ACCOUNTS
of chemical research

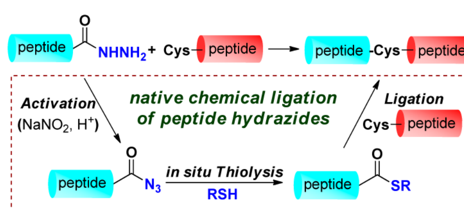
Development of New Thioester Equivalents for Protein Chemical Synthesis

JI-SHEN ZHENG, SHAN TANG, YI-CHAO HUANG, AND LEI LIU*
Tsinghua-Peking Center for Life Sciences, Key Laboratory of Bioorganic Phosphorus Chemistry & Chemical Biology (Ministry of Education), Department of Chemistry, Tsinghua University, Beijing 100084, China

RECEIVED ON JANUARY 17, 2013

CONSPECTUS

The chemical synthesis of proteins provides synthetic chemists with an interesting challenge and supports biological research through the generation of proteins that are not produced naturally. Although it offers advantages, studies of solid phase peptide synthesis have established limits for this technique: researchers can only prepare peptides up to 50 amino acids in length in sufficient yields and purity. Therefore, researchers have developed techniques to condense peptide segments to build longer polypeptide chains.



The method of choice for chemical synthesis of these longer polypeptides is convergent condensation of unprotected protein fragments by the native chemical ligation reaction in aqueous buffer. As researchers apply this strategy to increasingly difficult protein targets, they have needed to overcome diverse problems such as the requirement for a thiol-containing amino acid residue at the ligation site, the difficulty in synthesizing thioester intermediates under mild conditions, and the challenge of condensing multiple peptide segments with higher efficiency.

In this Account, we describe our research toward the development of new thioester equivalents for protein chemical synthesis. We have focused on a simple idea of finding new chemistry to selectively convert a relatively “low-energy” acyl group such as an ester or amide to a thioester under mild conditions. We have learned that this seemingly unfavorable acyl substitution process can occur by the coupling of the ester or amide with another energetically favorable reaction, such as the irreversible hydrolysis of an enamine or condensation of a hydrazide with nitrous acid. Using this strategy, we have developed several new thioester equivalents that we can use for the condensation of protein segments. These new thioester equivalents not only improve the efficiency for the preparation of the intermediates needed for protein chemical synthesis but also allow for the design of new convergent routes for the condensation of multiple protein fragments.

1. Introduction

Chemical synthesis of proteins enables a level of control over protein composition beyond that attainable by ribosome-dependent protein expression. Chemical synthesis not only facilitates the study of the structure–property relationship of protein functions at atomic resolution, but also holds promise for creating otherwise difficult-to-obtain proteins with usage in biomedicine.[1] Previous studies in the field of solid-phase peptide synthesis (SPPS) have established that only peptides containing less than about 50 amino acids can be reliably prepared with acceptable yields and purity.[1] Thus, for practical chemical synthesis of proteins, which typically contain polypeptide chains of much larger sizes, it is necessary to employ the strategy of peptide segment condensation.

A lesson from the early studies in this area is that the use of heavily protected peptides may cause many operational problems (e.g., solubility and intermediate characterization).[2,3] Accordingly a critical concept in protein chemical synthesis is the use of chemoselective condensation reactions to permit unambiguous coupling of unprotected peptides in aqueous buffers.[4] One such reaction is native chemical ligation (NCL) invented by Kent et al. (Figure 1) This method involves a chemoselective reaction between a peptide with a C-terminal thioester and a peptide with an N-terminal Cys to generate a peptide bond.[5] It allows efficient condensation of even very large peptide segments without causing complications such as racemization.

The advent of NCL greatly expands the capability of protein chemical synthesis.[6] However, an often encountered technical problem is how to prepare the peptide thioester intermediates. Peptide-thioesters can be prepared by Boc chemistry SPPS.[7] The HF used in the final deprotection step is

Published on the Web 05/23/2013 www.pubs.acs.org/accounts
10.1021/ar400012w © 2013 American Chemical Society

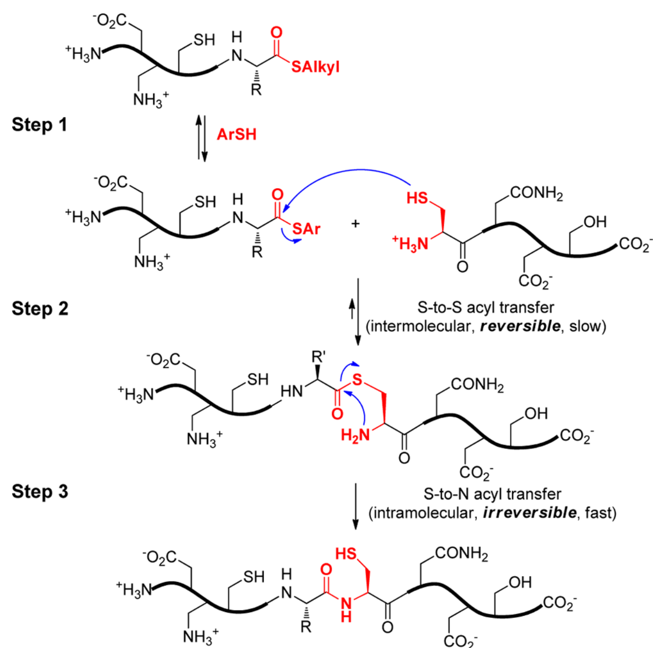


FIGURE 1. Native chemical ligation.

extremely hazardous, and for this reason the vast majority of researchers use Fmoc chemistry SPPS to prepare synthetic peptides. In addition, the strongly acidic HF deprotection is incompatible with the preparation of phosphopeptides and glycopeptides, post-translationally modified peptides that can be made by Fmoc SPPS.[8] Because of these considerations, it is important to develop practical and robust methods for the preparation of peptide-thioester equivalents for chemical protein synthesis.

Recently, several groups including us have developed new acyl derivatives to replace the thioester moieties required for NCL.[9] The present Account describes the studies mainly of our own group related to the development of new thioester equivalents. Our idea was to find a "low-energy" acyl functionality (i.e., an ester or amide) that can be selectively converted to a relatively "high-energy" thioester. Through the experiments, we learned that this task can be solved by the coupling of the seemingly unfavorable acyl substitution process with an independent "energy-downhill" reaction. With this idea in mind, we developed new thioester equivalents that can be used for protein chemical synthesis with more flexibility, especially for the convergent synthesis of proteins.[10]

2. Intramolecular Acyl Shift without External Activation

Thioesters are difficult to synthesize by Fmoc chemistry SPPS because of their low stability toward nucleophilic attack by the piperidine used to remove the Fmoc group at each stage of the synthesis.[8] To overcome this problem, one may propose to make a relatively more stable ester or amide first, and then convert it to a thioester right before use. Two strategies have been examined to promote the conversion of an ester or amide to a thioester. One requires the assistance of an intramolecular thiol moiety. The other involves the use of an external activating reagent.

