

An Advance in Proline Ligation

Shiying Shang,[†] Zhongping Tan,[†] Suwei Dong,[†] and Samuel J. Danishefsky^{*,†,‡}

[†]Laboratory for Bioorganic Chemistry, Sloan–Kettering Institute for Cancer Research, 1275 York Avenue, New York, New York 10065, United States

[‡]Department of Chemistry, Columbia University, 3000 Broadway, New York, New York 10027, United States

S Supporting Information

ABSTRACT: Native chemical ligation (NCL) is widely applicable for building proteins in the laboratory. Since the discovery of this method, many strategies have been developed to enhance its capability and efficiency. Because of the poor reactivity of proline thioesters, ligation at a C-terminal proline site is not readily accomplished. Here, we demonstrate that ligation at an N-terminal protein is feasible using the combined logic of NCL and metal-free dethylation (MFD).

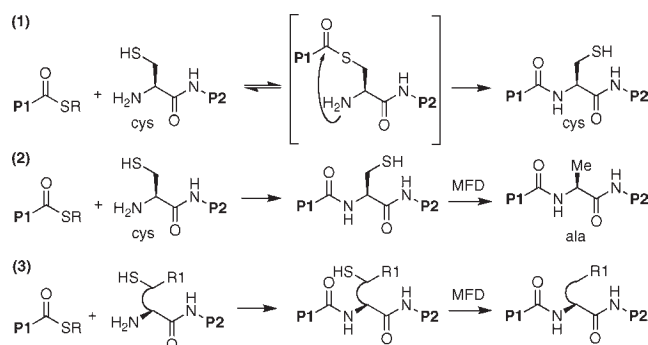
The groundbreaking discovery of native chemical ligation (NCL) by Kent and colleagues was a seminal event in the chemical synthesis of protein-based biopolymers with precisely defined structures.^{1,2} Typically, NCL involves the reaction between an unprotected peptidyl C-terminal thioester with another peptide bearing an N-terminal cysteine. Following rate-determining trans-thioacylation, a fast S→N acyl transfer presumably passing through a five-membered tetrahedral intermediate leads to product (Scheme 1, eq 1, Kent et al.).³

While the original discovery of NCL was based on a peptidyl N-terminal cysteine which appeared at the ligation site, it was soon extended to alanine ligation, via desulfurization of the thiol function by Raney Nickel reduction,⁴ to phenylalanine ligation, using nickel boride as a reductant,⁵ and to serine ligation, through cyanobromide-mediated cleavage of the methylcysteine precursor.⁶ The finding that high yielding metal-free dethylation is feasible in an aqueous medium and specific to SH groups constituted a major expansion of NCL (Scheme 1, eq 2). Not only did it enable alanine ligation in the presence of existing ACM protected cysteines (as well as methionines), it opened up an entirely new set of possibilities for ligation. *One could now incorporate a thiol containing unnatural amino acid at the N-terminus.*^{7,8} The combination of thiol and amino groups allows for simulations of the qualitative mechanism of cysteine based NCL. Metal-free dethylation (MFD) generates either a proteogenic or unnatural amino acid at the ligation site (Scheme 1, eq 3).^{7a}

Using this MFD enabled logic, we and others have demonstrated valine,^{9,10} leucine,^{11,12} threonine,¹³ and lysine ligations.^{14,15} Thus, MFD has brought with it a major expansion in the synthesis of reasonably sized proteins, using the powerful underlying logic of NCL.

The research described herein was directed to a remaining but important problematic issue, i.e. that of proline ligation.¹⁶ It turns out that, for reasons that have been elegantly interpreted,

Scheme 1. NCL, Desulfurization, and Extended NCL^a



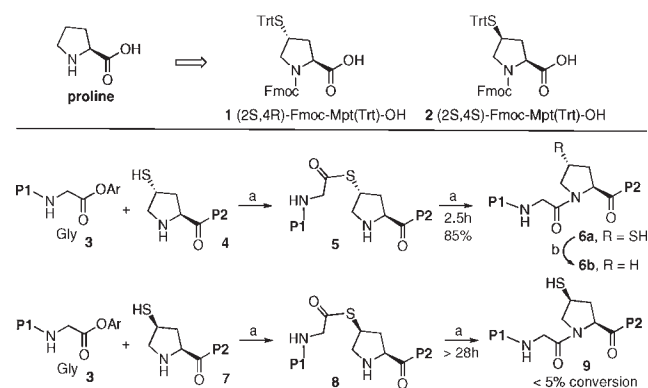
^a P1 and P2 are peptide fragments. R can be alkyl or aryl. R1 can be varied according to the amino acid structure.

