SCIENCE CHINA Chemistry

Recent advances in the preparation of Fmoc-SPPS-based peptide thioester and its surrogates for NCL-type reactions

Hongxing Li and Suwei Dong

Citation: SCIENCE CHINA Chemistry **60**, 201 (2017); doi: 10.1007/s11426-016-0381-1

View online: http://engine.scichina.com/doi/10.1007/s11426-016-0381-1

View Table of Contents: http://engine.scichina.com/publisher/scp/journal/SCC/60/2

Published by the Science China Press

Articles you may be interested in

Amides as surrogates of aldehydes for C–C bond formation: amide-based direct Knoevenagel-type condensation reaction and related reactions
SCIENCE CHINA Chemistry **63**, 11 (2020);

Recent advances in proteolysis and peptide/protein separation by chromatographic strategies
SCIENCE CHINA Chemistry **53**, 685 (2010);

Recent advances in the iron-catalyzed cycloaddition reactions
Chinese Science Bulletin **57**, 2338 (2012);

Synthesis of M(biN)_nCl_2 by Solid-Solid Reactions
Chinese Science Bulletin **39**, 349 (1994);

Recent advances in new-type molecular switches
SCIENCE CHINA Chemistry **58**, 916 (2015);

SCIENCE CHINA
Chemistry

• REVIEWS •

February 2017 Vol.60 No.2: 201–213
doi: 10.1007/s11426-016-0381-1

Recent advances in the preparation of Fmoc-SPPS-based peptide thioester and its surrogates for NCL-type reactions

Hongxing Li & Suwei Dong*

State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing 100191, China

Received September 14, 2016; accepted November 3, 2016; published online December 27, 2016

Solid phase peptide synthesis (SPPS) based on Fmoc chemistry has become a commonly used technique in peptide chemistry, as it can be easily conducted using automated machine, and not requiring highly toxic HF in comparison to Boc-SPPS. With the fast development in the emerging field of protein chemical synthesis, many efforts have been endeavored aiming to find more efficient methods for preparing peptide fragments required in ligation reactions. This review briefly summarizes recent advances in the engineering and modification of Fmoc-SPPS-derived peptides, which can be used as the N-terminal fragments in a native chemical ligation (NCL) or NCL-type ligation reactions.

native chemical ligation, Fmoc-SPPS, peptide thioester, protein chemical synthesis

Citation: Li H, Dong S. Recent advances in the preparation of Fmoc-SPPS-based peptide thioester and its surrogates for NCL-type reactions. *Sci China Chem*, 2017, 60: 201–213, doi: 10.1007/s11426-016-0381-1

1 Introduction

As a milestone discovery in the field of peptide/protein chemical synthesis, native chemical ligation (NCL, Figure 1) introduced by Kent and coworkers [1] has become one of the most frequently used techniques nowadays. As a result, the preparation of the required thioesters (or equivalents) in a racemization-free manner has been a routinely pursued task for researchers who use NCL in their studies. While solid phase peptide synthesis (SPPS) ensures the integrity of stereochemistry in the synthesized peptide chain [2], the use of thioesters as linkers on resin has been limited in Fmoc-based synthesis, due to the lability of such structures toward piperidine required for Fmoc-removal. Although several solutions have been attempted to keep the thioester linker intact, such as changing cocktails for Fmoc-deprotection [3] or using piperidine-resistant *tert*-butyl thiol linkers [4], most studies have been focusing on developing protocols that would generate the requisite peptide fragments in a more practical manner. Moreover, proper ligation handles would be favorably installed based on the materials generated from automated Fmoc-SPPS, which is believed to be more convenient and safer than using Boc-chemistry. Many creative strategies and methods have thus been developed, including the well-studied "safety-catch" sulfonamide linker (Figure 2) [5], and some of the literatures related to this topic have been very nicely reviewed in a number of account articles [6]. In this review, we intended to briefly summarize a number of recent advances regarding the preparation of peptidyl thioesters and the "masked" version, hoping to display a general view of some innovative strategies or ideas in this field.

*Corresponding author (email: dongs@hsc.pku.edu.cn)

© Science China Press and Springer-Verlag Berlin Heidelberg 2016

http://engine.scichina.com/doi/10.1007/s11426-016-0381-1

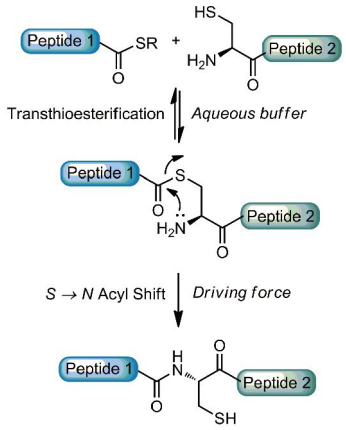

Figure 1 Native chemical ligation (NCL) developed by Kent and coworkers [1] (color online).

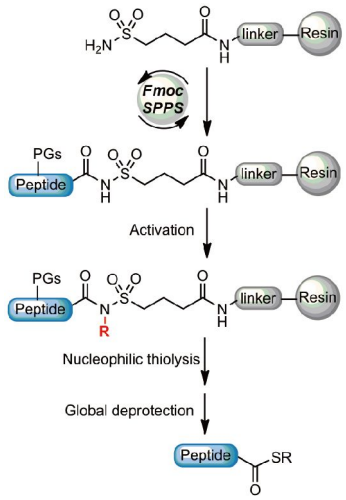

Figure 2 "Safety-catch" sulfonamide linker system (color online).

2 Direct thioesterification on side-chain protected peptidyl carboxylic acids

2.1 Condensation under fine-tuned conditions

Despite the tendency of racemization at the peptide C-termini while conducting activation under typical coupling conditions [7], efforts have been made to tune the conditions for jointing side-chain protected peptidyl carboxylic acids and mercaptans or thiol derivatives (Figure 3). One of the earliest examples is from the Sakakibara group [8], where they described an epimerization-free amide condensation method using EDCI/HOOBt as the coupling reagents. Accordingly, peptidyl thioesters were prepared from side-chain protected peptidyl carboxylate and an HCl salt of Xaa-thioester (Figure 3(b)). The use of free base form of EDCI was demonstrated to be essential in this protocol. Other coupling reagents, including DIC0 (diisopropylcarbodiimide) and PyBOP (benzotriazol-1-yl-*N*-oxytris-(pyrrolidino)phosphonium hexafluorophosphate), have also been utilized in preparing peptidyl *p*-acetamidothiophenol esters with significantly minimized epimerization [9].

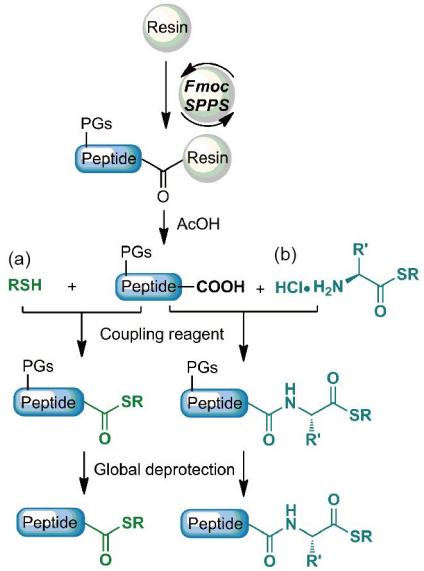

Figure 3 Direct condensation method for the preparation of peptide thioester derivatives. (a) Thioesterification between side-chain protected peptides and mercaptans; (b) amide-bond condensation using side-chain protected peptide and amino acid thioesters (color online).

Due to the diverse sequences of peptides, poor solubility may be observed in commonly used solvent such as dichloromethane (DCM). As an alternative, Flemer [10] reported an *in situ* procedure where the coupling reaction was conducted in the cocktail (1% trifluoroacetic acid (TFA)/DCM) that cleaves side chain-protected peptide from resin (Figure 4). The condensation reaction succeeded when excess amount of base (1.1 eq. of DIEA) was used to neutralize the acid, where improved solubility of peptides was observed.

In the case of more hydrophilic glycopeptides, particularly for ones with complex-type oligosaccharide attached, poor solubility of peptides in DCM and tetrahydrofuran (THF) [11] becomes an issue. To solve this problem, Kajihara's group [12] chose dimethylformamide (DMF) as solvent to explore the optimal condensation conditions for preparing glycopeptidyl thioesters from the corresponding side chain-protected glycopeptide carboxylic acid. After extensive experimentation, using PyBOP/DIEA and large excess of thiol (30 eq.) in DMF at –20 °C provided the best result. This optimal combination of coupling reagent and solvent was further confirmed later in Warriner's synthesis of an epimerization-prone peptide thioester [13].

