Synthetic Methods

DOI: 10.1002/ange.200803523

Native Chemical Ligation at Valine: A Contribution to Peptide and Glycopeptide Synthesis**

*Jin Chen, Qian Wan, Yu Yuan, Jianglong Zhu, and Samuel J. Danishefsky**

Our laboratory has been pursuing the total synthesis of naturally occurring glycoproteins bearing multiple oligosaccharide domains. Specifically, efforts are well underway to accomplish a de novo total synthesis of erythropoietin alpha (EPO) in homogeneous form.[1] Although a variety of peptide ligation strategies have been developed to facilitate the merger of large, complex peptide and glycopeptide fragments,[2–9] the need for highly efficient methodology continues to motivate the chemical community to develop more powerful strategies. Our pursuit of the total synthesis of homogeneous EPO, as well as other biologically active glycopeptides, has inspired new glycopeptide ligations.[10] To achieve our most complex goals, we must learn how to overcome the serious obstacles in joining glycopeptides in an iterative fashion.

Native chemical ligation (NCL), developed by Kent and co-workers,[3] constituted a fundamental advance, allowing the joining of two substantial peptide domains. An additional development provided by our group as well as others serves to extend the NCL method to the assembly of peptides bearing multiple sites of glycosylation.[10a,11] These methods currently require the presence of a cysteine residue at the N-terminus of the peptide coupling partner.

Application of the NCL method is often limited by the paucity of cysteine residues in naturally occurring proteins and glycoproteins. Several strategies have been investigated to circumvent the need for a cysteine in the target at the proposed ligation site. One approach involves appending an auxiliary thiol group to the N-terminal amino acid, and subsequent to ligation cleaving the auxiliary.[10b,c,4] This approach suffers from certain limitations, as the reaction may be inefficient at more hindered ligation sites and difficulties can arise at the stage of auxiliary removal.

A second means by which to circumvent the need for a resident cysteine residue involves the use of an amino acid surrogate containing a thiol moiety. After ligation the surrogate is converted into the desired amino acid. In this context, ligation at a serine site has been achieved by post-ligational conversion of cysteine into serine.[5] Similarly, a formal methionine ligation has been accomplished by homocysteine coupling, and subsequent post-ligational methylation.[6] Recently, Wong and co-workers reported a protocol for cysteine-free thioester ligation,[2,7] which allows for peptide assembly. Although lysine residues must be protected and the reaction usually takes more than 48 hours even at rather unhindered ligation sites, the method is quite promising.

In addition, two-step ligation/metal-based thiol reduction protocols, which formally serve to accomplish ligation at N-terminal alanine[8] and phenylalanine[9] residues, have been developed. However, these methods may not be compatible with a large range of functionalities, particularly sulfur-containing groups which are frequently present in peptide sequences. Alternatively, our group recently disclosed a mild and highly versatile free-radical cysteine reduction protocol, which tolerates all thiol-containing groups, as well as oligosaccharide domains.[10e]

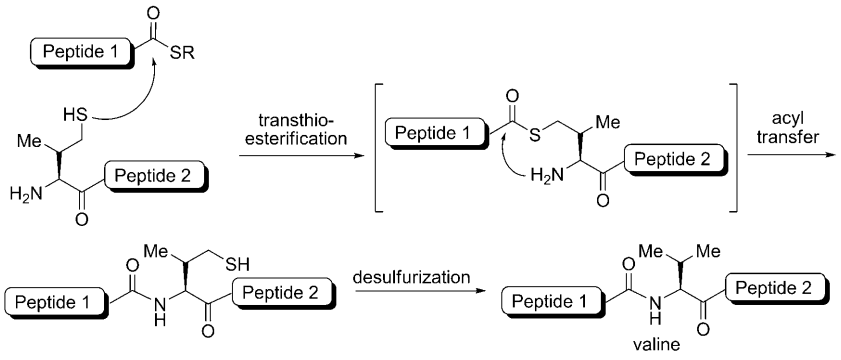
Theoretically, native chemical ligation could be achieved at any amino acid site, in the sense proposed by Yan and Dawson,[8] in which a sulfhydryl group is temporarily installed at a non-cysteine site. The implementation of this concept requires that the key mechanistic steps (i.e. *trans*-thioesterification and then S→N acyl transfer) be operative. Post-ligation desulfurization would then afford the desired peptide or glycopeptide adduct possessing the natural amino acid residue at the ligation site.

