N-Glycopeptide Synthesis

International Edition: DOI: 10.1002/anie.201605597
German Edition: DOI: 10.1002/ange.201605597

N-Linked Glycosyl Auxiliary-Mediated Native Chemical Ligation on Aspartic Acid: Application towards *N*-Glycopeptide Synthesis

*Hua Chai[+], Kim Le Mai Hoang[+], Minh Duy Vu, Kalyan Pasunooti, Chuan-Fa Liu, and Xue-Wei Liu**

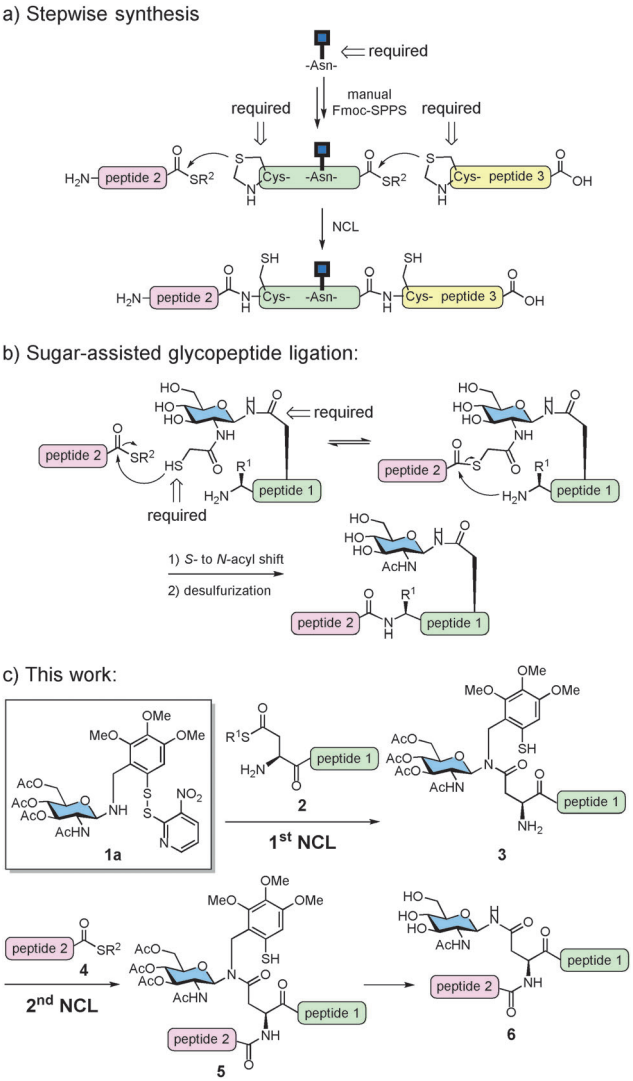
Abstract: *A practical approach towards N-glycopeptide synthesis using an auxiliary-mediated dual native chemical ligation (NCL) has been developed. The first NCL connects an N-linked glycosyl auxiliary to the thioester side chain of an N-terminal aspartate oligopeptide. This intermediate undergoes a second NCL with a C-terminal thioester oligopeptide. Mild cleavage provides the desired N-glycopeptide.*

Glycoproteins play vital roles in a wide array of biological activities, including cell–cell adhesion, cell differentiation, cell growth, and tumor metastasis.[1] Furthermore, a host of major illnesses, such as autoimmune diseases, infectious diseases, and cancer[2] are caused by erroneous glycosylation of proteins. Consequently, the development of new glycopeptide-based vaccines, diagnostics and therapeutics has gained significant interests, stimulating the study of glycopeptide synthesis.[3]

Native chemical ligation (NCL) is one of the most useful methods in protein synthesis. The NCL reaction enables an amide bond to be formed between the N-terminal cysteine residue of an unprotected oligopeptide and C-terminal thioester residue of a second oligopeptide, via a reversible *trans*-thioesterification step, followed by an irreversible intramolecular *S*- to *N*-acyl transfer. This permits the ligation of unprotected peptide chains, in aqueous solution at neutral pH, to form a longer peptide chain in near quantitative conversion without any further manipulation. Since early studies by Kent's group,[4] the potential of NCL has greatly expanded beyond cysteine to include other ligation sites, such as Ala,[5] Phe,[6] Val,[7] Thr,[8] and Lys[9] residues.