C-terminal proline thioesters are not particularly effective acyl donors, even in the context of cysteine-based NCL.^{17–19} While the corresponding C-terminal proline *p*-nitrophenyl esters are more effective,¹⁰ they too carry a serious liability, i.e. susceptibility to hydrolysis in a ligation context. We were thus led to ask whether, in light of the considerations discussed above, a proline-like moiety at the N-terminus of a peptide might serve as a ligation site. This led us to consider the possibility of utilizing an N-terminal thioproline at the ligation site. In so doing, we would be extending the mechanistic concepts of NCL in two major respects. First, the S→N acyl migration would have to occur even at a secondary amine in a ring context. Furthermore, the transfer would be occurring in the context of a bridged ring tetrahedral intermediate. To probe the feasibility of the idea, we took recourse to the commercially available diastereomeric 4-thioproline building blocks **1** and **2**.

We began by synthesizing peptides **4** and **7** incorporating N-terminal γ -thioprolines [(2*S*,4*R*)-Mpt and (2*S*,4*S*)-Mpt], respectively, by the solid-phase peptide synthesis (SPPS) method using the Fmoc chemistry. We went on to evaluate the ligation of both **4** and **7** with peptide **3**, presenting a C-terminal glycine residue. For this purpose, we took recourse to the surrogate acyl donor system bearing a 2-(ethylthiophenyl)ester, previously invented in our laboratory as a more durable thioester equivalent.^{20,21} The required thioester is generated *in situ*, as previously described, via unidirectional O→S acyl transfer.²¹ The reaction between **4** and **3** was initiated by the addition of the standard phosphate ligation buffer and was monitored by ultraperformance liquid

Received: May 10, 2011

Published: June 23, 2011

Scheme 2. Proline Ligation with Two Pro(SH) Diastereomers^a

^a Key: (a) 6 M Gn•HCl, 100 mM NaH₂PO₄, 50 mM TCEP, pH 7.5; (b) TCEP, VA-044, tBuSH, MeCN/H₂O, 37 °C, 10 min, 88%. P1: ALLVNSS-; P2: -WEPLN; Ar = 2-(ethylthio)phenyl.

chromatography (UPLC). As shown in Scheme 2, under standard ligation conditions, peptide 4 readily underwent trans-thioesterification with 3 to afford 5 (90% conversion),²² within 10 min at room temperature.²³ At a considerably slower rate, the thioester intermediate 5 underwent the S→N acyl transfer reaction to afford the desired peptide product in 85% yield within 2.5 h. By contrast, under identical reaction conditions, peptide 7, incorporating the diastereomeric γ -thioprolino, [(2S,4S)-Mpt], demonstrated very low proclivity toward ligation with peptide 3. Although the trans-thioesterification could be achieved as efficiently as in the case of 5, progression to peptide product 9 essentially failed. After 28 h, less than 5% conversion was observed. Instead, about 90% of thioester 8 was hydrolyzed to peptide 7.

This finding of a large difference is readily interpreted. Thus, in the case of the productive *trans* series, in the presumed tetrahedral S→N acyl transfer intermediate, the large COP₂ substructure is *exo*. By contrast, in the unproductive *cis* series, based on 2, the large COP₂ moiety would have to be contained in a highly hindered *endo* disposition.

Having established an efficient and practical γ -thioprolino ligation protocol, albeit only with a terminal glycine, we sought to probe the versatility of this protocol in the context of varying the nature of the C-terminal amino acid residue. Toward this end, we prepared several peptide esters, which would allow us to evaluate the effects of steric hindrance on peptide bond formation. In order to avoid any potential complications associated with residues not involved in amide bond formation, these peptide esters only differ by a side chain at the C-terminus. In accordance with previous NCL studies, the rate and efficiency of ligation were found to be dependent on the steric hindrance of the C-terminal residue.^{9,10,12,16} Under standard conditions, peptide 10 and 12 underwent rapid trans-thioesterification with the peptide 4 to generate the corresponding thioester intermediates. Although more time was required to complete the S→N acyl transfer, good yields were obtained (entries 1 and 2). As expected, the coupling reactions at the valine and proline sites were significantly less efficient (entries 3 and 4). Interestingly, the factors that caused the low conversion in these two reactions are different. In the case of valine, the trans-thioesterification step was fast and apparently efficient. However, the resulting thioester intermediate was insufficiently

