J|A|C|S COMMUNICATIONS

Published on Web 09/03/2009

Dual Native Chemical Ligation at Lysine

Renliang Yang,[†] Kalyan Kumar Pasunooti,[‡] Fupeng Li,[†] Xue-Wei Liu,*[,‡] and Chuan-Fa Liu*[,†]

Division of Chemical Biology and Biotechnology, School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, and Division of Chemistry and Biological Chemistry, School of Physical and Mathematical Sciences, Nanyang Technological University, 1 Nanyang Walk, Singapore 637616

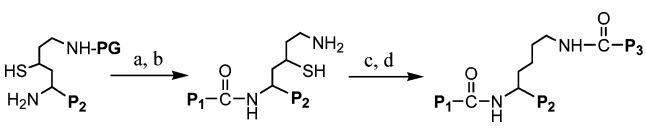
Received July 3, 2009; E-mail: xuewei@ntu.edu.sg; cfliu@ntu.edu.sg

Protein chemical synthesis through chemoselective peptide ligation is becoming an increasingly important enabling technology in the study of proteins.[1] A particularly useful ligation method is native chemical ligation (NCL), which involves the reaction of a thioester peptide with another cysteinyl peptide to form an Xaa−Cys bond at the ligation junction.[2,3]

Since cysteine is an uncommon amino acid in naturally occurring proteins, several strategies have been developed to expand the application scope of NCL. One strategy uses the appending of a removable thiol-containing auxiliary group on the N-terminal amine to mediate ligation.[4] Recently, this strategy was employed to introduce a ubiquitin moiety through ligation at a Gly−Gly junction on the side chain of Lys.[5] A drawback of this approach is that the ligation reaction exhibits very slow kinetics because of steric hindrance imposed by the resultant secondary amine; thus, its practical utility is limited only to unhindered ligation junctions. Ligation at Met is made possible through the use of homocysteine, which is converted to methionine postligationally by S-methylation.[6] Another strategy employs desulfurization to convert the post-NCL cysteine to alanine.[7] Similarly, other amino acids containing a surrogate thiol have been used in NCL followed by desulfurization, making phenylalanine[8] and valine[9] accessible as ligation sites.

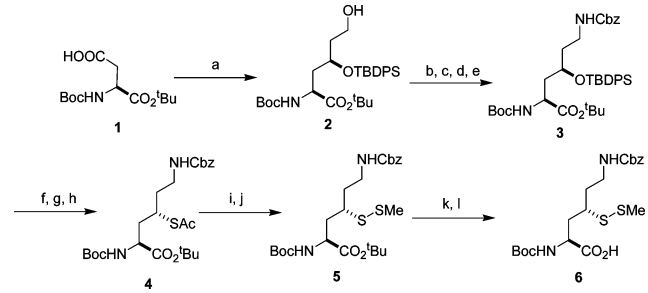
Herein, we report a dual native chemical ligation strategy through the use of a 4-HS-lysine residue (Scheme 1). Lysine is a frequently occurring amino acid in proteins. In addition, the ε-amino group of lysine provides a platform for post-translational protein modifications, such as methylation, acetylation, and ubiquitination or sumoylation. The side-chain amine of lysine is also often used as an anchor point for labeling proteins with a biophysical or biochemical tag. We realized that a single thiol group introduced on the γ-carbon of lysine would mediate ligation at both the α- and ε-amines, in each case via a six-membered-ring reaction intermediate, to form native and isopeptide bonds, respectively (Scheme 1). Therefore, by using a proper protecting group, one can perform two consecutive ligation steps via the same thiol to synthesize proteins containing specifically modified (e.g., ubiquitinated or biotinylated) lysine residues.