Herein, a strategy for implementing the logic of native chemical ligation at valine residues is described. Valine is a rather abundant amino acid, with approximately a 6.6% frequency in nature, (compared to 1.7% for cysteine). There are two potential valine surrogates, β-thiol-containing valine (penicillamine) and a γ-thiol valine. We envisioned that the γ-thiol valine, which contains a more reactive primary thiol group, would serve as a more suitable precursor. As illustrated in Scheme 1, an N-terminal thiol-modified valine derivative, installed on peptide 2, would react with the peptide 1 thioester by *trans*-thioesterification. The resultant thioester-linked intermediate would then undergo a rapid intramolecular acyl transfer, thus creating an amide bond. Radical-based desulfurization would serve to remove the thiol moiety and provide the desired peptide with valine at the ligation site.

[*] Prof. S. J. Danishefsky
Department of Chemistry, Columbia University
3000 Broadway, New York, NY 10027 (USA)
Fax: (+1) 212-772-8691
E-mail: s-danishefsky@mskcc.org
Dr. J. Chen, Dr. Q. Wan, Dr. Y. Yuan, Dr. J. Zhu, Prof. S. J. Danishefsky
Laboratory for Bioorganic Chemistry
Sloan-Kettering Institute for Cancer Research
1275 York Avenue, New York, NY 10065 (USA)

[**] Support for this research was provided by the National Institutes of Health (CA28824). We thank Dr. George Sukenick, Hui Fang, and Sylvi Rusli of the Sloan-Kettering Institute's NMR core facility for mass spectral and NMR spectroscopic analysis (SKI core grant no.: CA02848). We would like to express our appreciation to Rebecca Wilson for her assistance in preparing the manuscript.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200803523.

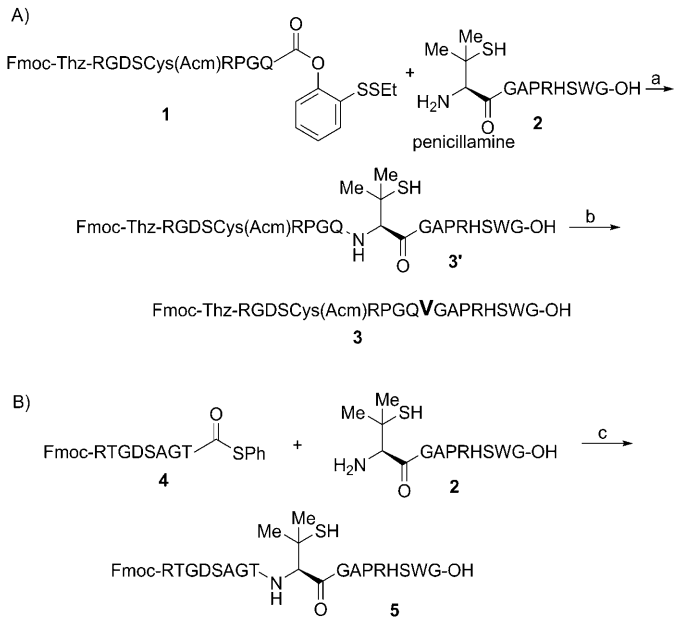


Scheme 1. Native chemical ligation at valine.

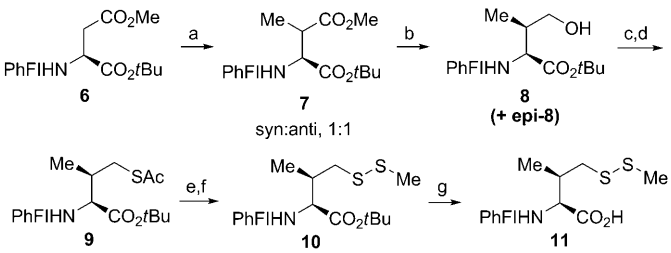
Penicillamine (**2**; Scheme 2), containing a β-thiol moiety, is a commercially available valine surrogate. As shown in Scheme 2, penicillamine was introduced to the N-terminus of the peptide as the acyl acceptor, and the NCL was performed under standard conditions. Because of the tertiary nature of the thiol group, ligation proceeded rather slowly. Though the ligation yield was reasonable in a case in which glutamine was present at the C-terminus (Scheme 2A), when the more bulky threonine residue was incorporated at the C-terminus the reaction became prohibitively slow (Scheme 2B). During our preparation of this communication, Seitz and co-workers reported a ligation culminating in valine, using penicillamine as the valine precursor.[12] They successfully achieved peptide bond-formation when glycine, histidine, methionine, and leucine were presented at the C-terminus. On the basis of their results and our own (see below) for the case of peptides bearing less hindered C-terminal amino acids, penicillamine can serve in fostering valine ligation. However, our primary goal was that of developing a logic which would encompass more generally applicable ligations including certain β-branched C-terminal acyl donors. Our successes in this regard are described below.

