Glycopeptide Synthesis

DOI: 10.1002/ange.201000384

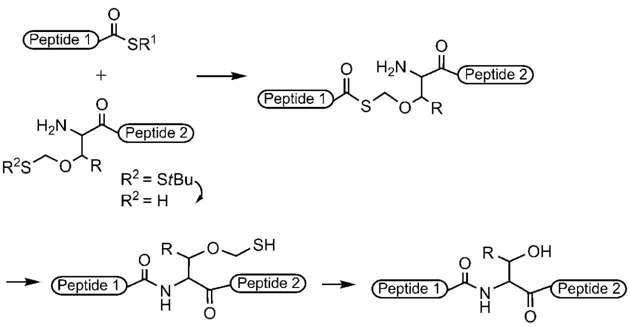
The Mercaptomethyl Group Facilitates an Efficient One-Pot Ligation at Xaa-Ser/Thr for (Glyco)peptide Synthesis**

Hironobu Hojo, Chinatsu Ozawa, Hidekazu Katayama, Akiharu Ueki, Yuko Nakahara, and Yoshiaki Nakahara**

The native chemical ligation (NCL) method is generally used for the synthesis of (glyco)proteins.[1] It efficiently condenses peptide segments without the protection of functional groups, except for the terminal carboxyl group of the carboxyl component by the thioester group. Thus, the segments can be easily prepared by the solid-phase method or recombinant DNA technology. Recently, the preparation of peptide thioesters by the recombinant DNA technique has also been realized by intein technology.[2] The major problem in NCL is that a cysteine residue is required at the ligation site to achieve chemoselective coupling. Since the natural abundance of the cysteine residue in proteins is low (1.5%), the ligation site is severely limited by this problem. Thus, much effort has been made to develop thiol auxiliary groups to overcome this limitation.[3]

In initial studies,[3a–d,f–i] the α-amino group was used to anchor thiol auxiliary groups, which were cleaved after the ligation. However, this strategy might lead to incomplete ligation as a result of steric hindrance, and thus has been mainly used where the ligation site contains at least one glycine residue. The other strategies used thiol groups at the side-chain functional groups, which were subsequently converted to proteinogenic amino acid forms. In these strategies the ligation seems to proceed efficiently, as the acyl group in the thioester intermediate is transferred to the primary amino group through five-, six-, or even larger-membered rings using the thiol group on carbohydrate moieties.[3i,m] These strategies also require additional reaction for the removal of auxiliary groups, which is mainly carried out by metal-based reduction. However, the reaction might not be compatible when applied to the synthesis of peptides having various functional groups, especially sulfur-containing groups. Recently, Wan and Danishefsky developed the free-radical reduction of thiol groups, which will be useful for (glyco)peptide synthesis by the NCL method.[3o]

Herein, we focused on the use of a mercaptomethyl group on the side-chain hydroxy groups of serine and threonine, which is a thiohemiacetal that is generally regarded as too labile to be used as an auxiliary group. In return, an additional deprotection step for this auxiliary group would not be required, as it is spontaneously hydrolyzed after the ligation. The use of the hydroxy groups of serine and threonine residues provides far more candidates for the ligation site, since these amino acids appear in proteins more frequently (ca. 12% in total) than cysteine. A general route for the new method is shown in Scheme 1. The mercaptomethyl group is generated in situ by the reduction of a disulfide-protected parent peptide. Then the ligation with N-terminal thioester is performed, followed by the *S*- to *N*-acyl shift via a seven-membered ring. After ligation, the mercaptomethyl group is hydrolyzed without any extra deprotection steps. If this concept functions well, all reactions can be performed in one pot, thus making this method efficient and practical. To examine the usefulness of the new strategy, contulakin-G (pGlu-Ser-Glu-Glu-Gly-Gly-Ser-Asn-Ala-Thr(GalNAc)-Lys-Lys-Pro-Tyr-Ile-Leu-OH, **1a**),[4] a glycopeptide toxin isolated from *Conus geographus* venom, and human calcitonin were synthesized.



