DOI: 10.1002/anie.201303073

Glycoprotein Synthesis

Chemoenzymatic Synthesis of the Immunoglobulin Domain of Tim-3 Carrying a Complex-Type N-Glycan by Using a One-pot Ligation**

*Yuya Asahina, Shigehiro Kamitori, Toshifumi Takao, Nozomu Nishi, and Hironobu Hojo**

Ligation methods enable the efficient synthesis of (glyco)-proteins with over 100 amino acid residues. At present, the native chemical ligation (NCL) reaction[1] is used almost exclusively owing to its high efficiency derived from the chemoselective coupling between the thioester of one segment with the N-terminal cysteine residue of the other segment. This realizes the protection-free ligation of the segments, which facilitates the preparation of the peptide segments by solid-phase peptide synthesis (SPPS). However, owing to the low abundance of cysteine in proteins (ca. 1.5%), the requirement of the cysteine residue at the ligation site often limits the application of the NCL method. To overcome this problem, thiol auxiliary groups as well as ligations that do not require cysteine have been developed,[2] but every method has advantages and disadvantages.

Since 1991, we have been developing a facile method for (glyco)protein synthesis based on the direct aminolysis of the N- and S-protected peptide thioester by the other segment.[3] Various (glyco)proteins have been synthesized by using the method.[4] Originally, the ligation reaction was performed by activation of the alkyl thioester by silver ions. We recently found that the reaction proceeds without the addition of silver ions when more reactive aryl thioesters are used.[4c,d] These Ag^+-promoted and Ag^+-free ligation methods have the advantage that, in principle, the ligation can be performed at any site, which enables a flexible synthetic plan. Nevertheless, as glycine residues exist abundantly in proteins (average abundance of glycine is about 7%), we seldom meet difficulty in finding appropriate ligation positions for Gly-Xaa, where epimerization-free ligation can be performed.

However, further improvements are still required to increase the general utility of our method. In the recent syntheses of (glyco)polypeptides, we used an azido group for the amino protection of the segments.[5] However, owing to its hydrophobic nature, the azido-protected thioester sometimes retains a low solubility in aqueous acetonitrile solution, which makes its purification by reversed-phase (RP) HPLC difficult. Thus, we attempted to use a more hydrophilic N-protecting group and selected the isonicotinyloxycarbonyl (*i*Noc)[6] group; it is hydrophilic and stable during the deprotection by trifluoroacetic acid (TFA), purification and the ligation steps, and it can be removed by a treatment with Zn in acetic acid.

Another issue about our method is that the sequential ligation requires the purification of the intermediates. In previous syntheses, sequential ligation in the C- to N-terminal direction was carried out.[4b,c] In this route, the 9-fluorenylmethoxycarbonyl (Fmoc) group, used for the N-terminal protection of the intermediate, has to be removed by piperidine after the ligation. To perform the second ligation reaction, piperidine has to be eliminated by purification; otherwise the remaining piperidine would consume the acid component. In 2006, Bang et al. realized a sequential one-pot native chemical ligation in the reverse N-to-C direction by using the difference in the reactivity of aryl and alkyl thioesters.[7] Chen et al. also demonstrated the sequential ligation in organic solvents using *ortho*-mercaptophenol as an aryl thioester precursor in the synthesis of small glycopeptides.[8] If the reactivity of the alkyl and aryl thioester can be regulated by silver ions, the one-pot sequential ligation can also be realized in our method, which further facilitates the synthesis.

Tim-3 is a membrane protein located on the surface of Th-1 cells and regulates the Th-1-mediated immune response.[9] Recently, it has been shown that galectin-9, one of the animal lectins, induces apoptosis of the Th-1 cells by binding to the N-glycan of the Tim-3 immunoglobulin (Ig) domain (Figure 1) and regulates the Th-1-cell-mediated immunity. However, the precise mechanism of action of galectin-9 is not yet known. To precisely study the interaction between the Tim-3 Ig domain and galectin-9, the Ig domain carrying a distinct carbohydrate structure is required. Recently, the transfer of a carbohydrate chain using an endoglycosidase, such as *endo*-β-*N*-acetylglucosaminidase from *Mucor hiemalis* (Endo-M),[10] was applied successfully in the synthesis of homogeneous glycoproteins.[11] We also succeeded in the chemoenzymatic synthesis of saposin C carrying a complex-type carbohydrate chain by using the

