ORGANIC LETTERS

2010
Vol. 12, No. 8
1724−1727

Salicylaldehyde Ester-Induced Chemoselective Peptide Ligations: Enabling Generation of Natural Peptidic Linkages at the Serine/Threonine Sites

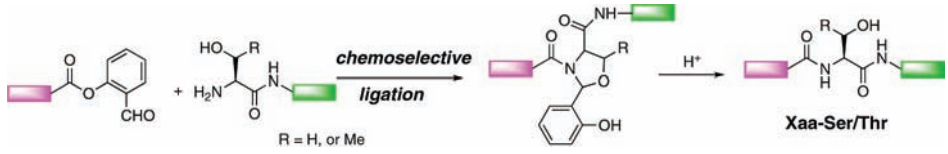
Xuechen Li,* Hiu Yung Lam, Yinfeng Zhang, and Chun Kei Chan

Department of Chemistry, The University of Hong Kong, Pokfulam Road, Hong Kong, China

xuechenl@hku.hk

Received February 6, 2010

ABSTRACT



A serine/threonine-based chemoselective ligation method is described. It uses an *O*-salicylaldehyde ester at the C-terminus, reacting with N-terminal serine or threonine to realize peptide ligations. The utility of the *O*-salicylaldehyde ester enables the rapid coupling reaction and the production of an *N,O*-benzylidene acetal intermediate, which is readily converted into natural peptidic linkages (Xaa-Ser/Thr) at the ligation site.

In addition to recombinant DNA technology to provide polypeptides/proteins, chemical synthesis has dramatically contributed to the exploration of the relationship of protein structure to function. Moreover, with the rapid emergence of middle-sized peptides (between 20 and 100 amino acids) as therapeutics,[1] such as fuzeon, synthetic peptide chemistry has attracted a lot of attention. Merrifield's linear solid-phase peptide synthesis (SPPS) has provided a general tool to prepare polypeptides.[2] However, to chemically synthesize a large polypeptide using linear solid-phase peptide synthesis is very costly and sometimes even impossible. Thus, the development of a method to achieve a convergent synthesis using smaller polypeptide fragments becomes critical in terms of both minimizing the cost of production of peptide therapeutics[1,3] and realizing chemical synthesis of proteins.[4] In contrast with conventional fragment coupling, chemoselective ligations using partially protected or unprotected peptide fragments would be advantageous to achieve an efficient convergent synthesis. Use of unprotected peptides helps circumvent the difficulty inherent to classical peptide coupling reactions due to limited solubility; thus, it will increase coupling yields and lead to easy purification and characterization. The key issue in the development of peptide coupling methods using unprotected peptide fragments is the availability of a chemoselective reaction to specifically and unambiguously join peptides through the C-terminus of a peptide and the N-terminus of a second peptide. Using a chemoselective reaction to join two peptide segments through formation of *an unnatural (i.e., nonpeptide) backbone* structure at the ligation site has permitted the facile preparation of a wide range of backbone-modified artificial peptides and proteins.[5] However, there are limited methods available by which to accomplish peptide ligation through formation

(1) (a) Mark, V. *Chem. Eng. News* **2005**, *83*, 17–24. (b) Bruckdorfer, T.; Marder, O.; Albericio, F. *Curr. Pharm. Biotech.* **2004**, *5*, 29–43.

(2) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.

(3) Fuzeon was chemically manufactured using a three-fragement assembly strategy. (a) Bray, B. L. *Nat. Rev. Drug Discovery* **2003**, *2*, 587–593. (b) Albericio, F. *Curr. Opin. Chem. Biol.* **2004**, *8*, 211–221.

(4) Nilsson, B. L.; Soellner, M. B.; Raines, R. T. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91–118.

(5) For a comprehensive review of chemoselective ligation methods, see: Hackenberger, C. P. R.; Schwarzer, D. *Angew. Chem., Int. Ed.* **2008**, *47*, 10030–10074.