2.1. Intramolecular *O*-to-*S* Acyl Shift. In 2004, Botti et al. reported that a peptide carboxyethylester bearing a free mercaptan in the β position may be in equilibrium with its thioester isomer (Figure 2).[11] This thioester intermediate can react with an N-terminal Cys peptide, generating the desired ligation product. Botti's finding can be explained by the involvement of an *O*-to-*S* acyl shift reaction occurring via a five-member ring intermediate. In the same year, Danishefsky et al. also reported the *O*-to-*S* acyl shift with peptide *ortho*-mercaptophenyl esters.[12]

We were interested by the *O*-to-*S* acyl shift method in the beginning of our study.[13] Recognizing that Botti's results had limitations (e.g., 20% of the peptide carboxyethylester was hydrolyzed during the ligation), we systematically screened the performance of different peptide esters bearing a thiol functionality (Figure 2). We focused on peptide alkylesters because peptide aryl esters were not found stable to NCL conditions.[14] Our experiments revealed that the presence of a substituent on the carbon adjacent to the ester *O* atom is beneficial to the *O*-to-*S* acyl shift because in its absence the ester is readily hydrolyzed, an observation also reported by Botti et al.[11] Furthermore, change of the side chain from an electron-withdrawing group (i.e., $-$CONHR) to an electron-donating one (i.e., $-CH_2OR$) can reduce the extent of hydrolysis (e.g., from 20% to 9%). Using the improved peptide ester capable of *O*-to-*S* acyl shift, we synthesized the cyclotide kalata B1 (a plant-derived circular protein) through intramolecular NCL.[15] Nevertheless, our work still has not eliminated the side reaction of ester hydrolysis which is a problem waiting for further studies.[13]

2.2. Intramolecular *N*-to-*S* Acyl Shift. The intramolecular *N*-to-*S* acyl shift converts a relatively inert amide to a more reactive thioester. Usually this reaction is reversible and thermodynamically unfavorable. As a result, acids as well as an excessive amount of external thiols are needed to push the equilibrium favoring the thioester formation (Figure 3a). Several recent studies have shown that the above strategy can be effective for the generation of thioesters.[16,17] However, the yield and reaction speed of this method remains to be improved, particularly for the peptide thioesters with a sterically crowded C-terminus.

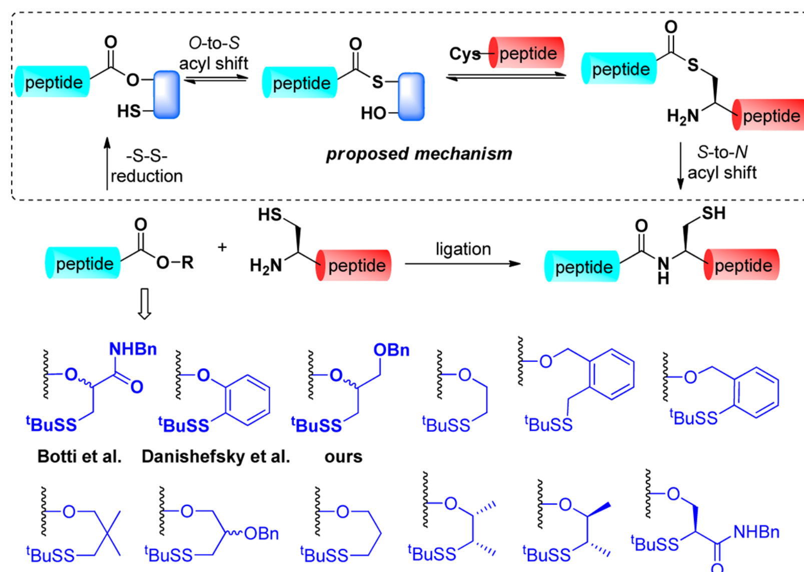
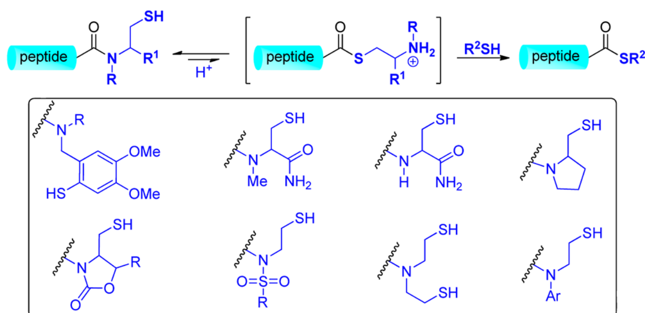
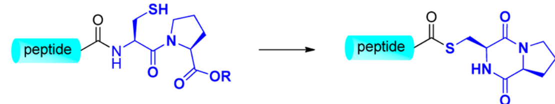


FIGURE 2. Methods based on intramolecular *O*-to-*S* acyl shift.

(a) Acid-mediated thioesterification



(b) The diketopiperazine strategy



(c) The enamide strategy

FIGURE 3. Methods based on intramolecular *N*-to-*S* acyl shift.

We hypothesized that the reversibility of the *N*-to-*S* acyl shift was a critical problem of the previous methods. A possible solution is to incorporate into the amide a functional group that can instantly deactivate the amino group as soon as it is generated from the acyl shift. Aimoto and Kawakami proposed an interesting method to capture the liberated amino group in the *N*-to-*S* acyl shift through diketopiperazine formation (Figure 3b).[17] We developed an alternative approach to prevent the reversibility of the *N*-to-*S* acyl shift by attaching a vinyl group to the amide nitrogen atom (Figure 3c).[18] The precursor of the *N*-to-*S* acyl shift is an enamide which is stable to the repeated treatment with piperidine in Fmoc SPPS. The target thioester is then produced through an intramolecular *N-to-S* acyl shift in the trifluoroacetic acid (TFA) cleavage media, which generates an enamine intermediate that is irreversibly hydrolyzed by water in the same cleavage cocktail.

The acid-promoted hydrolysis of an enamine to a carbonyl compound (which is an "energy-downhill" reaction) constitutes the key driving force for the irreversible *N-to-S* acyl shift. Note that the ketone group produced in the above process was not found to cause any problem for the NCL reaction in our experiments.[18] Even peptide thioesters with a sterically hindered C-terminus (e.g., Ile or Val) can be generated by this method. Only a routine Fmoc SPPS procedure is used without any postchain-assembly treatment, a feature that may be useful for the synthesis of peptide thioesters with sensitive post-translational modifications. To demonstrate the effectiveness of this method we used it to synthesize a model protein (human cytochrome oxidase 17). A problem of this approach is that the synthesis of the enamide-containing amino acid is laborious, calling for the design of simpler structures for irreversible *N*-to-*S* acyl shift.

