

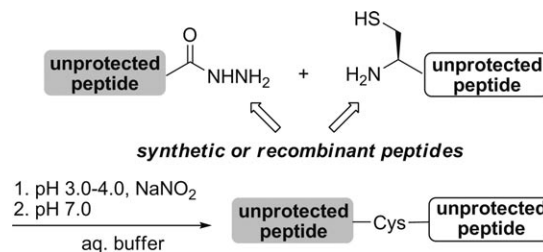
Protein Chemical Synthesis by Ligation of Peptide Hydrazides**

Ge-Min Fang, Yi-Ming Li, Fei Shen, Yi-Chao Huang, Jia-Bin Li, Yun Lin, Hong-Kui Cui, and Lei Liu*

Dedicated to Professor Ronald Breslow on the occasion of his 80th birthday

Protein chemical synthesis can overcome the potential limitations of protein expression and produce proteins with predesigned changes and modifications with atomic precision.^[1] Key to the success of modern protein chemical synthesis is the chemoselective ligation reaction. The most successful ligation method is the native chemical ligation developed by Kent et al.^[2] It involves a chemoselective reaction between a C-terminal peptide thioester and an N-terminal cysteine (Cys).^[3] Both synthetic and recombinant peptides can be used in native chemical ligation.^[2,4] The utility of native chemical ligation has been demonstrated by the total and semisynthesis of a variety of proteins, which enables the broad application of synthetic chemistry to the study of protein biology.^[5] Although native chemical ligation is a transformative method, its applicability can sometimes be limited in two respects: 1) peptide thioesters remain challenging to synthesize with Fmoc chemistry;^[6] 2) convergent synthesis of larger proteins without using protecting groups requires an “orthogonal” amide-forming ligation chemistry, that is, one that is compatible with the use of native chemical ligation.^[7]

In the present study we describe the ligation of peptide hydrazides that is complementary to native chemical ligation (Scheme 1). This method involves a chemoselective reaction between a C-terminal peptide hydrazide and a Cys-peptide to yield a native peptide bond. Notably, peptide hydrazides can be readily prepared through either Boc- or Fmoc-based solid-phase peptide synthesis (SPPS).^[8] In addition, peptide hydrazide can be obtained through biological expression. Thus both total and semisynthesis of proteins can be achieved with ligation of peptide hydrazides. More importantly, the ligation of peptide hydrazides enables sequential ligation in the N-to-C direction.^[9] It should be pointed out that peptide hydrazides have been used in protein chemistry since the beginning of the



Scheme 1. Ligation of peptide hydrazides.

field.^[10] However, the previous methods with peptide hydrazides are not chemoselective and necessitate protections at Lys and Cys residues. The new ligation method described here bypasses the requirement of protection groups, which is a critical advance.^[10]

Our study began with the ligation between model peptides H-Leu-Tyr-Arg-Ala-Tyr-NHNH₂ (**1a**) and H-Cys-Lys-Tyr-Met-His-OH (**2**).^[11] The ligation involves two steps that are carried out in a one-pot fashion. In the first step the two peptides (1.5 and 2.0 mM in final concentration, respectively) were added together to the aqueous phosphate (0.2 M) buffer containing 6.0 M guanidinium chloride (Gn-HCl). At a low pH (3.0–7.0) and –10°C, the oxidant (10 mM in final concentration) was added to the ligation mixture presumably producing a peptide azide. After 20 min, a thiol compound was added (100 mM in final concentration) and the pH value was adjusted to 7.0 to initiate the second step. The second step was allowed to proceed for 2 h at RT before the yield was determined by HPLC. As shown in Table 1, organic oxidants (entries 1 and 2) do not provide good results, whereas NaNO₂ can successfully stimulate the ligation to reach a high yield (entries 3 and 5). The optimal pH value for the oxidation is 3.0–4.0 and MPAA (4-mercaptophenylacetic acid)^[3] is important for mediating the ligation.

With the optimized conditions in hand, we investigated the scope of the ligation for 17 amino acids (Table 2). For three other amino acids (Xaa = Gln, Asp, and Asn) we have not been able to prepare the peptide hydrazides because intramolecular cyclization occurs between the hydrazide and C-terminal side-chain amide or acid group in SPPS (see the Supporting Information).^[12] In all cases, the first oxidation step was conducted at pH 3.0 for 20 min, while for most of the C-terminal amino acids the second step can reach completion with a high yield (84–99 %) in 2 h. Notably, unprotected Ser, Thr, Tyr, His, Lys, and even Cys are well compatible with the ligation. We did not observe any oxidized byproducts for Met and Trp either. For more sterically hindered amino acids (e.g.

