ORGANIC LETTERS

2009
Vol. 11, No. 22
5270−5273

# A One-Pot Approach to Neoglycopeptides using Orthogonal Native Chemical Ligation and Click Chemistry

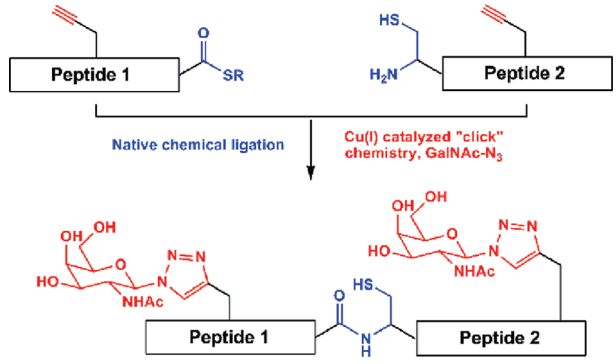
**Dong Jun Lee,[†] Kalyaneswar Mandal,[‡] Paul W. R. Harris,[†] Margaret A. Brimble,*[,†] and Stephen B. H. Kent*[,‡]**

*Department of Chemistry, The University of Auckland, 23 Symonds Street, Auckland, New Zealand, and Department of Chemistry, Department of Biochemistry & Molecular Biology, Institute for Biophysical Dynamics, The University of Chicago, 929 East 57th Street, Chicago, Illinois 60637*

*skent@uchicago.edu; m.brimble@auckland.ac.nz*

**Received September 24, 2009**

## ABSTRACT




**The powerful combination of native chemical ligation and click chemistry has been used to affect a one-pot synthesis of neoglycopeptides from propargyl-containing peptides using GalNAc-$N_3$ as the glycan component. A versatile chemical toolkit for the fully convergent synthesis of neoglycoproteins using click chemistry, native chemical ligation, and kinetically controlled ligation is thus demonstrated.**

Glycosylation is the most common and complex post-translational modification of proteins.[1−3] Glycoproteins are essential in many biological processes including immune defense, cell growth, and inflammation.[1] For these reasons, there is considerable interest in the structure−function relationships of glycoproteins, the investigation of which requires the systematic, controlled variation of glycoprotein structure.[2−5] Recombinant techniques for expression of glycoproteins are problematic because glycosylation is organism-specific and artificial cultivation conditions may result in incorrect or inconsistent glycosylation patterns.[2−5] Glycoproteins of defined chemical structure can be prepared by total or semisynthesis based on modern chemical ligation methods,[6] and chemistry can then be used to control both the site and structure of glycosylation.[5,7,8]

[†] The University of Auckland.
[‡] The University of Chicago.

(1) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683–720.
(2) Seitz, O. *ChemBioChem* **2000**, *1*, 214–246.
(3) Brocke, C.; Kunz, H. *Biorg. Med. Chem.* **2002**, *10*, 3085–3112.
(4) Grogan, M. J.; Pratt, M. R.; Marcaurelle, L. A.; Bertozzi, C. R. *Annu. Rev. Biochem.* **2002**, *71*, 593–634.
(5) Gamblin, D. P.; Scanlan, E. M.; Davis, B. G. *Chem. Rev.* **2009**, *109*, 131–163.
(6) Kent, S. B. H. *Chem. Soc. Rev.* **2009**, *38*, 338–351.