3. Activation through an External Chemoselective Reaction

An alternative approach for the conversion of an ester or amide to a thioester involves the use of an external activating reagent to promote the thiolysis of the peptide C-terminus. One famous example for this approach is the "safety catch"

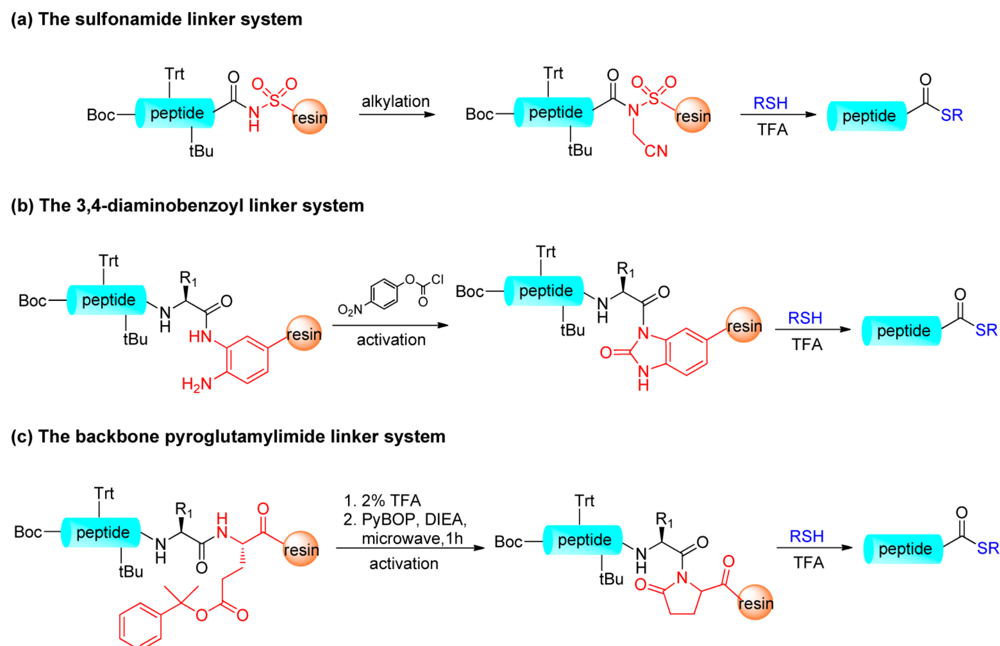


FIGURE 4. Thioester synthesis through activation by an external reagent.

linker system whose peptidyl-sulfonamide bond can be alkylated after the peptide chain assembly (Figure 4a).[19] The resulting *N*-alkyl-*N*-acyl sulfonamide is susceptible to attack by thiols to provide a fully protected peptide thioester, which is then treated with TFA for global deprotection. Using the analogous principle, Dawson and Blanco-Canosa recently developed the 3,4-diaminobenzoyl linker system in which the peptide thioester was obtained by thiolysis of C-terminal peptide N-acylurea generated via an acylation-promoted ring closure (Figure 4b).[20] In addition, Jensen et al. developed a thioester synthesis method that involves the activation of the peptide C-terminus via the formation of a backbone pyroglutamyl imide (Figure 4c).[21]

A shared feature of the above methods is that the activation/thiolysis process was carried out with protected peptides usually on the resin. This may cause technical difficulty in monitoring the progress of the transformation. To overcome this problem, we proposed to find a way to activate the peptide C-terminal acyl group in solution without any protecting group. The challenge behind this idea is that the activation/thiolysis process must be highly chemoselective to tolerate the many functional groups present in unprotected peptides. After many failed attempts, we found that the condensation reaction between a hydrazide and nitrous acid can satisfy our design requirements.

3.1. Peptide Hydrazides as Thioester Equivalents. The condensation of a hydrazide with nitrous acid produces an acyl azide. The latter can be aminolyzed by an amine to generate an amide bond. Due to this reason, hydrazides have been used for the coupling of peptide fragments since the beginning of peptide chemistry.[22a] Hofmann and several other earlier pioneers used the azide condensation method to prepare peptide segments from proteins such as ribonuclease T1.[22b,c] Unfortunately, the azide condensation method necessitates protections at Lys and Cys residues and use of organic solvents. Moreover, acyl azides tend to undergo Curtius degradation which further complicates their use.

We looked at the condensation of a hydrazide with nitrous acid from a different angle. Instead of using the resulting azide directly for amide coupling, we recognized the potential of the conversion of a hydrazide to an azide as a tool for chemoselective activation/thiolysis of peptide C-terminus (Figure 5).[23] Indeed, when we treated an unprotected peptide hydrazide (ca. 1 mM in aqueous guanidinium chloride solution) with $NaNO_2$ (ca. 10 mM) at pH 3−4 and −10 °C, we found that the conversion of the hydrazide to the corresponding acyl azide can be completed cleanly and quickly in about 20 min. We then added an external thiol (usually MPAA[24] (4-mercaptophenylacetic acid) at ca. 100 mM) to the reaction mixture and adjusted the pH value to 7.0. Almost instantly, the peptide azide was converted to the peptide thioester that can be directly used for NCL without isolation.

The above process is compatible with 17 amino acids at the C-terminus. For the other proteinogenic three amino acids (Gln, Asp, and Asn), the conversion was not successful because of the intramolecular cyclization of their own side

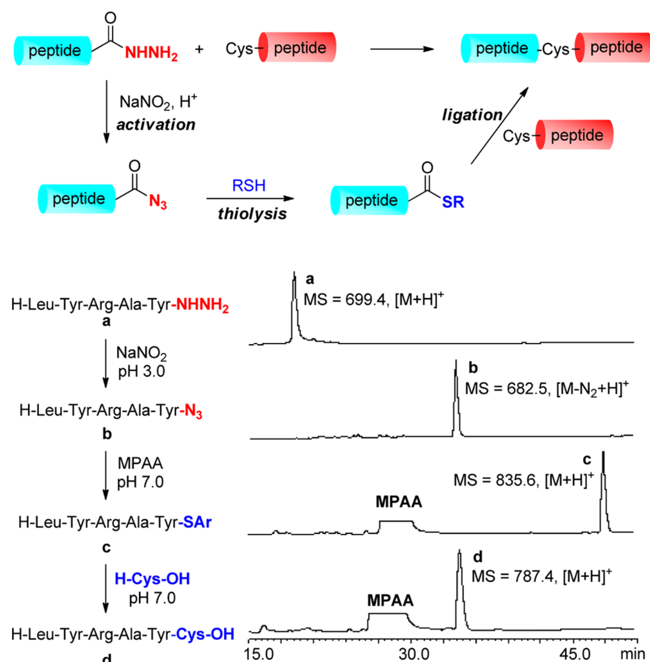


FIGURE 5. Conversion of a peptide hydrazide to a peptide thioester and use in NCL.

chains. Importantly, the brief treatment with $NaNO_2$ at pH 3–4 and −10 °C did not cause any oxidation for Met and Trp.[23] Furthermore, unprotected Ser, Thr, Tyr, His, Lys, Asp, Glu, and even Cys were found to be compatible with the above process.[23] As for Lys, its side-chain amino group does not react with nitrous acid at pH 3–4, while at pH 7 this amino group is also much less reactive than the external thiol. As for Cys, its side chain may be oxidized by HNO_2 to form nitrothioite (RSNO) and disulfide (RSSR). Fortunately, both nitrothioite and disulfide can be cleanly reduced by the external thiol added in the second step to the original thiol. Thus, although the acyl azide condensation method had been known, our study advanced it into a chemoselective method for the activation/thiolysis of peptide C-terminus. Finally, it is important to mention that our tests with the above procedure showed very little (<1%) racemization at the C-terminal amino acid.[23]

To test the utility of the new method, we synthesized a 42-mer peptide antibiotic trifolitoxin[Ala23Cys] through the ligation of two peptide fragments (i.e., [1–22]-$NHNH_2$ and [Cys23–42]). A one-pot protocol was used for the ligation, in which the peptide hydrazide fragment was treated with $NaNO_2$ at pH 3 in the first step, and then with the Cys-peptide and MPAA at pH 7 in the second step. The desired product was obtained in 97% HPLC yield without any intermediate isolation.[23] Note that trifolitoxin[Ala23Cys] contains nine Lys residues and one Cys residue. These residues were not protected as they do not interfere with the ligation.