Even though NCL is a mature field, NCL in the context of glycoprotein synthesis is progressing at a lesser pace. A survey[3e,10] of current methods of synthesizing *N*-glycoprotein discussed the challenges of incorporating glycans into the established workflow of peptide synthesis. SPPS methods can provide glycopeptides about 20 amino acids long. Beyond this, sporadic aggregations and side reactions often make further elongations impractical.[11] Hence, additional ligation sites are required for further NCL with other fragments to obtain the final glycoprotein. Given the relatively low abundance of cysteine in nature (ca. 1.7%), having this residue at a strategically useful location is not always feasible.

In a stepwise synthesis (Scheme 1a), a natural cysteine or a synthetic sulfur-containing amino acid in the vicinity of glycan-linked-Asn residue is required. Otaka's group recently reported a thiolated asparagine as ligation linker for NCL.[12] The glycan is usually of full-length or could be introduced



Scheme 1. Contemporary *N*-glycopeptide synthesis.

[*] H. Chai,[+] Dr. K. Le Mai Hoang,[+] M. D. Vu, Prof. Dr. X.-W. Liu
School of Physical and Mathematical Sciences
Nanyang Technological University
21 Nanyang Link, Singapore 637371 (Singapore)
E-mail: xuewei@ntu.edu.sg
Dr. K. Pasunooti, Prof. Dr. C.-F. Liu
School of Biological Sciences, Nanyang Technological University
60 Nanyang Drive, Singapore 637551 (Singapore)
[+] These authors contributed equally to this work.

Supporting information for this article can be found under: http://dx.doi.org/10.1002/anie.201605597.

through enzymatic *trans*-glycosylation after the ligation step. In the sugar-assisted ligation approach (Scheme 1b), Wong's group[13] demonstrated that a non-cysteine terminated glycan-linked peptide fragment can be used to join another standard thioester peptide. This allows less fragmentation of the peptide synthetic plan, accelerating the assembly of the *N*-glycopeptide. However, the sulfur appendage needs to be installed on the sugar unit, followed by glycosylation with Asn residue. Preparation of this crucial species is no trivial task which requires laborious work and careful planning, especially with more complex glycan chains.[13c] The modified sugar is also resistant to enzymatic *trans*-glycosylation and an additional desulfurization step is required. Understandably, the sulfhydryl group only participated in ligating the two peptide fragments and not in the glycosylation of the amino acid itself. Nevertheless, the method elegantly overcomes the requirement for cysteine residue at the ligation junction.

We studied the scope of NCL with modified glycosyl amine **1a** (Scheme 1c). During our preliminary study, the trimethoxybenzyl auxiliary was found to be the most proficient in the ligation and in its removal, in line with previous studies.[14] The 3-nitro-2-pyridinesulfenyl group is well-known for rapid disulfide exchange under NCL condition.[15] This scaffold is therefore capable of performing ligation to produce the glycan-linked N-terminal asparagine oligopeptide **3**, followed by a second ligation with C-terminal thioester oligopeptide **4** to provide the glycopeptide **5**. Subsequently, the auxiliary is readily cleaved following treatment with TFA and deacetylation to furnish glycopeptide **6**.

We conducted our first ligation of compound **1a** with oligopeptide sequence H-(COSPh)D**A**TGVS-CONH$_2$ **12a** in a phosphate buffer (pH 7.5) at 37 °C. Reaction progress was monitored by reverse-phase HPLC, and a ligation conversion of more than 70% was observed after 1 hour. Optimization of additives and buffer revealed the best ligation efficiency was achieved when the reaction was conducted in a phosphate buffer of pH 8.0 with 6 M Gn·HCl, 60 mM TCEP, and 1% *v*/*v* MeSNa as additives at 37 °C. The desired ligation product **13a** was formed in 92% conversion within 1 hour with less than 3% of peptide racemization at N-terminal Asp residue (Figure 1). A control experiment was carried out by mixing thioesters **12a** and additives in our buffer condition. Less than 3% racemization was observed after 12 hours and only 7% racemization was detected with 72 hours of incubation time.