[*] Y. Asahina, Prof. Dr. H. Hojo
Department of Applied Biochemistry
Institute of Glycoscience
Tokai University
Kanagawa 259-1292 (Japan)
E-mail: hojo@keyaki.cc.u-tokai.ac.jp
Prof. Dr. S. Kamitori, Prof. Dr. N. Nishi
Life Science Research Center, Kagawa University
Kagawa, 761-0793 (Japan)
Prof. Dr. T. Takao
Institute for Protein Research, Osaka University
Osaka 565-0871 (Japan)

[**] We thank Prof. Dr. Yoon-Sik Lee of Seoul National University and Dr. Sun-Jong Ryoo of BeadTech (Seoul, Korea) for providing core shell resin. This work was supported by a grant-in-aid for scientific research from the Ministry of Education, Sports, Science and Technology of Japan (23380065). We thank Tokai University for their support with a grant-in-aid for high technology research.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201303073.

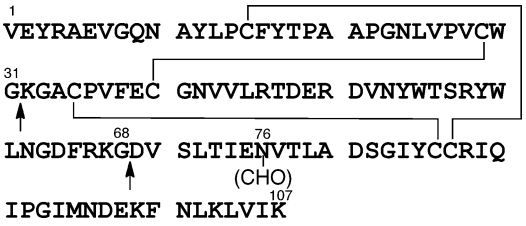


Figure 1. Structure of the Tim-3 immunoglobin (Ig) domain **1**. The arrows indicate the sites of the ligation. CHO denotes the N-linked carbohydrate.

Endo-M mutant (glycosynthase).[11 h] This method was now applied to the synthesis of the Tim-3 Ig domain carrying a homogeneous complex-type N-glycan **1**.

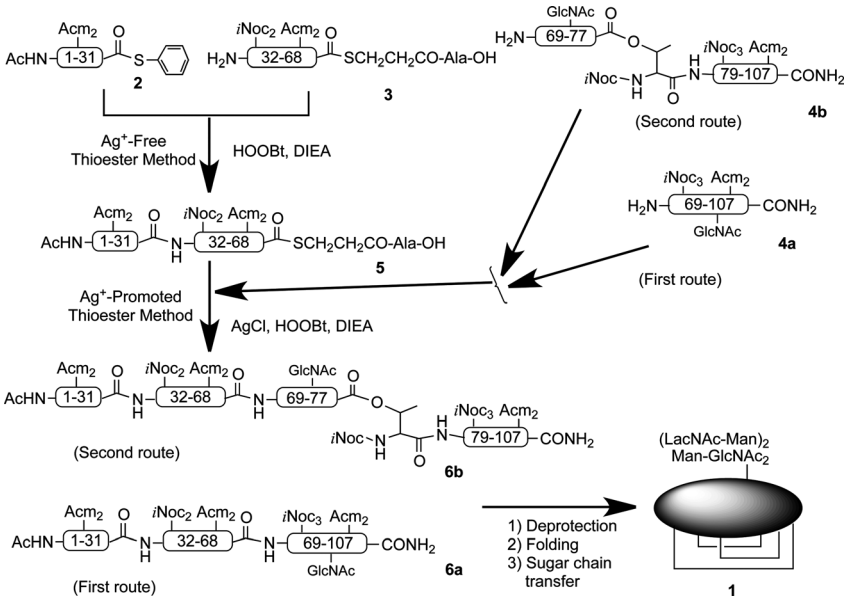
The synthetic route of the target glycoprotein **1** is shown in Scheme 1. The entire sequence of the Tim-3 Ig domain (1–107) was divided into three segments, (1–31) **2**, (32–68) **3**, (69–107) **4a** (or **4b** in the second route), and the sequential ligation in the N-to-C-terminal direction was performed using aryl and alkyl thioesters. After removal of the protecting groups and the disulfide bond formation, enzymatic transfer of the complex-type sugar was performed to obtain the desired product **1**.