3.2. Chemical Synthesis of Peptide Hydrazides. The use of peptide hydrazides is a modified version of NCL with an in situ generation of a peptide thioester from a peptide hydrazide. One important advantage of this method is that peptide hydrazides can be readily prepared through either Boc- or Fmoc-SPPS. In our initial studies, we synthesized peptide hydrazides through Fmoc SPPS on the Wang resin modified with a hydrazinecarboxylate linker. Later, we recognized that the *N*,*N*′-diacylhydrazine moiety may react with an activated Fmoc-Gly complicating the purification step at the end of the synthesis. To solve this problem we changed to the use of 2-Cl-(Trt)-$NHNH_2$ resin prepared through the treatment of 2-Cl-(Trt)-Cl resin with 5% hydrazine in DMF (Figure 6a).[25] Due to the protective effect of the bulky 2-Cl-Trt group, this resin can be used to make peptide hydrazides with higher purity.

We tested the application of the hydrazide-based method for making cyclic peptides (Figure 6b).[25] Through the use of 2-Cl-(Trt)-Cl resin we can readily prepare linear peptides bearing a C-terminal hydrazide and an N-terminal Cys. Upon treatment with $NaNO_2$ and an external thiol, these peptides can be smoothly cyclized with good isolated yields (18–63%). Relatively large cyclic peptides such as cyclotides (with ca. 30 amino acids) can be efficiently prepared with this hydrazide-based method. Moreover, by using the hydrazide-based method we can efficiently synthesize all-*L* cyclic tetrapeptides, which are highly strained molecules and have been challenging synthetic targets.[26]

3.3. Preparation of Peptide Hydrazides by Biological Expression. Unlike other thioester equivalents, peptide hydrazides have a unique feature that they can be generated via biological expression. Up to now we have examined two different approaches for the biological expression of peptide hydrazides. The first approach[23] is based on the expressed protein ligation (EPL) technology invented by Muir.[27] In the standard EPL, a recombinant protein thioester is produced via thiolysis of an intein fusion protein, which is then used for the NCL with a Cys-peptide to generate the semisynthetic protein. We found that by replacement of the external thiol in the thiolysis step with NH_2NH_2, we can readily obtain the protein hydrazide from the same intein fusion protein (Figure 7a). The reaction rate as well as yield of the hydrazinolysis process is usually higher than that of thiolysis.[23] By using the above approach we synthesized the microtubule-associated protein light chain 3 (LC3). After the fusion protein of LC3[Pro2-Thr118] and C-terminal gyrA intein-chitin

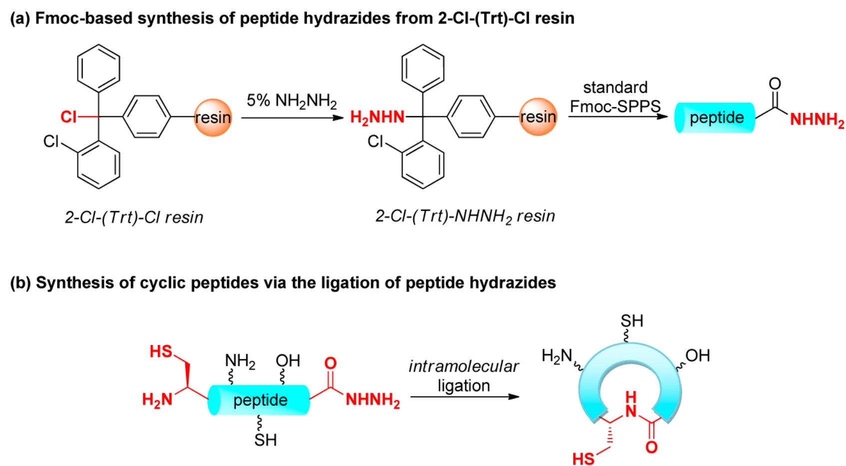


FIGURE 6. Intramolecular NCL of peptide hydrazides to produce cyclic peptides.

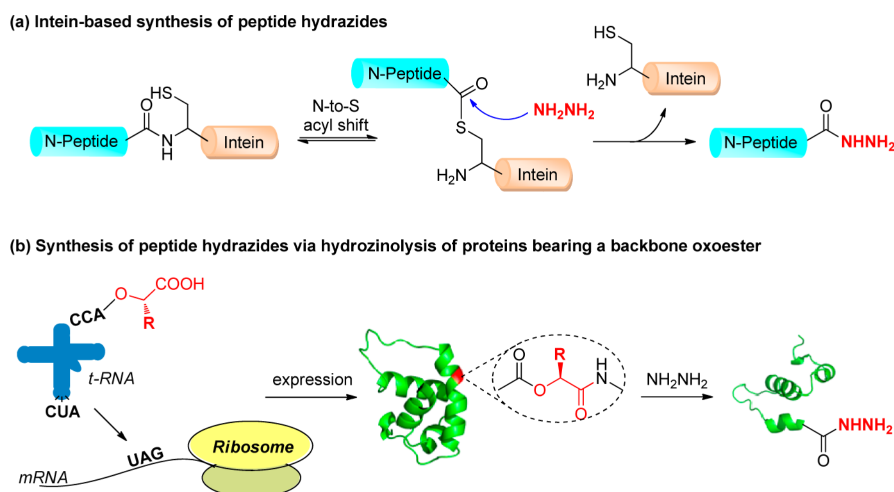


FIGURE 7. Preparation of peptide hydrazides by biological expression.

binding domain loaded on chitin affinity column, a cleavage buffer containing 8% NH_2NH_2 was added to produce LC3-[Pro2-Thr118]-$NHNH_2$. The resulting protein hydrazide was then reacted with MPAA to generate the protein thioester and then reacted with a synthetic Cys-peptide to obtain the C-terminal modified LC3.

Another approach for the biological expression of protein hydrazides involves the use of a protein bearing a site-specifically incorporated oxo-ester in its backbone (Figure 7b).[28] This approach is based on the genetic code expansion technology[29] pioneered by Schultz et al. In our implementation, we employed an evolved mutant pyrrolysyl-tRNA synthetase (i.e., ACPK-RS) that is compatible with both prokaryotic and eukaryotic cells to express recombinant proteins containing backbone oxo-esters. Subsequently, the oxo-ester containing proteins can be treated with aqueous NH_2NH_2 to generate the desired protein hydrazides. By using the *E. coli* acid chaperone HdeA as a model protein, we showed that the above approach can be used to prepare proteins in a semisynthetic manner with correctly folded structures and full biological activity.[28] Note that the above protein oxo-ester/hydrazide approach is a new version of EPL that does not involve protein splicing. Our results also indicated that the new approach may supplement the intein-based expressed protein ligation method by allowing for a more flexible selection of ligation site.