To examine the effect of different residues at the ligation junction, a series of oligopeptides (**12a**–**12h**) with thioesterified aspartic acid on the N-terminal were subjected to the optimized conditions. As seen from Table 1, all native chemical ligation proceeded smoothly and high to excellent conversions could be achieved. Slightly lower conversions were observed when Ile, Leu, Val, and Phe were present next to the ligation site, which is most likely due to increase in steric bulk. Liu's group[16] recently developed a facile method for in situ formation of peptide thioester from a C-terminal hydrazide, allowing the convergent synthesis of several proteins.[17] In our experiment, the hydrazide peptide **12i** gave the first ligation product in excellent conversion (entry 9).

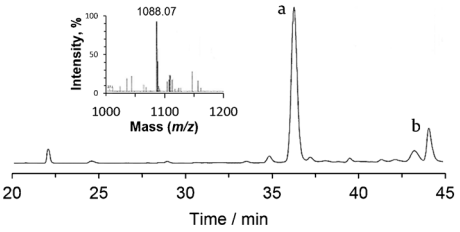
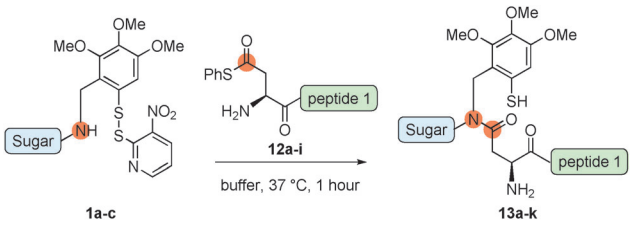


Figure 1. Analytical RP-HPLC ($\lambda = 280$ nm) of the first ligation (Table 1, entry 1). Reaction at 1 h: peak a: ligation product **13a** with calcd $M = 1087.4$ Da, obs. $M = 1088.1$ Da (see inset, ESI-MS); peak b: excess of **1a** (disulfide-bond cleaved).

Table 1: First native chemical ligation of oligopeptide thioester **12** and sugar auxiliaries **1a**–**c**.[a]



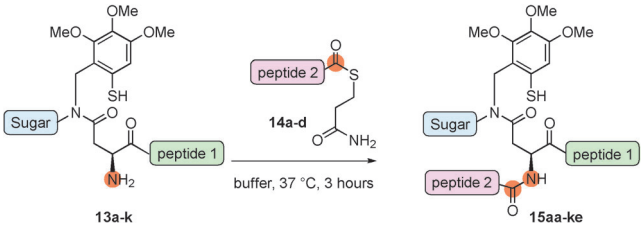
Entry	Peptide 1[b]	Product	Conv [%][c]
1[d]	**12a** A**T**GVS-CONH$_2$	**13a**	92
2[d]	**12b** G**T**GVS-CONH$_2$	**13b**	91
3[d]	**12c** I **T**GVS-CONH$_2$	**13c**	89
4[d]	**12d** L**T**GVS-CONH$_2$	**13d**	90
5[d]	**12e** F**T**GVS-CONH$_2$	**13e**	87
6[d]	**12f** S**T**GVS-CONH$_2$	**13f**	94
7[d]	**12g** T**T**GVS-CONH$_2$	**13g**	91
8[d]	**12h** V**T**GVS-CONH$_2$	**13h**	90
9[d]	**12i** G**T**GVS-CONHNH$_2$	**13i**	93
10[e]	**12i** G**T**GVS-CONHNH$_2$	**13j**	87
11[f]	**12i** G**T**GVS-CONHNH$_2$	**13k**	81

[a] Conditions: 6 M Gn·HCl, 60 mM TCEP, 1% MeSNa, pH 8.0, 37 °C, 1 h. [b] ratio of **1** to **12** is 1.3:1. Amino acids adjacent to aspartic acid are highlighted in bold. [c] Conversion based on RP-HPLC analysis upon complete consumption of **1**. [d] **1a**: Glc(OAc)NAc. [e] **1b**: unprotected GlcNAc. [f] **1c**: unprotected maltotriose.