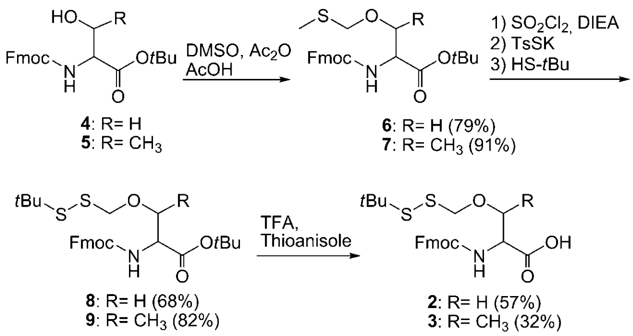
Scheme 1. Novel ligation reaction at the Xaa-Ser/Thr site facilitated by the mercaptomethyl group. Xaa = any amino acid. R = H or CH_3, $R^1 = C_6H_5$ or $C_6H_4CH_2COOH$.

The synthesis of Fmoc-Ser/Thr (Fmoc = 9-fluorenylmethoxycarbonyl) units carrying a protected mercaptomethyl group was easily accomplished in three steps by following the procedure of Semenyuk et al.,[5] as shown in Scheme 2. The commercially available Fmoc-Ser/Thr *t*Bu esters (**4** and **5**)

[*] Prof. Dr. H. Hojo, C. Ozawa, Dr. H. Katayama, Dr. A. Ueki, Dr. Y. Nakahara, Prof. Dr. Y. Nakahara
Department of Applied Biochemistry
Institute of Glycoscience, Tokai University
1117 Kitakaname, Hiratsuka, Kanagawa 259-1292 (Japan)
Fax: (+81) 463-50-2075
E-mail: hojo@keyaki.cc.u-tokai.ac.jp
yonak@keyaki.cc.u-tokai.ac.jp

[**] This work was supported by a grant-in-aid for creative scientific research from the Japan Society for the Promotion of Science (17GS0420), and in part by a grant-in-aid for scientific research from the Ministry of Education, Sports, Science, and Technology of Japan (20380069). We thank Tokai University for their support with a grant-in-aid for high technology research. We also thank Prof. Hiromichi Nagasawa and Shinji Nagata of Tokyo University for sequence analysis and Tsuyoshi Ohira of Kanagawa University for providing Glu-C. Xaa = any amino acid.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201000384.

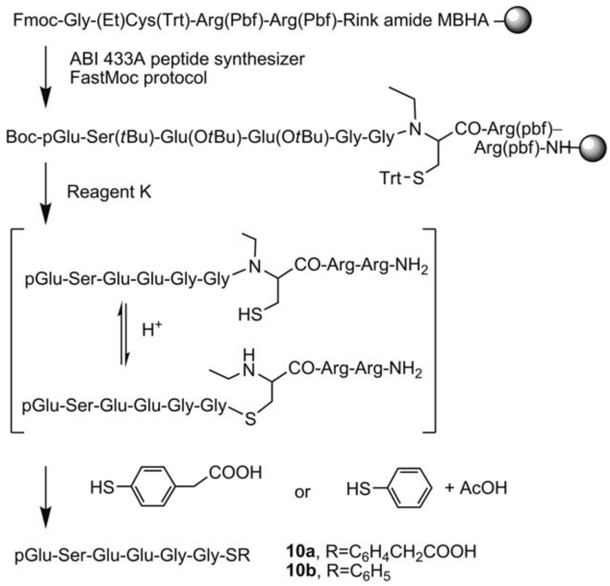


Scheme 2. Synthetic route for the Ser/Thr units **2** and **3** carrying the thiol auxiliary group.

were methylthiomethylated by the Pummerer rearrangement using dimethyl sulfoxide (DMSO), acetic anhydride (Ac_2O), and acetic acid (AcOH). The products were chlorinated by SO_2Cl_2 in diisopropylethylamine (DIEA), then treated with potassium thiotosylate (TsSK) and *t*BuSH, followed by trifluoroacetic acid (TFA) to give Fmoc-Ser/Thr carrying the disulfide-protected auxiliary group (**2** and **3**).