4. Condensation of Multiple Fragments

Typical proteins consist of several hundred amino acids whereas synthetic peptides usually contain less than 50 residues. Thus an important challenge for protein chemical synthesis is the efficient condensation of multiple peptide segments.[10,30] To maximize the convergence of the condensation process, both the sequential ligations in the C-to-N and N-to-C directions are needed. The C-to-N sequential ligation can be achieved by protecting Cys as Thz

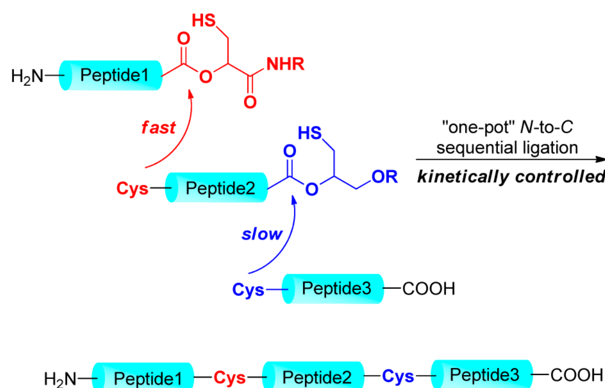


FIGURE 8. Kinetically controlled ligation of peptide oxo-esters.

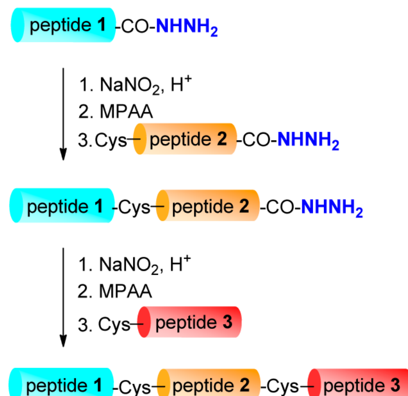


FIGURE 9. N-to-C sequential NCL of peptide hydrazides.

(L-thiazolidine-4-carboxylic acid).[10a,31] On the other hand, it is less straightforward to achieve the N-to-C sequential ligation due to the difficulty of protecting a thioester functionality.

4.1. Kinetically Controlled Ligation of Peptide oxo-Esters. An elegant method for the N-to-C sequential ligation is the kinetically controlled ligation (KCL) invented by Kent et al., which provides an effective solution for the convergent synthesis of proteins.[10a] The basic idea of KCL is that a peptide-thioarylester is much more reactive than a thioalkylester, so that in the absence of any added thiol a peptide-thioarylester can selectively react with a Cys-peptide-thioalkylester without affecting the thioalkylester group.

We were interested in the use of thioester equivalents in KCL.[13,32] We observed that the *O*-to-*S* acyl shift reaction with −CONHR side chain adjacent to the *O* atom is faster than that with $-CH_2OR$ by ca. 30-fold. This large reactivity difference allowed for carrying out KCL solely with peptide oxo-esters. As shown in Figure 8, three peptide segments can be sequentially ligated in the N-to-C direction. This strategy was tested for the synthesis of [V15A]crambin containing 46 amino acids. The left segment was first treated with the middle segment at 15 °C. After the ligation between these two segments was completed, the right segment was added into the reaction mixture and the ligation temperature was increased to 40 °C. Through this one-pot procedure, the target protein was obtained in 59% HPLC yield. Unfortunately, the side reaction of ester hydrolysis limits the utility of this approach.

4.2. N-to-C Sequential NCL of Peptide Hydrazides. KCL sometimes does not proceed smoothly because the dual reactivity of the middle segment is sensitive to the steric hindrance of the ligation sites.[10b,32b,33] To solve this problem, it may be necessary to completely "protect" the thioester of the middle segment to prevent any unwanted reaction.[6b,30c,34] Such protection means that the thioester surrogate should remain inert in the first ligation step. After the protecting group is removed, the intramolecular *N*-to-*S* acyl shift takes place and provides the transient thioester needed for the next ligation. A problem of this approach is that the yield and speed of the intramolecular *N*-to-*S* acyl shift may vary depending on the steric crowding of the peptide C-terminus.

We were interested in the use of peptide hydrazides in the N-to-C sequential ligation.[23] We found that peptide hydrazides are inert under the NCL conditions. Accordingly a Cys-peptide-$NHNH_2$ can react only at its N-terminus in a ligation reaction, whose product is still a peptide hydrazide capable of subsequent activation and ligation (Figure 9). To demonstrate this idea, we synthesized a model protein CssII (bearing 66 amino acids) through an N-to-C sequential ligation of four peptide fragments. The left fragment CssII-[Lys1-Gly11]-$NHNH_2$ was treated with $NaNO_2$ and MPAA to produce the corresponding thioarylester. This thioarylester was ligated with CssII[Cys12-Tyr24]-$NHNH_2$ to produce CssII[Lys1-Tyr24]-$NHNH_2$ that can be activated again. By repeating this strategy the remaining two fragments were successfully assembled into the target product.

4.3. Convergent Protein Synthesis with Peptide Hydrazides. Fully convergent synthesis is the most effective for the preparation of proteins by chemical means. Kent and other pioneers have examined a number of approaches for achieving fully convergent chemical protein synthesis.[10] In our own study, we thought that the ability of peptide hydrazides for the N-to-C sequential ligation may enable a new approach for convergent protein chemical synthesis.[35] For this purpose we first examined the use of peptide hydrazides in the C-to-N sequential ligation. Unfortunately, the Thz residue was found unstable during the activation of peptide hydrazides at pH 3−4. To overcome this problem,

