ORGANIC LETTERS

2012
Vol. 14, No. 1
374–377

N-to-C Sequential Ligation Using Peptidyl *N,N*-Bis(2-mercaptoethyl)amide Building Blocks

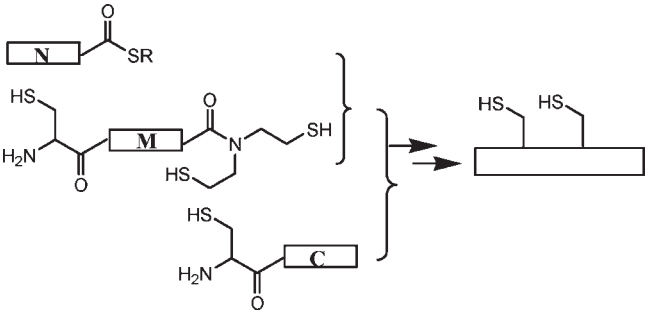
Renliang Yang, Wen Hou, Xiaohong Zhang, and Chuan-Fa Liu*

Structural Biology and Biochemistry Division, School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551

**cfliu@ntu.edu.sg*

Received November 22, 2011

ABSTRACT



A novel N- to C-terminus sequential chemical ligation approach has been developed for protein synthesis. Key to this strategy is the relative stability of the *N,N*-bis(2-mercaptoethyl)amide (BMEA) to the conventional conditions of native chemical ligation. We have also found a new thiol additive for the BMEA-mediated ligation reaction. The usefulness of this approach was demonstrated in the syntheses of a medium-sized peptide and ubiquitin.

The advent of solid-phase peptide synthesis[1] (SPPS) and chemical ligation methods[2,3] has revolutionized the practice of peptide and protein chemical synthesis and dramatically promoted biological structure–function study. For the total synthesis of an average-sized protein, it is often necessary to conduct sequential or convergent ligation of multiple fragments.[4] Sequential ligation can be done either from the C- to N-terminus or in the reverse direction. For instance, when performing sequential ligation from C to N, one needs to protect the N-terminal cysteine residue of all the internal thioester fragments to prevent the undesired self-ligation or cyclization reactions.[5] The deprotection and the following purification steps after each round of ligation introduce extra labors and lower the overall yield. While the ligation from C to N is straightforward, the ligation in the reverse direction is more challenging. To ensure that the N-to-C sequential ligation works, the electrophilic C-terminal moiety of the middle Cys-peptide should remain intact during the first ligation and therefore will be available for the next ligation step. Based on this

(1) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149.

(2) For Reviews, see: (a) Aimoto, S. *Biopolymers* **1999**, *51*, 247. (b) Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923. (c) Tam, J. P.; Yu, Q.; Miao, Z. *Biopolymers* **1999**, *51*, 311. (d) Kent, S. B. H. *Chem. Soc. Rev.* **2009**, *38*, 338. (e) Hackenberger, C. P. R; Schwarzer, D. *Angew. Chem., Int. Ed.* **2008**, *47*, 10030. (f) Payne, R. J.; Wong, C.-H. *Chem. Commun.* **2010**, *46*, 21.

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

(4) Examples of multisegment synthesis of large proteins: (a) Tam, J. P.; Yu, Q.; Yang, J. L. *J. Am. Chem. Soc.* **2001**, *123*, 2487. (b) Durek, T.; Torbeev, V. Y.; Kent, S. B. H. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4846. (c) Shimko, J. C.; North, J. A.; Bruns, A. N.; Poirier, M. G.; Ottesen, J. J. *J. Mol. Biol.* **2011**, *408*, 187. (d) Becker, C. F. W.; Hunter, C. L.; Seidel, R.; Kent., S. B. H.; Goody, R. S.; Engelhard, M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5075. (e) Lahiri, S.; Brehs, M.; Olschewski, D.; Becker, C. F. W. *Angew. Chem., Int. Ed.* **2011**, *50*, 3988.