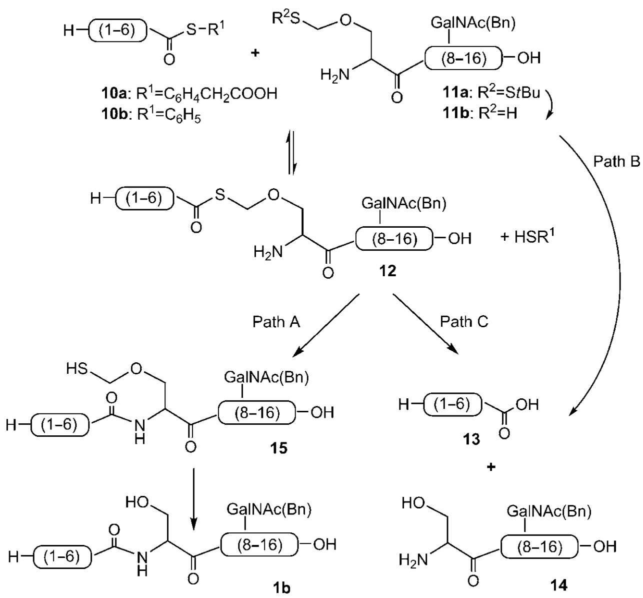
The serine unit **2** was used for the synthesis of contulakin-G through ligation at Gly^6–Ser^7. The C-terminal segment (H-Ser(CH_2SS*t*Bu)-Asn-Ala-Thr(GalNAcBn)-Lys-Lys-Pro-Tyr-Ile-Leu-OH (Bn = benzyl), **11a**) was prepared by the conventional Fmoc strategy introducing compound **2** as the final amino acid. It was essential to use a thiol-free TFA cocktail for deprotection to minimize the decomposition of the mercaptomethyl group to less than 15 %.

N-Terminal thioesters **10a** and **10b** were prepared by the Fmoc method using our *N*-alkylcysteine (NAC)-assisted thioesterification method,[6] as shown in Scheme 3. Starting from Fmoc-Gly-(Et)Cys(Trt)-Arg(Pbf)-Arg(Pbf)-NH-resin (Trt = trityl; Pbf = 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl),[6f] the peptide chain was elongated by adopting the Fmoc strategy. The obtained resin was treated with Reagent K,[7] and the crude peptide was dissolved in aqueous CH_3CN containing 6 M urea in the presence of 4-mercapto-phenylacetic acid (MPAA)[8] or thiophenol and AcOH. The NAC moiety was converted to the corresponding thioester within 24 h. The crude peptide was purified by reversed-phase (RP) HPLC to give peptide thioesters **10a** and **10b**.



Scheme 3. Synthetic route for the N-terminal peptide thioesters. MBHA = 4-methylbenzhydrylamine resin, Boc = *tert*-butoxycarbonyl.

The ligation was first conducted using thioester **10a** and glycopeptide **11a** in sodium phosphate (pH 7.0) containing 6 M guanidine HCl and MPAA. Glycopeptide **11a** was quickly reduced by MPAA to yield **11b**, which was ligated with thioester **10a** (Scheme 4). However, the product **1b** was not obtained. Under these conditions, only a small amount of the intermediate **12** was detected by HPLC. This might be



Scheme 4. Pathways of the ligation between peptide thioesters and glycopeptide carrying the thiol auxiliary group, and side reactions.

because of an equilibrium between the intermediate **12** and the peptides **10a** and **11b**, which is shifted to the latter by the excess MPAA. Since the *S*- to *N*-acyl shift (Scheme 4, path A) proceeds through the seven-membered ring, the reaction was too slow to be productive. As a result, peptides **10a** and **11b** were gradually hydrolyzed to produce nonreactive peptides **13** and **14** (Scheme 4, path B).

To suppress the undesirable reverse reaction, the ligation between **10a** and **11a** was conducted without the external MPAA (Table 1, entry 1). Guanidine HCl was also excluded to simplify the buffer composition. The reduction of the disulfide bond in **11a** was achieved with triscarboxyethyl-phosphine (TCEP). Under these conditions, about half of the generated glycopeptide **11b** was converted to intermediate **12** within 10 min (Supporting Information, Figure S1a). However, the formation of **12** did not become quantitative even after 30 min, which indicated that the intermediate **12** and the liberated MPAA are still in equilibrium with peptides **10a** and **11b**. The relative content of the desired peptide **1b** after

Table 1: Yield of the product **1b** by the ligation of peptide thioester with **11a**.