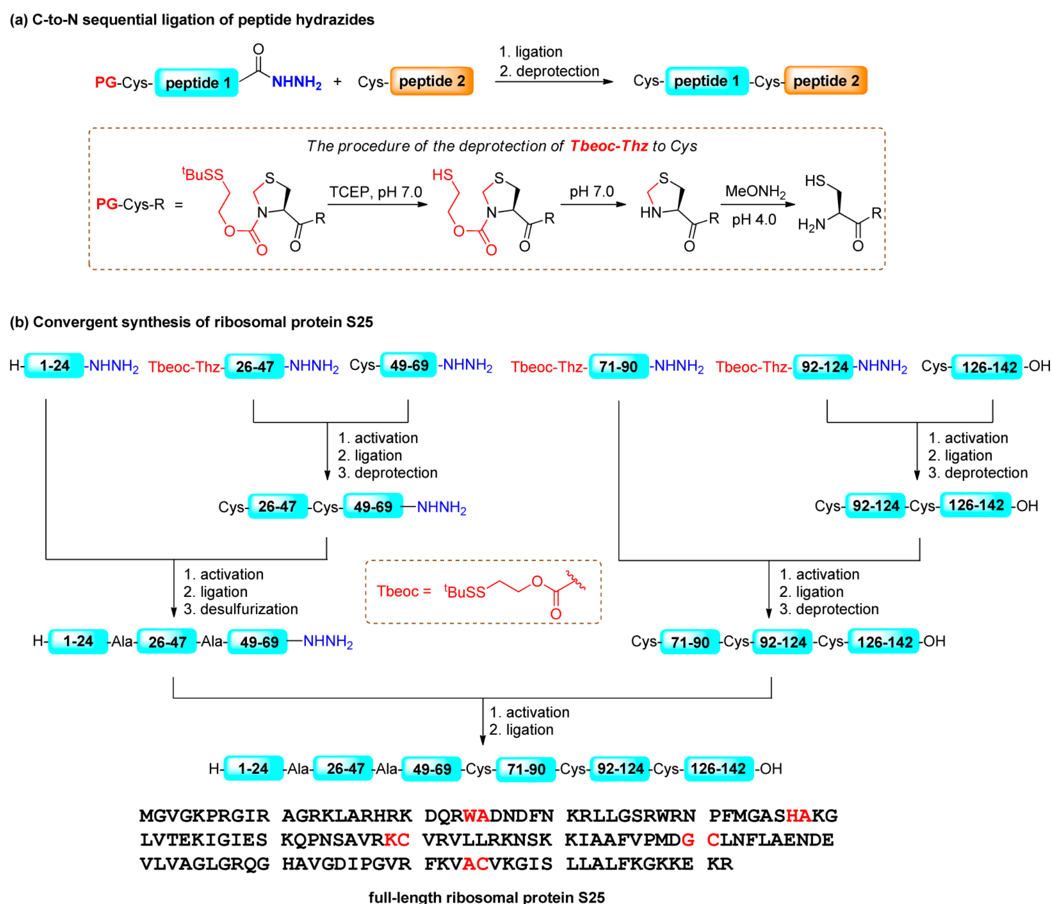


FIGURE 10. Convergent protein synthesis with peptide hydrazides.

we developed the 2-(*tert*-butyldisulfanyl)ethyloxycarbonyl (Tbeoc) protected Thz, with which the C-to-N sequential NCL of peptide hydrazides can be readily accomplished (Figure 10). The detailed process is as follows: First, Tbeoc-Thz-peptide1-$NHNH_2$ is activated by $NaNO_2$ at pH 3−4 and then ligated with Cys-peptide2 at pH 7 in the presence of MPAA. In this step, Tbeoc-Thz is converted into Thz so that the ligation product is Thz-peptide1-peptide2. Second, through the treatment with $MeONH_2 \cdot HCl$, Thz-peptide1-peptide2 can be readily converted into Cys-peptide1-peptide2 that is ready for the next ligation.

Having confirmed that both the N-to-C and C-to-N sequential ligations can be accomplished with the hydrazide-based approach, we designed a convergent route for the total synthesis a 142 residue protein, ribosomal protein S25 (RpS25). We divided the target protein into two halves, each containing three peptide segments. For the assembly of the right half, three peptides, that is, Tbeoc-Thz-[Val71-Gly90]-$NHNH_2$, Tbeoc-Thz-[Leu92-Ala124]-$NHNH_2$, and [Cys125-Arg142], were prepared by using the hydrazine-Trt(2-Cl) resin and Fmoc chemistry. Their C-to-N sequential ligation was conducted in the following sequence: First, Tbeoc-Thz-[Leu92-Ala124]-$NHNH_2$ was activated by $NaNO_2$ at pH 3 and then reacted with [Cys125-Arg142] by NCL at pH 7 in the presence of MPAA. The Thz-peptide was treated with $MeONH_2$ at pH 4 to generate [Cys91-Arg142]. Second, Tbeoc-Thz-[Val71-Gly90]-$NHNH_2$ was activated by $NaNO_2$ and then reacted with [Cys91-Arg142] by NCL in the presence of MPAA. After treatment with $MeONH_2 \cdot HCl$, we obtained [Cys70-Arg142].

For the synthesis of the left half, we also used the hydrazine-Trt(2-Cl) resin and Fmoc chemistry to prepare three peptide segments, that is, [Met1-Trp24]-$NHNH_2$, Tbeoc-Thz-[Asp26-His47]-$NHNH_2$, and [Cys48-Lys69]-$NHNH_2$. Their C-to-N sequential ligation was conducted as below: First, Tbeoc-Thz-[Asp26-His47]-$NHNH_2$ was activated by $NaNO_2$ at pH 3 and reacted with [Cys48-Lys69]-$NHNH_2$ by NCL in the presence of MPAA. The Thz-peptide was then treated with $MeONH_2 \cdot HCl$ to generate [Cys25-Lys69]-$NHNH_2$. Second, [Met1-Trp24]-$NHNH_2$ was activated by $NaNO_2$ and then reacted with [Cys25-Lys69]-$NHNH_2$ by NCL in the presence of MPAA to generate the ligation product [Met1-Lys69]-$NHNH_2$.

By using the free radical desulfurization conditions we desulfurized Cys25 and Cys48, leaving the hydrazide unit completely intact. This step converted Cys25 and Cys48 into Ala25 and Ala48 belonging to the native sequence of RpS25.

With the left and right halves in hand, we carried out the oxidation to the azide and the final native chemical ligation to obtain the target protein. The full-length RpS25 was successfully characterized with the ESI/MS and CD spectroscopy. Thus, our results demonstrated that the NCL of peptide hydrazides can be carried out in a convergent fashion. This convergent synthetic route is expected to provide a new approach for the preparation of proteins from multiple peptide segments.

5. Conclusions

From a philosophy point of view, chemical synthesis of proteins is a "grand challenge" in chemical sciences because it tests the limits of synthetic chemistry.[36] From a practical point of view, chemical synthesis of proteins that cannot be produced biologically is also an irreplaceable technology to tackle emerging questions in biology and medicine. The development of new chemistry is always important to continuously improve the power and efficiency of protein chemical synthesis.

In the Account, we described our research efforts on the development of thioester equivalents for protein chemical synthesis. We focused on the idea of converting a relatively "low-energy" acyl group (i.e., an ester or amide) to a thioester. Such a seemingly unfavorable acyl substitution process can be accomplished by the coupling with another "energy-downhill" reaction, such as the irreversible hydrolysis of an enamine or condensation of a hydrazide with nitrous acid. These methods, together with the new approaches recently developed by the other groups,[19−21,37] are expected to enable the tackling of more biological questions with chemically synthesized proteins.

This study was supported by the National Basic Research Program of China (973 program, No. 2013CB932800), the National Natural Science Foundation of China (Grants No. 20932006), National Science Fund for Distinguished Young Scholars, and the Specialized Research Fund for the Doctoral Program of Higher Education (Grant No. 20120002130004).

BIOGRAPHICAL INFORMATION

Ji-Shen Zheng received his B.S. degree in 2006. Afterwards he studied with Prof. Qing-Xiang Guo at University of Science and Technology of China and received Ph.D. degree in 2012. He is currently a postdoctoral scholar with Prof. Lei Liu at Tsinghua University, where he studies the development of new approaches for protein chemical synthesis.

Shan Tang was born and raised in Hubei, China. She graduated in Chemistry (B.S. 2011) from Tsinghua University. She is currently conducting her doctoral research on chemical synthesis of proteins in the laboratory of Prof. Lei Liu.

Yi-Chao Huang received his B.S. in chemical biology from Tsinghua University in 2011. Currently, he is pursuing his Ph.D. with Prof. Lei Liu. His research interest focuses on semisynthesis of proteins with posttranslational modifications.

Prof. Lei Liu obtained his B.S. degree from University of Science and Technology of China. After obtaining his Ph.D. degree from Columbia University in 2004 under the supervision of Prof. Ronald Breslow, he conducted postdoctoral research in the laboratory of Prof. Chi-Huey Wong at Scripps Research Institute. He has been teaching at Tsinghua University since 2007. He is interested in chemically synthesized proteins.