(5) (a) Bang, D.; Kent, S. B. H. *Angew. Chem., Int. Ed.* **2004**, *43*, 2534. (b) Bang, D.; Kent, S. B. H. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5014. (c) Ueda, S.; Fujita, M.; Tamamura, H.; Fujii, N.; Otaka, A. *ChemBioChem* **2005**, *6*, 1983.

(6) (a) Tan, X.-H.; Zhang, X; Yang, R; Liu, C.-F. *ChemBioChem* **2008**, *9*, 1052. (b) Li, X.; Lam, H. Y.; Zhang, Y.; Chan, C. K. *Org. Lett.* **2010**, *12*, 1724.

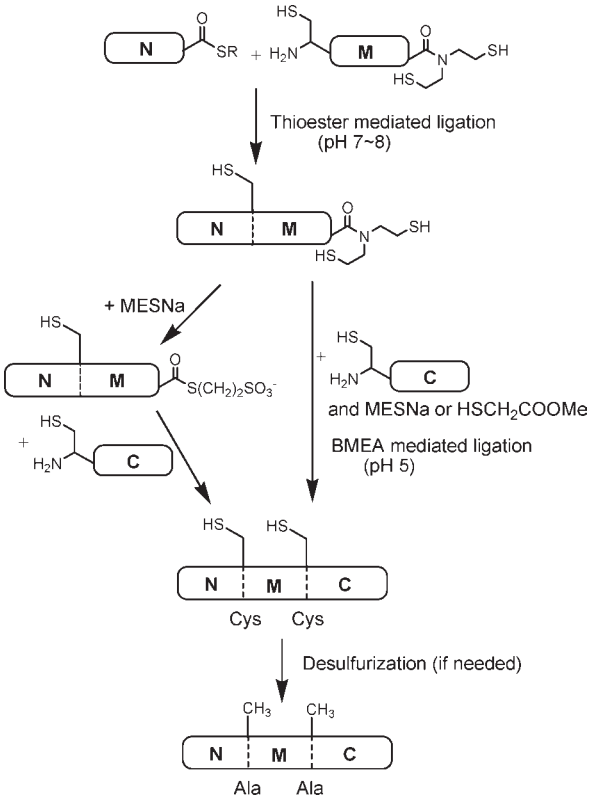
10.1021/ol2031284
Published on Web 12/16/2011

principle, many N-to-C sequential ligation methods were developed by using two or more orthogonal ligation chemistries.[4a,6] When the thioester-Cys ligation method was used as the sole chemistry for N-to-C sequential ligation, protein assembling was conducted either by using two thioesters with different reactivity as in kinetically controlled ligations[7] or by using thioester precursors.[8] The thioester precursor approach may require additional manipulations to convert the precursors to thioesters after each round of ligation, although the thioesters may also be formed in situ to avoid extra purification steps.[8]

Herein, we introduce a novel N-to-C sequential ligation approach based on the differential reactivity between a regular peptidyl thioester and *N,N*-bis(2-mercaptoethyl)-amide (BMEA) which was recently developed by our group[9] and Melnyk's group.[10] In the previous studies performed by the two groups, it was found that peptidyl BMEA could be converted to a thioester through N-to-S acyl transfer.[9−11] Under mild acidic conditions (pH 4−6), peptidyl BMEA can (i) be converted to a thioester by exchanging with a thiol and (ii) ligate directly with a Cys-peptide through an in situ formed thioester.[9−11] Under alkaline conditions, it remains in its amide form with relatively low reactivity toward a Cys-peptide. We have also shown that BMEA-mediated ligation tolerates most amino acids (except the much hindered β-branched amino acids) at the C-terminal position.[9] Based on these observations, we realize that, when a peptide thioester reacts with a Cys-peptidyl BMEA at neutral or basic pH, conventional thioester-Cys ligation will be dominant and the BMEA moiety will remain intact during the reaction. Therefore, we propose a novel three-segmental N-to-C sequential ligation strategy. As shown in Scheme 1, the first step is the ligation reaction between the N-terminal thioester peptide and the middle Cys-peptidyl BMEA segment. The second step is the BMEA mediated ligation between the product of the first step and the C-terminal Cys-peptide.