The β-*N*-glycosidic bond of **13i** remained intact based on 1D-TOCSY NMR analysis (see the Supporting Information). Evaluation of the reaction with unprotected GlcNAc and unprotected maltotriose as glycan auxiliaries showed excellent conversions without compatibility issue (entries 10, 11). However, we found unprotected sugar auxiliary more prone to hydrolysis compared to peracetylated form. Anomerization of the glycosidic benzyl amine was observed to occur at room temperature after several hours. Nevertheless, glycopeptide **13j** and **13k** were stable during isolation and purification. We recommend carrying out the first ligation reaction immediately after furnishing the auxiliary from unprotected glycosylamine.

We then investigated the second native chemical ligation of glycopeptide **13** with a C-terminal oligopeptide thioester. Under similar conditions, glycopeptide **13a** was treated with oligopeptide thioester H-IST**V**S-COS(CH$_2$CH$_2$CONH$_2$) **14a** (Table 2, entry 1). A final ligation conversion of 78% was

Table 2: Second native chemical ligation of oligopeptide thioester **14** and glycopeptide **13**.[a]



Entry	Peptide 1[b]	Peptide 2[b]	Product	Conv [%][c]
1[d]	**13a** ATGVS-$CONH_2$	**14a** H-ISTV**S**	**15aa**	78
2[d]	**13a** ATGVS-$CONH_2$	**14b** H-ISTV**A**	**15ab**	82
3[d]	**13a** ATGVS-$CONH_2$	**14c** H-ISTV**G**	**15ac**	78
4[d]	**13c** I TGVS-$CONH_2$	**14a** H-ISTV**S**	**15ca**	72
5[d]	**13c** I TGVS-$CONH_2$	**14c** H-ISTV**G**	**15cc**	74
6[d]	**13e** FTGVS-$CONH_2$	**14a** H-ISTV**S**	**15ea**	65
7[d]	**13e** FTGVS-$CONH_2$	**14c** H-ISTV**G**	**15ec**	66
8[d]	**13f** STGVS-$CONH_2$	**14b** H-ISTV**A**	**15fb**	71
9[d]	**13h** VTGVS-$CONH_2$	**14a** H-ISTV**S**	**15ha**	74
10[d]	**13f** STGVS-$CONH_2$	**14d** H-ISTV**F**	**15fd**	54
11[d]	**13i** GTGVS-$CONHNH_2$	**14e** H-Thz-STV**G**	**15ie**	75
12[e]	**13j** GTGVS-$CONHNH_2$	**14e** H-Thz-STV**G**	**15je**	68
13[f]	**13k** GTGVS-$CONHNH_2$	**14e** H-Thz-STV**G**	**15ke**	50

[a] Conditions: 6 M Gn·HCl, 60 mM TCEP, 1% MeSNa, pH 8.0, 37 °C, 3 h. [b] Ratio of **13** to **14** is 1:3. Amino acids adjacent to asparagine residue are highlighted in bold. [c] Conversion based on RP-HPLC analysis upon complete consumption of **13**. [d] Glc(OAc)NAc. [e] Unprotected GlcNAc. [e] Unprotected maltotriose.