Table 1. Substrate Scope of Proline Ligation^a

entry	N-terminal peptide	C-terminal peptide	Ligation product
1			 8h, 69%
2			 13h, 63%
3			 23h, ~15% conversion
4			 8h, no product
5			 3h, 73%
6			 10h, 66%

^a Reaction conditions: 6 M Gn•HCl, 100 mM NaH₂PO₄, 50 mM TCEP, pH 7.5. P1: ALLVNSS-; P2: -WEPLN; P3: -PPWEPLN; Ar = 2-(ethylthio)phenyl.

reactive to allow for reasonable rates of acyl transfer. By contrast, the proline peptide ester 16 was found to exhibit remarkably diminished levels of reactivity in the trans-thioacylation step. Only about 5% of 16 was converted to the thioester intermediate, which was rapidly hydrolyzed under attempted ligation conditions to peptide 4.

Next, in order to probe the origin of the observed low reactivity of the proline peptide ester 16, an additional control ligation was performed between peptide 4 and proline ester 18. Under the standard conditions, 4 and 18 underwent rapid trans-thioesterification and acyl transfer to furnish adduct 19 in 73% yield within 3 h (entry 5). This result agreed well with the observations of Kent et al., suggesting that the presence of a peptidyl group on the α -amine of the C-terminal proline is the dominant factor affecting the reactivity of the proline peptide ester in NCL.¹⁷

Furthermore, the versatility of the proline ligation protocol was demonstrated in the synthesis of the proline-rich peptide 21. In the event, peptide 20, containing 50% proline, was coupled with the glutamine peptide ester 10. The result is shown in Table 1, entry 6. Within 10 h, a 66% isolated yield of 21 was obtained.

Finally, we demonstrated the ability to efficiently convert the γ -thioprolino to the natural amino acid proline in the context of a

complex peptide. Upon exposure to our standard metal-free desulfurization conditions, the γ -thioprolines containing peptides **11**, **13**, **19**, and **21** were readily converted to the target peptide products in less than 10 min.^{7a,10,12,24}

Since both native chemical ligation and direct aminolysis are possible options for the synthesis of large peptides and small proteins, we also directly compared these two methods for synthesizing desulfurized **11**.^{25,26} The direct aminolysis reaction was conducted in DMSO using modified Blake–Aimoto conditions (Supporting Information, Figure S4).²⁷ At room temperature, it proceeded well to give the coupling product in a slightly higher yield. However, because of the difficulties related to the preparation and purification of partially protected peptide fragments²⁸ and the risk of epimerization of the residue on the C-terminal side of the condensation site,²⁷ in many cases, proline ligation would be a better choice for coupling two synthetic fragments at the proline site.

In summary, an efficient and broadly useful two-stage proline ligation protocol has been developed. The likely basis for the different ability of the peptide esters to undergo proline ligation has been rationalized. The proline ligation approach provides a significant advance in the field of protein chemical synthesis and will likely find broad application in the study of proline-rich proteins and glycoproteins. It is of note that proline-rich proteins frequently participate in diverse signaling pathways and serve many crucial biological functions.^{29,30} The unique conformational properties of proline inserts provide the molecular basis for highly discriminatory recognition during the formation of multi-protein signaling complexes. Understanding how these interactions contribute to the stability of the complex should prove valuable in the rational design of peptide mimics that could potentially be used to disrupt the interfaces between proline-rich proteins and their binding partners.³¹ The methodology developed herein is expected to facilitate the preparation of proline-rich polypeptide probe structures, thus enhancing opportunities for detailed analysis of protein–protein interactions.

■ ASSOCIATED CONTENT

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

■ AUTHOR INFORMATION

Corresponding Author

s-danishefsky@ski.mskcc.org

■ ACKNOWLEDGMENT

Support for this research was provided by the National Institute of Health (CA28824 to S.J.D.). We thank Rebecca Lambert for valuable discussions. We also thank Dr. George Sukenick, Hui Fang, and Sylvi Rusli of SKI's NMR core facility for mass spectral and NMR assistance and Laura Wilson and Jason Chan for assistance with the preparation of the manuscript.