Entry	Peptide thioester		Buffer[a]	Relative content [%][b]	
				1b	**14**
1	pESEEGG-MPAA	**10a**	A	27	73
2	pESEEGG-SPh	**10b**	A	42	58
3	pESEEGG-SPh	**10b**	B	47	53
4	pESEEGG-SPh	**10b**	C	37	15[c]
5	pESEEGG-SPh	**10b**	D	80	20
6	pESEEGG-SPh	**10b**	E	60	40
7	pESEEGG-SPh	**10b**	F	77	23
8	pESEEGL-SPh	**10c**	D	69	31

[a] Buffer A: 0.1 M sodium phosphate containing 15 mM TCEP (pH 7.0); buffer B: 1:1 mixture of buffer A and CH_3CN; buffers C and D: buffer A, which after 10 min was diluted (×10) with DMF and 10% AcOH in DMF, respectively; buffer E: buffer A containing 6 M guanidine HCl (pH 7.0), which was diluted (×10) with 10% AcOH in DMF after 2 h of ligation; buffer F: buffer A containing 6 M guanidine HCl (pH 7.0) in the presence of ether, followed by the removal of ether and dilution with 10% AcOH in DMF (×10) after 1 h of ligation. [b] The value shows the ratio of the area of peptide **1b** or **14** to that of the total area of peptides derived from C-terminal peptides upon HPLC. [c] The remaining 48% is acyl shift products to the side-chain amino group.

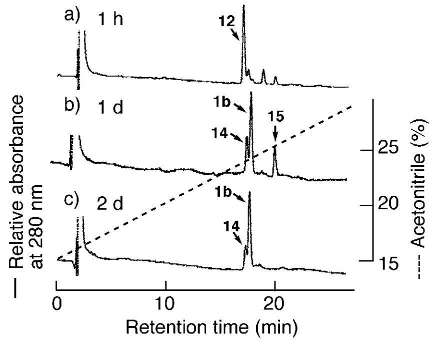


Figure 1. RP-HPLC profiles of the ligation to obtain glycopeptide **1b** under the conditions of Table 1, entry 7: a) 1 h ligation in buffer A containing 6 M guanidine HCl (pH 7.0) in the presence of ether; b) overnight reaction after the addition of 10% AcOH in DMF, c) 2 days reaction time. Elution conditions: column, Mightysil RP-18 GP (4.6×150 mm, Kanto, Japan); eluent: acetonitrile/water/TFA 85:15:0.1→70:30:0.1 over 30 min, flow rate 1 mL min^{-1}.

overnight reaction was 27% (Supporting Information, Figure S1b). Based on these results, the MPAA thioester seemed to be unsuitable for this ligation. In contrast, the ligation with the thiophenyl ester **10b** gave the intermediate **12** in higher yield within 10 min and the product **1b** was obtained in an increased yield after overnight reaction (Table 1, entry 2; Supporting Information, Figure S1d). The addition of CH_3CN had little effect on the yield of the product **1b**, as shown in Table 1, entry 3.

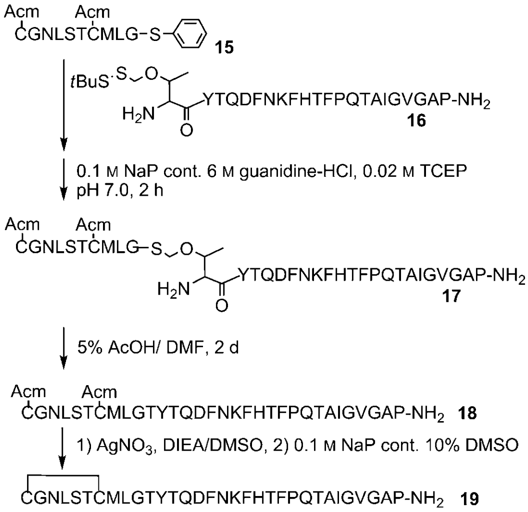
To prevent the hydrolysis of the intermediate **12** (Scheme 4, path C), the ligation mixture derived from **10b** and **11a** was diluted with DMF after 10 min, when a sufficient amount of the intermediate **12** was formed. Although the hydrolysis was reduced to less than 20%, about 50% of the acyl shift products to the side-chain amino groups of lysine was obtained, which might be because of the reduced protonation state of the side-chain amino groups in DMF (Table 1, entry 4; Supporting Information, Figure S1e). However, this side reaction was almost completely suppressed by the addition of acetic acid to DMF, and the desired product was successfully obtained after overnight reaction at a relative content of 80% (Table 1, entry 5; Supporting Information, Figure S1f).