Given the greater reactivity and diminished steric hindrance of a primary thiol relative to a tertiary thiol, we felt that it was worthwhile to develop a protocol for the efficient synthesis of a modified valine surrogate with a thiol group installed at its γ position. Drawing from the earlier work of Rapoport and Wolf[13] in which a γ-hydroxy valine was obtained from L-aspartic acid, we modified the route thereby enabling the preparation of a γ-thiol valine derivative (**11**), albeit as two diastereomers. Either epimer of **11** could, in principle, be used for our purposes since the removal of the thiol group after the ligation would eliminate the chirality at the β position. Thus, as outlined in Scheme 3, the protected L-aspartic acid diester **6** was prepared from Fmoc-Asp-O*t*Bu (see the Supporting Information for details). The PhFl (9-(9-phenylfluorenyl)) group was used to block the α center, providing exclusively β-alkylated compound **7**. Regioselective reduction of **7** furnished alcohols **8** and *epi*-**8** as a mixture of diastereomers (1:1), which were readily separated by chromatography. After mesylation and treatment with the DBU salt of thioacetic acid, acetylated thiol **9** was in hand. The latter was advanced to the target compound, **11**, in a straightforward manner, as shown.

Having the required thiol-containing amino acids in hand, we next sought to probe the versatility of our newly developed protocol. A variety of substrates incorporating a range of relevant functionalities was evaluated (Table 1). In the cases of the less hindered C-terminal amino acids, such as Gln (Table 1, entry 1) and Phe (Table 1, entry 2), the ligation

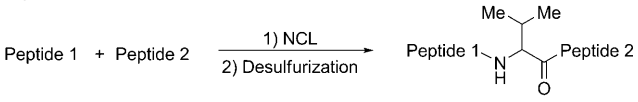


Scheme 2. Reagents and conditions: A. a) buffer at pH 6.5 (6.0 M Gn·HCl, 188.8 mM Na_2HPO_4), TCEP, RT, 9 h, 66% yield; b) TCEP, VA-044, *t*BuSH, 37 °C, 3 h, 79% yield. B. c) buffer at pH 6.5 (6.0 M Gn·HCl, 188.8 mM Na_2HPO_4), TCEP, RT, 10 h, 4% yield based on LC-MS trace. VA-044 = 2,2′-azobis-[2-(2-imidazolin-2-yl)propane] dihydrochloride. TCEP = tris(2-carboxyethyl)phosphine.

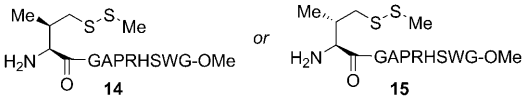


Scheme 3. Reagents and conditions: a) KHMDS, MeI, THF, −78 °C, 3 h, 97%; b) DIBAL-H, THF, −35 °C, 1 h, 83% (combined yield for **8** and *epi*-**8**); c) MsCl, Et_3N, CH_2Cl_2, 0 °C, 1 h; d) AcSH, DBU, DMF, RT, 16 h, 73% in 2 steps; e) 1 N NaOH, MeOH, 0 °C, 10 min; f) MMTS, Et_3N, CH_2Cl_2, RT, 30 min, 86% in 2 steps; g) HCl in EtOAc, RT, 82%. KHMDS = potassium 1,1,1,3,3,3-hexamethyldisilazane; DIBAL-H = diisobutylaluminum hydride; Ms = methylsulfonyl; Ac = acetyl; DBU = 1,8-diazabicyclo[5.4.0]undec-7-ene; MMTS = *S*-methyl methanethiosulfonate.