In general, the direct condensation approaches have the advantage of easy operation and requiring the least amount of derivatization on amino acids, resins, and peptide fragments.

http://engine.scichina.com/doi/10.1007/s11426-016-0381-1

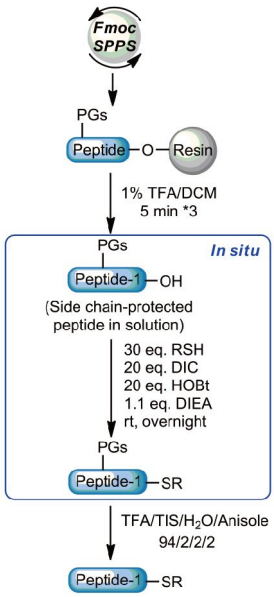

Figure 4 *In situ* condensation approach after peptide cleavage from resin [10] (color online).

However, the potential difficulties in determining epimerization within peptides, in particular large peptides, creates uncertainty when using these methods.

2.2 Glycinyl/prolyl thioesters or surrogates in NCL

To circumvent possible epimerization in thioesterification, two amino acids are considered ideal sites for direct thioesterification on the side chain-protected peptide carboxylate: One is glycine, which contains no stereogenic center, and the other one is proline, which is not prone to epimerize due to the disfavored rigid structure of the forming intermediate (Figure 5) [14]. Many efforts have been attempted to take advantages of these two unique amino acids and use them as the connection sites in peptide synthesis. For instance, peptides with C-terminal Pro or Gly residues can be activated directly and used in fragment condensations [15]. Along the same line, peptides with C-terminal pseudoprolines were successfully utilized in the condensation reactions to generate a 32-mer glycopeptide, which is a required fragment in Unverzagt's synthetic attempt toward glycosylated ribonuclease C [16].

Apparently, peptide glycinyl/prolyl thioesters would be ideal building blocks for NCL reactions, as the preparation of the corresponding fragments should have the convenience of easy preparation under typical coupling conditions without any concerns of epimerization. However, while peptide glycinyl thioesters have been utilized in chemical synthesis of several proteins [17], prolyl thioesters were suggested to be extremely unreactive under typical NCL conditions [18]. To expand the utility of prolyl thioester or its surrogates in NCL, a number of research groups have explored possible solutions. In 2008, taking notes from the work by Bodanszky

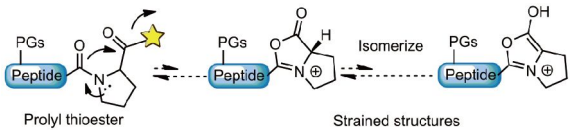

Figure 5 Epimerization-free properties of peptides with C-terminal proline (color online).

[19] applying *p*-nitrophenyl esters to the synthesis of small peptides in a direct aminolysis manner, the Danishefsky group [20] utilized the *p*-nitrophenyl ester in an NCL-like process to connect two peptides (Figure 6(a)), where the prolyl *p*-nitrophenyl ester was found to be sufficiently activated and react with the cysteinyl peptide to generate ligated product in 50% yield, albeit in relatively long reaction time (15 h).

Later in 2011, Durek *et al.* [21] applied the α-selenoester as a more effective acyl donor to realize Pro-Cys ligation under aqueous buffer conditions. Large excess of cysteinyl peptide (5 eq.) was used in the reaction leading to nearly full conversion within only 2 h. However, the "branched" side products could not be eliminated due to the much higher reactivity of selenoester in comparison to the starting thioesters (Figure 6(b)). As the catalyst diphenyldiselenide (DPDS) in buffer could not reverse the unproductive side product, the applicability of prolyl selenoester in protein chemical synthesis has been limited.

In Melnyk's studies of bis(2-sulfanylethyl)amino (SEA) amide in ligation reactions (*cf.* Section 4), they noticed that the prolyl SEA amide was more reactive and more resistant to Xaa-Pro deletion side reactions in comparison to the corresponding 3-mercaptopropionic acid (MPA) and thiazolidine thioester (TT) (Figure 6(c)) [22]. Although the reaction requires only 100 mM of thiol additive mercaptophenylacetic acid (MPAA), the relatively harsh reaction condition and inevitable Xaa-Pro byproducts call for further improvement to make this protocol more suitable for large protein synthesis.

Besides these prolyl thioester surrogates, strategies to activate alkyl thioesters were also investigated. A recent example is from Otaka's group [23] in 2014, where they carefully adjusted the recipe of NCL buffer and successfully realized ligation affording product with native Pro-Cys residues in the presence of MPAA (250 mM) at 50 °C. However, side products with Xaa-Pro deletion were also observed, presumably due to the formation of an arginyl diketopiperazine intermediate (Figure 7). Nevertheless, this protocol has been successfully applied in the synthesis of a 37-mer peptide, kaliotoxin, and more recently a 162-residue *S*-monoglycosylated GM2-activator protein (GM2AP) [24].

Based on these previous studies, it is obvious that tuning the reactivity of prolyl thioester intermediate is crucial for the ligation rate. On the basis of the rational that the poor electrophilicity of prolyl thioester may be resulted from the

http://engine.scichina.com/doi/10.1007/s11426-016-0381-1

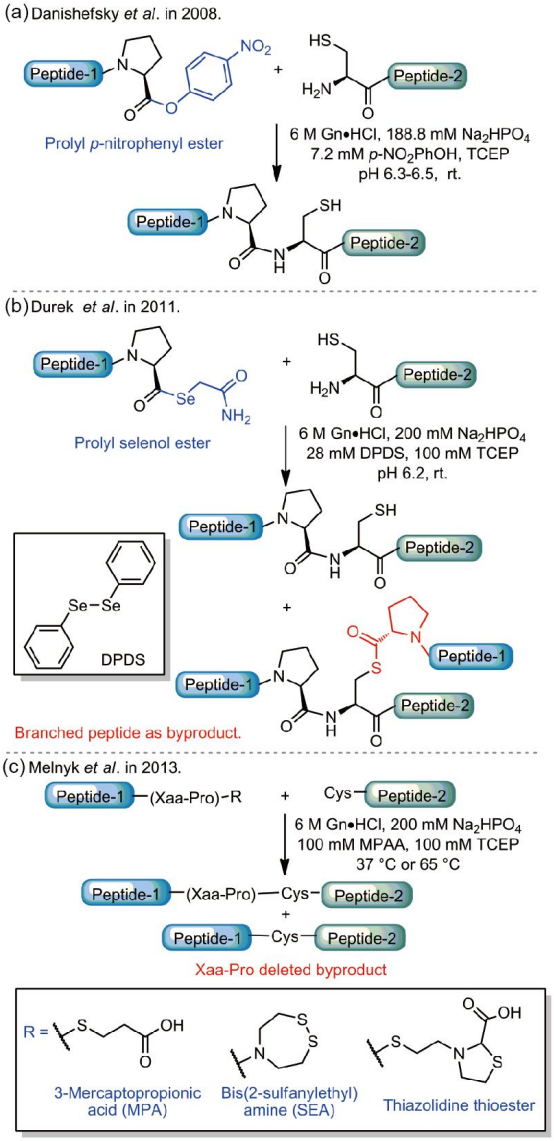

Figure 6 Ligations using thioester surrogates at the Pro-Cys site (color online).

n→π* orbital interaction [25] and steric hindrance of the *N*-carbonyl of proline [18b], a heterobicyclo[2.2.1]septane structure was designed to improve the reactivity [26]. It was proposed that by precluding the *N*-carbonyl oxygen/thioester carbonyl n→π* interaction, and possessing a ring strain releasing tendency (Figure 8), the thioester carbonyl within such structure should exhibit increased electrophilicity. Accordingly, ligation reactions were conducted in NCL buffer without thiol additives (e.g. MPAA) and desired ligation of two peptides was observed, providing good to excellent yields within 8 h. It is noteworthy that using thiol additive-free condition, such "internal-activation" strategy allows for "one-pot" sequential ligation-desulfurization at the native Pro-Ala sites. The utility of this protocol was demonstrated by synthesizing a proline-rich region of Wilms tumor protein 1 (WT1) by merging two segments using Pro-Cys ligation

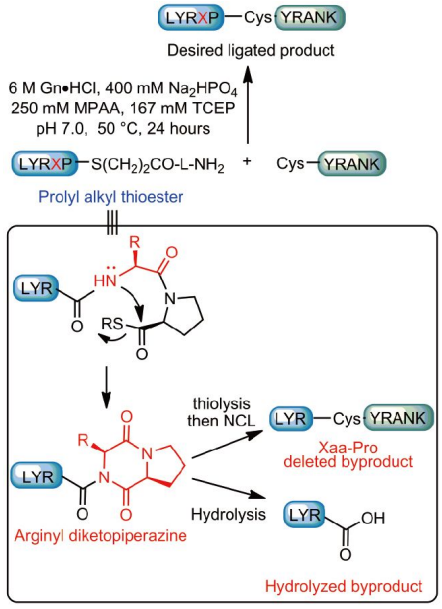

Figure 7 Ligation at Pro-Cys sites using alkyl prolyl thioester [23] (color online).