10.1021/ol1003109
Published on Web 03/16/2010

of a *natural peptide linkage* at the ligation site. Clearly, cysteine-based native chemical ligation (NCL),[6] developed by Kent and co-workers, is the first and the best method in this regard, and it has become no doubt the most powerful method in synthetic peptide chemistry. Cysteine-based NCL features a thio capture between an N-terminal cysteine and a C-terminal thioester, as a transthioesterification step which is highly chemoselective, followed by a rapid 1,4 *S*→*N* acyl transfer to afford a natural Xaa-Cys peptidic linkage (Xaa represents any amino acid). Its efficiency, ease of operation, and chemoselectivity (in the presence of any unprotected amino acid) are very attractive to its users/practitioners; thus, cysteine-based NCL has been widely used for chemically synthesizing many proteins.[7]

However, the rare presence of the cysteine residue (1.4% content in proteins) has limited the utility of NCL (most peptide pharmaceuticals rarely contain an internal cysteine[1]). To address this issue, people have extensively searched for alternative "native chemical ligation" methods at other amino acid sites over the past decade.[8] We are interested in the development of ligations at serine[9] and threonine sites. We took note that the cysteine-NCL achieving the chemoselectivity lies in that the N-terminal cysteine can differentiate itself from other inner unprotected amino acid functional groups with its bifunctionality: 1,2-mercaptoamine. On the other hand, the capture−rearrangement chemical ligation does not involve activating the carboxyl group; thus, it is effective to overcome the racemization problem of the conventional segment condensation method. Following the same logic, it is conceivable that serine and threonine possess great promise for achieving chemoselective ligation due to their 1,2-hydroxyamine bifunctionality. We were attracted by the imine-induced proximity acyl-transfer approach. Such a strategy was originally introduced by Kemp[10] and fully developed by Tam to ligate a C-terminal glycolaldehyde peptide with another peptide containing a Cys, Thr, or Ser residue at the N-terminus to furnish a coupled product with a pseudoproline structure (thiazolidine or oxazolidine) at the ligation site.[11] The method achieved great chemoselectivity, but the ligation was extremely slow[11] when threonine or serine was used at the N-terminus, and more importantly, the conversion from the pseudoproline structure into natural peptidic bonds was not achieved.[11] Therefore, although these works have conceptually demonstrated the utility of weakly activated alkylesters as acyl donors for chemoselective peptide ligation at serine and threonine sites, they are unable to provide a convergent synthesis of proteins or polypeptides with the *natural* peptide sequence.

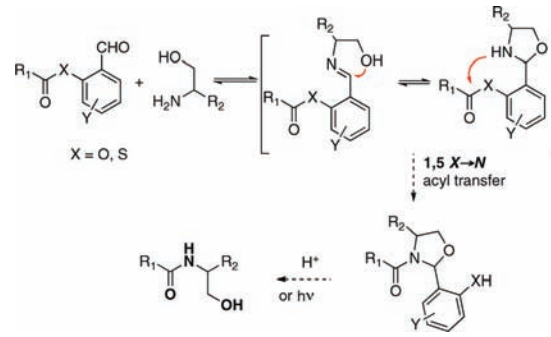


Figure 1. Proposed Ser/Thr-based chemical ligation resulting in a natural peptide bond at the ligation site.

We believe that, with further extensive investigations, a Ser/Thr-based chemical ligation resulting in natural Xaa-Ser/Thr linkage at the ligation site using such an imine capture−rearrangement strategy can be realized. We conceived of the idea of using a two-step strategy (Figure 1). The first step involves the amine group of the N-terminal serine or threonine reversibly reacting with the aldehyde group of the C-terminus to form an imine, followed by the cyclization from the α-hydroxyl group of the N-terminal serine or threonine. Next, an *O*→*N* acyl transfer affords a stable acetal intermediate. Other nucleophiles such as lysine may also react with the aldehyde group, but the reaction is reversible and unable to generate a stable product. The second step is to remove the formed acetal group to afford the natural peptide bond. In order to develop a *practical* chemoselective peptide ligation based on the above strategy, two questions must be addressed: (1) whether either the imine capture step or the acyl transfer step can be accelerated (it is not known which step is the rate-determining step) and (2) whether the formed pseudoproline moiety[12] can be readily transformed into natural peptide linkages. The second question is more challenging and critical because the removal conditions should ideally be simple without any sophisticated chemistry and compatible with other functionalities in peptides or even oligosaccharides in glycopeptides.

Among many candidates, we tentatively identified the (thio)salicylaldehyde ester as a potential functional donor at

(6) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776–779.

(7) (a) Kent, S. B. H. *Chem. Soc. Rev.* **2009**, *38*, 338–351. (b) Muralidharan, V.; Muir, T. W. *Nature Methods* **2006**, *3*, 429–438. (c) Cotton, G. J.; Muir, T. W. *Chem. Biol.* **1999**, *6*, R247–R256.