The formation of the intermediate **12** was also tested in the presence of 6 M guanidine HCl (Table 1, entry 6). However, the reaction became slower and incomplete. The relative content of the product **1b** after 2 h of ligation followed by *S–N* shift for 2 days in AcOH/DMF was decreased to 60%. Then, the ligation was conducted in the presence of ether with vortexing to extract a part of the generated thiophenol and make the equilibrium more productive (Table 1, entry 7). As a result, the intermediate **12** was obtained in good yield within 1 h, as shown in Figure 1a. After 2 days, the product **1b** was obtained at a relative content of 77%, which is comparable to that of Table 1, entry 5 (Figure 1c). The yield of isolated **1b** obtained by RP-HPLC purification was 45%. Thus, these conditions can be generally used when segments retain low solubility. Under the organic and acidic conditions, the thiol auxiliary group was fairly stable: it took approximately 2 days for complete removal. Since the *S*- to *N*-acyl shift itself is completed within 6 h, the dilution of the mixture with neutral buffer at this stage can accelerate the overall reaction time. The native peptide bond between Gly^6 and Ser^7 in glycopeptide **1b** was confirmed by sequence analysis of the Glu-C digest of **1b** (see the Supporting Information). Glycopeptide **1b** was easily converted to the final product, contulakin-G (**1a**), by low-acidity trifluoromethanesulfonic acid (TfOH) treatment to remove a benzyl group.[9]

The peptide thioester **10c** with a C-terminal Leu residue was also ligated with peptide **11a**. As shown in Table 1 (entry 8), the desired Leu^6-contulakin-G (**1c**) was successfully obtained at a 69% relative content and 31% yield of isolated product, thus showing the applicability of this method to ligation with a sterically demanding amino acid.

The stability of the mercaptomethyl group on the glycopeptide **11b** was examined in buffer A containing 6 M guanidine HCl at pH 6, 7, and 8 (see the Supporting Information). The half-lives were about 4, 2.5, and 2 h, respectively, which might be sufficient periods to perform the initial transthioesterification step of the ligation at these pH values.

The method was further applied to the synthesis of human calcitonin using Thr derivative **3**, as shown in Scheme 5. Peptide thioester **15** and peptide **16** were prepared according to the same procedure as that for contulakin-G. The ligation was performed under the conditions of Table 1, entry 7. Within 2 h at room temperature, the formation of the intermediate **17** was maximum. After dilution with AcOH/DMF, the solution was left undisturbed for 2 days. The desired peptide **18** was obtained at a relative content of 66% and 40% yield of isolated product.[10] Considering that this is a total yield of the ligation and the removal of the auxiliary group, the value is acceptable compared with those of other auxiliary-mediated ligation reactions.[3] The removal of the



Scheme 5. Synthetic route for human calcitonin. Acm = acetamidomethyl.

Acm group and disulfide bond formation successfully gave calcitonin **19** (see the Supporting Information).

In conclusion, novel ligation at the naturally abundant Xaa-Ser/Thr site using the mercaptomethyl group as a thiol auxiliary was demonstrated by the syntheses of contulakin-G and human calcitonin. The ease of the synthesis of the Ser/Thr unit carrying the thiol auxiliary group as well as its autocleavable property make this method versatile for the synthesis of various (glyco)proteins. Currently, this method is being applied to glycoprotein synthesis.

Received: January 22, 2010

Keywords: glycopeptides · ligation · peptide thioesters · synthetic methods

[1] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776–779; b) P. E. Dawson, S. B. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923–960.

[2] T. W. Muir, D. Sondhi, P. A. Cole, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6705–6710; b) T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249–289.

