. Clin. Oncol.* **1988**, *18*, 203−216.
(8) Ohyabu, N.; Hinou, H.; Matsushita, T.; Izumi, R.; Shimizu, H.; Kawamoto, K.; Numata, Y.; Togame, H.; Takemoto, H.; Kondo, H.; Nishimura, S.-I. *J. Am. Chem. Soc.* **2009**, *131*, 17102−17109.
(9) Tanaka, S.; Hattori, N.; Ishikawa, N.; Shoda, H.; Takano, A.; Nishino, R.; Okada, M.; Arihiro, K.; Inai, K.; Hamada, H.; Yokoyama, A.; Kohno, N. *Int. J. Cancer* **2012**, *130*, 377−387.
(10) (a) Matsushita, T.; Takada, W.; Igarashi, K.; Naruchi, K.; Miyoshi, R.; Garcia-Martin, F.; Amano, M.; Hinou, H.; Nishimura, S.-I. *Biochim. Biophys. Acta, Gen. Subj.* **2014**, *1840*, 1105−1116. (b) Coelho, H.; Matsushita, T.; Artigas, G.; Hinou, H.; Canada, F. J.; Lo-Man, R.; Leclerc, C.; Cabrita, E. J.; Jiménez-Barbero, J.; Nishimura, S.-I.; Garcia-Martin, F.; Marcelo, F. *J. Am. Chem. Soc.* **2015**, *137*, 12438−12441.
(11) Rangappa, R.; Artigas, G.; Miyoshi, R.; Yokoi, Y.; Hayakawa, S.; Garcia-Martin, F.; Hinou, H.; Nishimura, S.-I. *MedChemComm* **2016**, *7*, 1102−1122.
(12) Lloyd, K. O.; Burchell, J.; Kudryashov, V.; Yin, B. W. T.; Taylor-Papadimitriou, J. *J. Biol. Chem.* **1996**, *271*, 33325−33334.
(13) (a) Müller, S.; Alving, K.; Peter-Katalinic, J.; Zachara, N.; Gooley, A. A.; Hanisch, F.-G. *J. Biol. Chem.* **1999**, *274*, 18165−18172. (b) Sihlbom, C.; van Dijk Härd, I.; Lidell, M. E.; Noll, T.; Hansson, G.; Bäckström, M. *Glycobiology* **2009**, *19*, 375−381. (c) Miura, Y.; Kato, K.; Takegawa, Y.; Kurogochi, M.; Furukawa, J.; Shinohara, Y.; Nagahori, N.; Amano, M.; Hinou, H.; Nishimura, S.-I. *Anal. Chem.* **2010**, *82*, 10021−10029.
(14) (a) Matsushita, T.; Hinou, H.; Kurogochi, M.; Shimizu, H.; Nishimura, S.-I. *Org. Lett.* **2005**, *7*, 877−880. (b) Naruchi, K.; Hamamoto, T.; Kurogochi, M.; Hinou, H.; Shimizu, H.; Matsushita, T.; Fujitani, N.; Kondo, H.; Nishimura, S.-I. *J. Org. Chem.* **2006**, *71*, 9609−9621.
(15) Fumoto, M.; Hinou, H.; Ohta, T.; Ito, T.; Yamada, K.; Takimoto, A.; Kondo, H.; Shimizu, H.; Inazu, T.; Nakahara, Y.; Nishimura, S.-I. *J. Am. Chem. Soc.* **2005**, *127*, 11804−11818.
(16) Matsushita, T.; Nagashima, I.; Fumoto, M.; Ohta, T.; Yamada, K.; Shimizu, H.; Hinou, H.; Naruchi, K.; Ito, T.; Kondo, H.; Nishimura, S.-I. *J. Am. Chem. Soc.* **2010**, *132*, 16651−16656.