(8) For methods by post-modifying the cysteine-NCL product, see: (a) Yan, L. Z.; Dawson, P. E. *J. Am. Chem. Soc.* **2001**, *123*, 526–533. (b) Wan, Q.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 9248–9252. (c) Okamoto, R.; Kajihara, Y. *Angew. Chem., Int. Ed.* **2008**, *47*, 5402–5406. (d) Canne, L. E.; Bark, S. J.; Kent, S. B. H. *J. Am. Chem. Soc.* **1996**, *118*, 5891–5896. (e) Offer, J.; Boddy, C. N. C.; Dawson, P. E. *J. Am. Chem. Soc.* **2002**, *124*, 4642–4646. (f) Wu, B.; Chen, J.; Warren, J. D.; Chen, G.; Hua, Z.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2006**, *45*, 4116–4125. (g) Crich, D.; Banerjee, A. *J. Am. Chem. Soc.* **2007**, *129*, 10064–10065. (h) Haase, C.; Rohde, H.; Seitz, O. *Angew. Chem., Int. Ed.* **2008**, *47*, 6807–6810. (i) Chen, J.; Wan, Q.; Yuan, Y.; Zhu, J.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2008**, *47*, 8521–8524. Other significant non-cysteine NCL-derived chemoselective ligations include traceless Staudinger ligation and decarboxylative amide ligation. (j) Nisson, B. L.; Kiessling, L. L.; Taines, R. T. *Org. Lett.* **2000**, *2*, 1939–1941. (k) Tam, A.; Soellner, M. B.; Raines, R. T. *J. Am. Chem. Soc.* **2007**, *129*, 11421–11430. (l) Bode, J. W.; Fox, R. M.; Baucom, K. D. *Angew. Chem., Int. Ed.* **2006**, *45*, 1248–1252. For a sugar-assisted ligation, see: (m) Brik, A.; Yang, Y. Y.; Ficht, S.; Wong, C.-H. *J. Am. Chem. Soc.* **2006**, *128*, 5626–5627.

(9) A "serine" ligation was reported deriving from post-modification of a cystein-NCL product; see ref 7c.

(10) Kemp, D. S. *Biopolymer* **1981**, *20*, 1793–1804. (b) For a comprehensive review, see: Coltart, D. M. *Tetrahedron* **2000**, *56*, 3449–3491.

(11) (a) Liu, C.-F.; Tam, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 6584–6588. (b) Liu, C.-F.; Tam, J. P. *J. Am. Chem. Soc.* **1994**, *116*, 4149–4153. (c) Tam, J. P.; Miao, Z. *J. Am. Chem. Soc.* **1999**, *121*, 9013–09022.

(12) Serine and threonine pseudoprolines in the form of $\psi^{Me,Me}$ pro were originally developed by Mutter and widely used in Fmoc SPPS: Haack, T.; Mutter, M. *Tetrahedron Lett.* **1992**, *33*, 1589–1592.

the C-terminus (Figure 1). We hypothesized that, when the benzaldehyde ester is used, either the hydroxyl addition to the imine or O→N acyl transfer would proceed faster. However, as the penalty, the acyl transfer has to progress through an unfavored 1,5-acyl transfer. More importantly, the formed *N*,*O*-benzylidene acetal group is expected to be readily removed under acidic conditions. Potentially, the reactivity of the coupling step and removal step could be tuned by introducing certain substituents on the phenyl ring of the salicylaldehyde. Installation of photoreactive groups onto the phenyl ring of the salicylaldehyde ester, such as a nitro group, would likely enable the removal of the *N*,*O*-benzylidene acetal under UV light.

We began by testing the hypothesis with the commercially available salicylaldehyde. As a model study, we prepared FmocAla **1** with a salicylaldehyde ester at the C-terminus (Figure 2). To our delight, the coupling between **1** and serine benzyl ester **2** proceeded smoothly in pyridine/acetic acid[11] (1:1, mol/mol), resulting in 85% conversion in 30 min and a quantitative conversion in 2 h to afford *N*,*O*-benzylidene acetal intermediate **3**.[13] Furthermore, after evaporation of the solvent (pyridine/acetic acid), the crude reaction mixture was directly subjected to the next step. The removal of the acetal group can be readily achieved under acidic conditions (TFA/H_2O/*i*-Pr_3SiH, 5 min) with quantitative conversion to afford the "natural" Ala-Ser dipeptide **4**. Such conditions are generally used in peptide synthesis and even glycoprotein synthesis.[14]

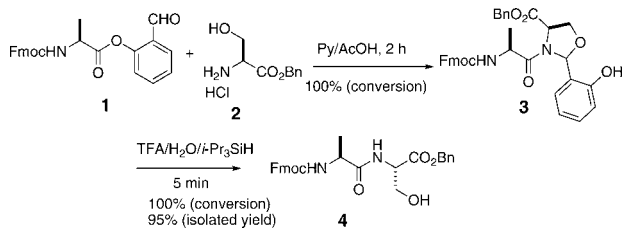


Figure 2. Model study.