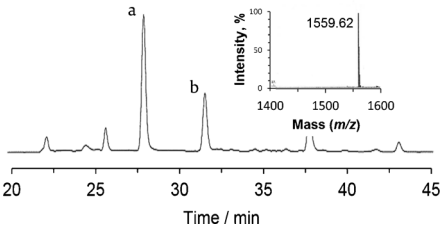
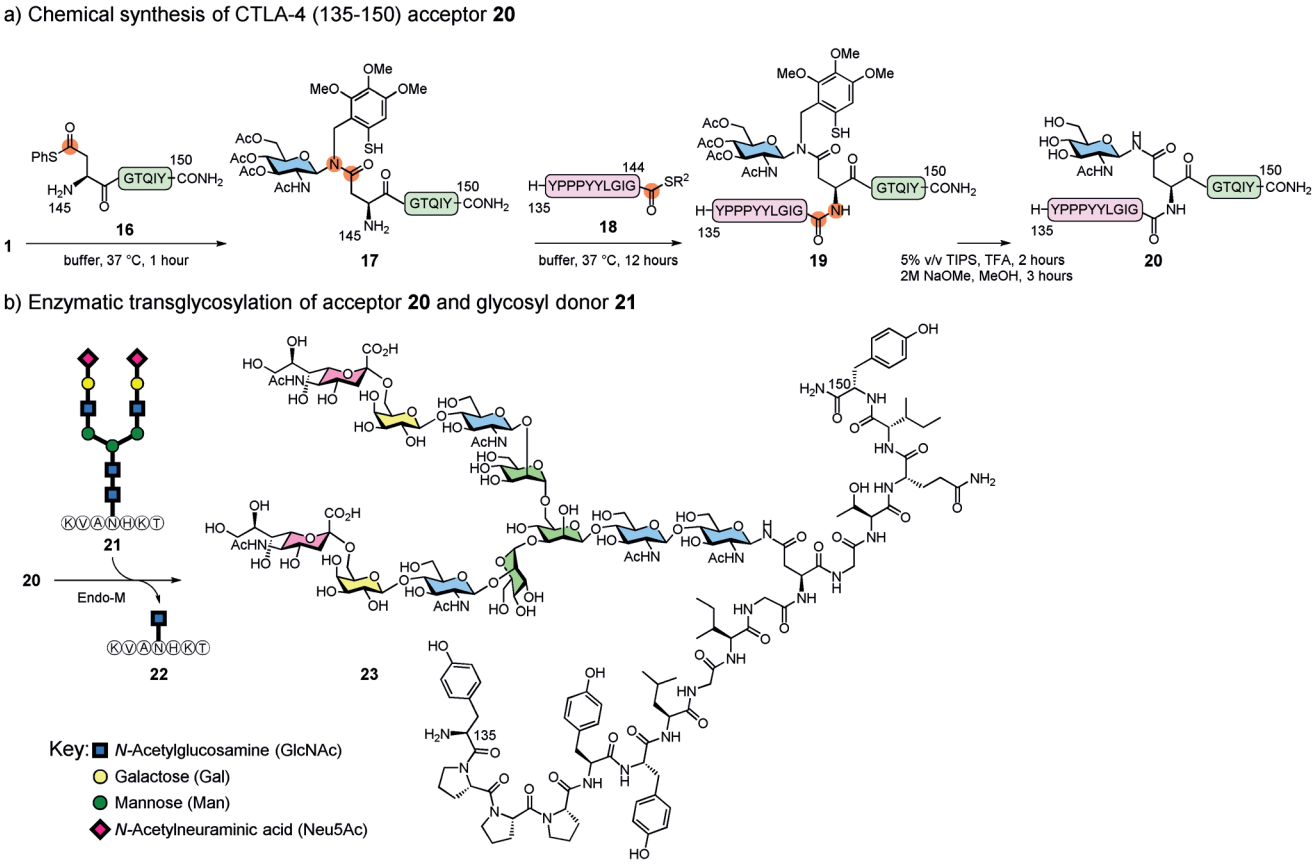


Figure 2. Analytical RP-HPLC ($\lambda = 280$ nm) of second ligation (Table 2, entry 1). Reaction at 3 h: peak a: exchanged MeS thioester of **14a**; peak b: ligation product **15aa** with calcd $M = 1558.7$ Da, obs. $M = 1559.6$ Da (see inset, ESI-MS).