FOOTNOTES

*To whom correspondence should be addressed.
The authors declare no competing financial interest.

REFERENCES

1 (a) Kent, S. B. H. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **2009**, *38*, 338–351. (b) Nilsson, B. L.; Soellner, M. B.; Raines, R. T. Chemical synthesis of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91–118.

2 (a) Kung, Y. T.; Du, Y. C.; Huang, W. T.; Chen, C. C.; Ke, L. T. Total synthesis of crystalline bovine insulin. *Sci. Sinica* **1965**, *14*, 1710–1716. (b) Nishiuchi, Y.; Inui, T.; Nishio, H.; Bodi, J.; Kimura, T.; Tsuji, F. I.; Sakakibara, S. Chemical synthesis of the precursor molecule of the Aequorea green fluorescent protein, subsequent folding, and development of fluorescence. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 13549–13554. (c) Tam, J. P.; Xu, J.; Eom, K. D. Methods and strategies of peptide ligation. *Biopolymers,* **2001**, *60*, 194–205.

3 Kent, S. B. H. Chemical synthesis of peptides and proteins. *Annu. Rev. Biochem.* **1988**, *57*, 957–989.

4 Schnolzer, M.; Kent, S. B. H. Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease. *Science* **1992**, *256*, 221–225.

5 Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. Synthesis of proteins by native chemical ligation. *Science* **1994**, *266*, 776–779.

6 (a) McGinty, R. K.; Kim, J.; Chatterjee, C.; Roeder, R. G.; Muir, T. W. Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* **2008**, *453*, 812–816. (b) Kumar, K. S.; Bavikar, S. N.; Spasser, L.; Moyal, T.; Ohayon, S.; Brik, A. Total chemical synthesis of a 304 amino acid K48-linked tetraubiquitin protein. *Angew. Chem., Int. Ed.* **2011**, *50*, 6137–6141.

7 (a) Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. In situ neutralization in Boc-chemistry solid phase peptide synthesis. *Int. J. Pept. Res. Ther.* **2007**, *13*, 31–44. (b) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10068–10073.

8 (a) Shin, Y.; Winans, K. A.; Backes, B. J.; Kent, S. B. H.; Ellman, J. A.; Bertozzi, C. R. Fmoc-based synthesis of peptide-$^{\alpha}$thioesters: application to the total chemical synthesis of a glycoprotein by native chemical ligation. *J. Am. Chem. Soc.* **1999**, *121*, 11684–11689. (b) Huse, M.; Holford, M. N.; Kuriyan, J.; Muir, T. W. Semisynthesis of hyperphosphorylated type I TGFβ receptor: addressing the mechanism of kinase activation. *J. Am. Chem. Soc.* **2000**, *122*, 8337–8338.

9 (a) Canne, L. E.; Bark, S. J.; Kent, S. B. H. Extending the applicability of native chemical ligation. *J. Am. Chem. Soc.* **1996**, *1118*, 5891–5896. (b) Yan, L. Z.; Dawson, P. E. Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **2001**, *123*, 526–533. (c) Wan, Q.; Danishefsky, S. J. Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angew. Chem., Int. Ed.* **2007**, *46*, 9248–9252. (d) Li, J.; Cui, H. K.; Liu, L. Peptide ligation assisted by an auxiliary attached to amidyl nitrogen. *Tetrahedron Lett.* **2010**, *51*, 1793–1796. (d) Rohde, H.; Seitz,

O. Ligation-desulfurization: a powerful combination in the synthesis of peptides and glycopeptides. *Biopolymers* **2010**, *94*, 551–559.

10 Previous studies on convergent protein synthesis: (a) Bang, D.; Pentelute, B. L.; Kent, S. B. H. Kinetically controlled ligation for the convergent chemical synthesis. *Angew. Chem., Int. Ed.* **2006**, *45*, 3985–3988. (b) Torbeev, V. Y.; Kent, S. B. H. Convergent chemical synthesis and crystal structure of a 203 amino acid "covalent dimer" HIV-1 protease enzyme molecule. *Angew. Chem., Int. Ed.* **2007**, *46*, 1667–1670. (c) Liu, S.; Pentelute, B. L.; Kent, S. B. H. Convergent chemical synthesis of [lysine(24,38,83)] human erythropoietin. *Angew. Chem., Int. Ed.* **2012**, *51*, 993–999. (d) Avital-Shmilovici, M.; Mandal, K.; Gates, Z. P.; Phillips, N. B.; Weiss, M. A.; Kent, S. B. H. Fully convergent chemical synthesis of ester insulin: determination of the high resolution X-ray structure by racemic protein crystallography. *J. Am. Chem. Soc.* **2013**, *135*, 3173–3185.

11 Botti, P.; Villain, M.; Manganiello, S.; Gaertner, H. Native chemical ligation through in situ O to S acyl shift. *Org. Lett.* **2004**, *6*, 4861–4864.

12 Warren, J. D.; Miller, J. S.; Keding, S. J.; Danishefsky, S. J. Toward fully synthetic glycoproteins by ultimately convergent routes: a solution to a long-standing problem. *J. Am. Chem. Soc.* **2004**, *126*, 6576–6578.

13 Zheng, J. S.; Cui, H. K.; Fang, G. M.; Xi, W. X.; Liu, L. Chemical protein synthesis by kinetically controlled ligation of peptide O-esters. *ChemBioChem* **2010**, *11*, 511–515.

14 Fang, G. M.; Cui, H. K.; Zheng, J. S.; Liu, L. Chemoselective ligation of peptide phenyl esters with N-terminal cysteines. *ChemBioChem* **2010**, *11*, 1061–1065.

15 Zheng, J. S.; Chang, H. N.; Shi, J.; Liu, L. Chemical synthesis of a cyclotide via intramolecular cyclization of peptide O-esters. *Sci. China Chem.* **2012**, *55*, 64–69.

16 (a) Kawakami, T.; Sumida, M.; Nakamura, K.; Vorherr, T.; Aimoto, A. Peptide thioester preparation based on an N-S acyl shift reaction mediated by a thiol ligation auxiliary. *Tetrahedron Lett.* **2005**, *46*, 8805–8807. (b) Ollivier, N.; Behr, J. B.; El-Mahdi, O.; Blanpain, A.; Melnyk, O. Fmoc solid-phase synthesis of peptide thioesters using an intramolecular N,S-acyl shift. *Org. Lett.* **2005**, *7*, 2647–2650. (c) Nagaike, F.; Onuma, Y.; Kanazawa, C.; Hojo, H.; Ueki, A.; Nakahara, Y.; Nakahara, Y. Efficient microwave-assisted tandem N- to S-acyl transfer and thioester exchange for the preparation of a glycosylated peptide thioester. *Org. Lett.* **2006**, *8*, 4465–4468. (d) Ohta, Y.; Itoh, S.; Shigenaga, A.; Shintaku, S.; Fujii, N.; Otaka, A. Cysteine-derived S-protected oxazolidinones: potential chemical devices for the preparation of peptide thioesters. *Org. Lett.* **2006**, *8*, 467–470. (e) Kang, J.; Richardson, J. P.; Macmillan, D. 3-Mercaptopropionic acid-mediated synthesis of peptide and protein thioesters. *Chem. Commun.* **2009**, 407–409. (f) Ollivier, N.; Dheur, J.; Mhidia, R.; Blanpain, A.; Melnyk, O. Bis(2-sulfanylethyl)amino native peptide ligation. *Org. Lett.* **2010**, *12*, 5238–5241. (g) Hou, W.; Zhang, X.; Li, F.; Liu, C. F. Peptidyl N,N-bis(2-mercaptoethyl)-amides as thioester precursors for native chemical ligation. *Org. Lett.* **2011**, *13*, 386–389. (h) Sharma, R.; Tam, J. P. Tandem Thiol Switch Synthesis of Peptide Thioesters via N-S Acyl Shift on Thiazolidine. *Org. Lett.* **2011**, *13*, 5176–5179.