[*] G.-M. Fang, Y.-M. Li, F. Shen, Y.-C. Huang, J.-B. Li, Y. Lin, H.-K. Cui, Prof. Dr. L. Liu

Key Laboratory of Bioorganic Phosphorus Chemistry & Chemical Biology (Ministry of Education)
Department of Chemistry, Tsinghua University
Beijing 100084 (China)
Fax: (+86) 10-6277-1149
E-mail: lliu@mails.tsinghua.edu.cn
Homepage: <http://chem.tsinghua.edu.cn/liugroup/>

[**] This work was supported by NSFC (no. 20802040 and no. 20932006) and the national “973” grants from the Ministry of Science and Technology (no. 2011CB965300).

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201100996>.

Table 1: Optimization of the ligation conditions.^[a]

H-Leu-Tyr-Arg-Ala-Tyr-NHNH ₂ 1a (1.5 mM)		H-Cys-Lys-Tyr-Met-His-OH 2 (2.0 mM)		
1. Oxidant (10 mM), pH		2. Thiol (100 mM), pH 7.0		H-Leu-Tyr-Arg-Ala-Tyr-Cys-Lys-Tyr-Met-His-OH 3a
Entry	pH of oxidation	Oxidant	Thiol	Yield [%]
1	3.0	<i>t</i> BuONO	MPAA	57
2	3.0	isopentyl nitrite	MPAA	32
3	3.0	NaNO ₂	MPAA	96
4	3.0	NaNO ₂	MESNa	36
5	4.0	NaNO ₂	MPAA	95
6	5.0	NaNO ₂	MPAA	0
7	6.0	NaNO ₂	MPAA	0
8	7.0	NaNO ₂	MPAA	0

[a] The first oxidation step was conducted at pH 3.0–7.0 and –10 °C in 6 M Gn-HCl aqueous solution for 20 min. The second step was conducted at pH 7.0 and RT after the addition of thiol (100 mM). MES = 2-mercaptoethane sulfonate.

Table 2: Scope of the ligation of peptide hydrazides.^[a]

Entry	Xaa (1)	Ligation time [h]	Yield [%]
1	Tyr (1a)	2.0	94
2	Ala (1b)	2.0	94
3	Met (1c)	2.0	91
4	Phe (1d)	2.0	96
5	Gly (1e)	2.0	99
6	His (1f)	2.0	93
7	Arg (1g)	2.0	95
8	Ser (1h)	2.0	91
9	Thr (1i)	5.0	95
10	Cys (1j)	2.0	96
11	Glu (1k)	2.0	84
12	Lys (1l)	2.0	90
13	Trp (1m)	2.0	92
14	Leu (1n)	3.0	90
15	Val (1o)	7.0	86
16	Ile (1p)	7.0	86
17	Pro (1q)	9.0	84

[a] The oxidation step was conducted at pH 3.0 and –10 °C in 6 M Gn-HCl aqueous solution for 20 min. The second step was conducted at pH 7.0 and RT after the addition of MPAA (100 mM).

Thr, Val, Ile, and Pro), the second step requires longer time (5–9 h) but the ligation yields are still good (84–95 %). To test whether the ligation involves racemization at the C-terminal amino acid, we compared the ligations of benzyl-protected Bz-Gly-(D)Xaa-NHNH₂ and Bz-Gly-(L)Xaa-NHNH₂ with H-Cys-OH for three C-terminal amino acids (Xaa = Ala, Phe, Cys) (see the Supporting Information). We also compared the ligations of H-Leu-Tyr-Ala-Ala-(D)Tyr-NHNH₂ and H-Leu-Tyr-Ala-Ala-(L)Tyr-NHNH₂ with H-Cys-OH (see the Supporting Information). The results showed that the extent of racemization is always less than 1 %. Moreover, we conducted competition experiments by ligating **1a** with a mixture of H-Lys-OH (20 equiv) and H-Cys-OH (1 equiv). Only H-Cys-OH was found to participate the ligation reaction, whereas H-Lys-OH remained inert (see the Supporting Information).