10.1021/ol902131n CCC: $40.75 
**Published on Web 10/20/2009**

To date, glycoconjugate mimetics, or neoglycoconjugates, containing a variety of unnatural linkages between the carbohydrate and aglycone moieties have also been explored.[9,10] This mimicry allows rapid and convenient access to a wide variety of neoglycopeptides and neoglycoproteins that may also maintain, or even enhance, the biological activity of the natural glycoproteins. One of the most studied of these unnatural linkages is the triazole ring formed via 1,3-dipolar cycloaddition (click chemistry) of an organic azide to a terminal alkyne.[11−14] There are a number of extensive reviews on this topic.[15−18] Previous work by Rutjes et al.[19] made use of protected building blocks to synthesize a series of triazole-linked glycosyl amino acids and dipeptides. Danishefsky's group[20] and Walsh et al.[21] have independently reported the union of unprotected carbohydrates and peptides containing a limited range of side chains using click chemistry.[22] Macmillan's group reported that the triazole linkage itself was compatible with conditions used in native chemical ligation (6 M guanidine HCl, 0.3 M $Na_2PO_4$ buffer pH 8.0, 10 mM tris(2-carboxyethyl)phosphine, 1% w/v sodium 2-mercaptoethanesulfonate (MESNA)).[23] Davis et al.[24,25] have utilized click chemistry to demonstrate the diversity of post-translational chemical protein modification. Semisynthetic lipoproteins have also been synthesized using click chemistry by Moroder et al.[26] Although the ligation of peptides using click chemistry has been reported,[27] somewhat surprisingly, a detailed investigation of reaction conditions for click reactions, at useful scale with isolation of pure products and full product characterization, between sugar azides and alkyne-containing unprotected peptides that contain all 20 amino acids found in proteins has not been carried out. Furthermore, a one-pot approach that combines native chemical ligation and click chemistry to yield neoglycopeptides has not been explored. Combining these two chemistries would provide a powerful tool to synthesize neoglycopeptides as it avoids low yielding intervening HPLC purifications steps.

Herein we report our studies on the use of click chemistry to attach an azido sugar to propargyl-containing unprotected peptides. The compatibility of the key click reaction with the thiazolidine (Thz-), Cys-, -thioester, and acetamidomethyl (Acm) moieties used in total protein synthesis by thioester-mediated amide-forming chemical ligation methods and most importantly, for the first time, the use of orthogonal native chemical ligation and copper(I)-mediated alkyne−azide cycloaddition reactions in a one-pot approach demonstrate a convenient method to access highly sophisticated assemblies of neoglycopeptides efficiently (Figure 1).

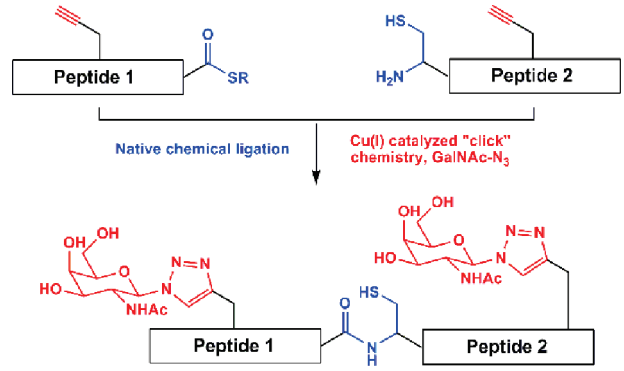



**Figure 1.** Proposed one-pot native chemical ligation and click chemistry.

To illustrate this strategy, we designed and synthesized three propargyl-containing model peptides using in situ Boc chemistry solid phase peptide synthesis (SPPS) (Figure 2).[28]

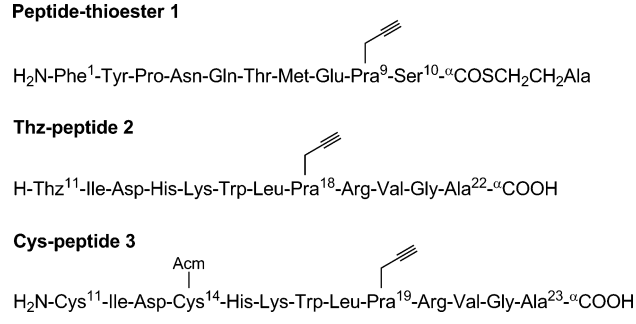
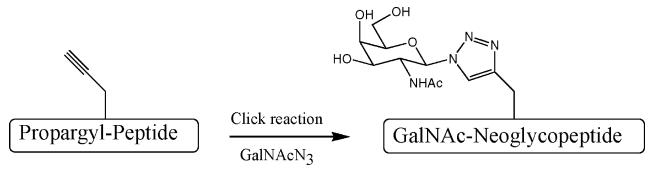



**Figure 2.** Synthesized peptides **1**, **2**, and **3**.

These three peptides contained all 20 genetically encoded amino acids found in proteins, a thioester moiety (peptide **1**), a Thz- moeity (peptide **2**), and a Cys(Acm) (peptide **3**).