To test our proposal, we first synthesized a 46-residue model peptide through ligating three short peptides. Peptide **2**, the middle Cys-peptidyl BMEA segment, was synthesized as previously reported.[9] For the first ligation step,

Scheme 1. N-to-C Sequential Ligation Scheme Using the Combination of NCL and BMEA Mediated Ligation



1 and **2** were reacted under normal conditions. 15 mM of **1** and 5 mM of **2** were dissolved in ligation buffer containing 6 M guanidine hydrochloride (Gdn•HCl), 0.2 M phosphate, 50 mM tris(2-carboxyethyl)phosphine (TCEP), and 2% v/v thiophenol, pH 7.0. The ligation reaction was monitored with HPLC. After 4 h at room temperature, the reaction was completed with only minor side reactions (Figure 1A). The ligation product was purified and subjected to the next ligation step with peptide **3**. The second ligation was performed under the optimal conditions for BMEA-mediated ligation. The above ligation product (5 mM) was reacted with peptide **3** (15 mM) in a buffer containing 6 M Gdn•HCl, 0.2 M NaOAc, 50 mM TCEP, and 0.2 M sodium 2-mercaptoethanesulfonate (MESNa), pH 5.0. The ligation was completed under microwave irradiation within 15 h (Figure 1C) with a $>90\%$ yield based on HPLC analysis. Therefore, this model study demonstrated the feasibility of our N-to-C sequential ligation strategy.

Next, to test the general utility of our strategy, we applied this approach to the synthesis of a small protein, ubiquitin. Ubiquitin is a highly conserved protein with 76 amino acids and can be linked to the lysine side chain of another ubiquitin or other proteins in a process called ubiquitination. The assembly of the ubiquitin polypeptide

(7) (a) Bang, D.; Pentelute, B. L.; Kent, S. B. H. *Angew. Chem., Int. Ed.* **2006**, *45*, 3985. (b) Zheng, J. S.; Cui, H.- K.; Fang, G.-M.; Xi, W.-X.; Liu, L. *ChemBioChem* **2010**, *11*, 511.

(8) (a) Kawakami, T.; Aimoto, S. *Tetrahedron Lett.* **2007**, *48*, 1903. (b) Shigenaga, A.; Sumikawa, Y.; Tsuda, S.; Sato, K.; Otaka, A. *Tetrahedron* **2010**, *66*, 3290. (c) Sato, K.; Shigenaga, A.; Tsuji, K.; Tsuda, S.; Sumikawa, Y.; Sakamoto, K.; Otaka, A. *ChemBioChem* **2011**, *12*, 1840. (d) Ding, H.; Shigenaga, A.; Sato, K.; Morishita, K.; Otaka, A. *Org. Lett.* **2011**, *13*, 5588.

(9) (a) Hou, W.; Yang, R.; Li, F.; Zhang, X.; Liu, C.-F. In *Peptides: Chemistry, Biology and Pharmaceutical Sciences. Proceedings of the 11th Chinese International Peptide Symposium (CPS-2010)*; Lanzhou, China, July 5−8, 2010; Wang, R., Wang, K.-R., Eds.; Chemical Industry Press: Beijing; pp 93−96. (b) Hou, W.; Zhang, X.; Li, F.; Liu, C.-F. *Org. Lett.* **2011**, *13*, 386.

(10) (a) Ollivier, N.; Dheur, J.; Mhidia, R.; Blanpain, A.; Melnyk, O. *Org. Lett.* **2010**, *12*, 5238. (b) Dheur, J.; Ollivier, N.; Vallin, A.; Melnyk, O. *J. Org. Chem.* **2011**, *76*, 3194.