This experiment further confirmed that the ligation of peptide hydrazides is chemoselective towards the N-terminal Cys.

To understand the mechanism of the ligation of peptide hydrazides, we analyzed the intermediates of the ligation between **1a** and H-Cys-OH (Figure 1). We found that **1a** was

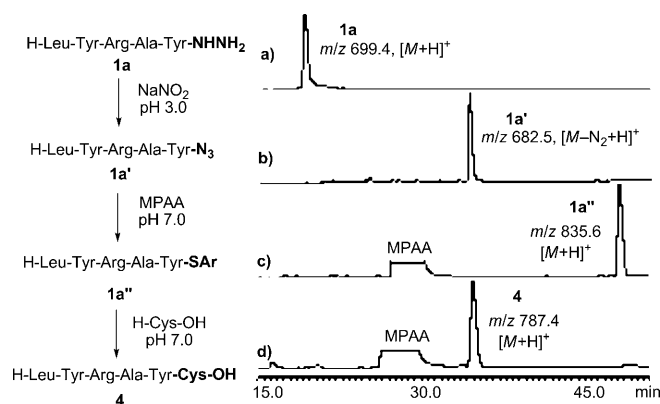


Figure 1. Intermediates of the ligation peptide hydrazides. Analytical HPLC traces (215 nm) for the reaction solution: a) before addition of NaNO₂; b) after addition of NaNO₂, *t* = 20 min, pH 3.0; c) *t* = 20 min, pH has been adjusted to 7.0 and MPAA added; d) after addition of H-Cys-OH, 2 h, pH 7.0.

cleanly oxidized by NaNO₂ to the peptide azide **1a'** in less than 20 min at pH 3.0. After MPAA had been added and the pH value adjusted to 7, **1a'** was immediately converted to the peptide thioester **1a''**. The thioester then reacted with H-Cys-OH chemoselectively by native chemical ligation to give the desired ligation product **4**. Note that the N-terminal amino group of **1a** was not protected during the process. Similarly, the Lys side chain does not require protection (Table 2). These amino groups do not react with either the peptide azide or peptide thioester at pH 3.0–7.0. Note that the Cys side chain also does not need protection in the ligation of peptide hydrazides. It has been known that the thiol group (RSH) may be oxidized by HNO₂ to form nitrothioite (RSNO) and disulfide (RSSR).^[13] Fortunately, both nitrothioite and disulfide can be cleanly reduced by other thiols to the original thiol.^[14] In the ligation of peptide hydrazides, the addition of excess MPAA not only promotes the thioester exchange reaction, but also reduces any oxidized Cys residues. Thus, although the acyl azide condensation method is known,^[10] the present protocol advances it into a truly chemoselective ligation method.

Like peptide thioesters used in native chemical ligation, peptide hydrazides can be readily prepared with the Boc-based SPPS.^[8] However, unlike peptide thioesters that are not stable to piperidine,^[6] peptide hydrazides can be directly synthesized with the Fmoc-based SPPS without any difficulty.^[11] Thus we expect the ligation of peptide hydrazides to be particularly useful for the synthesis of phosphorylated and glycosylated proteins,^[15] where the peptide thioesters have to be prepared through Fmoc chemistry. As an initial test of this idea, we successfully synthesized H-Leu-Tyr-Arg-Ser[β-D-Glc(OAc)₄]-Ala-NHNH₂ (**5**). This peptide was readily ligated

with **2** showing that the ligation of peptide hydrazides is compatible with glycopeptides.

Application of the ligation of peptide hydrazides to model protein synthesis is shown in Figure 2. In this example, the 42-

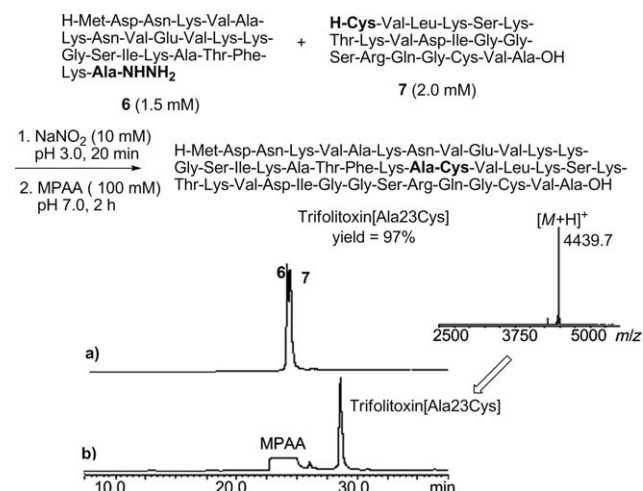


Figure 2. Synthesis of trifolitin[Ala23Cys]. Analytical HPLC traces (215 nm) for the reaction mixture a) before ligation and b) after ligation. The final product was characterized by MALDI-TOF mass spectrometry.