The synthesis of each segment was carried out by solid-phase synthesis. The N-terminal peptide **2** was prepared as an aryl thioester by Fmoc SPPS using the described *N*-alkylcysteine-assisted thioesterification method.[12] In contrast, the intermediate alkyl thioester **3** was synthesized by the fast *tert*-butoxycarbonyl (Boc) SPPS[13] using a stable thioester linker.[14] To examine the usefulness of the *i*Noc group for the protection of side chain amino groups, all the Lys residues were introduced using Boc-Lys(*i*Noc)-OH. The preparation of the glycopeptide **4a** was also performed by using Boc chemistry. Asn[76] was introduced using the *O*-benzyl-protected Boc-Asn(GlcNAc)-OH to realize the enzymatic transfer of the complex-type sugar to this GlcNAc after the chain assembly.

Segments **2**, **3**, and **4a** were soluble in aqueous acetonitrile when supplemented by guanidine HCl (Gu·HCl) and could be purified by RP HPLC. In a control experiment, peptide **4c** (see the Supporting Information), the form of **4a** in which the Lys side chains are carrying azido groups, was synthesized. However, the solubility of segment **4c** under the RP-HPLC purification conditions is extremely low and it was not eluted from the HPLC column at all (data not shown). These data demonstrate the suitability of the *i*Noc group, which is the amino protecting group in the thioester method.

For the ligation following the first route, the three segments were condensed (Scheme 1). Segments **2** and **3** were condensed by using the Ag^+-free ligation method,[4c,d] keeping the alkyl thioester of the segment **3** inactive. The reaction proceeded efficiently without significant side reactions and was almost complete within 12 h (Figure 2). Thus, perfect chemoselectivity between the alkyl and aryl thioesters was achieved in the absence of silver ions. Without isolation of the intermediate **5**, the segment **4a** was added to the reaction mixture and the second ligation was carried out in a one-pot manner using AgCl as an activator. After 6 h, the second reaction also proceeded efficiently to give the protected Ig domain **6a** (Figure 2).

The deprotection and folding of the glycoprotein **6a** were carried out as shown in Scheme 2. The *i*Noc groups were removed using zinc dust in an aqueous solution containing mercaptopropionic acid (MPA) and 6 M Gu·HCl. The MPA was used as acid as well as the Zn^{2+} scavenger. Removal of the Acm groups was then carried out by using silver ions. These reactions proceeded efficiently to give the completely deprotected product **8a** in good purity. Finally, the folding and the disulfide bond formation was carried out in the redox buffer



Scheme 1. One-pot synthesis of the Ig domain of Tim-3 (**1**). HOOBt = 3,4-dihydro-3-hydroxy-4-oxo-1,2,3-benzotriazine, Acm = acetamidomethyl, DIEA = diisopropylethylamine.

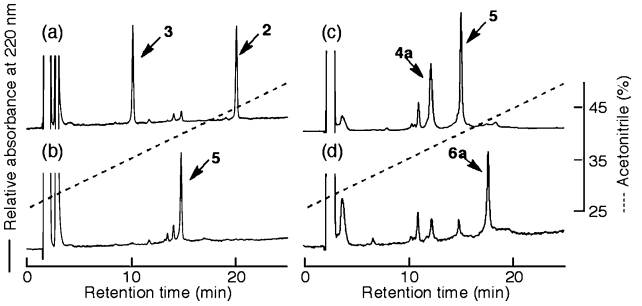
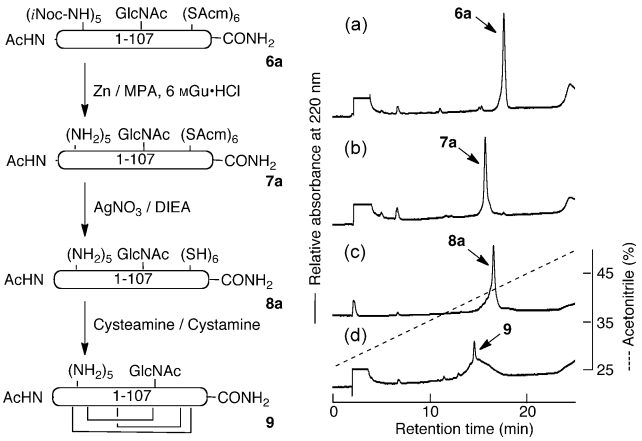


Figure 2. RP HPLC profile of the ligation reaction; a, b) reaction mixture for the synthesis of peptide **5** at $t = 0$ (a) and $t = 12$ h (b); c, d) reaction mixture for the synthesis of peptide **6a** at $t = 0$ (c) and $t = 6$ h (d). In each case, a linear gradient starting from 25% acetonitrile at $t_{ret} = 0$ min to 50% acetonitrile at $t_{ret} = 25$ min was applied.