(11) For an excellent review on the use of N-to-S acyl transfer in thioester synthesis, see: Kang, J.; Macmillan, D. *Org. Biomol. Chem.* **2010**, *8*, 1993.

(12) Bang, D.; Makhatadze, G. I.; Tereshko, V.; Kossiakoff, A. A.; Kent, S. B. H. *Angew. Chem., Int. Ed.* **2005**, *44*, 3852.

Peptide **1**: H-**ADKRAHHNALERKRRDHA**-$S(CH_2)_2CONH_2$
Peptide **2**: H-**CDSFHSLRDSY**-$N(CH_2CH_2SH)_2$
Peptide **3**: H-**CLKPLHEKDSES**$(_P)$**GGGKD**-NH_2
S(p): phosphoserine

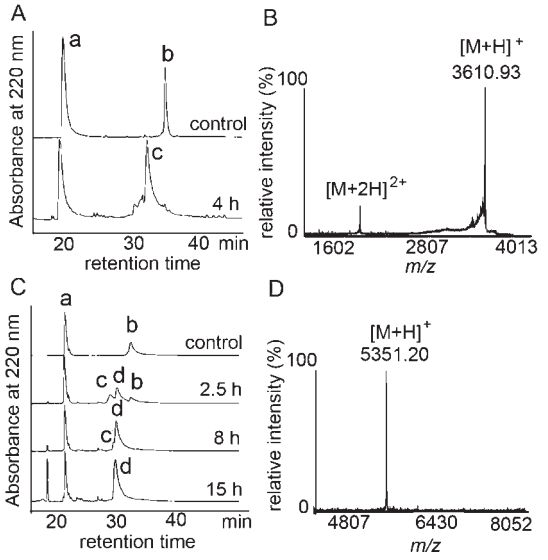


Figure 1. Model study of the N-to-C sequential ligation strategy. (A) C18 analytical HPLC analysis of the ligation reaction between peptide **1** and **2**. HPLC conditions: 0% to 40% of buffer B in buffer A in 40 min. Peak a: peptide **1**; peak b: peptide **2**; peak c: ligation product of **1** and **2**, H-**ADKRAHHNALERKRRDHACDSFHSLRDSY**-$N(CH_2CH_2SH)_2$ or H-$\mathbf{AY_{29}}$-BMEA. (B) The MALDI-TOF MS of H-$\mathbf{AY_{29}}$-BMEA. $[M + H]^+$ found: 3610.93, molecular mass calcd: 3609.71. (C) C18 analytical HPLC analysis of the ligation reaction between H-$\mathbf{AY_{29}}$-BMEA and peptide **3**. HPLC conditions: 0% to 50% of buffer B in buffer A in 50 min. Peak a: peptide **3**; peak b: H-$\mathbf{AY_{29}}$-BMEA; peak c: H-$\mathbf{AY_{29}}$-MES; peak d: ligation product, H-**ADKRAHHNALERKRRDHA CDSFHSLRDSYCLKPLHEKDSES**$(_P)$**GGGKD**-NH_2 or H-$\mathbf{AD_{46}}$-NH_2. (D) The MALDI-TOF MS of H-$\mathbf{AD_{46}}$-NH_2. $[M + H]^+$ found: 5351.20, molecular mass calcd: 5349.50.

Peptide **4**: H-**LQIFVKTLTGKTITLEVEPSDTIENVK**-$S(CH_2)_2CONH_2$
Peptide **5**: H-**CKIQDKEGIPPDQQRLIF**-$N(CH_2CH_2SH)_2$
Peptide **6**: H-**CGKQLEDGRTLSDYNIQKESTLHLVL-RLRGG**-OH