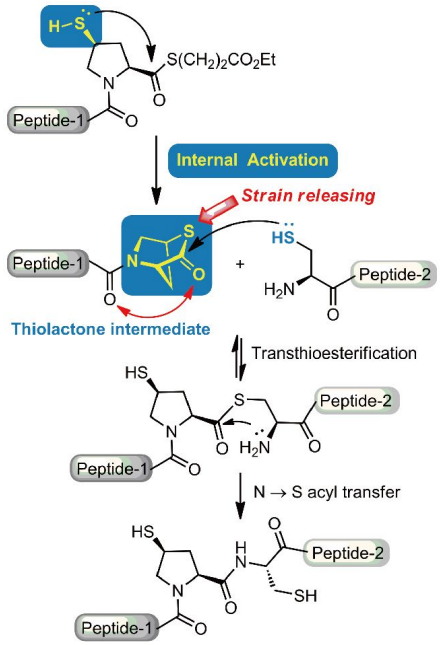

Figure 8 Internal activation strategy to realize Pro-Cys ligation [26] (color online).

site followed by metal-free desulfurization [27].

Regarding the mechanism of this thioprolyl thioester-mediated ligation, besides the investigations conducted by Dong *et al.* [26], a recent computational study provides further supports to clarify the origin of increased reactivity of the bicyclic thiolactone structure [28]. It was suggested that entropy effect facilitates the formation of thiolactone intermediate, and among all factors, the release of ring strain is the main reason to effect the observed high reactivity.

http://engine.scichina.com/doi/10.1007/s11426-016-0381-1

3 Peptide α-thioester formation based on *O*→*S* acyl transfer

Since the intrinsic lability of thioester makes the development of Fmoc-SPPS-compatible protocols a tough problem, researchers turned to other more stable structure, e.g. oxo-esters, and made proper modification to effect a post-generation of thioester after SPPS. One major type of such "masked thioesters" is based on the *O*→*S* acyl transfer process [6b]. In these studies on using thiol-containing alkyl esters as precursors to thioesters, 2-hydroxy-3-mercapto-propionic acid (HMP) is a commonly used motif, which is first introduced to peptide synthesis by Botti *et al.* [29] in 2004 (Figure 9(a)). This strategy possesses advantages including mild reaction conditions and faster rate of *in situ* thioester generation, but the incorporation of HMP on resins was still suffering from either low efficiency of the diazotization/hydrolysis steps [29], or complicated synthetic operations [30]. In 2013, this approach was further optimized by a Lilly research team, where they derivatized the HMP as *S*-trityl protection species in one step (Figure 9(b)), followed by the installation on resin and elongation of peptide chain using Fmoc chemistry [31]. As the linkage between peptide and HMP is an ester bond, using 2-methylpiperidine (2-MP) for Fmoc removal, instead of commonly used piperidine, is essential for obtaining optimal yield in some cases. Finally upon TFA deprotection, HMP peptides could be obtained in decent yields, which were further transformed to the corresponding thioester in a guanidine buffer containing 10% of MPA.

Besides the derivatization of HMP, others have also utilized thiophenol ester [32], or other mercapto alkylesters [33] as thioester precursors. However, in all cases, the hydrolyzed byproducts were observed representing the major limitation of this strategy. Moreover, as the ester C–O bond is still susceptible under nucleophilic attack during Fmoc-SPPS, more attentions have thus been turning to more stable amide linkages.

4 Peptide α-thioester formation based on *N*→*S*/*Se* acyl transfer

In comparison to the use of *O*→*S* acyl migration strategy, selective activation of the desired amide motif at the C-termini becomes a challenge, considering the presence of other amides all over the peptide chain. In contrast to the relatively stable secondary amide, it was noticed that certain tertiary amide, e.g. *N*-alkylcystiene, could undergo an *N*→*S* acyl shift under suitable reaction conditions. Such motif provides vast opportunities for developing peptidyl amide-based thioester precursors [6b,6d,6e].

A representative example is the *N*-alkyl cysteine (NAC) system developed by Hojo *et al.* [34] in 2007, which is fully compatible with Fmoc-based SPPS, and can be readily converted to thioesters *in situ* after proper modifications. For instance, in 2015, Hojo *et al.* [35] utilized a C-terminal carboxyl group to help facilitate fast *N*→*S* acyl transfer (Figure 10), thus allowing for the direct application of the tertiary amide precursor to NCL reactions.

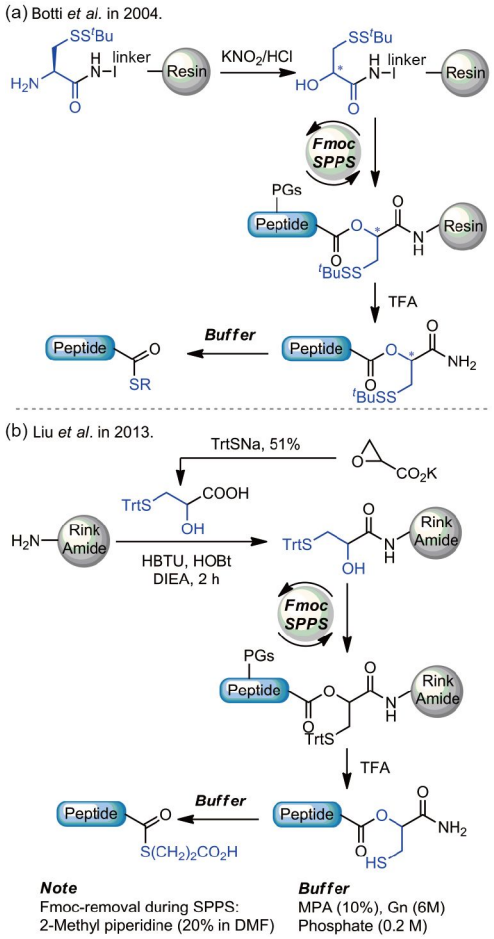

Figure 9 Strategies using Fmoc-SPPS compatible Hmp-based oxo-esters as peptide thioester precursors (color online).

Similarly in 2016, Aucagne *et al.* [36] designed an NAC derivative to incorporate a phenol moiety on the *N*-alkyl chain as a proton donor (Figure 11), mimicking the function of *N*-protonated histidine residue within class 1 inteins during the process of *in vivo* protein splicing [37]. After SPPS, the resulting peptidyl *N*-(2-hydroxybenzyl)cysteine (*N*-Hnb-Cys) derivative could be utilized directly in an NCL reaction, where the fast *N*→*S* acyl transfer facilitated peptidyl thioesters *in situ* after disulfide cleavage by Tris(2-carboxyethyl)phosphine (TCEP). It is noteworthy that the *N*-Hnb-Cys(S^{t}Bu)-containing peptides could be obtained directly using Fmoc-SPPS without any other post-synthetic steps. The synthesis of two cysteine-rich peptides, muscarinic toxin 7 (MT7, 65 residues) and big defensin 1 (Cg-BigDef1, 93 residues), demonstrated the applicability of such crypto-thioester protocol in protein chemical synthesis.

Besides NAC derivatives, other α-thioester precursors,

http://engine.scichina.com/doi/10.1007/s11426-016-0381-1

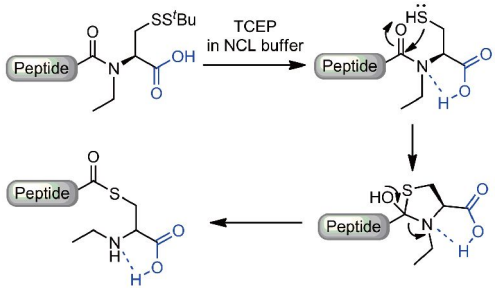

Figure 10 Cysteinyl carboxylic acid promoted *N*→*S* acyl transfer [35] (color online).

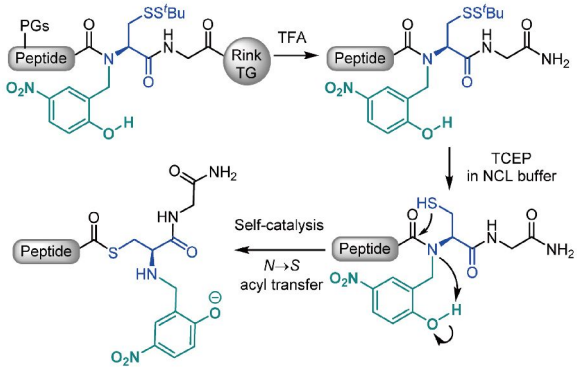

Figure 11 Self-catalyzed *N*-(2-hydroxybenzyl)cysteine device as crypto-thioester in NCL reactions [36] (color online).