Scheme 2. Synthetic route to glycopolypeptide **9** (first route). In (a–d) a linear gradient starting from 25% acetonitrile at $t_{ret}=0$ min to 50% acetonitrile at $t_{ret}=25$ min was applied.

overnight at 4°C. However, the folding efficiency to **9** was moderate (Scheme 2d), which seemed to be due to the low solubility of the peptide **8a**.

To improve the solubility of the glycopolypeptide **8a**, we examined the introduction of an O-acylisopeptide structure (second route), which enables an increase in the solubility of the peptide segments and can be easily converted to the native peptide bond under neutral conditions.[15] We successfully applied this method previously to the chemoenzymatic synthesis of the hydrophobic glycoprotein saposin C by using NCL.[10h] In the present synthesis, the isopeptide structure was introduced to the C-terminal segment **4a**, which retained the lowest solubility among **2**, **3**, and **4a** during the RP HPLC purification. In contrast to the synthesis of saposin C, where the amino group of the isopeptide moiety was free and thus the O-to-N acyl shift occurred during the ligation, in this synthesis the *i*Noc protecting group was installed at this site to keep the isopeptide structure even after the ligation. The preparation of peptide **4b** having an isopeptide structure at Val^{77}-Thr^{78} was carried out by introducing *i*Noc-Thr(Fmoc-Val)-OH during the solid-phase synthesis. The solubility of peptide **4b** was significantly higher than that of glycopeptide **4a**, thereby resulting in about a four times higher yield of isolated peptide **4b** than that of **4a**.

The one-pot ligation was carried out as shown in Scheme 1 (second route). The reaction also proceeded efficiently and the glycopolypeptide **6b** was successfully obtained (Figure SI1 in the Supporting Information). After removal of the *i*Noc and Acm groups, the completely deprotected peptide **8b**, which still has an isopeptide structure at Val^{77}-Thr^{78} was obtained. The folding and the disulfide bond formation were then carried out. In contrast to glycopolypeptide **8a**, **8b** easily dissolved in 6 M urea. The solution was then diluted with the redox buffer and kept at 4°C overnight. The desired product **9** was obtained in a significantly higher purity than obtained in the first route, thus demonstrating the importance of complete solubilization prior to the disulfide bond formation (Figure 3). To confirm that the isopeptide bond was converted to the amide bond, the glycopolypeptide **8b** was dissolved in the same folding buffer without cystamine and cysteamine and kept at 4°C. As shown in Figure SI3 in the Supporting Information, the peptide **8b** was converted to peptide **8a** with a half-life of about 30 min.

The transglycosylation reaction was carried out according to the synthesis of saposin C (Scheme 3).[10h] The Ig domain

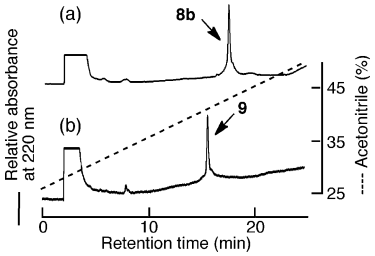
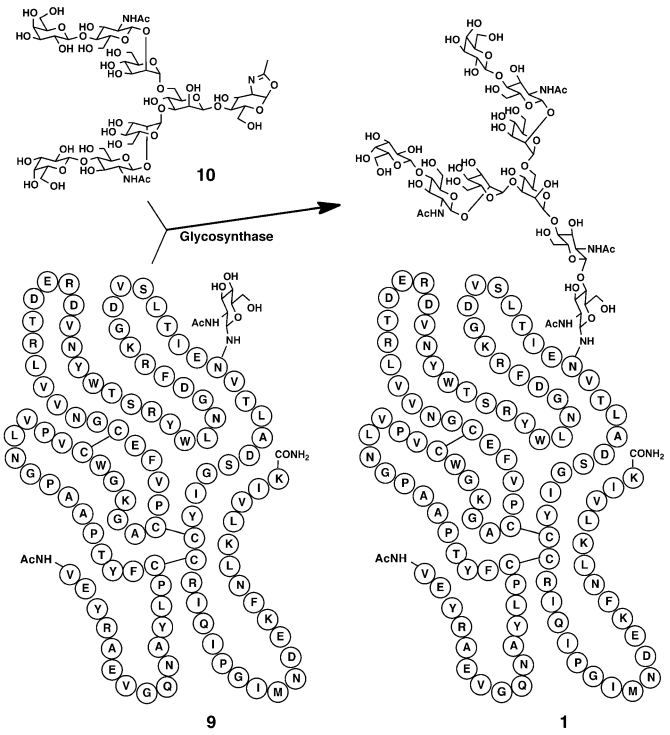