[3] a) L. E. Canne, S. J. Bark, S. B. H. Kent, *J. Am. Chem. Soc.* **1996**, *118*, 5891–5896; b) D. W. Low, M. G. Hill, M. R. Carrasco, S. B. H. Kent, P. Botti, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6554–6559; c) T. Kawakami, K. Akaji, S. Aimoto, *Org. Lett.* **2001**, *3*, 1403–1405; d) P. Botti, M. R. Carrasco, S. B. H. Kent, *Tetrahedron Lett.* **2001**, *42*, 1831–1833; e) L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526–533; f) J. Offer, C. N. C. Boddy, P. E. Dawson, *J. Am. Chem. Soc.* **2002**, *124*, 4642–4646; g) T. Kawakami, S. Aimoto, *Tetrahedron Lett.* **2003**, *44*, 6059–6061; h) D. Macmillan, D. W. Anderson, *Org. Lett.* **2004**, *6*, 4659–4662; i) 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; j) A. Brik, Y. Y. Yang, S. Ficht, C. H. Wong, *J. Am. Chem. Soc.* **2006**, *128*, 5626–5627; k) A. Saporito, D. Marasco, A. Chambery, P. Botti, S. M. Monti, C. Pedone, M. Ruvo, *Biopolymers* **2006**, *83*, 508–518; l) B. L. Pentelute, S. B. H. Kent, *Org. Lett.* **2007**, *9*, 687–690; m) Y.-Y. Yang, S. Ficht, A. Brik, C.-H. Wong, *J. Am. Chem. Soc.* **2007**, *129*, 7690–7701; n) D. Crich, A. Banerjee, *J. Am. Chem. Soc.* **2007**, *129*, 10064–10065; o) Q. Wan, S. J. Danishefsky, *Angew. Chem.* **2007**, *119*, 9408–9412; *Angew. Chem. Int. Ed.* **2007**, *46*, 9248–9252; p) 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; q) R. Okamoto, Y. Kajihara, *Angew. Chem.* **2008**, *120*, 5482–5486; *Angew. Chem. Int. Ed.* **2008**, *47*, 5402–5406; r) C. Haase, H. Rohde, O. Seitz, *Angew. Chem.* **2008**, *120*, 6912–6915; *Angew. Chem. Int. Ed.* **2008**, *47*, 6807–6810; s) J. Chen, P. Wang, J. Zhu, Q. Wan, S. J. Danishefsky, *Tetrahedron* **2010**, *66*, 2277–2283.

[4] A. G. Craig, T. Norberg, D. Griffin, C. Hoeger, M. Akhtar, K. Schmidt, W. Low, J. Dykert, E. Richelson, V. Navarro, J. Mazella, M. Watkins, D. Hillyard, J. Imperial, L. J. Cruz, B. M. Olivera, *J. Biol. Chem.* **1999**, *274*, 13752–13759.

[5] A. Semenyuk, A. Foldesi, T. Johansson, C. Estmer-Nilsson, P. Blomgren, M. Brannvall, L. A. Kirsebom, M. Kwiatkowski, *J. Am. Chem. Soc.* **2006**, *128*, 12356–12357.

[6] a) H. Hojo, Y. Onuma, Y. Akimoto, Y. Nakahara, Y. Nakahara, *Tetrahedron Lett.* **2007**, *48*, 25–28; b) C. Ozawa, H. Katayama, H. Hojo, Y. Nakahara, Y. Nakahara, *Org. Lett.* **2008**, *10*, 3531–3533; c) H. Katayama, H. Hojo, T. Ohira, Y. Nakahara, *Tetrahedron Lett.* **2008**, *49*, 5492–5494; d) H. Hojo, Y. Murasawa, H. Katayama, T. Ohira, Y. Nakahara, Y. Nakahara, *Org. Biomol. Chem.* **2008**, *6*, 1808–1813; e) Y. Nakahara, I. Matsuo, Y. Ito, R. Ubagai, H. Hojo, Y. Nakahara, *Tetrahedron Lett.* **2010**, *51*, 407–410; f) H. Katayama, H. Hojo, I. Shimizu, Y. Nakahara, Y. Nakahara, *Org. Biomol. Chem.* **2010**, *8*, 1966–1972.

[7] D. S. King, C. G. Fields, G. B. Fields, *Int. J. Pept. Protein Res.* **1990**, *36*, 255–266.

[8] D. Bang, B. L. Pentelute, Z. P. Gates, S. B. H. Kent, *Org. Lett.* **2006**, *8*, 1049–1052.

[9] a) J. P. Tam, W. F. Heath, R. B. Merrifield, *J. Am. Chem. Soc.* **1986**, *108*, 5242–5251; b) H. Hojo, E. Haginoya, Y. Matsumoto, Y. Nakahara, K. Nabeshima, B. P. Toole, Y. Watanabe, *Tetrahedron Lett.* **2003**, *44*, 2961–2964; c) H. Hojo, Y. Matsumoto, Y. Nakahara, E. Ito, Y. Suzuki, M. Suzuki, A. Suzuki, Y. Nakahara, *J. Am. Chem. Soc.* **2005**, *127*, 13720–13725.

[10] The ligation at Val–Thr, Leu–Thr, and Tyr–Thr sites was also tested. The first one was unproductive, but the latter two were useful for ligation (see the Supporting Information).