Table 1: NCL at valine by ligation and subsequent free-radical desulfurization.[a]



Peptide 2:



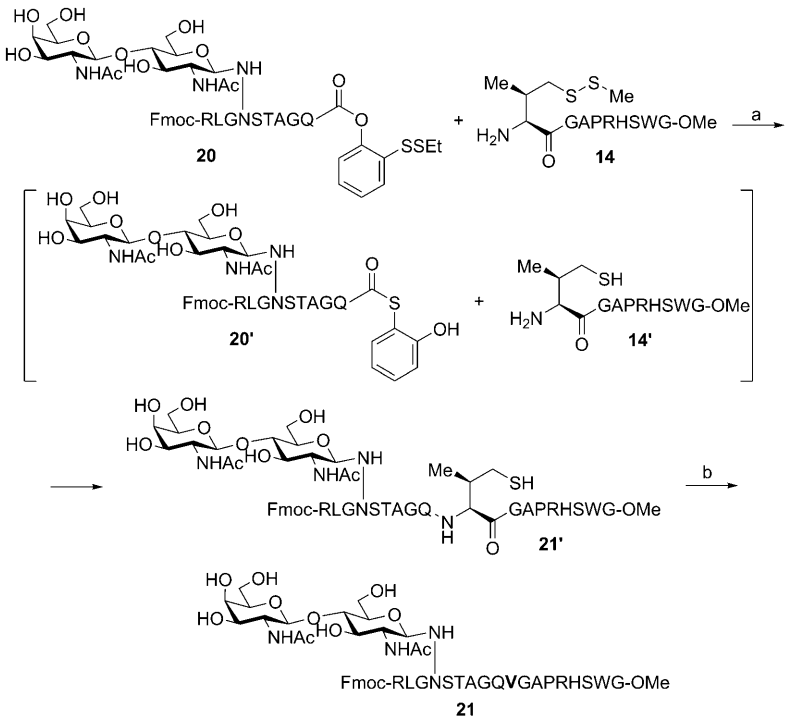
Entry	Peptide 1	Peptide 2	Ligation Product/Yield[b]/Time	Desulfurization Product/Yield[b]/Time
1	FmocThzRGDSCys(Acm)RPGQ-O-(SSEt-Ph) **1**	**14**	FmocThzRGDSCys(Acm)RPGQ**V**GAPRHSWG-OMe **16'** / 78% / 1 h	FmocThzRGDSCys(Acm)RPGQVGAPRHSWG-OMe **16** / 84% / 3 h
2	FmocKYDSRGF-O-OPh-*p*CN **12**	**14**	FmocKYDSRGF**V**GAPRHSWG-OMe **17'** / 80% / 1 h	FmocKYDSRGFVGAPRHSWG-OMe **17** / 81% / 3 h
3	FmocRTGDSAGT-O-SPh **4**	**14**	FmocRTGDSAGT**V**GAPRHSWG-OMe **18'** / 87% / 4 h	FmocRTGDSAGTVGAPRHSWG-OMe **18** / 89 % / 3 h
4	FmocRTGDSAGT-O-SPh **4**	**15**	FmocRTGDSAGT**V**GAPRHSWGOMe **18''** / 87% / 9 h	FmocRTGDSAGTVGAPRHSWG-OMe **18** / 90% / 3 h
5	FmocVRSYTAGP-O-OPh-*p*NO_2 **13**	**14**	FmocVRSYTAGP**V**GAPRHSWG-OMe **19'** / 55% / 10 h	FmocVRSYTAGPVGAPRHSWG-OMe **19** / 98% / 3 h

[a] For detailed reaction conditions, please see the Supporting Information. [b] Yield of the isolated product. Fmoc = 9-fluorenylmethoxycarbonyl.