Figure 3. RP HPLC profile of the folding of glycopolypeptide **8b** at $t=0$ (a) and $t=12$ h (b). In (a, b) a linear gradient starting from 25% acetonitrile at $t_{ret}=0$ min to 50% acetonitrile at $t_{ret}=25$ min was applied.

carrying GlcNAc **9** and the synthetic sugar oxazoline **10** (50 equiv to **9**), which was prepared according to the previously described procedure,[11h] were reacted in a phosphate buffer containing 2 M urea (pH 7.0) in the presence of glycosynthase. The solution was kept at room temperature for 6 h. RP HPLC analysis of the reaction mixture showed the appearance of a peak of a slightly hydrophilic compound (Figure SI4 in the Supporting Information). The matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass analysis of this peak showed a value that corresponds well with the theoretical one of the desired



Scheme 3. Enzymatic transfer of a complex-type sugar.

product, thus indicating the success of the synthesis (see the Supporting Information). The transfer efficiency was about 30%.

To prove the correct folding, the disulfide pairings of product **9** were determined (see the Supporting Information). Owing to the lack of appropriate enzymatic cleavage sites in the Ig domain, five sequential enzyme digestions were performed. Among the two major fragments, one proved to have a disulfide bond at Cys^{29}-Cys^{40}, as shown by mass analysis. The disulfide bond pairs in the other fragment were determined by comparing the retention time on HPLC with that of synthetic samples as well as by MALDI MS/MS analysis. The results proved the disulfide bonds at Cys^{15}-Cys^{87} and Cys^{35}-Cys^{86}.

The secondary structure of the Ig domain was next characterized by CD spectroscopy. The products **1** and **9** showed typical spectra of the Ig-like β-sheet-rich structure, indicating the correct folding of the products (Figure 4).

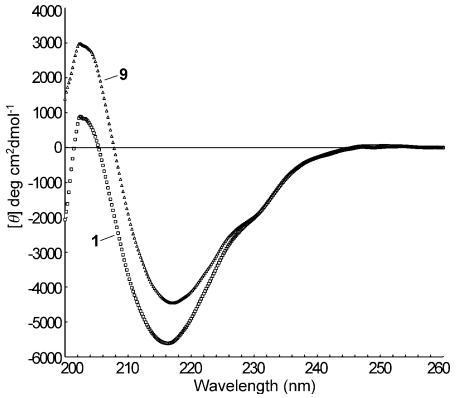


Figure 4. CD spectra of glycoproteins **1** (□) and **9** (△).

The immunoreactivity of product **9** was analyzed by using a sandwich ELISA assay with anti-Tim-3 antibodies (Figure SI7 in the Supporting Information). About 60% of the product **9** was recognized by the antibodies, thus indicating that it is folded to form correct epitope structures. The decreased ratio of the recognition might be due to its reduced size in peptide and/or carbohydrate chains compared to the entire Tim-3 protein.

In conclusion, we successfully performed the one-pot sequential ligation using the aryl and alkyl thioesters to obtain the Tim-3 Ig domain carrying GlcNAc. The *i*Noc group proved to be favorable for the amino group protection in this ligation. The introduction of the O-acylisopeptide structure markedly enhanced the solubility of the unfolded glycoprotein, which enabled the efficient folding. The obtained glycoprotein was transformed into its complex-type homogeneous glycoform by using the enzymatic sugar chain transfer by a glycosynthase. The interaction between the Tim-3 Ig domain and galectin-9 is under investigation.