obtained after 3 hours (Figure 2). Tandem mass spectrometry was obtained to confirm the integrity of product **15aa** (see the Supporting Information). Modifications including varying pH or changing additives did not further improve the conversion. The condition was applied to the coupling of some glycopeptides **13** with various thioesters peptide (**14a–c**) and moderate to high conversions were observed (entries 1–9). The second ligation was tested with branched C-terminal sequence H-ISTV**F**-COS($CH_2CH_2CONH_2$) **14d** (entry 10). The drop in conversion suggested a higher influence of sterically demanding residue on ligation efficacy. The thioester **14e** with Thz-protected cysteine at the N-*t*erminal was subjected to second NCL with three different glycopeptides **13i-13k** (entries 11–13, respectively). A moderate second ligation with malto-



Scheme 2. Synthesis of glycoprotein CTLA-4 (135–150) segment.

triose glycopeptide **13k** was observed, probably due to the increase in steric bulk of the sugar moiety. The β-*N*-glycosidic bond of **15ie** remained intact based on 1D-TOCSY NMR analysis (see the Supporting Information).

Encouraged by these results, we attempted to adopt our method into the synthesis of a homogeneous glycoprotein. The cytotoxic T-lymphocyte-associated antigen-4 (CTLA-4)[18] was chosen as our target molecule and we present herein the synthesis of sialyl-glycopeptide CTLA-4 fragment (135–150) **23** via our approach (Scheme 2). The CTLA-4 segment (145–150) **16** was first prepared by Fmoc SPPS and ligated with **1** to give glycopeptide **17** in 93% conversion under our conditions. **17** was subjected to a second NCL with C-terminal thioesterified CTLA-4 (135–144) fragment 18 for 12 hours to afford glycopeptide **19** in 64% conversion. We decided to introduce the complex-type sialyloligosaccharide via *endo*-M catalyzed *trans*-glycosylation.[19] Finally, deprotection of glycopeptide **19** followed by incubation of resulting acceptor **20** with glycosyl donor **21** for 12 hours at 37 °C gave the desired glycoprotein CTLA-4 (135–150) fragment **23** in 52% overall conversion after purification by RP-HPLC (Figure 3). Preparation of the remaining fragments of CTLA-4 are underway in our laboratory and will be reported in due course.

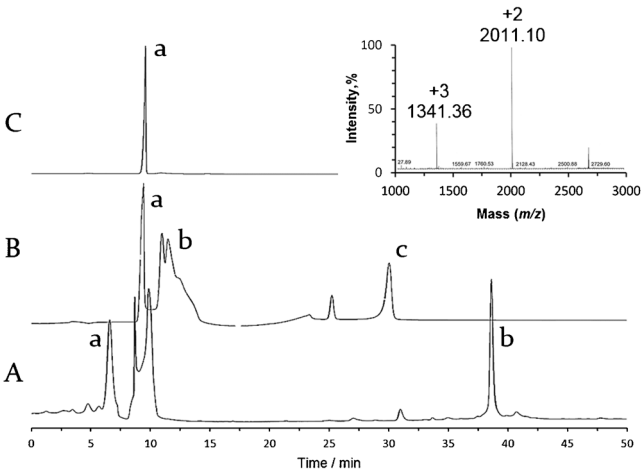


Figure 3. Analytical RP-HPLC ($\lambda = 280$ nm) of *trans*-glycosylation reaction: A) Reaction at 1 h: peak a: glycosyl donor **21**; peak b: glycosyl acceptor **20**. B) Reaction at 12 h: peak a: product **23**; peak b: *endo*-M; peak c: product **22**; C) purified **23** with calcd M = 4021.10 Da, obs. M = 4022.1 Da (see inset, ESI-MS).

In conclusion, we demonstrated an attractive and practical procedure to prepare *N*-glycopeptide that does not require cysteine residue, allowing for straightforward glycosylation and ligation at the aspartic acid terminal. The dual-NCL capability of a single scaffold offers higher conversion rate and atomic economy. Since we are not relying on sequence specific or sugar-linked peptide fragments, the method should provide more flexibility for complex glycopeptide synthesis.

Acknowledgements

Financial support from Nanyang Technological University (RG6/13 and RG132/14) and the Ministry of Education, Singapore (MOE 2013-T3-1–002) are gratefully acknowledged.