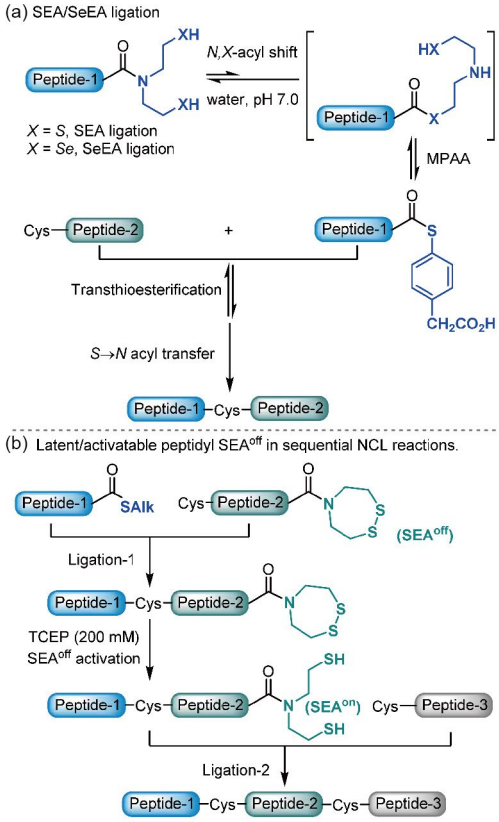

Figure 12 Utilization of SEA/SeEA and SEA^{off}/SeEA^{off} in sequential NCL reactions (color online).

such as the SEA system [38], have also been evaluated. The incorporation of SEA motif to the C-termini of peptides as thioester precursors in NCL, also known as "SEA ligation", was introduced by the Melnyk group [38a] and the Liu group [38b]independently around the same time in 2011 (Figure 12(a), X=S). Later on, Melnyk *et al.* [39] further utilized the cyclic disulfide form of SEA (SEA^{off}) as a latent thioester in one-pot, three-segments, sequential N-to-C peptide ligations, during which the peptidyl SEA^{off} is reduced/activated upon TCEP treatment after completion of the first ligation with peptidyl alkyl thioester (Figure 12(b)). In a similar manner, the selenium analogue of SEA, SeEA, has also been proven to be compatible, with even more potent reactivity [40]. Peptidyl SeEA^{off} is the first reported latent thioester surrogate in *N*→*Se* acyl transfer system, and the differentiation of reactivities in SEA/SeEA systems is controlled by reductants, providing an alternative to kinetically controlled ligation (KCL) in protein chemical synthesis.

In 2011, one convenient protocol was developed by Liu's research group [41], where a premade enamide-containing *N*,*N*-dialkylamide derivative was incorporated to the peptide C-termini (Figure 13(a)). Upon TFA treatment, the corresponding peptide thioester was produced through *N*→*S* acyl shift followed by irreversible hydrolysis of the enamine intermediate. More recently, Zheng *et al.* [42] reported an improved method to solve the previously existed problems, such as the laborious synthesis of enamide-containing amino acid derivatives, and relatively low efficiency of thioester formation. Structurally, by moving the solid phase linker away from the thiol-containing *N*-alkyl chain (Figure 13(b)), the newly developed enamine linker can be prepared in gram-scale, possessing better reactivity that effects faster rate of *N*→*S* acyl migration.

Besides the above examples based on *N*,*N*-dialkylamide structures, investigations on thiol-containing *N*-aryl-*N*-alkyl amide system have also been conducted, mainly from the Otaka group [43]. For instance, they utilized a crypto-thioester, peptidyl *N*-sulfanylethylanilide (SEAlide), to generate thioester under acidic incubation or in the presence of phosphate salts (Figure 14(a, b)) [44]. Recently in 2016 [45], based on the previous studies, they developed peptidyl *N*-sulfanylethylcoumarinyl amide (SECmide) (Figure 14(c)), which was found to perform better under mild acidic condition to form the corresponding thioester in an epimerization-free manner. 4-Mercaptobenzylphosphonic acid (MBPA), possessing both phosphate and thiol moieties within the structure, was proven to be a useful dual catalyst for both peptidyl SEAlide and SECmide. However, these crypto-thioesters exhibited decreased efficiency in the cases of hindered C-terminal amino acid residues, where elevated temperature (e.g. 50 °C for phenylalanine) was required to

http://engine.scichina.com/doi/10.1007/s11426-016-0381-1

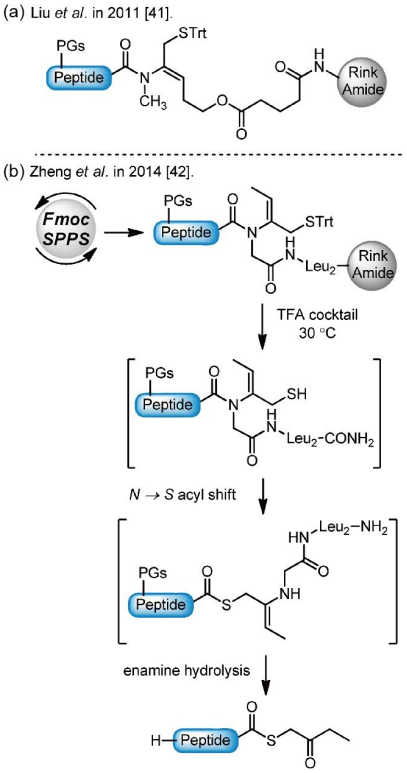

Figure 13 C-terminal amino acid enamide linker in Fmoc-SPPS and the thioester formation based on *N*→*S* acyl shift (color online).

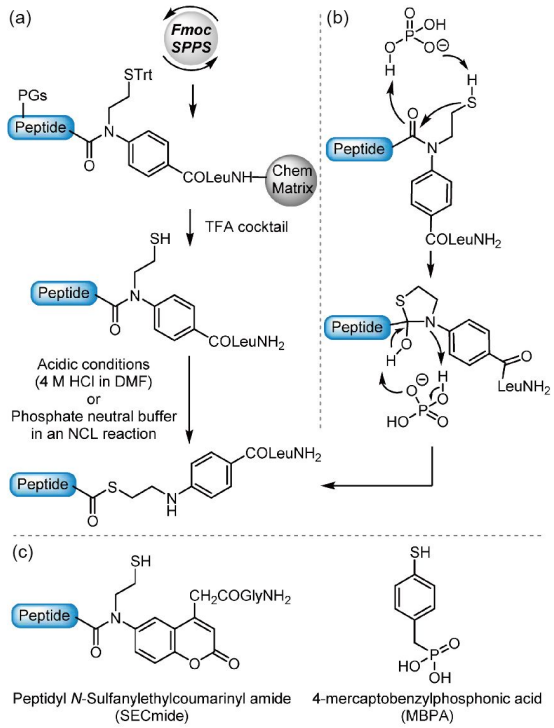

Figure 14 Peptidyl SEAlide developed by Otaka's group. (a) SEAlide undergoes *N*→*S* acyl transfer in acidic conditions or neutral buffer containing phosphate salts [44]; (b) proposed mechanism for the phosphate catalysis in buffer [44]; (c) peptidyl SECmide and dually functional promotor MBPA [45] (color online).

ensure successful NCL reactions.

A more recent development based on SEAlide peptide was the incorporation of a photo-cleavable 6-nitroveratryl (NV) group on the thiol moiety (Figure 15) [46]. This latent activation strategy facilitates efficient sequential ligations of multi-fragments in one-pot, and its applicability was demonstrated by a 41-mer peptide, SNX-482, which is a potent R-type Ca^{2+} channels inhibitor.

5 Recent advances of peptide Nbz in ligation reactions

In 2008, inspired by the reported conversion of *o*-aminoanilides to the corresponding aromatic *N*-acyl urea derivatives as an activated acyl donor in organic synthesis [47], Dawson and co-worker [48] extended this *N*-acylurea activation strategy to peptide synthesis (Figure 16(a)). In a typical procedure, peptidyl diaminobenzoyl carboxylate (Dbz) on resin is acylated and converted into *N*-acyl-benzimidazolinone (Nbz) on solid phase, and subsequent cleavage from resin using TFA cocktail yields the free peptide-Nbz derivative. The preparation of peptide-Nbz on solid phase significantly reduces racemization at the peptidyl C-terminal amino acids, and affords the desired peptide segments in good to excellent yield. Furthermore, peptidyl Nbz can be converted to the corresponding peptide thioesters through rapid thiolysis in buffer solutions, which ensures either direct usage in NCL reactions or storage for a long period of time. Utilization of this thioester surrogate in protein chemical synthesis has been demonstrated in the synthesis of a 59