Received: April 12, 2013
Revised: May 29, 2013
Published online: July 19, 2013

Keywords: glycoproteins · ligations · peptide thioester · synthesis design

[1] 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.

[2] a) L. E. Canne, S. J. Bark, S. B. H. Kent, *J. Am. Chem. Soc.* **1996**, *118*, 5891–5896; b) J. Offer, P. E. Dawson, *Org. Lett.* **2000**, *2*, 23–26; c) D. W. Low, M. G. Hill, M. R. Carrasco, S. B. H. Kent, P. Botti, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6554–6559; d) T. Kawakami, K. Akaji, S. Aimoto, *Org. Lett.* **2001**, *3*, 1403–1405; e) P. Botti, M. R. Carrasco, S. B. H. Kent, *Tetrahedron Lett.* **2001**, *42*, 1831–1833; f) L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526–533; g) J. Offer, C. N. C. Boddy, P. E. Dawson, *J. Am. Chem. Soc.* **2002**, *124*, 4642–4646; h) T. Kawakami, S. Aimoto, *Tetrahedron Lett.* **2003**, *44*, 6059–6061; i) D. Macmillan, D. W. Anderson, *Org. Lett.* **2004**, *6*, 4659–4662; j) B. Wu, J. Chen, J. D. Warren, G. Chen, Z. Hua, S. J. Danishefsky, *Angew. Chem.* **2006**, *118*, 4222–4231; *Angew. Chem. Int. Ed.* **2006**, *45*, 4116–4125; k) A. Brik, S. Ficht, Y. Y. Yang, C. S. Bennett, C. H. Wong, *J. Am. Chem. Soc.* **2006**, *128*, 5626–5627; l) A. Saporito, D. Marasco, A. Chambery, P. Botti, S. M. Monti, C. Pedone, M. Ruvo, *Biopolymers* **2006**, *83*, 508–518; m) B. L. Pentelute, S. B. H. Kent, *Org. Lett.* **2007**, *9*, 687–690; n) Y. Y. Yang, S. Ficht, A. Brik, C. H. Wong, *J. Am. Chem. Soc.* **2007**, *129*, 7690–7701; o) D. Crich, A. Banerjee, *J. Am. Chem. Soc.* **2007**, *129*, 10064–10065; p) Q. Wan, S. J. Danishefsky, *Angew. Chem.* **2007**, *119*, 9408–9412; *Angew. Chem. Int. Ed.* **2007**, *46*, 9248–9252; q) C. Chatterjee, R. K. McGinty, J. P. Pellois, T. W. Muir, *Angew. Chem.* **2007**, *119*, 2872–2876; *Angew. Chem. Int. Ed.* **2007**, *46*, 2814–2818; r) R. Okamoto, Y. Kajihara, *Angew. Chem.* **2008**, *120*, 5482–5486; *Angew. Chem. Int. Ed.* **2008**, *47*, 5402–5406; s) C. Haase, H. Rohde, O. Seitz, *Angew. Chem.* **2008**, *120*, 6912–6915; *Angew. Chem. Int. Ed.* **2008**, *47*, 6807–6810; t) G. L. Thomas, R. J. Payne, *Chem. Commun.* **2009**, 4260–4262; u) H. Hojo, C. Ozawa, H. Katayama, A. Ueki, Y. Nakahara, Y. Nakahara, *Angew. Chem.* **2010**, *122*, 5446–5449; *Angew. Chem. Int. Ed.* **2010**, *49*, 5318–5321.

[3] a) H. Hojo, S. Aimoto, *Bull. Chem. Soc. Jpn.* **1991**, *64*, 111–117; b) S. Aimoto, *Biopolymers* **1999**, *51*, 247–265.

[4] a) H. Hojo, E. Haginoya, Y. Matsumoto, Y. Nakahara, K. Nabeshima, B. P. Toole, Y. Watanabe, *Tetrahedron Lett.* **2003**, *44*, 2961–2964; b) H. Hojo, Y. Matsumoto, Y. Nakahara, E. Ito, Y. Suzuki, M. Suzuki, A. Suzuki, Y. Nakahara, *J. Am. Chem. Soc.* **2005**, *127*, 13720–13725; c) H. Hojo, Y. Murasawa, H. Katayama, T. Ohira, Y. Nakahara, Y. Nakahara, *Org. Biomol. Chem.* **2008**, *6*, 1808–1813; d) H. Katayama, H. Hojo, I. Shimizu, Y. Nakahara, Y. Nakahara, *Org. Biomol. Chem.* **2010**, *8*, 1966–1972.