(7) Pratt, M. R.; Bertozzi, C. R. *Chem. Soc. Rev.* **2005**, *34*, 58–68.
(8) Davis, B. G. *Angew. Chem., Int. Ed.* **2009**, *48*, 4674–4678.
(9) Specker, D.; Wittmann, V. *Top. Curr. Chem.* **2007**, *267*, 65–107.
(10) Nicotra, F.; Cipolla, L.; Peri, F.; La Ferla, B.; Redaelli, C. *Adv. Carbohydr. Chem. Biochem.* **2007**, *61*, 353–398.
(11) Toenøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057–3064.
(12) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596–2599.
(13) Huisgen, R. *Chem. Ber.* **1967**, *100*, 2494–2507.
(14) Huisgen, R. *Pure Appl. Chem.* **1989**, *61*, 613–628.
(15) Meldal, M.; Tornoe, C. W. *Chem. Rev.* **2008**, *108*, 2952–3015.
(16) Kolb, H. C.; Sharpless, K. B. *Drug Discovery Today* **2003**, *8*, 1128–1137.
(17) Bock, V. D.; Hiemstra, H.; van Maarseveen, J. H. *Eur. J. Org. Chem.* **2006**, *2006*, 51–68.
(18) Lutz, J. *Angew. Chem., Int. Ed.* **2007**, *46*, 1018–1025.
(19) Rutjes, F. P. J.; van Delft, F. L.; Blaauw, R. H.; Quaedflieg, P. J. L. M.; Keereweer, A. R.; Groothuys, S.; Kuijpers, B. H. M. *Org. Lett.* **2004**, *6*, 3123–3126.
(20) Danishefsky, S. J.; Chen, G.; Chen, J.; Wan, Q. *J. Org. Chem.* **2006**, *71*, 8244–8249.
(21) Walsh, C. T.; Lin, H. *J. Am. Chem. Soc.* **2004**, *126*, 13998–14003.
(22) For reviews dedicated to click neoglycoconjugates, see: (a) Dondoni, A. *Chem. Asian J.* **2007**, *2*, 700–708. (b) Field, R. A.; Nepogodiev, S. A.; Dedola, S. *Org. Biomol. Chem.* **2007**, *5*, 1006–1017.
(23) Macmillan, D.; Blanc, J. *Org. Biomol. Chem.* **2006**, *4*, 2847–2850.
(24) Davis, B. G.; van Kasteren, S. I.; Kramer, H. B.; Gamblin, D. P. *Nat. Protocols* **2007**, *2*, 3185–3194.
(25) Davis, B. G.; van Kasteren, S. I.; Kramer, H. B.; Jensen, H. H.; Campbell, S. J.; Kirkpatrick, J.; Oldham, N. J.; Anthony, D. C. *Nature* **2007**, *446*, 1105–1109.
(26) Moroder, L.; Musiol, H.; Dong, S.; Kaiser, M.; Bausinger, R.; Zumbusch, A.; Bertsch, U. *ChemBioChem* **2005**, *6*, 625–628.
(27) Xiao, J.; Tolbert, T. J. *Org. Lett.* **2009**, *11*, 4144–4147.
(28) Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Res. Ther.* **2007**, *13*, 31–44.