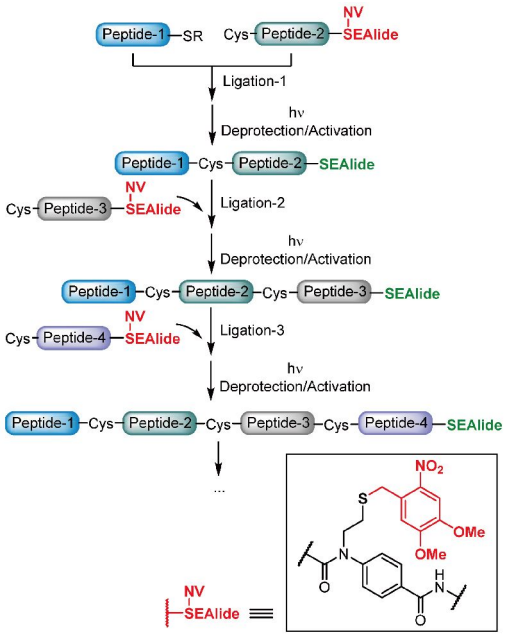

Figure 15 Photo-caged latent peptidyl SEAlide in sequential multi-fragment ligations in one-pot (color online).

http://engine.scichina.com/doi/10.1007/s11426-016-0381-1

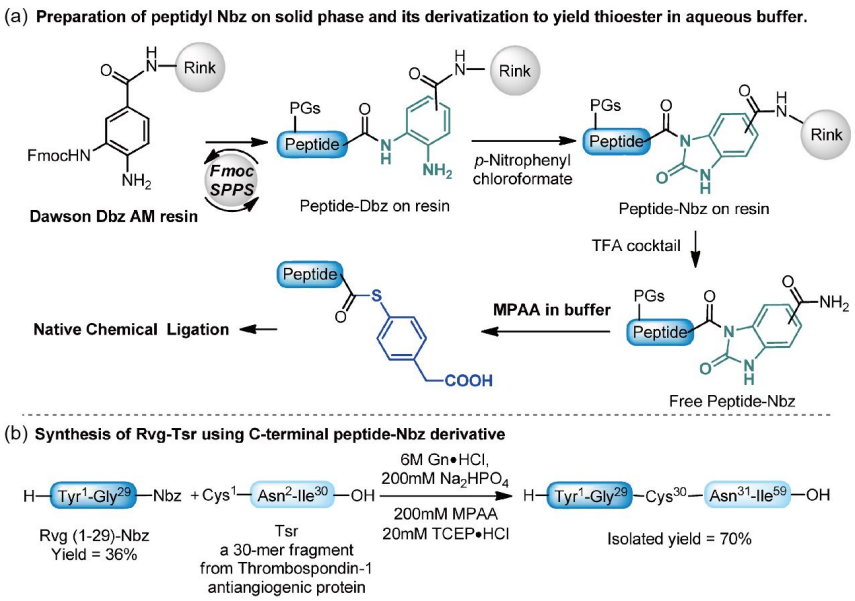

Figure 16 Peptidyl Nbz as thioester surrogate developed by Dawson and co-workers [48] (color online).

amino acid polypeptide, Rvg(1-29)-Tsr (Figure 16(b)), as well as a number of proteins and glycoproteins [49].

One major advantage of using Nbz as thioester surrogate is the good compatibility in Fmoc-SPPS. However, it was noticed that uncontrollable over-acylation of Dbz species may occur as a side reaction that diminishes the formation of the desired peptide-Nbz species, particularly for some glycine-rich sequences (Figure 17(a)) [50,51]. Accordingly, more carefully controlled Fmoc-SPPS conditions are required [49a], and several optimization strategies have thus been investigated. For instance, Ottesen *et al.* [51] utilized orthogonal allyloxycarbonyl (Alloc) group to protect Dbz, which can be deprotected under palladium (0)-catalyzed conditions after SPPS and subsequently converted to Nbz smoothly (Figure 17(b)). This protocol was successfully applied to the preparation of a 44-mer peptide H3M as Nbz derivative, which was further utilized in the total chemical synthesis of modified histone H3 [52].

In 2015, Dawson and co-workers [53] utilized an *o*-amino(methyl)aniline (MeDbz) moiety as the second-generation *N*-acylurea linker (Figure 17(c)). Compared to the first-generation Dbz linker, the methylated *para*-amine is more resistant toward acylation during SPPS even under microwave irradiation at 90 °C, yet maintaining sufficient reactivity toward *p*-nitrochloroformate to be activated afterwards. Such modification renders the *N*-acylurea activation strategy more robust and useful in peptide synthesis, e.g. two difficult cysteine-rich cyclotides, Kalata B1 and MCoTI-II.

Instead of transforming the peptidyl Dbz to the corresponding Nbz moiety, an alternative activation method is using $NaNO_2$-mediated oxidation developed by Liu *et al.* [54]. In such a protocol, peptide *o*-aminoanilide was cleaved from resin using TFA (Figure 18), followed by the treatment of $NaNO_2$ in buffered conditions to form peptide benzotriazole, which could further exchange with excess thiols to generate more reactive thioesters. Based on this new activation strategy, using benzotriazole-type crypto thioester, they successfully synthesized Histone H2B (type 1-M) from five fragments following N-to-C direction [54].

Besides the racemization-free feature and compatibility towards SPPS, another important advantage of using peptide *o*-aminoanilide lies in the carboxyl group on Dbz, where the incorporation of useful auxiliaries becomes possible. For instance, the attachment of an Arg_6-tag facilitated the synthesis of several hydrophobic protein regions [54,55], and use of His_6-tag simplified purification during multiple ligation steps [56]. In all these cases the efficiency of Fmoc-SPPS and NCL was greatly improved.

6 Peptide hydrazide in ligation reactions

While acyl azide has been utilized in amide bond formation to prepare peptide since 1904 [57], the idea of using peptidyl azide in peptide ligation has not been revealed until 2011, where the Liu group [58] took advantage of easily prepared peptide hydrazide and turned it into peptide-azide under oxidation-activation in mild ligation buffer. This versatile strategy has been proven to have multiple advantages, including easy and efficient preparation of peptide hydrazide via SPPS (Figure 19(a)), free of racemization during the activation and ligation in buffer (Figure 19(b)), compatibility to the bio-expression system, etc. Moreover, as peptide hydrazide remains inert in NCL, it becomes orthogonal to peptidyl thioesters, thus allowing for ligations performed in an N-to-C direction. To date, several glycoproteins have been successfully prepared using this peptide hydrazide strategy [59], and the

http://engine.scichina.com/doi/10.1007/s11426-016-0381-1

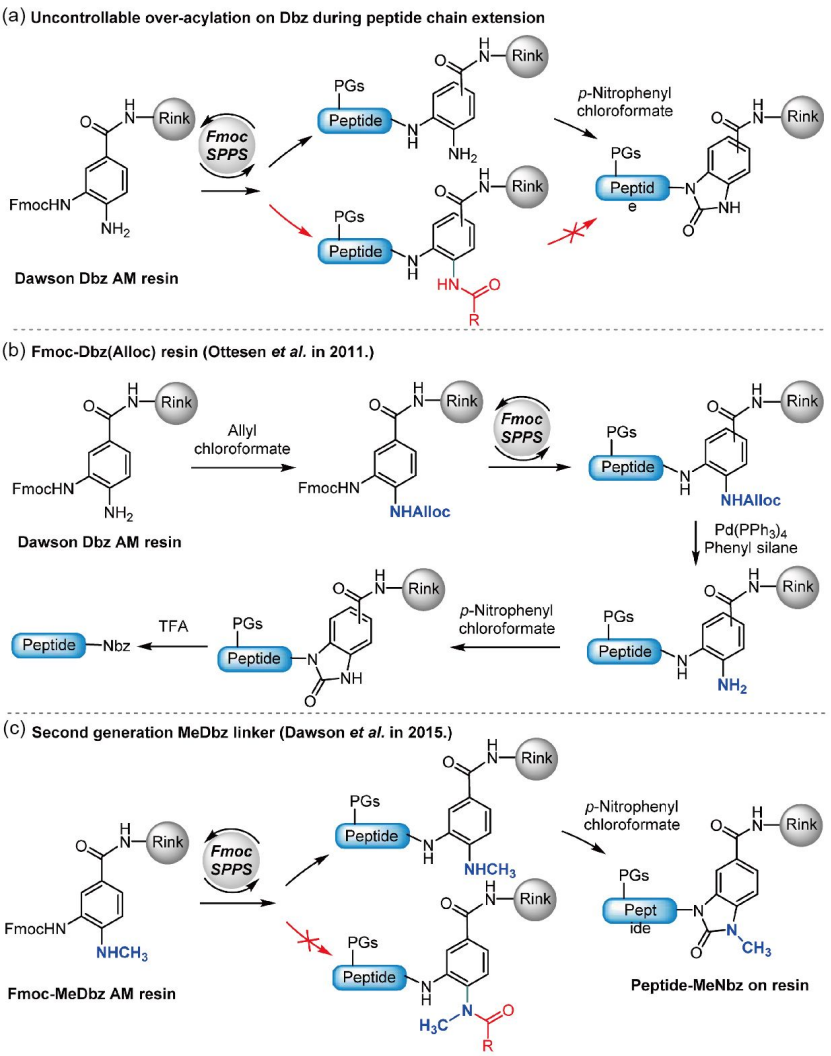

Figure 17 Uncontrollable over-acylation on Dbz and modified protocols (color online).