Scheme 1. General Scheme for Dual Chemical Ligation: (a) First Ligation with P_1-COSR; (b) Deprotection; (c) Second Ligation with P_3-COSR; (d) Desulfurization



Scheme 2 shows the synthesis of the protected γ-thiol-substituted lysine analogue **6**. Essentially, the method of Guichard and coworkers[10] was adopted to prepare a 4-hydroxylysine derivative of which the 4-hydroxyl group was subsequently converted to a thiol. Starting from (*S*)-Boc-Asp-O^{t}Bu **1**, the O-silyl-protected 1,3-diol **2** was prepared in enantiopure form through a key δ-lactam intermediate using known literature procedures.[10] The terminal alcohol in **2** was mesylated and further converted to an azide.[10] The azide was then reduced by catalytic hydrogenation to an amine, which was protected using CbzCl to afford **3**. After deprotection of the silyl ether in **3** with TBAF, mesylation of the secondary alcohol followed by nucleophilic substitution of thioacetate afforded **4** in good yields. Following saponification, the thiol was protected in disulfide form using MMTS to furnish **5** according to literature methods.[8,9b] Acidolytic deprotection with TFA and further protection with Boc anhydride yielded **6**, the protected form of 4-mercaptolysine ready for use in peptide synthesis.

Scheme 2. Synthesis of **6**[a]



[a] Reagents and conditions: (a) ref 10; (b) MsCl, DIPEA, 0 °C; (c) NaN_3, DMF, 80 °C, 83% yield over two steps; (d) H_2, Pd/C, ethyl acetate, rt, quantitative yield; (e) CbzCl, $NaHCO_3$, 2:1 dioxane/water, 0 °C, 81%; (f) TBAF, THF, 0 °C, 77%; (g) MsCl, DIPEA, 0 °C; (h) CH_3COSK, DMF, 40 °C, 70% over two steps; (i) NaOH, MeOH, rt; (j) (*S*)-methyl methanethiosulfonate (MMTS), triethylamine, CH_2Cl_2, rt, 50% over two steps; (k) 95% TFA, H_2O, rt; (l) Boc_2O/TEA, MeOH, rt, 78% over two steps.

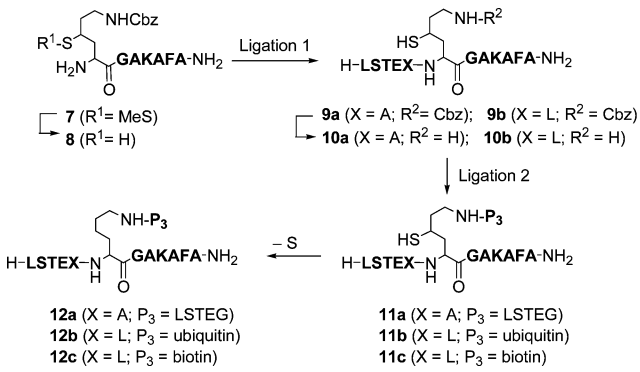
The 4-mercaptolysine derivative was introduced as the N-terminal residue of the small peptide **7** with Cbz remaining on its ε-amine. Peptide **7** was directly used in ligation reactions, as the free-thiol form **8** could be readily generated in situ under the reducing conditions of the ligation reaction (Scheme 3).

Robust ligation was observed when **7** was subjected to reaction with the thioester peptide H-**LSTEA**-COSR at pH 8. After reaction for 1 h at 37 °C, the 4-mercaptolysyl peptide limiting reactant was completely consumed, and the ligation product **9a** was obtained in ∼90% yield based on HPLC analysis. Ligation with another thioester peptide with a larger C-terminal Leu residue, H-**LSTEL**-COSR, also proceeded efficiently (Figure 1), yielding the ligation product **9b** in 92% yield by HPLC analysis after reaction for 1 h at 37 °C, indicating that the ligation reaction was not affected by the bulkiness of Leu in either reaction rate or yield. The Cbz group in **9** was then removed using a cocktail containing TFMSA in order to expose the ε-NH_2 of 4-mercaptolysine for the second ligation. First, the purified peptide **10a** was reacted with a small peptide thioester, H-**LSTEG**-COSR. Similar to the first ligation at the α-amine, the second ligation proceeded very

[†] Division of Chemical Biology and Biotechnology.
[‡] Division of Chemistry and Biological Chemistry.