Keywords: aspartic acid · auxiliary mediation · CTLA-4 · dual native chemical ligation · *N*-glycopeptide synthesis

[1] a) R. Apweiler, H. Hermjakob, N. Sharon, *Biochim. Biophys. Acta Gen. Subj.* **1999**, *1473*, 4–8; b) R. G. Spiro, *Glycobiology* **2002**, *12*, 43R–56R; c) R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683–720.

[2] a) A. Varki, *Glycobiology* **1993**, *3*, 97–130; b) L. H. Miller, M. F. Good, G. Milon, *Science* **1994**, *264*, 1878–1883; c) D. H. Dube, C. R. Bertozzi, *Nat. Rev. Drug Discovery* **2005**, *4*, 477–488; d) C. N. Scanlan, J. Offer, N. Zitzmann, R. A. Dwek, *Nature* **2007**, *446*, 1038–1045.

[3] a) A. Dove, *Nat. Biotechnol.* **2001**, *19*, 913–917; b) L. Liu, C. S. Bennett, C.-H. Wong, *Chem. Commun.* **2006**, 21–33; c) H. Hojo, Y. Nakahara, *Biopolymers* **2007**, *88*, 308–324; d) R. J. Payne, C.-H. Wong, *Chem. Commun.* **2010**, *46*, 21–43; e) C. Unverzagt, Y. Kajihara, *Chem. Soc. Rev.* **2013**, *42*, 4408–4420.

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

[5] a) L.-Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526–533; b) Q. Wan, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **2007**, *46*, 9248–9252; *Angew. Chem.* **2007**, *119*, 9408–9412.

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

[7] a) C. Haase, H. Rohde, O. Seitz, *Angew. Chem. Int. Ed.* **2008**, *47*, 6807–6810; *Angew. Chem.* **2008**, *120*, 6912–6915; b) J. Chen, Q. Wan, Y. Yuan, J.-L. Zhu, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **2008**, *47*, 8521–8524; *Angew. Chem.* **2008**, *120*, 8649–8652.

[8] J. Chen, P. Wang, J.-L. Zhu, Q. Wan, S. J. Danishefsky, *Tetrahedron* **2010**, *66*, 2277–2283.

[9] a) R.-L. Yang, K. K. Pasunooti, F.-P. Li, X.-W. Liu, C.-F. Liu, *J. Am. Chem. Soc.* **2009**, *131*, 13592–13593; b) K. K. Pasunooti, R.-L. Yang, S. Vedachalam, B. K. Gorityala, C.-F. Liu, X.-W. Liu, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6268–6271; c) R.-L. Yang, K. K. Pasunooti, F.-P. Li, X.-W. Liu, C.-F. Liu, *Chem. Commun.* **2010**, *46*, 7199–7201.

[10] a) T. Buskas, S. Ingale, G.-J. Boons, *Glycobiology* **2006**, *16*, 113R–136R; b) C. Bello, S. Wang, L. Meng, K. W. Moremen, C. F. W. Becker, *Angew. Chem. Int. Ed.* **2015**, *54*, 7711–7715; *Angew. Chem.* **2015**, *127*, 7823–7828.

[11] M. Murakami, T. Kiuchi, M. Nishihara, K. Tezuka, R. Okamoto, M. Izumi, Y. Kajihara, *Sci. Adv.* **2016**, *2*, e1500678.

[12] K. Sato, K. Kitakaze, T. Nakamura, N. Naruse, K. Aihara, A. Shigenaga, T. Inokuma, D. Tsuji, K. Itoh, A. Otaka, *Chem. Commun.* **2015**, *51*, 9946–9948.

[13] a) A. Brik, Y.-Y. Yang, S. Ficht, C.-H. Wong, *J. Am. Chem. Soc.* **2006**, *128*, 5626–5627; b) S. Ficht, R. J. Payne, A. Brik, C.-H. Wong, *Angew. Chem. Int. Ed.* **2007**, *46*, 5975–5979; *Angew. Chem.* **2007**, *119*, 6079–6083; c) C. S. Bennett, S. M. Dean, R. J. Payne, S. Ficht, A. Brik, C.-H. Wong, *J. Am. Chem. Soc.* **2008**, *130*, 11945–11952.