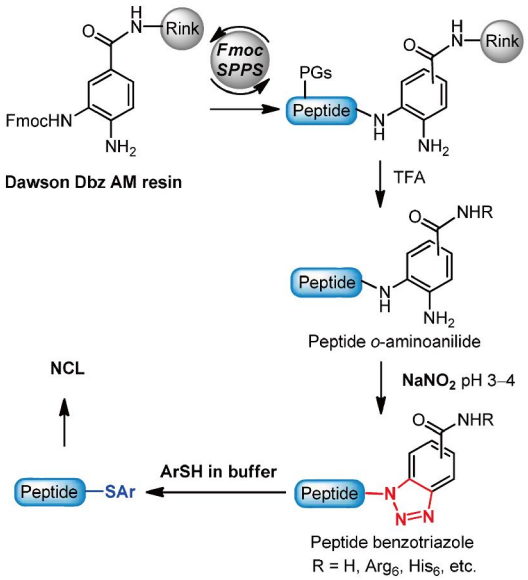

Figure 18 $NaNO_2$-mediated activation of peptide-Dbz [54] (color online).

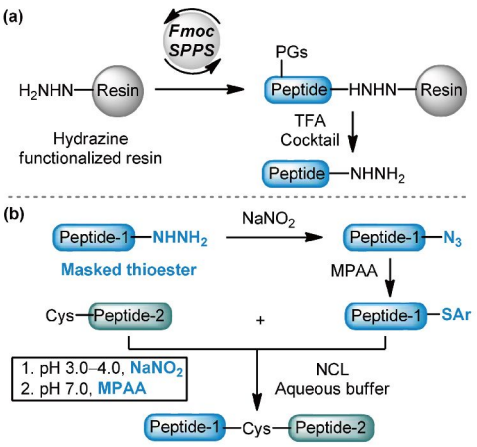

Figure 19 Peptide hydrazide as a "masked" thioester developed by Liu *et al.* [58]. (a) Preparation of peptide-hydrazide using hydrazine-functionalized resin; (b) use of peptide-hydrazide in NCL (color online).

http://engine.scichina.com/doi/10.1007/s11426-016-0381-1

hydrazinolysis process has been utilized to facilitate the development of several other strategies to prepare peptide thioesters [60]. This powerful strategy has been nicely summarized and reviewed in several review articles [6b,6f], thus the detail is not discussed further here.

7 Other approaches to thioesters or thioester surrogates

Besides the previous strategies, several other Fmoc-SPPS-compatible approaches leading to thioesters are also worth mentioning. In the studies toward synthesis of thioesters through direct Fmoc-SPPS, Crich *et al.* [61] utilized a thioglycinamide linker to connect peptide chain and resin (Figure 20). After *S*-alkylation of the thioamide, the resulting thioimide can be subsequently cleaved from resin under the treatment of TFA cocktail, which generates peptidyl thioester simultaneously. Unfortunately during the *S*-alkylation step the use of diazabicyclo[5.4.0]undec-7-ene (DBU) lead to epimerization of C-terminal phenylalanine residue, thus this strategy is still limited to the preparation of peptidyl glycinyl thioesters.

Another interesting approach is from the Michael group, where a photo-activated *N*-acyl-7-nitroindoline moiety was used as the thioester precursor (Figure 21) [62]. In their optimized protocols [63], after cleavage from resin, the peptidyl acyl-7-nitroindoline was exposed under UV irradiation in the presence of HOBt/HOSu, producing the corresponding activated peptide ester. Upon treatment with excess PhSH and TFA cocktail, the desired peptide phenylthioester was obtained in good efficiency even in the case of a 21-amino acid sequence. Although the application of this method in preparing longer peptide fragment has yet to be investigated, the idea of combining photo-activation and peptide thioester

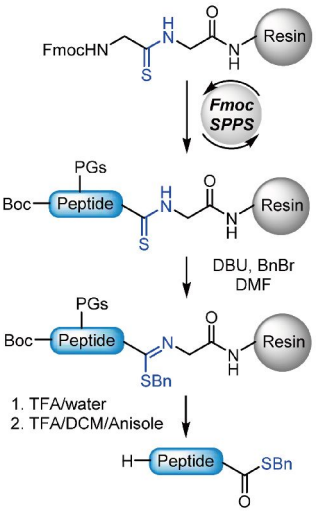

Figure 20 Direct production of thioester from solid phase synthesis by Crich *et al.* in 2011 [61] (color online).

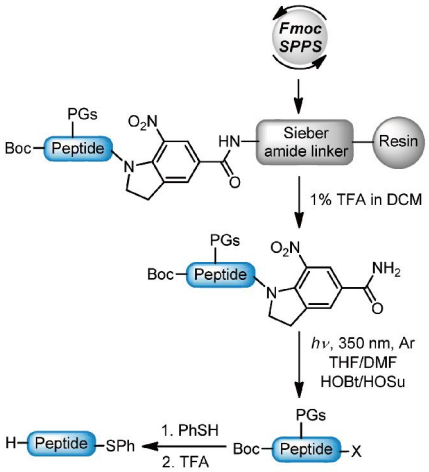

Figure 21 Photochemical synthesis of thioesters via a photo-labile moiety developed by Michael *et al.* [63] (color online).

formation may promote more innovative strategies in peptide synthesis.

Taking notes from the manipulation of peptidyl thioester through expression method using intein system [64], Kajihara, Okamoto and co-corkers [65] developed a new type of thioester surrogate, peptide-*N*-acetylguanidine, which can be prepared from a Cys-containing peptide through *S*-thiocarbonylation and followed by the replacement with *N*-acetylguanidine (Figure 22(a)). The required free peptide sequence with one additional C-terminal Cys incorporated could be obtained from either biological expression or chemical synthesis. The generated peptidyl *N*-acetylguanidine has been demonstrated to be less reactive than alkyl thioesters in ligation reactions, even in the presence of thiol additives (e.g. MPAA), which provides the possibility of using such peptide derivatives in KCL.

Due to the lack of chemoselectivity in the *S*-thiocarbonylation reaction, the first generation of *N*-acetylguanidine-based strategy could not be applied to free Cys-containing sequences until a solid phase derivatization strategy was developed in 2016 [66]. In this second-generation protocol, the thiol group of C-terminal Cys was orthogonally protected as disulfide form before the completion of peptide extension (Figure 22(b)), which can be further manipulated on resin and cleaved directly using TFA cocktail. With such a site-selective modification, a semi-synthesis of glycosylated interleukin-13 was accomplished [67].

Other than the cysteine residue, serine has also been identified as a suitable site for activation. In 2016, inspired by the Pessi method [5a] based on Kenner linker [68], a concise and straightforward method was developed by the Raj group [69] to form an activated cyclic urethane structure from an orthogonally protected serine residue (Figure 23), which can be easily transformed into thioester through thiolysis and cleaved from resin simultaneously. Additionally, threonine

http://engine.scichina.com/doi/10.1007/s11426-016-0381-1

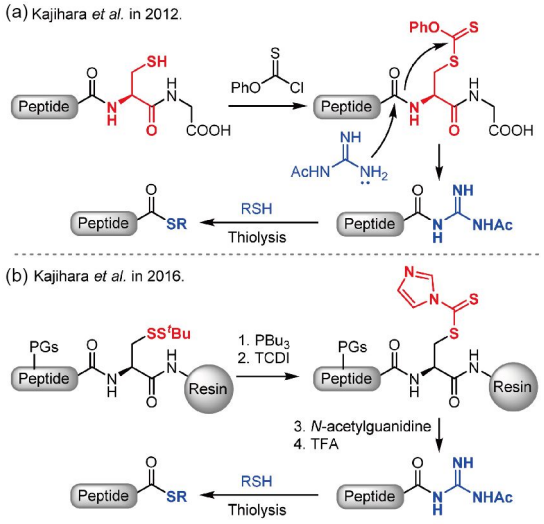

Figure 22 Peptidyl *N*-acetylguanidine as thioester precursor (color online).

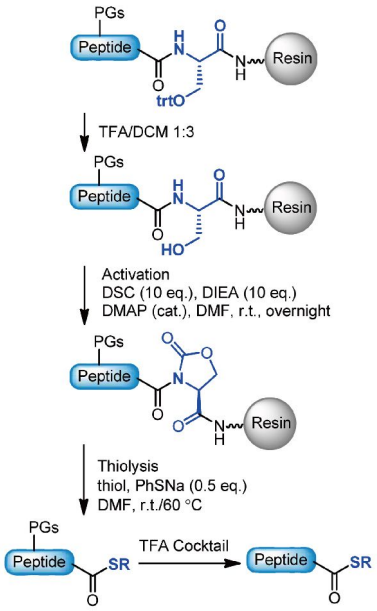

Figure 23 Cyclic urethane activation strategy developed by Raj *et al*. in 2016 [69] (color online).

and cysteine were also proved to be applicable activation sites under their optimized conditions. However, the requirement of TFA treatment after the thiolysis step adds an extra step to the whole process, which is the major disadvantage of related strategies. Nevertheless, the synthesis of a 29-mer fragment of rabies virus glycoprotein (Rvg) demonstrated the utility of this promising protocol in peptide synthesis.