[5] H. Katayama, H. Hojo, T. Ohira, Y. Nakahara, *Tetrahedron Lett.* **2008**, *49*, 5492–5494.

[6] D. F. Veber, W. J. Paleveda, Jr., Y. C. Lee, R. Hirschmann, *J. Org. Chem.* **1977**, *42*, 3286–3288.

[7] D. Bang, B. L. Pentelute, S. B. H. Kent, *Angew. Chem.* **2006**, *118*, 4089–4092; *Angew. Chem. Int. Ed.* **2006**, *45*, 3985–3988.

[8] G. Chen, Q. Wan, Z. Tan, C. Kan, Z. Hua, K. Ranganathan, S. J. Danishefsky, *Angew. Chem.* **2007**, *119*, 7527–7531; *Angew. Chem. Int. Ed.* **2007**, *46*, 7383–7387.

[9] E. Cao, X. Zang, U. A. Ramagopal, A. Mukhopadhaya, A. Fedorov, E. Fedorov, W. D. Zenchek, J. W. Lary, J. L. Cole, H. Deng, H. Xiao, T. P. DiLorenzo, J. P. Allison, S. G. Nathenson, S. C. Almo, *Immunity* **2007**, *26*, 311–321.

[10] a) K. Yamamoto, S. Kadowaki, M. Fujisaki, H. Kumagai, T. Tochikura, *Biosci. Biotechnol. Biochem.* **1994**, *58*, 72 – 77; b) K. Yamamoto, S. Kadowaki, J. Watanabe, H. Kumagai, *Biochem. Biophys. Res. Commun.* **1994**, *203*, 244 – 252.

[11] a) M. Mizuno, K. Haneda, R. Iguchi, I. Muramoto, T. Kawakami, S. Aimoto, K. Yamamoto, T. Inazu, *J. Am. Chem. Soc.* **1999**, *121*, 284 – 290; b) B. Li, Y. Zeng, S. Hauser, H. Song, L. X. Wang, *J. Am. Chem. Soc.* **2005**, *127*, 9692 – 9693; c) H. Ochiai, W. Huang, L. Wang, *J. Am. Chem. Soc.* **2008**, *130*, 13790 – 13803; d) Y. Wei, C. Li, W. Huang, B. Li, S. Strome, L. X. Wang, *Biochemistry* **2008**, *47*, 10294 – 10304; e) M. Umekawa, W. Huang, B. Li, K. Fujita, H. Ashida, L. X. Wang, K. Yamamoto, *J. Biol. Chem.* **2008**, *283*, 4469 – 4479; f) W. Huang, C. Li, B. Li, M. Umekawa, K. Yamamoto, X. Zhang, L. X. Wang, *J. Am. Chem. Soc.* **2009**, *131*, 2214 – 2223; g) M. Umekawa, T. Higashiyama, Y. Koga, T. Tanaka, M. Noguchi, A. Kobayashi, S. Shoda, W. Huang, L. Wang, H. Ashida, K. Yamamoto, *Biochim. Biophys. Acta Gen. Subj.* **2010**, *1800*, 1203 – 1209; h) 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 – 9446.

[12] H. Hojo, Y. Onuma, Y. Akimoto, Y. Nakahara, Y. Nakahara, *Tetrahedron Lett.* **2007**, *48*, 25 – 28.

[13] M. Schnölzer, P. Alewood, A. Jones, D. Alewood, S. B. H. Kent, *Int. J. Pept. Protein Res.* **1992**, *40*, 180 – 193.

[14] H. Hojo, Y. D. Kwon, Y. Kakuta, S. Tsuda, I. Tanaka, K. Hikichi, S. Aimoto, *Bull. Chem. Soc. Jpn.* **1993**, *66*, 2700 – 2706.

[15] Y. Sohma, M. Sasaki, Y. Hayashi, T. Kimura, Y. Kiso, *Chem. Commun.* **2004**, 124 – 125.