**Table 1.** Synthesis of Neoglycopeptides via Click Reaction of Peptides **1** and **3** with GalNAc-$N_3$




| entry | peptide[a] | catalyst | solvent | temperature | time | product yield (%)[b] |
|---|---|---|---|---|---|---|
| 1 | **3** | 20 mM CuI/30 mM DIEA | DMF | 25 °C | >12 h | <40 |
| 2 | **3** | 20 mM $CuSO_4$/30 mM NaAsc | MeCN/$H_2O$ 1:1 | 25 °C | 4 h | >85 |
| 3 | **3** | 20 mM $CuSO_4$/30 mM TCEP | 6 M GnHCl/0.2 M $Na_2HPO_4$, pH 7 | 25 °C | 3 h | >95 |
| 4 | **1** | 20 mM $CuSO_4$/30 mM NaAsc | MeCN/$H_2O$ 1:1 | 35 °C[c] | 7 h | >85 |
| 5 | **1** | 20 mM $CuSO_4$/30 mM TCEP | 6 M GnHCl/0.2 M $Na_2HPO_4$, pH 7 | 35 °C[c] | 5 h | >95 |

[a] Reactions were performed on 3 mM concentration of peptide using GalNAc-$N_3$ (5 mM). [b] Based on integration of HPLC traces, not isolated. [c] No reaction was observed after 1 h at 25 °C, at which point the temperature was increased to 35 °C.

*N*-Boc-L-propargylglycine (Boc-L-Pra-OH) was synthesized using established methods and was readily incorporated into the synthetic peptides. The propargyl group was entirely compatible with in situ Boc SPPS conditions, commonly used for the synthesis of peptide thioesters as well as the final HF treatment for deprotection and cleavage of the peptide from the resin support. *N*-Acetylgalactosamine azide (GalNAc-$N_3$) was chosen as the azide-containing sugar component. GalNAc is a common and well-studied motif for *O*-glycans found in higher eukaryotes,[29] thus it was deemed a useful neoglycan for our model study.

With the propargyl-containing peptides in hand, we set out to investigate the use of click chemistry on peptides **1** and **3** containing unprotected side chains (Table 1). After extensive investigation, it was found that a ratio of 20 mM of Cu(I) "catalyst" to 3 mM of unprotected peptide was optimal for the click reaction to occur within an acceptable time period (hours). It was also found that in situ reduction of $CuSO_4$ by NaAsc or TCEP proved to be a much better catalyst than the use of the CuI/DIEA system. Data for the click product from these two peptides are in Supporting Information, Figures S2 and S3. The successful addition of GalNAc-$N_3$ to peptides **1** and **3** demonstrated that click reactions were compatible with all the unprotected side chain functionalities, as well as with an N-terminal Cys residue, a Cys(Acm), and a peptide-thioester moiety, all of which are important functionalities for the synthesis of proteins using native chemical ligation.

The Thz- group is also important for the synthesis of large proteins by native chemical ligation.[30−32] To test whether Thz- to Cys- conversion was occurring under optimized click conditions, the Thz-peptide **2** was reacted with GalNAc-$N_3$ in the presence of $CuSO_4$ and TCEP. Both the Thz- to Cys-conversion and the click reaction were essentially complete after 1 h at room temperature (see Supporting Information). Thus, the click reaction cannot be used to directly make a Thz-neoglycopeptide from a Thz-(propargyl)peptide. However, if necessary, the resulting Cys-neoglycopeptide can subsequently be converted to the Thz-neoglycopeptide with exquisite specificity simply by addition of formaldehyde.[30]

The results reported above show that click reactions with GalNAc-$N_3$ can be successfully carried out on unprotected propargyl-containing peptides that contain any of the 20 amino acids found in proteins, in the presence of high concentrations of $CuSO_4$ and TCEP, in 6 M GnHCl/0.2 M $Na_2HPO_4$ buffer at pH 7. These are the same solvent conditions used for native chemical ligation. It has become standard practice to use TCEP as an additive in native chemical ligation reactions, to avoid the formation of mixed disulfides between the (4-carboxymethyl)thiophenol catalyst (mercaptophenylacetic acid, MPAA) and cysteine-containing reactant/product peptides.[33] We therefore postulated that it should be possible to devise conditions under which consecutive native chemical ligation and click reactions could be carried out in a one-pot fashion, without intervening isolations.