8 Conclusions

Research on protein chemical synthesis has been an emerging and essential area in the field of chemical biology, which brings new vitality to the traditional peptide science. Researchers have made tremendous efforts, and successfully developed innovative methodologies, not only NCL, but also Staudinger ligation [70], serine/threonine ligation (STL) [71], α-ketoacid-hydroxylamine (KAHA) ligation [72], etc. There have been a number of new developments of these strategies [73], which could not be reviewed in details in this account due to the limited space. Nevertheless, these protocols, and many more advancements in the field in the foreseeable future, will continue bringing us inspirations on finding better ways to synthesize large biological molecules in a more efficient, precise, and controllable manner, helping us solve difficult problems in the fields of biology and medicine using the irreplaceable chemical tools.

Acknowledgments This work was supported by the Peking University Health Science Center (BMU20130354), State Key Laboratory of Natural and Biomimetic Drugs, the National Recruitment Program of Global Youth Experts (1000 Plan), and the National Natural Science Foundation of China (21502005).

Conflict of interest The authors declare no conflict of interest.

1 Dawson PE, Muir TW, Clark-Lewis I, Kent SBH. *Science*, 1994, 266: 776–779
2 Merrifield RB. *J Am Chem Soc*, 1963, 85: 2149–2154
3 (a) Li X, Kawakami T, Aimoto S. *Tetrahedron Lett*, 1998, 39: 8669–8672; (b) Clippingdale AB, Barrow CJ, Wade JD. *J Peptide Sci*, 2000, 6: 225–234; (c) Bu X, Xie G, Law CW, Guo Z. *Tetrahedron Lett*, 2002, 43: 2419–2422; (d) Hasegawa K, Sha YL, Bang JK, Kawakami T, Akaji K, Aimoto S. *Lett Peptide Sci*, 2001, 8: 277–284
4 Raz R, Rademann J. *Org Lett*, 2011, 13: 1606–1609
5 (a) Ingenito R, Bianchi E, Fattori D, Pessi A. *J Am Chem Soc*, 1999, 121: 11369–11374; (b) Shin Y, Winans KA, Backes BJ, Kent SBH, Ellman JA, Bertozzi CR. *J Am Chem Soc*, 1999, 121: 11684–11689; (c) Burlina F, Morris C, Behrendt R, White P, Offer J. *Chem Commun*, 2012, 48: 2579–2581
6 (a) Mende F, Seitz O. *Angew Chem Int Ed*, 2011, 50: 1232–1240; (b) Zheng JS, Tang S, Huang YC, Liu L. *Acc Chem Res*, 2013, 46: 2475–2484; (c) Thomas F. *J Pept Sci*, 2013, 19: 141–147; (d) Kawakami T. Protein ligation and total synthesis I. In: Liu L, Ed. *Peptide Thioester Formation via an Intramolecular N to S Acyl Shift for Peptide Ligation*. Switzerland: Springer International Publishing, 2014. 107–135; (e) Melnyk O, Agouridas V. *Curr Opin Chem Biol*, 2014, 22: 137–145; (f) Huang YC, Fang GM, Liu L. *Natl Sci Rev*, 2016: 1–10; (g) Huang Y, Liu L. *Sci China Chem*, 2015, 58: 1779–1781; (h) Pedersen SL, Jensen K. Instruments for automated peptide synthesis. In: Jensen KJ, Tofteng Shelton P, Pedersen SL, Eds. *Peptide Synthesis and Applications. Vol. 1047. Methods in Molecular Biology*. Totowa: Humana Press, 2013. 215–224
7 Jones JH, Witty MJ. *J Chem Soc Perkin Trans 1*, 1979, 3203
8 Sakakibara S. *Biopolymers*, 1995, 37: 17–28
9 von Eggelkraut-Gottanka R, Klose A, Beck-Sickinger AG, Beyermann M. *Tetrahedron Lett*, 2003, 44: 3551–3554
10 Flemer Jr S. *J Pept Sci*, 2009, 15: 693–696
11 Mezo AR, Cheng RP, Imperiali B. *J Am Chem Soc*, 2001, 123: 3885–3891