The phenolic ester has been known to directly condense with amines to afford amides.[15] To rule out this direct condensation possibility and evaluate the chemoselectivity, two convincing competition experiments were carried out. As outlined in Figure 3, we compared the reactivities between an alanine and a serine derivative when coupling with *O*-salicylaldehyde ester **1** and the reactivities between a thiophenol ester and a salicyladehyde ester when coupling with serine **2**. In both cases, only the *N*,*O*-benzylidene acetal intermediate derived from coupling between Fmoc-Ala-salicylaldehyde ester and serine was observed in almost quantitative yield. Impressively, the thiophenol ester is a rather reactive acyl donor and has been used for direct aminolysis, but it is completely inert under this condition and does not compete with the serine reacting with the salicylaldehyde ester.

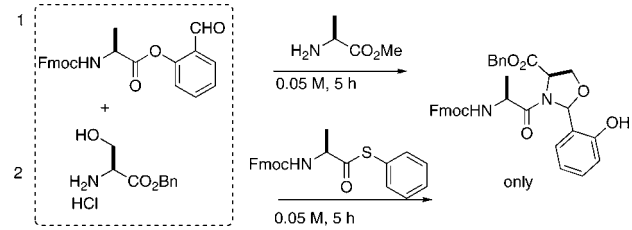


Figure 3. Chemoselectivity studies.

With this success in hand, we next explored the scope of the coupling reaction between N-terminal serine or threonine and the C-terminal salicylaldehyde ester of some rather hindered β-branched amino acid sites (e.g., Val, Pro, Ile, and Thr). These β branched amino acids when used at the C-terminus generally retard the peptide coupling step dramatically; thus, most known ligation methods require a prolonged time for completion at these amino acid sites[16] (i.e., 48 h for cysteine-NCL) or are limited to less hindered amino acids at the C-terminal site. However, these amino acids possessing a salicylaldehyde ester at the C-terminus react with both serine and threonine derivatives surprisingly rapidly, resulting in >70% conversion after 30 min and completion within 5 h.[17] Treatment with (TFA/H_2O/*i*-Pr_3SiH) for 5 min gave rise to peptides with natural peptide bonds at the ligation site (Table 1). Tripeptide Val-Thr-Ser-OBn (entry 5) was isolated in diastereomerically pure form, without detection of the epimerized diastereomer by LCMS and NMR.[18]

To demonstrate the feasibility of this ligation strategy in peptide segment ligation using unprotected peptides, a two-step sequential ligation was carried out between the *O*-salicylaldehyde ester peptide **5**[19] and unprotected peptide **6** possessing an N-terminal serine (Figure 4). With only one purification, the desired peptide **7** with a natural peptide bond (Thr-Ser) at the ligation site was obtained in 72% yield. It is noteworthy that peptide **7** possesses a thiophenol ester at the C-terminus. The thiophenol ester, being a rather reactive acyl donor in chemical ligation, is completely inert and fully

(13) The *N*,*O*-benzylidene acetal intermediate between BocNH-Ala-salicylaldehyde ester and H_2N-Ser-OMe was fully characterized and studied; see the Supporting Information.

(14) Yamamoto, N.; Tanabe, Y.; Okamoto, R.; Dawson, P. E.; Kajihara, Y. *J. Am. Chem. Soc.* **2008**, *130*, 501–510.

(15) (a) Chen, G.; Wan, Q.; Tan, Z.; Kan, C.; Hua, Z.; Ranganathan, K.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 7383–7387. (b) Yuan, Y.; Chen, J.; Wan, Q.; Tan, Z.; Chen, G.; Kan, C.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2009**, *131*, 5432–5437.

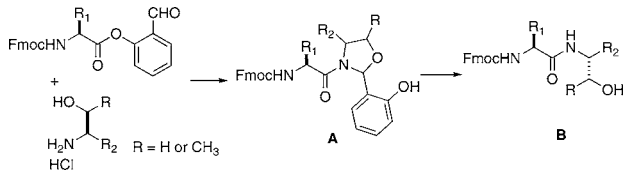
(16) Hackeng, T. M.; Griffin, J. F.; Dawson, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *96*, 10068–10073.

(17) NH_2-Thr-Phe-OEt dipeptide was used here. Under the reaction conditions, the formation of a diketopiperazine was not observed.