proceeded very rapidly (1 h). By comparison, in the analogous case of penicillamine as an acyl acceptor (**2**), ligation with the same peptide (**1**) proceeded at a much slower rate (9 h; Scheme 2A). As the C-terminus became sterically more hindered, (Table 1, entries 3 and 4), the reaction rate dropped correspondingly, but the ligations were still completed within a reasonable time range and in good yield. Table 1, entry 3, records an example of a non-cysteine driven NCL reaction in which both peptides present β-branched amino acids (Thr and Val) the ligation site. In contrast, the analogous reaction of substrate **4** with penicillamine (**2**) displayed much lower reactivity (Scheme 2B); only 4% product was formed in 10 hours. A range of C-terminal esters was investigated, including the standard thiophenyl ester (Table 1, entries 3 and 4) as well as an *ortho*-thiophenolic ester, developed in our laboratory to allow glycopeptide ligation (Table 1, entry 1).[10a] We recently described a novel direct oxo-ester ligation protocol, which circumvents the need for a thioester intermediate, and allows ligation at sterically hindered C-terminal sites.[14] In this context, we were pleased to find that both the C-terminal *p*-cyanophenyl ester (Table 1, entry 2), and the *p*-nitrophenyl ester (Table 1, entry 5) readily participated in our two-step valine ligation protocol. In addition, the combination of a C-terminal *p*-nitrophenyl ester and an N-terminal γ-thiol valine furnished the ligation at a Pro-Val site (Table 1, entry 5), which is extremely sterically demanding. Additionally, substrates incorporating unprotected lysine residues (Table 1, entry 2), as well as thiazolidine and cysteine(Acm) moieties (Table 1, entry 1) were examined to probe the limits of our system. In each of these cases, ligation and subsequent reduction proceeded smoothly, providing the desired peptide adducts in moderate to good overall yield.

Interestingly, it was found that although both γ-thiol valine epimers (i.e. **11** and *epi*-**11**) participate in the ligation protocol, **11** is not markedly more reactive than *epi*-**11**, as evidenced by a comparison of entries 3 and 4 in Table 1. Thus, when **11** serves as the N-terminal amino acid, as in **14**, ligation proceeds in 4 hours to yield the product in 87% yield. An analogous ligation using **15**, which incorporated *epi*-**11** as the N-terminal amino acid, was somewhat slower, presumably because of a 1,2-*cis* steric interaction between the β-methyl group and the amide moiety during the S→N acyl transfer step. That the difference in rates is relatively modest, suggests that the rate-determining step in the coupling is the initial intermolecular transthioesterification rather than the intramolecular S→N acyl transfer.

We next investigated the efficiency of the glycopeptide coupling employing γ-thiol valine **11** as a key partner. Thus, glycopeptide **20**, with an N-linked glycan, and a C-terminal *ortho*-thiophenolic ester was prepared and coupled with peptide **14** (Scheme 4). The ligation proceeded smoothly and provided adduct **21′** in 1 hour. Subsequent thiol reduction generated desired product **21** in nearly quantitative conversion (Figure 1).



Scheme 4. Reagents and conditions: a) buffer at pH 6.5 (6.0 M Gn·HCl, 188.8 mM Na_2HPO_4), TCEP, RT, 30 min, 90% yield; b) TCEP, VA-044, *t*BuSH, 37 °C, 3 h, 89% yield.

In summary, an efficient and broadly useful two-step valine ligation protocol has been developed. The γ-thiol valine (**11**), containing a primary thiol group, serves as a valine surrogate. The combination of the NCL method and the metal-free reduction provides an efficient, two-step protocol by which to accomplish a formal valine ligation. This method represents an important extension of existing NCL techniques, and should be valuable in our ongoing programs to bring synthesis to bear in domains which had previously been considered to be inaccessible to chemistry-based interventions.[15]

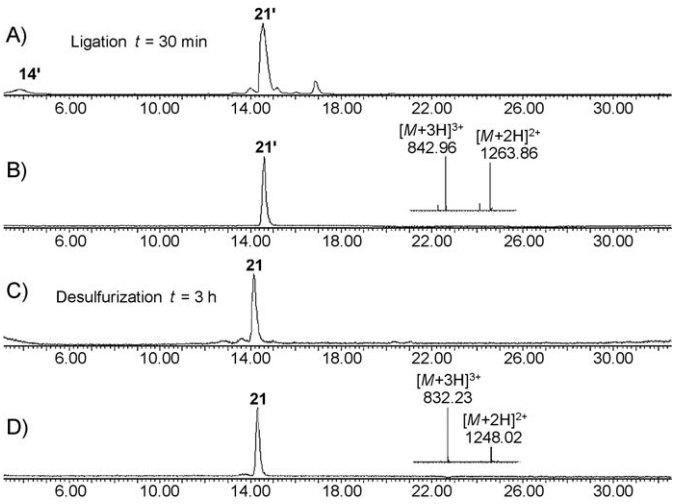


Figure 1. Synthesis of glycopeptide **21**. LC-MS trace of: A) Ligation between glycopeptide **20** and peptide **14** after 30 min, **21′** is the ligation product. B) **21′** after purification by HPLC methods; observed mass $[M+2H]^{2+}=1263.86$, $[M+3H]^{3+}=842.96$. C) Desulfurization of **21′** after 3 h; **21** is the desulfurization product. D) **21** after purification by HPLC methods; observed mass $[M+2H]^{2+}=1248.02$, $[M+3H]^{3+}=832.23$.