10.1021/ja905491p CCC: $40.75

Scheme 3. Demonstration of Peptide Dual Ligation[a]



[a] See the Supporting Information for experimental details and characterization data.

efficiently, reaching completion in 1.5 h to give the product **11a** in 92% yield by HPLC analysis. A control experiment performed with the thioester peptide H-**LSTEA**-COSR and lysyl peptide H-**KGAKAFA**-NH_2 for 2 h showed no detectable amount (<2%) of direct aminolysis at either the α- or ε-amine. This indicated the vital role of the γ-SH group in mediating ligation at both the α- and ε-amino groups of the 4-mercoptolysine residue. Next, to verify whether our method can be used for site-specific peptide ubiquitination, peptide **10b** was reacted with a large (76 amino acids) ubiquitin thioester, ubi(1−76)-CO-$SCH_2CH_2SO_3^-$ [ubi(1−76)-MES], generated by thiolysis of a ubiquitin−intein fusion protein.[11] As shown in Figure 2, ligation with the ubiquitin thioester was complete within only 45 min, giving a very clean product in >90% yield based on HPLC analysis. The efficiency of the reaction in our dual NCL scheme makes it a particularly viable method for the synthesis of complex protein conjugates such as ubiquitinated proteins for the functional elucidation of such post-translational modifications on lysine.

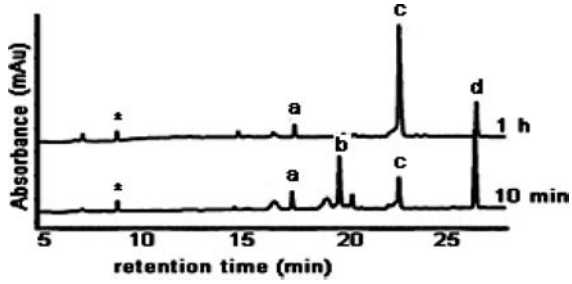


Figure 1. HPLC monitoring of ligation between H-**LSTEL**-COSR and peptide **8** at (bottom) 10 min and (top) 1 h. Gradient: 0−60% buffer B in 30 min. Peak a, H-**LSTEL**-COSR; peak b, peptide **8**; peak c, ligation product peptide **9b**; peak d, H-**LSTEL**-COSBz; peak *, H-**LSTEL**-OH. Ligation conditions: 28 mM H-**LSTEL**-COSR, 20 mM **8**, 6 M Gdn·HCl, 0.2 M phosphate, 60 mM TCEP, 1% benzyl mercaptan, pH 8.0, 37 °C.

To generate native Lys at the ligation junction, we first tried the Raney nickel-mediated desulfurization method.[7] Desulfurization of **11a** gave only a moderate yield of 44% (based on quantitative HPLC analysis) after 8 h of reaction. Moreover, this metal-based method did not work for **11b**, which is much larger and more complex than **11a**. We then tried the recently developed free-radical desulfurization approach.[12] Desulfurization of **11b** using VA-044 reached completion within 5 h to give the final ubiquitinated peptide **12b**, and the conversion was near quantitative based on MS analysis (see the Supporting Information).

To further test whether our method can be used for specific biotinylation on the lysine side chain, peptide **10b** was reacted with a biotin thioester. The reaction was complete in 3 h and gave the biotinyl ligation product

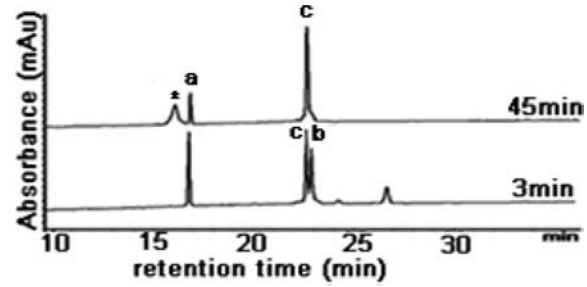


Figure 2. HPLC monitoring of ligation between peptide **10b** and ubi(1−76)-MES at (bottom) 3 and (top) 45 min. Gradient: 0−36% buffer B in 18 min, 36−45% in 18 min. Peak a, peptide **10b**; peak b, ubi(1−76)-MES; peak c, ligation product **11b**; peak *, unidentified. Ligation conditions: 6.6 mM **10b**, 1.2 mM ubi(1−76)-MES, 6 M Gdn·HCl, 0.2 M phosphate, 60 mM TCEP, pH 8.0, 37 °C.

11c in 90% yield by HPLC analysis. Free-radical desulfurization afforded the final biotinylated peptide **12c** in 80% yield based on HPLC analysis.