Peptide-thioester **1** (3 mM) and Cys-peptide **3** (3 mM) were dissolved in 6 M GnHCl/0.2 M $Na_2HPO_4$ (200 $\mu$L) with TCEP (40 mM) and degassed with helium for 30 min. MPAA (20 mM) was added, and the pH was adjusted to 7.0. Ligated product was detected within 5 min (Figure 3A), and after overnight reaction the ligation was complete (Figure 3B). $CuSO_4$ (20 mM) was added to the reaction mixture, followed by GalNAc-$N_3$ (10 mM). Within 2 h at room temperature more than 50% of the starting ligated peptide was converted into the desired double triazole product, and two different single triazole products (from reaction at $Pra^9$ and $Pra^{19}$) were also observed (Figure 3C). After 6 h, reaction was complete, and the desired double triazole ligated neoglycopeptide product was readily purified by reverse-phase HPLC (Figure 3D and 3E). The desired product, a 23

(29) Haase, C.; Seitz, O. *Top. Curr. Chem.* **2007**, *267*, 1–36.
(30) Rose, K.; Vizzavona, J.; Villain, M. *Chem. Biol.* **2001**, *8*, 673–679.
(31) Bang, D.; Kent, S. B. H. *Angew. Chem., Int. Ed.* **2004**, *43*, 2534–2538.
(32) Durek, T.; Torbeev, V. Y.; Kent, S. B. H. *Proc. Natl. Acad. Sci.* **2007**, *104*, 4846–4851.
(33) Johnson, E. C.; Kent, S. B. H. *J. Am. Chem. Soc.* **2006**, *128*, 6640–6646.