http://engine.scichina.com/doi/10.1007/s11426-016-0381-1

12 Kajihara Y, Yoshihara A, Hirano K, Yamamoto N. *Carbohyd Res*, 2006, 341: 1333–1340
13 Nagalingam A, Radford S, Warriner S. *Synlett*, 2007, 2007: 2517–2520
14 Williams MW, Young GT. *J Chem Soc*, 1963, 881
15 Barlos K, Gatos D. *Biopolymers*, 1999, 51: 266–278
16 Heinlein C, Varón Silva D, Tröster A, Schmidt J, Gross A, Unverzagt C. *Angew Chem Int Ed*, 2011, 50: 6406–6410
17 (a) Asahina Y, Komiya S, Ohagi A, Fujimoto R, Tamagaki H, Nakagawa K, Sato T, Akira S, Takao T, Ishii A, Nakahara Y, Hojo H. *Angew Chem Int Ed*, 2015, 54: 8226–8230; (b) Ding H, Shigenaga A, Sato K, Morishita K, Otaka A. *Org Lett*, 2011, 13: 5588–5591; (c) Aussedat B, Fasching B, Johnston E, Sane N, Nagorny P, Danishefsky SJ. *J Am Chem Soc*, 2012, 134: 3532–3541; (d) Fernández-Tejada A, Vadola PA, Danishefsky SJ. *J Am Chem Soc*, 2014, 136: 8450–8458
18 (a) Hackeng TM, Griffin JH, Dawson PE. *Proc Natl Acad Sci USA*, 1999, 96: 10068–10073; (b) Pollock SB, Kent SBH. *Chem Commun*, 2011, 47: 2342–2344
19 Bodanszky M. *Nature*, 1955, 175: 685
20 Wan Q, Chen J, Yuan Y, Danishefsky SJ. *J Am Chem Soc*, 2008, 130: 15814–15816
21 Durek T, Alewood PF. *Angew Chem Int Ed*, 2011, 50: 12042–12045
22 Raibaut L, Seeberger P, Melnyk O. *Org Lett*, 2013, 15: 5516–5519
23 Nakamura T, Shigenaga A, Sato K, Tsuda Y, Sakamoto K, Otaka A. *Chem Commun*, 2014, 50: 58–60
24 Nakamura T, Sato K, Naruse N, Kitakaze K, Inokuma T, Hirokawa T, Shigenaga A, Itoh K, Otaka A. *ChemBioChem*, 2016, 17: 1986–1992
25 (a) Hinderaker MP, Raines RT. *Protein Sci*, 2003, 12: 1188–1194; (b) Hodges JA, Raines RT. *Org Lett*, 2006, 8: 4695–4697; (c) Choudhary A, Kamer KJ, Powner MW, Sutherland JD, Raines RT. *ACS Chem Biol*, 2010, 5: 655–657
26 Gui Y, Qiu L, Li Y, Li H, Dong S. *J Am Chem Soc*, 2016, 138: 4890–4899
27 Wan Q, Danishefsky SJ. *Angew Chem Int Ed*, 2007, 46: 9248–9252
28 Ali Shah MI, Xu ZY, Liu L, Jiang YY, Shi J. *RSC Adv*, 2016, 6: 68312–68321
29 Botti P, Villain M, Manganiello S, Gaertner H. *Org Lett*, 2004, 6: 4861–4864
30 George EA, Novick RP, Muir TW. *J Am Chem Soc*, 2008, 130: 4914–4924
31 Liu F, Mayer JP. *J Org Chem*, 2013, 78: 9848–9856
32 Warren JD, Miller JS, Keding SJ, Danishefsky SJ. *J Am Chem Soc*, 2004, 126: 6576–6578
33 (a) Zheng JS, Cui HK, Fang GM, Xi WX, Liu L. *ChemBioChem*, 2010, 11: 511–515; (b) Fang GM, Cui HK, Zheng JS, Liu L. *Chem Eur J Chem Bio*, 2010, 11: 1061–1065
34 Hojo H, Onuma Y, Akimoto Y, Nakahara Y, Nakahara Y. *Tetrahedron Lett*, 2007, 48: 25–28
35 Asahina Y, Nabeshima K, Hojo H. *Tetrahedron Lett*, 2015, 56: 1370–1373
36 Terrier VP, Adihou H, Arnould M, Delmas AF, Aucagne V. *Chem Sci*, 2016, 7: 339–345
37 Shah NH, Muir TW. *Chem Sci*, 2014, 5: 446–461
38 (a) Ollivier N, Dheur J, Mhidia R, Blanpain A, Melnyk O. *Org Lett*, 2010, 12: 5238–5241; (b) Hou W, Zhang X, Li F, Liu CF. *Org Lett*, 2011, 13: 386–389
39 Ollivier N, Vicogne J, Vallin A, Drobecq H, Desmet R, El Mahdi O, Leclercq B, Goormachtigh G, Fafeur V, Melnyk O. *Angew Chem Int Ed*, 2012, 51: 209–213
40 (a) Raibaut L, Drobecq H, Melnyk O. *Org Lett*, 2015, 17: 3636–3639; (b) Raibaut L, Cargoët M, Ollivier N, Chang YM, Drobecq H, Boll E, Desmet R, Monbaliu JCM, Melnyk O. *Chem Sci*, 2016, 7: 2657–2665
41 Zheng JS, Chang HN, Wang FL, Liu L. *J Am Chem Soc*, 2011, 133: 11080–11083
42 Zheng JS, Chen X, Tang S, Chang HN, Wang FL, Zuo C. *Org Lett*, 2014, 16: 4908–4911
43 Otaka A, Sato K, Shigenaga A. Protein ligation and total synthesis II. In: Liu L, Ed. *Chemical Synthesis of Proteins Using N-sulfanylethylanilide Peptides, Based on N−S Acyl Transfer Chemistry*. Switzerland: Springer International Publishing, 2014. 33–56
44 Tsuda S, Shigenaga A, Bando K, Otaka A. *Org Lett*, 2009, 11: 823–826
45 Eto M, Naruse N, Morimoto K, Yamaoka K, Sato K, Tsuji K, Inokuma T, Shigenaga A, Otaka A. *Org Lett*, 2016, 18: 4416–4419
46 Aihara K, Yamaoka K, Naruse N, Inokuma T, Shigenaga A, Otaka A. *Org Lett*, 2016, 18: 596–599
47 (a) Zacharie B, Sauvé G, Penney C. *Tetrahedron*, 1993, 49: 10489–10500; (b) Pascal R, Chauvey D, Sola R. *Tetrahedron Lett*, 1994, 35: 6291–6294
48 Blanco-Canosa JB, Dawson PE. *Angew Chem Int Ed*, 2008, 47: 6851–6855
49 (a) Tiefenbrunn TK, Blanco-Canosa J, Dawson PE. *Biopolymers*, 2010, 94: 405–413; (b) Siman P, Blatt O, Moyal T, Danieli T, Lebendiker M, Lashuel HA, Friedler A, Brik A. *ChemBioChem*, 2011, 12: 1097–1104; (c) Fauvet B, Butterfield SM, Fuks J, Brik A, Lashuel HA. *Chem Commun*, 2013, 49: 9254–9256; (d) Okamoto R, Mandal K, Sawaya MR, Kajihara Y, Yeates TO, Kent SBH. *Angew Chem Int Ed*, 2014, 53: 5194–5198
50 White PD, Behrendt R. *J Pep Sci*, 2010, 16: 71–72
51 Mahto SK, Howard CJ, Shimko JC, Ottesen JJ. *ChemBioChem*, 2011, 12: 2488–2494
52 Shimko JC, North JA, Bruns AN, Poirier MG, Ottesen JJ. *J Mol Biol*, 2011, 408: 187–204
53 Blanco-Canosa JB, Nardone B, Albericio F, Dawson PE. *J Am Chem Soc*, 2015, 137: 7197–7209
54 Wang JX, Fang GM, He Y, Qu DL, Yu M, Hong ZY, Liu L. *Angew Chem Int Ed*, 2015, 54: 2194–2198
55 (a) Sato T, Saito Y, Aimoto S. *J Peptide Sci*, 2005, 11: 410–416; (b) Johnson ECB, Kent SBH. *Tetrahedron Lett*, 2007, 48: 1795–1799
56 (a) Bang D, Kent SBH. *Proc Natl Acad Sci USA*, 2005, 102: 5014–5019; (b) Gordon WR, Bang D, Hoff WD, Kent SBH. *Bioorgan Med Chem*, 2013, 21: 3436–3442
57 (a) Curtius T. *J Prakt Chem*, 1904, 70: 57–72; (b) Yanaihara N, Yanaihara C, Dupuis G, Beacham J, Camble R, Hofmann K. *J Am Chem Soc*, 1969, 91: 2184–2185; (c) Felix AM, Merrifield RB. *J Am Chem Soc*, 1970, 92: 1385–1391; (d) Romovacek H, Dowd SR, Kawasaki K, Nishi N, Hofmann K. *J Am Chem Soc*, 1979, 101: 6081–6091
58 (a) Fang GM, Li YM, Shen F, Huang YC, Li JB, Lin Y, Cui HK, Liu L. *Angew Chem Int Ed*, 2011, 50: 7645–7649; (b) Fang GM, Wang JX, Liu L. *Angew Chem Int Ed*, 2012, 51: 10347–10350; (c) Zheng JS, Tang S, Qi YK, Wang ZP, Liu L. *Nat Protoc*, 2013, 8: 2483–2495; (d) Tang S, Si YY, Wang ZP, Mei KR, Chen X, Cheng JY, Zheng JS, Liu L. *Angew Chem Int Ed*, 2015, 54: 5713–5717
59 (a) Li YM, Yang MY, Huang YC, Li YT, Chen PR, Liu L. *ACS Chem Biol*, 2012, 7: 1015–1022; (b) Reif A, Siebenhaar S, Tröster A, Schmälzlein M, Lechner C, Velisetty P, Gottwald K, Pöhner C, Boos I, Schubert V, Rose-John S, Unverzagt C. *Angew Chem Int Ed*, 2014, 53: 12125–12131; (c) Chang HN, Liu BY, Qi YK, Zhou Y, Chen YP,

http://engine.scichina.com/doi/10.1007/s11426-016-0381-1

Pan KM, Li WW, Zhou XM, Ma WW, Fu CY, Qi YM, Liu L, Gao YF. *Angew Chem Int Ed*, 2015, 54: 11760–11764; (d) Murakami M, Kiuchi T, Nishihara M, Tezuka K, Okamoto R, Izumi M, Kajihara Y. *Sci Adv*, 2016, 2: e1500678
60 (a) Tsuda Y, Shigenaga A, Tsuji K, Denda M, Sato K, Kitakaze K, Nakamura T, Inokuma T, Itoh K, Otaka A. *ChemOpen*, 2015, 4: 448–452; (b) Miyajima R, Tsuda Y, Inokuma T, Shigenaga A, Imanishi M, Futaki S, Otaka A. *Biopolymers*, 2016, 106: 531–546
61 Sharma I, Crich D. *J Org Chem*, 2011, 76: 6518–6524
62 Hogenauer TJ, Wang Q, Sanki AK, Gammon AJ, Chu CHL, Kaneshiro CM, Kajihara Y, Michael K. *Org Biomol Chem*, 2007, 5: 759–762
63 Pardo A, Hogenauer TJ, Cai Z, Vellucci JA, Castillo EM, Dirk CW, Franz AH, Michael K. *ChemBioChem*, 2015, 16: 1884–1889
64 Muir TW, Sondhi D, Cole PA. *Proc Natl Acad Sci USA*, 1998, 95: 6705–6710
65 Okamoto R, Morooka K, Kajihara Y. *Angew Chem Int Ed*, 2012, 51: 191–196
66 Okamoto R, Isoe M, Izumi M, Kajihara Y. *J Pept Sci*, 2016, 22: 343–351
67 Okamoto R, Kimura M, Ishimizu T, Izumi M, Kajihara Y. *Chem Eur J*, 2014, 20: 10425–10430
68 Kenner GW, McDermott JR, Sheppard RC. *J Chem Soc D*, 1971, 636
69 Elashal HE, Sim YE, Raj M. *Chem Sci*, 2016, doi: 10.1039/C6SC02162J
70 van Berkel SS, van Eldijk MB, van Hest JCM. *Angew Chem Int Ed*, 2011, 50: 8806–8827
71 Lee CL, Li X. *Curr Opin Chem Biol*, 2014, 22: 108–114
72 (a) Rohrbacher F, Wucherpfennig TG, Bode JW. Protein ligation and total synthesis II. In: Liu L, Ed. *Chemical Protein Synthesis with the KAHA Ligation*. Switzerland: Springer International Publishing, 2015. 1–31; (b) Harmand TJ, Murar CE, Bode JW. *Nat Protoc*, 2016, 11: 1130–1147
73 Thuaud F, Rohrbacher F, Zwicky A, Bode JW. *Helv Chim Acta*, 2016, 99: 868–894

http://engine.scichina.com/doi/10.1007/s11426-016-0381-1