(18) The epimerization issue was also studied using Fmoc-Ala-Phe(**L**)-salicylaldehyde ester and Fmoc-Ala-Phe(**D**)-salicylaldehyde ester; see the Supporting Information.

(19) Peptide salicylaldehyde esters could be synthesized using Fmoc SPPS chemistry with the protocol developed by Dawson for the preparation of peptide thioesters: Blanco-Canosa, J. B.; Dawson, P. E. *Angew. Chem., Int. Ed.* **2008**, *47*, 6851–6855.

Table 1. Scope of the Reaction between Serine/Threonine and Salicylaldehyde Esters[a]



	R_1	R_2	**A**[b] (%)	**A**[c] (%)	**B**[d]	%
1	Ala	Ser-OBn	85	100	A-S	95
2	Val	Ser-OBn	80	100	V-S	98
3	Pro	Ser-OBn	71	100	P-S	99
4	Ile	Ser-OBn	70	100	I-S	95
5	Val-Thr	Ser-OBn	76	100	VT-S	97
6	Ala	Thr-Phe-OEt	76	100	A-TF	99
7	Val	Thr-Phe-OEt	73	100	V-TF	96
8	Pro	Thr-Phe-OEt	75	100	P-TF	98
9	Ile	Thr-Phe-OEt	80	100	I-TF	98
10	Val-Thr	Thr-Phe-OEt	74	100	VT-TF	92
11	Lys(TFA)[e]	Ser-OMe	90		K-S	86
12	Ala	Ala-OMe	NR[f]	NR		
13	Ala	Phe-OEt	NR	NR		

[a] All reactions were performed at the concentration of around 0.05 M. [b] The reaction conversion at 30 min. Five microliters of sample from the reaction mixture was taken, diluted with acetonitrile/water, and directly injected into the LCMS system; the conversion is calculated on the basis of the consumption of the salicylaldehyde. [c] Conversion at 5 h. [d] The deprotection step is performed between 5 and 10 min. [e] Derived from the treatment of Fmoc-Lys(Boc)-O-salicylaldehyde ester with TFA. [f] NR: no reaction.

orthogonal to this serine/threonine based chemical ligation. This outcome adds high value to this ligation protocol because the C-terminal thioester peptide is ready for further elongation from the N terminus to the C terminus.

In summary, we reported herein a new method utilizing Ser and Thr at the N-terminus to achieve a chemoselective ligation resulting in natural Xaa−Ser and Xaa−Thr bonds at the ligation site. Serine and threonine are well represented (12.7%) in proteins, and many synthetic peptide pharmaceuticals contain internal serine or threonine, such as forteo, fuzeon, nesiritide, and corticoliberin; thus, this Ser/Thr-induced chemoselective peptide ligation method should find great applications in synthetic peptide chemistry and peptide pharmaceutical manufacturing. Next, we will apply this method to convergently synthesize the antiretroviral drug fuzeon and glycopolypeptides.

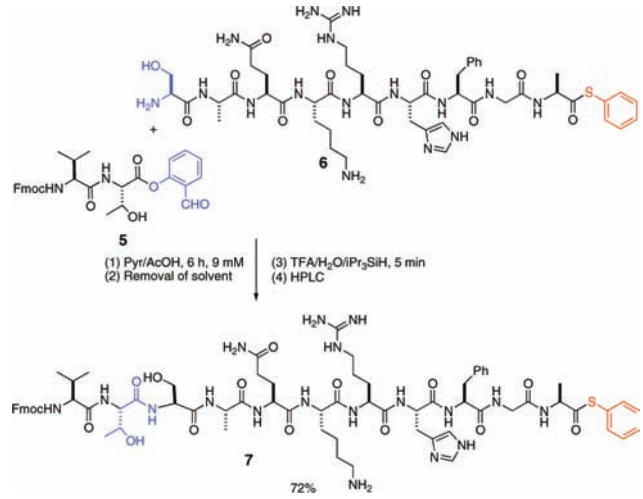


Figure 4. Serine ligation with a unprotected peptide possessing a thiophenyl ester at the C terminus.

Acknowledgment. We thank the Department of Chemistry and Faculty of Science of The University of Hong Kong for providing a startup fund and a Unversity Development Fund for purchasing a LCMS. We thank Ms. Bonnie Yan at The University of Hong Kong for the help with recording NMR spectra.

Supporting Information Available: Synthetic methods and characterization. This material is available free of charge via the Internet at http://pubs.acs.org.

OL1003109