mer peptide antibiotic^[16] trifolitin[Ala23Cys] (i.e. Ala23 was mutated to Cys) was made by ligating two synthetic peptide fragments (Figure 2). The ligation site was chosen to be Ala22–Cys23. By using the optimized one-pot protocol shown in Table 2, we found that the ligation proceeds cleanly to provide the desired product in a high yield (97%). Note that trifolitin[Ala23Cys] contains nine Lys residues and one Cys residue. These residues do not interfere with the ligation of peptide hydrazides.

We want to point out that the ligation of peptide hydrazides provides complementary utility that is not easily achievable with native chemical ligation, that is, the sequential ligation in the N-to-C direction.^[9] In the previous studies the sequential ligation of multiple peptide fragments was usually conducted in the C-to-N direction by introducing a protecting group at the N-terminal Cys group.^[17] The N-to-C sequential ligation, on the other hand, is mostly carried out by using the kinetically controlled ligation (KCL) strategy.^[9] Using the ligation of peptide hydrazides we designed the new N-to-C sequential ligation approach shown in Figure 3. The key idea is to conduct the oxidation step for the first peptide hydrazide before adding the second peptide that possesses both an N-terminal Cys and a C-terminal hydrazide. After the oxidation, the addition of MPAA not only converts the peptide azide to peptide thioester, but also eliminates the remaining NaNO₂ in the ligation mixture. Then, addition of the second peptide into the reaction mixture will only cause ligation to its N-terminal Cys residue, while its C-terminal hydrazide will survive the ligation.

The new N-to-C sequential ligation strategy was tested for the synthesis of the model protein CssII.^[18] This protein has 66 amino acid residues including six Lys and eight Cys residues.

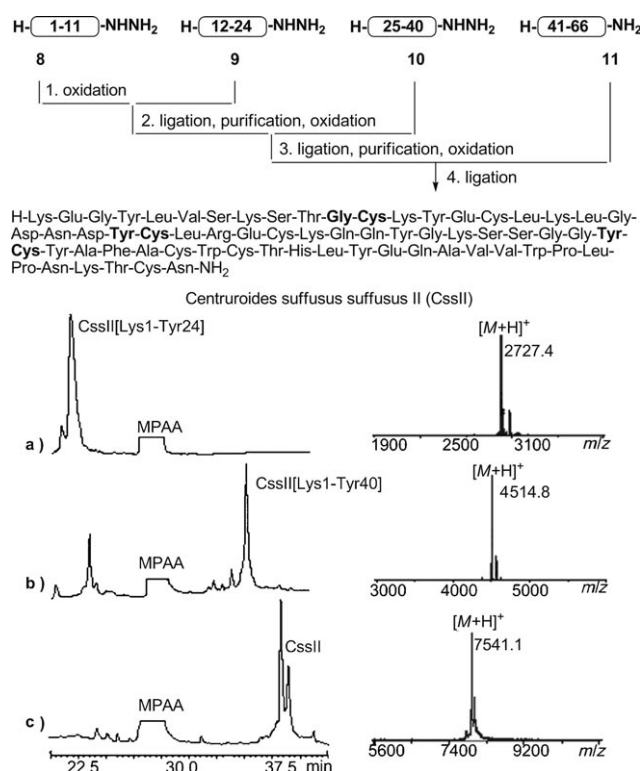


Figure 3. Total synthesis of CssII. Analytical HPLC traces (215 nm) for the reaction mixture a) after the first ligation, b) after the second ligation, and c) after the third ligation. More details are available in the Supporting Information.