[14] a) J. Offer, C. N. C. Boddy, P. E. Dawson, *J. Am. Chem. Soc.* **2002**, *124*, 4642–4646; b) B. Wu, J.-H. Chen, J. D. Warren, G. Chen, Z.-H. Hua, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **2006**, *45*, 4116–4125; *Angew. Chem.* **2006**, *118*, 4222–4231; c) J. Offer, *Biopolymers* **2010**, *94*, 530–541.

[15] a) R. Matsueda, S. Higashida, R. J. Ridge, G. R. Matsueda, *Chem. Lett.* **1982**, *11*, 921 – 924; b) M. S. Bernatowicz, R. Matsueda, G. R. Matsueda, *Int. J. Pept. Protein Res.* **1986**, *28*, 107 – 112.
[16] G.-M. Fang, Y.-M. Li, F. Shen, Y.-C. Huang, J.-B. Li, Y. Lin, H.-K. Cui, L. Liu, *Angew. Chem. Int. Ed.* **2011**, *50*, 7645 – 7649; *Angew. Chem.* **2011**, *123*, 7787 – 7791.
[17] a) G.-M. Fang, J.-X. Wang, L. Liu, *Angew. Chem. Int. Ed.* **2012**, *51*, 10347 – 10350; *Angew. Chem.* **2012**, *124*, 10493 – 10496; b) P. Siman, S. V. Karthikeyan, M. Nikolov, W. Fischle, A. Brik, *Angew. Chem. Int. Ed.* **2013**, *52*, 8059 – 8063; *Angew. Chem.* **2013**, *125*, 8217 – 8221; c) M. T. Weinstock, M. T. Jacobsen, M. S. Kay, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 11679 – 11684.
[18] N. Yamamoto, Y. Takayanagi, A. Yoshino, T. Sakakibara, Y. Kajihara, *Chem. Eur. J.* **2007**, *13*, 613 – 625.
[19] a) K. Haneda, T. Inazu, K. Yamamoto, H. Kumagai, Y. Nakahara, A. Kobata, *Carbohydr. Res.* **1996**, *292*, 61 – 70; b) L.-X. Wang, J. V. Lomino, *ACS Chem. Biol.* **2012**, *7*, 110 – 122.

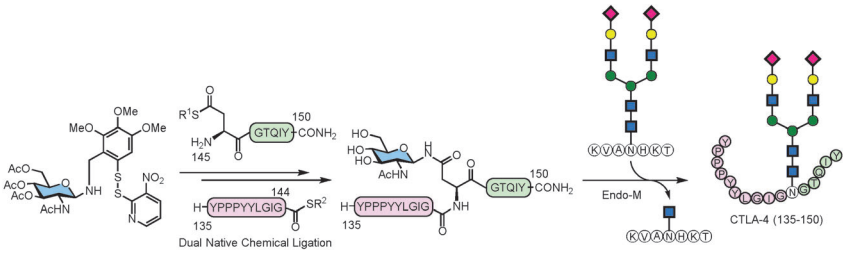
Received: June 9, 2016
Published online: ■■ ■■, ■■■■

Communications

N-Glycopeptide Synthesis

H. Chai, K. Le Mai Hoang, M. D. Vu, K. Pasunooti, C.-F. Liu, X.-W. Liu* _____ ■■■■–■■■■

N-Linked Glycosyl Auxiliary-Mediated Native Chemical Ligation on Aspartic Acid: Application towards *N*-Glycopeptide Synthesis



A practical approach towards *N*-glycopeptide synthesis using an auxiliary-mediated dual native chemical ligation (NCL) has been developed. The first NCL connects an *N*-linked glycosyl auxiliary to the thioester side chain of an *N*-terminal aspartate oligopeptide. This intermediate undergoes a second NCL with a *C*-terminal thioester oligopeptide. Mild cleavage provides the desired *N*-glycopeptide.