17 Kawakami, T.; Aimoto, S. Peptide ligation using a building block having a cysteinyl prolyl ester (CPE) autoactivating unit at the carboxy terminus. *Chem. Lett.* **2007**, *36*, 76–77.

18 Zheng, J. S.; Chang, H. N.; Wang, F. L.; Liu, L. Fmoc synthesis of peptide thioesters without post-chain-assembly manipulation. *J. Am. Chem. Soc.* **2011**, *133*, 11080–11083.

19 (a) Ingenito, R.; Bianchi, E.; Fattori, D.; Pessi, A. Solid phase synthesis of peptide C-terminal thioesters by Fmoc/t-Bu chemistry. *J. Am. Chem. Soc.* **1999**, *121*, 11369–11374.

20 Blanco-Canosa, J. B.; Dawson, P. E. An efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. *Angew. Chem., Int. Ed.* **2008**, *47*, 6851–6855.

21 Tofteng, A. P.; Sorensen, K. K.; Conde-Frieboes, K. W.; Hoeg-Jensen, T.; Jensen, K. J. Fmoc solid-phase synthesis of C-terminal peptide thioesters by formation of a backbone pyroglutamyl imide moiety. *Angew. Chem., Int. Ed.* **2009**, *48*, 7411–7414.

22 (a) Curtius, T. Verkettung von amidosäuren I. Abhandlung. *J. Prakt. Chem.* **1904**, *70*, 57–72. (b) Yanaihara, N.; Yanaihara, C.; Dupuis, G.; Beacham, J.; Camble, R.; Hofmann, K. Studies on polypeptides. XLII. Synthesis and characterization of seven fragments spanning the entire sequence of ribonuclease T1. *J. Am. Chem. Soc.* **1969**, *91*, 2184–2185. (c) Felix, A. M.; Merrifield, R. B. Azide solid phase peptide synthesis. *J. Am. Chem. Soc.* **1970**, *92*, 1385–1391.

23 Fang, G. M.; Li, Y. M.; Shen, F.; Huang, Y. C.; Li, J. B.; Lin, Y.; Cui, H. K.; Liu, L. Protein chemical synthesis by ligation of peptide hydrazides. *Angew. Chem., Int. Ed.* **2011**, *50*, 7645–7649.

24 Johnson, E. C. B.; Kent, S. B. H. Insights into the mechanism and catalysis of the native chemical ligation reaction. *J. Am. Chem. Soc.* **2006**, *128*, 6640–6646.

25 Zheng, J. S.; Tang, S.; Guo, Y.; Chang, H. N.; Liu, L. Synthesis of cyclic peptides and cyclic proteins via ligation of peptide hydrazides. *ChemBioChem* **2012**, *13*, 542–546.

26 Tang, S.; Zheng, J. S.; Yang, K.; Liu, L. Synthesis of cyclic tetrapeptides via ligation of peptide hydrazides. *Acta Chim. Sin.* **2012**, *70*, 1471–1476.

27 Muir, T. W.; Sondhi, D.; Cole, P. A. Expressed protein ligation: A general method for protein engineering. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705–6710.

28 Li, Y. M.; Yang, M. Y.; Huang, Y. C.; Li, Y. T.; Chen, P. R.; Liu, L. Ligation of expressed protein α-hydrazides via genetic incorporation of an α-hydroxy acid. *ACS Chem. Biol.* **2012**, *7*, 1015–1022.

29 Guo, J.; Wang, J.; Anderson, J. C.; Schultz, P. G. Addition of an alpha-hydroxy acid to the genetic code of bacteria. *Angew. Chem., Int. Ed.* **2008**, *47*, 722–725.

30 (a) Canne, L. E.; Ferre-D'Amare, A. R.; Burley, S. K.; Kent, S. B. H. Total chemical synthesis of a unique transcription factor-related Protein: cMyc-Max. *J. Am. Chem. Soc.* **1995**, *117*, 2998–3007. (b) Bang, D.; Chopra, N.; Kent, S. B. H. Total chemical synthesis of crambin. *J. Am. Chem. Soc.* **2004**, *126*, 1377–1383. (c) Ollivier, N.; Vicogne, J.; Vallin, A.; Drobecq, H.; Desmet, R.; El-Mahdi, O.; Leclercq, B.; Goormachtigh, G.; Fafeur, V.; Melnyk, O. A one-pot three-segment ligation strategy for protein chemical synthesis. *Angew. Chem., Int. Ed.* **2012**, *51*, 209–213.

31 Bang, D.; Kent, S. B. H. A one-pot total synthesis of crambin. *Angew. Chem., Int. Ed.* **2004**, *43*, 2534–2538.

32 (a) Wan, Q.; Danishefsky, S. J. Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angew. Chem., Int. Ed.* **2007**, *46*, 9248–9252. (b) Sato, K.; Shigenaga, A.; Tsuji, K.; Tsuda, S.; Sumikawa, Y.; Sakamoto, K.; Otaka, A. N-sulfanylethylanilide peptide as a crypto-thioester peptide. *ChemBioChem* **2011**, *12*, 1840–1844.

33 Lee, J.; Kwon, Y.; Pentelute, B. L.; Bang, D. Use of model peptide reactions for the characterization of kinetically controlled ligation. *Bioconjugate Chem.* **2011**, *22*, 1645–1649.

34 Kawakami, T.; Aimoto, S. Sequential peptide ligation by using a controlled cysteinyl prolyl ester (CPE) autoactivating unit. *Tetrahedron Lett.* **2007**, *48*, 1903–1905.

35 Fang, G. M.; Wang, J. X.; Liu, L. Convergent chemical synthesis of proteins by ligation of peptide hydrazides. *Angew. Chem., Int. Ed.* **2012**, *51*, 10347–10350.

36 Kent, S. B. H.; Sohma, Y.; Liu, S.; Bang, D.; Pentelute, B.; Mandal, K. Through the looking glass: a new world of proteins enabled by chemical synthesis. *J. Pept. Sci.* **2012**, *18*, 428–436.

37 Recent reviews: (a) Pattabiraman, V. R.; Bode, J. W. Rethinking amide bond synthesis. *Nature* **2011**, *480*, 471–479. (b) Baibaut, L.; Ollivier, N.; Melnyk, O. Sequential native peptide ligation strategies for total chemical protein synthesis. *Chem. Soc. Rev.* **2012**, *41*, 7001–7015. (c) Siman, P.; Brik, A. Chemical and semisynthesis of posttranslationally modified proteins. *Org. Biomol. Chem.* **2012**, *10*, 5684–5697.