(17) (a) Lloyd-Williams, P.; Albericio, F.; Giralt, E. *Tetrahedron* **1993**, *49*, 11065−11133. (b) Krambovitis, E.; Hatzidakis, G.; Barlos, K. *J. Biol. Chem.* **1998**, *273*, 10874−10879.
(18) (a) Barlos, K.; Gatos, D.; Schäfer, W. *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 590−593. (b) Kuroda, H.; Chen, Y. − N.; Kimura, T.; Sakakibara, S. *Int. J. Pept. Protein Res.* **1992**, *40*, 294−299. (c) Quibell, M.; Packman, L. C.; Johnson, T. *J. Chem. Soc., Perkin Trans. 1* **1996**, 1219−1225. (d) Kuge, K.; Brack, A.; Fujii, N. *Chem. - Eur. J.* **2007**, *13*, 5617−5621.
(19) König, W.; Geiger, R. *Chem. Ber.* **1970**, *103*, 2024−2033.
(20) König, W.; Geiger, R. *Chem. Ber.* **1970**, *103*, 2034−2040.
(21) (a) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K.-I. *Anal. Chem.* **1997**, *69*, 5146−5151. (b) B'Hymer, C.; Montes-Bayon, M.; Caruso, J. A. *J. Sep. Sci.* **2003**, *26*, 7−19. (c) Nishiuchi, Y.; Nishio, H.; Ishimaru, M.; Kimura, T. *Int. J. Pept. Res. Ther.* **2007**, *13*, 119−128.
(22) (a) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776−779. (b) Hojo, H.; Matsumoto, Y.; Nakahara, Y.; Ito, E.; Suzuki, Y.; Suzuki, M.; Suzuki, A.; Nakahara, Y. *J. Am. Chem. Soc.* **2005**, *127*, 13720−13725. (c) Bennett, C. S.; Dean, S. M.; Payne, R. J.; Ficht, S.; Brik, A.; Wong, C.-H. *J. Am. Chem. Soc.* **2008**, *130*, 11945−11952. (d) Okamoto, R.; Souma, S.; Kajihara, Y. *J. Org. Chem.* **2009**, *74*, 2494−2501. (e) Gaidzik, N.; Kaiser, A.; Kowalczyk, D.; Westerlind, U.; Gerlitzki, B.; Sinn, H. P.; Schmitt, E.; Kunz, H. *Angew. Chem., Int. Ed.* **2011**, *50*, 9977−9981.
(23) (a) Shapira, R.; Austin, G. E.; Mirra, S. S. *J. Neurochem.* **1988**, *50*, 69−74. (b) Tomiyama, T.; Asano, S.; Furiya, Y.; Shirasawa, T.; Endo, N.; Mori, H. *J. Biol. Chem.* **1994**, *269*, 10205−10208.
(24) (a) Hiruma-Shimizu, K.; Hosoguchi, K.; Liu, Y.; Fujitani, N.; Ohta, T.; Hinou, H.; Matsushita, T.; Shimizu, H.; Feizi, T.; Nishimura, S.-I. *J. Am. Chem. Soc.* **2010**, *132*, 14857−14865. (b) Hayakawa, S.; Koide, R.; Hinou, H.; Nishimura, S.-I. *Biochemistry* **2016**, *55*, 776−787.
(25) (a) Tarp, M. A.; Clausen, H. *Biochim. Biophys. Acta, Gen. Subj.* **2008**, *1780*, 546−563. (b) Lakshminarayanan, V.; Thompson, P.; Wolfert, M. A.; Buskas, T.; Bradley, J. M.; Pathangey, L. B.; Madsen, C. S.; Cohen, P. A.; Gendler, S. J.; Boons, G.-J. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 261−266. (c) Gaidzik, N.; Westerlind, U.; Kunz, H. *Chem. Soc. Rev.* **2013**, *42*, 4421−4442. (d) Wilson, R. M.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2013**, *135*, 14462−14472.

DOI: 10.1021/jacs.6b04973
J. Am. Chem. Soc. 2016, 138, 8392−8395