■ REFERENCES

- (1) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776.
- (2) Tam, J. P.; Lu, Y. A.; Liu, C. F.; Shao, J. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 12485.

- (3) Johnson, E. C.; Kent, S. B. *J. Am. Chem. Soc.* **2006**, *128*, 6640.
- (4) Yan, L. Z.; Dawson, P. E. *J. Am. Chem. Soc.* **2001**, *123*, 526.
- (5) Crich, D.; Banerjee, A. *J. Am. Chem. Soc.* **2007**, *129*, 10064.
- (6) Okamoto, R.; Kajihara, Y. *Angew. Chem., Int. Ed.* **2008**, *47*, 5402.
- (7) (a) Wan, Q.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 9248. For earlier work on phosphorus derivative-mediated radical desulfurization, see: (b) Walling, C.; Rabinowitz, R. *J. Am. Chem. Soc.* **1957**, *79*, 5326. (c) Walling, C.; Basedow, O. H.; Savas, E. S. *J. Am. Chem. Soc.* **1960**, *82*, 2181.
- (8) A futuristic statement of this concept can be found in Botti, P.; Tchertchian, S. 2006, Patent No. WO2006133962.
- (9) Haase, C.; Rohde, H.; Seitz, O. *Angew. Chem., Int. Ed.* **2008**, *47*, 6807.
- (10) Chen, J.; Wan, Q.; Yuan, Y.; Zhu, J.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2008**, *47*, 8521.
- (11) Harpaz, Z.; Siman, P.; Kumar, K. S.; Brik, A. *ChemBioChem* **2010**, *11*, 1232.
- (12) Tan, Z.; Shang, S.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2010**, *49*, 9500.
- (13) Chen, J.; Wang, P.; Zhu, J. L.; Wan, Q.; Danishefsky, S. J. *Tetrahedron* **2010**, *66*, 2277.
- (14) Yang, R.; Pasunooti, K. K.; Li, F.; Liu, X. W.; Liu, C. F. *J. Am. Chem. Soc.* **2009**, *131*, 13592.
- (15) Kumar, K. S. A.; Haj-Yahya, M.; Olschewski, D.; Lashuel, H. A.; Brik, A. *Angew. Chem., Int. Ed.* **2009**, *48*, 8090.
- (16) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10068.
- (17) Pollock, S. B.; Kent, S. B. *Chem. Commun.* **2011**, *47*, 2342.
- (18) Hinderaker, M. P.; Raines, R. T. *Protein Sci.* **2003**, *12*, 1188.
- (19) Choudhary, A.; Raines, R. T. *Protein Sci.* **2011**, *20*, 1077.
- (20) Warren, J. D.; Miller, J. S.; Keding, S. J.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2004**, *126*, 6576.
- (21) Chen, G.; Warren, J. D.; Chen, J.; Wu, B.; Wan, Q.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2006**, *128*, 7460.
- (22) Thioester intermediate **5** could be observed by UPLC analysis and was differentiated from the N-linked product **6a** by the addition of MESNa, where the disappearance of **5** would be observed.
- (23) Canne, L. E.; Bark, S. J.; Kent, S. B. *J. Am. Chem. Soc.* **1996**, *118*, 5891.
- (24) Shang, S.; Tan, Z.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 5986.
- (25) Chen, G.; Wan, Q.; Tan, Z.; Kan, C.; Hua, Z.; Ranganathan, K.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 7383.
- (26) Tan, Z.; Shang, S.; Halkina, T.; Yuan, Y.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2009**, *131*, 5424.
- (27) Aimoto, S. *Curr. Org. Chem.* **2001**, *5*, 45.
- (28) Tan, Z.; Shang, S.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 4297.
- (29) Vanhoof, G.; Goossens, F.; Demeester, I.; Hendriks, D.; Scharpe, S. *FASEB J.* **1995**, *9*, 736.
- (30) Ball, L. J.; Kuhne, R.; Schneider-Mergener, J.; Oschkinat, H. *Angew. Chem., Int. Ed.* **2005**, *44*, 2852.
- (31) Kuriyan, J.; Cowburn, D. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 259.