In our synthesis the protein was divided into four fragments (8–11) and each fragment was prepared using Fmoc-based SPPS. To assemble the target, we first oxidized peptide **8** with NaNO₂ at pH 3.0 and –10 °C for 20 min. Then, MPAA was added to the reaction mixture, followed by the addition of peptide **9**. After the ligation was completed, the product CssII[Lys1–Tyr24] was purified from the reaction mixture. The purified CssII[Lys1–Tyr24] was used for the next rounds of ligation until the final product was obtained. The yields of the three ligations at the Gly11–Cys12, Tyr24–Cys25, and Tyr40–Cys41 sites are 92%, 85%, and 50%, respectively. The synthetic 66-residue polypeptide was characterized by MALDI-TOF mass spectrometry (Figure 3).

Finally, another important feature of the ligation of peptide hydrazides is that the key component of the method (i.e. peptide hydrazide) can be obtained through biological expression. This task is readily achieved by hydrazinolysis of the protein thioester intermediate in the standard expressed protein ligation (EPL) process.^[4] To test the application of the ligation of peptide hydrazides in protein semisynthesis, we chose the microtubule-associated protein light chain 3 (LC3) as the synthetic target.^[19] Native LC3 does not contain Cys, but to facilitate the ligation Phe119 was mutated to Cys. To obtain the protein hydrazide, the DNA sequence for the first 118 amino residues of LC3 was cloned into pTXB1 bacterial expression vector. The protein fused to C-terminal gyrA intein-chitin binding domain (CBD) was expressed in *E. coli* ER2566 cells and further loaded on chitin affinity column for

1–2 h for purification. LC3[Pro2–Thr118]-NHNH₂ was cleaved from the column during incubation with the normal MESNa cleavage buffer (pH 8.0) containing 8% NH₂NH₂. Then, the target protein was eluted with the cleavage buffer, concentrated by membrane ultrafiltration, and purified by RP-HPLC. The product LC3[Pro2–Thr118]-NHNH₂ was characterized by SDS-PAGE and ESI mass spectrometry (Figure 4).

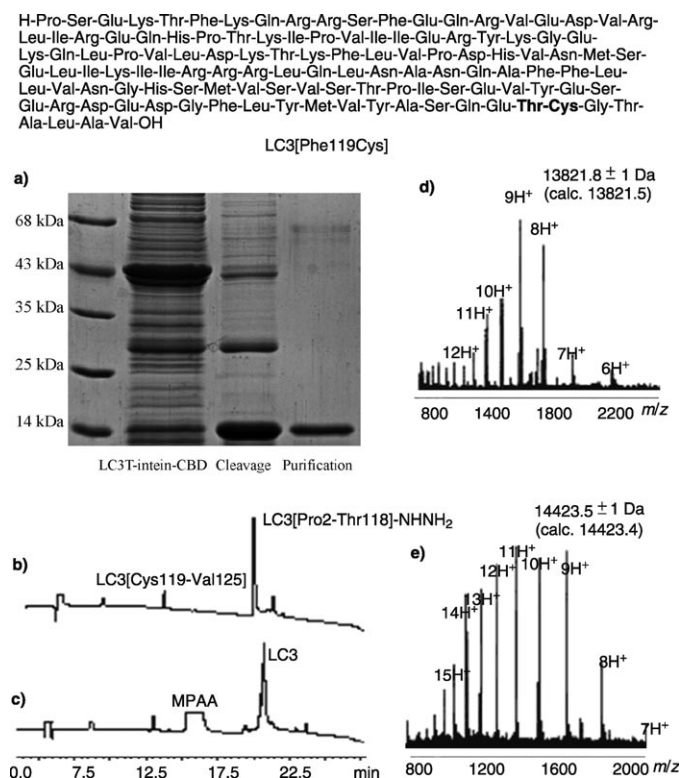


Figure 4. Semisynthesis of LC3[Phel19Cys]. a) SDS-PAGE analysis of the expression and cleavage of the protein hydrazide LC3[Pro2–Thr118]-NHNH₂. b,c) Analytical HPLC traces (215 nm) for the reaction mixture before (b) and after ligation (c). d) ESI mass spectrum for LC3[Pro2–Thr118]-NHNH₂. e) ESI mass spectrum for LC3[Phel19Cys]. More details can be found in the Supporting Information.