The above results show that in a unique one-stone-two-birds fashion, a γ-SH group on an N-terminal lysine mediates facile chemical ligation at both its α- and ε-amines. The unhindered nature of the two primary amines likely accounts for the robustness of the ligation reactions. If used without the second ligation step, our method would allow conventional linear NCL at Lys, a notable expansion of the application scope of NCL in view of the abundance of lysine in proteins. Through the use of the dual ligation scheme, it is possible to synthesize complex protein molecules that are acylated on specific lysine side chains. We believe that our method will be particularly useful for the chemical synthesis of lysine-rich and lysine-modified proteins.

Acknowledgment. We thank A*Star of Singapore (BMRC 08/1/22/19/588 to C.-F.L.) and Nanyang Technological University for financial support.

Supporting Information Available: Experimental procedures and compound characterization data. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) For reviews, see: (a) Dawson, P. E.; Kent, S. B. *Annu. Rev. Biochem.* **2000**, *69*, 923. (b) Tam, J. P.; Yu, Q.; Miao, Z. *Biopolymers* **1999**, *51*, 311. (c) Nilsson, B. L.; Soellner, M. B.; Raines, R. T. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91. (d) Hahn, M. E.; Muir, T. W. *Trends Biochem. Sci.* **2005**, *30*, 26. (e) Kent, S. B. H. *Chem. Soc. Rev.* **2009**, *38*, 338.
(2) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776.
(3) (a) Tam, J. P.; Lu, Y.-A.; Liu, C.-F.; Shao, J. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 12485. (b) Durek, T.; Torbeev, V. Y.; Kent, S. B. H. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4846. (c) Boerema, D. J.; Tereshko, V. A.; Kent, S. B. H. *Pept. Sci.* **2008**, *90*, 278. (d) Johnson, E. C. B.; Malito, E.; Shen, Y.; Rich, D.; Tang, W.; Kent, S. B. H. *J. Am. Chem. Soc.* **2007**, *129*, 11480. (e) Pentelute, B. L.; Gates, Z. P.; Dashnau, J. L.; Vanderkooi, J. M.; Kent, S. B. H. *J. Am. Chem. Soc.* **2007**, *130*, 9702.
(4) (a) Canne, L. E.; Bark, S. J.; Kent, S. B. H. *J. Am. Chem. Soc.* **1996**, *118*, 5891. (b) Low, D. W.; Hill, M. G.; Carrasco, M. R.; Kent, S. B. H.; Botti, P. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 6554. (c) Offer, J.; Boddy, C. N. C.; Dawson, P. E. *J. Am. Chem. Soc.* **2002**, *124*, 4642. (d) Wu, B.; Chen, J.; Warren, J. D.; Chen, G.; Hua, Z.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2006**, *45*, 4116. (e) Chen, G.; Warren, J. D.; Chen, J.; Wu, B.; Wan, Q.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2006**, *128*, 7460.
(5) Chatterjee, C.; McGinty, R. K.; Pellois, J.; Muir, T. W. *Angew. Chem., Int. Ed.* **2007**, *46*, 2814.
(6) (a) Tam, J. P.; Yu, Q. *Biopolymers* **1998**, *46*, 319. (b) Saporito, A.; Marasco, D.; Chambery, A.; Botti, P.; Monti, S. M.; Pedone, C.; Ruvo, M. *Biopolymers* **2006**, *83*, 508.
(7) Yan, L. Z.; Dawson, P. E. *J. Am. Chem. Soc.* **2001**, *123*, 526.
(8) Crich, D.; Banerjee, A. *J. Am. Chem. Soc.* **2007**, *129*, 10064.
(9) (a) Haase, C.; Rohde, H.; Seitz, O. *Angew. Chem., Int. Ed.* **2008**, *47*, 6807. (b) Chen, J.; Wan, Q.; Yuan, Y.; Zhu, J.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2008**, *47*, 8521.
(10) Marin, J.; Didierjean, C.; Aubry, A.; Casimir, J.-R.; Briand, J.-P.; Guichard, G. *J. Org. Chem.* **2004**, *69*, 130.
(11) Borodovsky, A.; Ovaa, H.; Kolli, N.; Gan-Erdene, T.; Wilkinson, K. D.; Ploegh, H. L.; Kessler, B. M. *Chem. Biol.* **2002**, *9*, 1149.
(12) Wan, Q.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 9248.

JA905491P