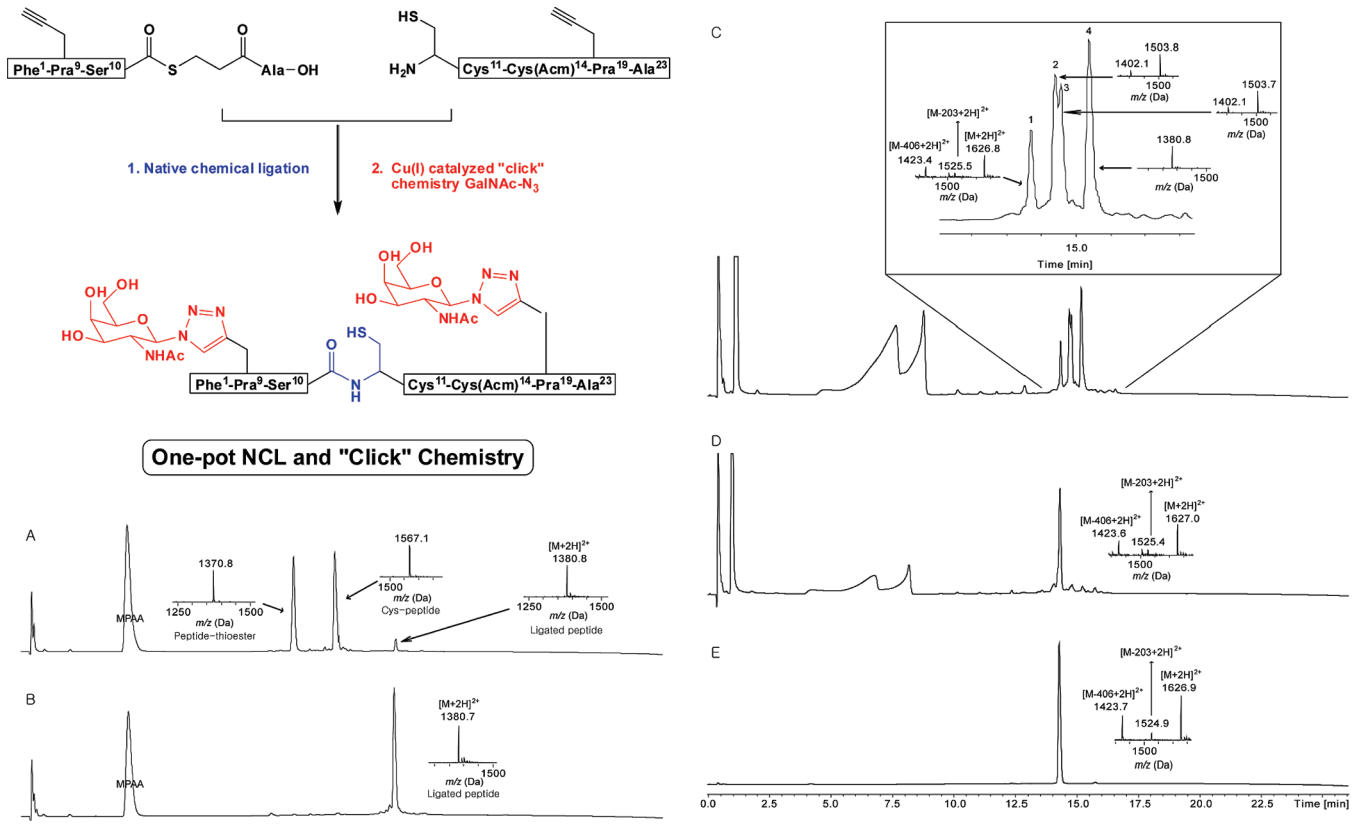



**Figure 3.** One-pot consecutive native chemical ligation plus click reactions to give a double triazole neoglycopeptide ligated product. Analytical HPLC profiles ($\lambda = 214$ nm) together with ESI MS data (inset). (A) $t = 5$ min. (B) $t =$ overnight reaction. At this stage, $CuSO_4$ (20 mM) and GalNAc-$N_3$ (10 mM) were added. (C) $t = 2$ h after beginning of click reaction. Peak 1 is the desired double triazole product; peaks 2 and 3 are single triazole products at different reaction sites; and peak 4 is the starting ligated peptide. (D) Total crude products at $t = 6$ h; formation of the double trizaole neoglycopeptide product was completed. (The broad, early eluting peaks in (C) and (D) are chromatography artifacts.) (E) Purified 23 residue double trizaole neoglycopeptide product from one-pot native chemical ligation plus click chemistry reactions.

residue neoglycopeptide, had an observed mass of 3252.0 ± 0.4 Da (theoretical mass 3252.6 Da (average isotope composition)).

In conclusion, we have described a novel combination of click chemistry and native chemical ligation for the fully convergent synthesis of neoglycopeptides. The two ligation chemistries are fully compatible, and can be carried out in any order. The optimized click conditions reported here are compatible with all 20 genetically encoded amino acids, N-terminal Cys, the peptide-thioester moiety, and a Cys(Acm) residue, all of which are used for chemical protein synthesis by native chemical ligation. Native chemical ligation and click chemistry have been successfully combined to affect an efficient one-pot synthesis of a neoglycopeptide containing two sugars, starting from three unprotected building blocks (two peptides and a sugar azide). These findings provide a versatile chemical toolkit for the fully convergent synthesis of neoglycoproteins using click chemistry, native chemical ligation, and kinetically controlled ligation.[34] Syntheses involving the separate functionalization of each peptide building block with distinct sugars or using more complex carbohydrate moieties are underway in our laboratories, using erythopoietin as an initial target.

**Acknowledgment.** We thank the Maurice Wilkins Centre for Molecular Discovery for financial assistance.

**Supporting Information Available:** Experimental data for syntheses of building blocks, model peptides **1**, **2**, and **3**, and the click chemistry and ligation studies (including HPLC traces). This material is available free of charge via the Internet at http://pubs.acs.org.

OL902131N

(34) Bang, D.; Pentelute, B. L.; Kent, S. B. H. *Angew. Chem., Int. Ed.* **2006**, *45*, 3985–3988.