Further experiments showed that MESNa is not necessary in the cleavage buffer. A neutral aqueous buffer (pH 7.0) containing 8% NH₂NH₂ was sufficient for the hydrazinolysis of the LC3[Pro2–Thr118]-intein-CBD fusion protein affording the hydrazide product in less than 3 h. LC3[Pro2–Thr118]-NHNH₂ was then ligated with the heptapeptide H-Cys-Gly-Thr-Ala-Leu-Ala-Val-OH (i.e. LC3[Cys119–Val125]) using the one-pot protocol shown in Table 2 to give the full-length LC3[Phel19Cys] in 80% yield within 5 h. Thus, by changing the thiol cleavage step to hydrazine cleavage, the intein-based EPL method can be readily adapted to produce recombinant protein hydrazides that are suitable for the ligation of peptide hydrazides. In comparison to protein thioesters, we found that protein hydrazides can be produced more quickly (cleavage time 3 h) and in higher yields.^[20] A protein hydrazide is also chemically more stable than a protein thioester and therefore, easier to handle. Finally, unlike a protein thioester, the C-

terminal hydrazide is compatible with native chemical ligation unless it is activated through oxidation. Therefore, a protein hydrazide may be chemically modified first at its N terminus or side chains before the final ligation at its C terminus.

In summary, we have developed a ligation of peptide hydrazides that is complementary to the native chemical ligation. A useful advantage of the new ligation method is that peptide hydrazides can be easily prepared through both Boc- and Fmoc-based SPPS. In addition, peptide hydrazides can be readily obtained through recombinant expression. Both total and semisynthesis of proteins can be achieved with the ligation of peptide hydrazides. Moreover, the ligation of peptide hydrazides enables sequential ligation in the N-to-C direction. It should be pointed out that the ligation of peptide hydrazides is, in essence, a modified version of native chemical ligation with an in situ generation of a peptide-thioester from a peptide-hydrazide. We expect that the combination of native chemical ligation and the ligation of peptide hydrazides may enable more convergent synthesis of proteins.

Received: February 9, 2011

Published online: June 6, 2011

Keywords: native chemical ligation · peptide hydrazides · protein chemical synthesis · protein design · proteins

- [1] For a recent comprehensive review, see: S. B. H. Kent, *Chem. Soc. Rev.* **2009**, 38, 338.
- [2] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, 266, 776.
- [3] E. C. B. Johnson, S. B. H. Kent, *J. Am. Chem. Soc.* **2006**, 128, 6640.
- [4] M. Vila-Perelló, T. W. Muir, *Cell* **2010**, 143, 191.
- [5] Recent examples: a) Y. Sohma, Q. X. Hua, J. Whittaker, M. A. Weiss, S. B. H. Kent, *Angew. Chem.* **2010**, 122, 5621; *Angew. Chem. Int. Ed.* **2010**, 49, 5489; b) Y. W. Wu, L. K. Oesterlin, K. T. Tan, H. Waldmann, K. Alexandrov, R. S. Goody, *Nat. Chem. Biol.* **2010**, 6, 534; c) J. C. Scheuermann, A. G. D. Alonso, K. Oktaba, N. Ly-Hartig, R. K. McGinty, S. Fraterman, T. W. Muir, J. Muller, *Nature* **2010**, 465, 243.
- [6] For a recent review, see: F. Mende, M. Beisswenger, O. Seitz, *Angew. Chem.* **2011**, 123, 1266; *Angew. Chem. Int. Ed.* **2011**, 50, 1232.
- [7] a) L. E. Canne, A. R. Ferré-D'Amaré, S. K. Burley, S. B. H. Kent, *J. Am. Chem. Soc.* **1995**, 117, 2998; b) C. F. Liu, C. Rao, J. P. Tam, *J. Am. Chem. Soc.* **1996**, 118, 307; c) B. L. Nilsson, L. L. Kiessling, R. T. Raines, *Org. Lett.* **2000**, 2, 1939; d) J. W. Bode, R. M. Fox, K. D. Baucom, *Angew. Chem.* **2006**, 118, 1270; *Angew. Chem. Int. Ed.* **2006**, 45, 1248; e) C. P. R. Hackenberger, D. Schwarzer, *Angew. Chem.* **2008**, 120, 10182; *Angew. Chem. Int. Ed.* **2008**, 47, 10030.
- [8] a) S. S. Wang, *J. Am. Chem. Soc.* **1973**, 95, 1328; b) J. K. Chang, M. Shimizu, *J. Org. Chem.* **1976**, 41, 3255.
- [9] a) L. E. Canne, P. Botti, R. J. Simon, Y. Y. Chen, E. A. Dennis, S. B. H. Kent, *J. Am. Chem. Soc.* **1999**, 121, 8720; b) D. Bang, B. L. Pentelute, S. B. H. Kent, *Angew. Chem.* **2006**, 118, 4089; *Angew. Chem. Int. Ed.* **2006**, 45, 3985; c) J. S. Zheng, H. K. Cui, G. M. Fang, W. X. Xi, L. Liu, *ChemBioChem* **2010**, 11, 511; d) Z. P. Tan, S. Y. Shang, S. J. Danishefsky, *Angew. Chem.* **2010**, 122, 9690; *Angew. Chem. Int. Ed.* **2010**, 49, 9500.