Received: July 19, 2008
Revised: August 20, 2008
Published online: October 2, 2008

Keywords: desulfurization · glycopeptides · native chemical ligation · peptides · valine

[1] a) P. Lai, R. Everett, F. Wang, T. Arakawa, E. Goldwasser, *J. Biol. Chem.* **1986**, *261*, 3116–3121; b) D. M. Ridley, F. Dawkins, E. Perlin, *J. Natl. Med. Assoc.* **1994**, *86*, 129–135.

[2] a) J. Blake, *Int. J. Pept. Protein Res.* **1981**, *17*, 273–274; b) S. Aimoto, N. Mizoguchi, H. Hojo, S. Yoshimura, *Bull. Chem. Soc. Jpn.* **1989**, *62*, 524–531; c) S. Aimoto, *Biopolymers* **1999**, *51*, 247–265.

[3] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776–779.

[4] D. W. Low, M. G. Hill, M. R. Carrasco, S. B. H. Kent, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6554–6559.

[5] R. Okamoto, Y. Kajihara, *Angew. Chem.* **2008**, *120*, 5482–5486; *Angew. Chem. Int. Ed.* **2008**, *47*, 5402–5406.

[6] a) J. P. Tam, Q. Yu, *Biopolymers* **1998**, *46*, 319–327; b) K. Pachamuthu, R. R. Schmidt, *Synlett* **2003**, 659–662; c) A. Saporito, D. Marasco, A. Chambery, P. Botti, C. Pedone, M. Ruvo, *Biopolymers* **2006**, *83*, 508–518.

[7] R. J. Payne, S. Fichet, W. A. Greenberg, C.-H. Wong, *Angew. Chem.* **2008**, *120*, 4483–4487; *Angew. Chem. Int. Ed.* **2008**, *47*, 4411–4415.

[8] L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526–533.

[9] D. Crich, A. Banerjee, *J. Am. Chem. Soc.* **2007**, *129*, 10064–10065.

[10] a) J. D. Warren, J. S. Miller, S. J. Keding, S. J. Danishefsky, *J. Am. Chem. Soc.* **2004**, *126*, 6576–6578; b) G. Chen, J. D. Warren, J. Chen, B. Wu, Q. Wan, S. J. Danishefsky, *J. Am. Chem. Soc.* **2006**, *128*, 7460–7462; c) 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; d) 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; e) Q. Wan, S. J. Danishefsky, *Angew. Chem.* **2007**, *119*, 9408–9412; *Angew. Chem. Int. Ed.* **2007**, *46*, 9248–9252.

[11] For reviews on glycopeptide and glycoprotein synthesis, see a) B. G. Davis, *Chem. Rev.* **2002**, *102*, 579–601; b) M. J. Grogan, M. R. Pratt, L. A. Marcaurelle, C. R. Bertozzi, *Annu. Rev. Biochem.* **2002**, *71*, 593–634; c) L. Liu, C. S. Bennett, C.-H. Wong, *Chem. Commun.* **2006**, 21–33.

[12] C. Haase, H. Rohde, O. Seitz, *Angew. Chem.* **2008**, *120*, 6912-6915; *Angew. Chem. Int. Ed.* **2008**, *47*, 6807-6810.

[13] a) H. Rapoport, J. P. Wolf, *J. Org. Chem.* **1989**, *54*, 3164–3173; b) P. Gmeiner, P. L. Feldman, M. Y. Chu-Moyer, H. J. Rapoport, *J. Org. Chem.* **1990**, *55*, 3068–3074.

[14] Q. Wan, J. Chen, Y. Yuan, S. J. Danishefsky, **2008**, submitted.

[15] R. M. Wilson, S. J. Danishefsky, *Pure Appl. Chem.* **2007**, *79*, 2189–2216.