- [10] a) T. Curtius, *J. Prakt. Chem.* **1904**, 70, 57; b) N. Yanaihara, C. Yanaihara, G. Dupuis, J. Beacham, R. Camble, K. Hofmann, *J. Am. Chem. Soc.* **1969**, 91, 2184; c) A. M. Felix, R. B. Merrifield, *J. Am. Chem. Soc.* **1970**, 92, 1385; d) H. Romovacek, S. R. Dowd, K. Kawasaki, N. Nishi, K. Hofmann, *J. Am. Chem. Soc.* **1979**, 101, 6081; e) For a related study on the Fmoc-based peptide thioester synthesis from an aryl hydrazine support, see: J. A. Camarero, B. J. Hackel, J. J. de Yoreo, A. R. Mitchell, *J. Org. Chem.* **2004**, 69, 4145.
- [11] To synthesize peptide hydrazides, we treated Wang resin with 4-nitrophenyl carbonochloridate. The resin was then reacted with hydrazine to load the hydrazinecarboxylate moiety. Next, by using the Fmoc protocols we prepared peptide hydrazides in good quality.
- [12] L. Campayo, B. Jimenez, T. Manzano, P. Navarro, *Synthesis* **1985**, 197.
- [13] L. Grossi, P. C. Montecchi, *J. Org. Chem.* **2002**, 67, 8625.
- [14] M. Kashiba-Iwatsuki, M. Yamagushi, M. Inoue, *FEBS Lett.* **1996**, 389, 149.
- [15] D. P. Gamblin, E. M. Scanlan, B. G. Davis, *Chem. Rev.* **2009**, 109, 131.
- [16] B. Y. Breil, P. W. Ludden, E. W. Triplett, *J. Bacteriol.* **1993**, 175, 3693.
- [17] a) A. Brik, E. Keinan, P. E. Dawson, *J. Org. Chem.* **2000**, 65, 3829; b) C. F. W. Becker, C. L. Hunter, R. Seidel, S. B. H. Kent, R. S. Goddy, M. Engelhard, *Proc. Natl. Acad. Sci. USA* **2003**, 100, 5075; c) D. Bang, N. Chopra, S. B. H. Kent, *J. Am. Chem. Soc.* **2004**, 126, 1377; d) E. C. B. Johnson, T. Durek, S. B. H. Kent, *Angew. Chem.* **2006**, 118, 3361; *Angew. Chem. Int. Ed.* **2006**, 45, 3283; e) D. Bang, S. B. H. Kent, *Proc. Natl. Acad. Sci. USA* **2005**, 102, 5014; f) E. C. B. Johnson, E. Malito, Y. Shen, D. Richi, W. J. Tang, S. B. H. Kent, *J. Am. Chem. Soc.* **2007**, 129, 11480; g) B. L. Pentelute, Z. P. Gates, J. L. Dashnau, J. M. Vanderkooi, S. B. H. Kent, *J. Am. Chem. Soc.* **2008**, 130, 9702.
- [18] M. F. Martin, L. G. G. Perez, M. E. Ayeb, C. Kopeyan, G. Bechis, E. Jover, H. Roach, *J. Biol. Chem.* **1987**, 262, 4452.
- [19] Y. Kabeya, N. Mizushima, T. Ueno, A. Yamamoto, T. Kirisako, T. Noda, E. Kominami, Y. Ohsumi, T. Yoshimori, *EMBO J.* **2000**, 19, 5720.
- [20] Through the MESNa cleavage we obtained 3–4 mg of LC3-thioester from 1 L of LB medium, whereas through the hydrazine cleavage we obtained 6–7 mg of LC3-hydrazide from 1 L of LB medium.