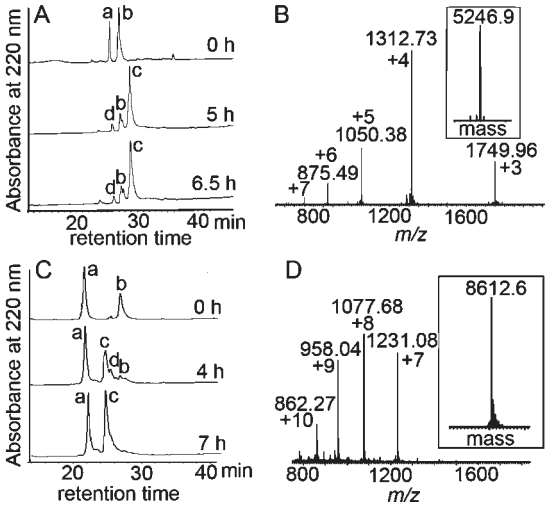


Figure 2. Synthesis of ubiquitin using N-to-C sequential ligation. (A) C18 analytical HPLC analysis of the ligation reaction between peptide **4** and **5**. HPLC conditions: 0% to 80% of buffer B in buffer A in 40 min. Peak a: peptide **4**; peak b: peptide **5**; peak c: ligation product of **4** and **5**, H-$\mathbf{LF_{45}}$-BMEA; peak d: self-cyclization (with the C-ter Lys side-chain amine) and hydrolysis product of **4**. (B) The raw and deconvoluted ESI-MS of H-$\mathbf{LF_{45}}$-BMEA. Molecular mass found: 5246.9, calcd: 5247.16. (C) C18 analytical HPLC analysis of the ligation reaction between H-$\mathbf{LF_{45}}$-BMEA and peptide **6** with methyl mercaptoacetate as thiol additive. HPLC conditions: 0% to 80% of buffer B in buffer A in 40 min. Peak a: peptide **6**; peak b: H- $\mathbf{LF_{45}}$-BMEA; peak c: ligation product; peak d: H-$\mathbf{LF_{45}}$-SCH_2COOMe. (D) The raw and deconvoluted ESI-MS of full length ubiquitin. Molecular mass found: 8612.6, calcd: 8610.8.

chain has previously been reported through sequential ligation using other established methods.[12,13]

To synthesize ubiquitin with our BMEA-enabled N-to-C sequential ligation strategy, three segements were synthesized by SPPS. Therefore, the middle segment, peptide **5**, was synthesized possessing C-terminal BMEA on a 2-chlorotrityl chloride resin using Fmoc chemistry.[9] For the first thioester-mediated ligation step, 3.5 mg of the N-terminal peptide thioester **4** (final concentration: 3.8 mM) and 2.3 mg of the middle segment **5** (3.4 mM) were dissolved in 300 μL of ligation buffer containing 6 M Gdn•HCl, 0.2 M phosphate, 20 mM TCEP, and 0.2 M MESNa, pH 8.0. As shown in Figure 2A, the ligation was completed within 5 h at room temperature. No BMEA-originated side reactions were observed. Further incubation for 1.5 h did not change the HPLC profile in any significant way, indicating that the BMEA moeity was stable under these conditions. After purification and lyophilization, 2.5 mg of ligation product H-$\mathbf{LF_{45}}$-BMEA were obtained (isolated yield 47%).

For the second step of ligation, we had two options (Scheme 1). The first option is to convert the peptidyl BMEA to thioester and perform ligation between the isolated thioester and the C-terminal Cys-peptide **6**. The second option is to let the peptidyl BMEA react directly with peptide **6** via the in situ generated thioester. We first explored the first option. To convert H-$\mathbf{LF_{45}}$-BMEA to thioester, about 3 mg of H-$\mathbf{LF_{45}}$-BMEA were dissolved in 300 μL of buffer (containing 6 M Gdn•HCl, 0.2 M NaOAc, 0.2 M MESNa and 40 mM TCEP, pH 5.0) and irradiated with a low-power microwave. After 10 h, the majority of the BMEA was converted to the thioester form with a small amount of hydrolysis product (Supporting Information, Figure S7). 1.5 mg of H-$\mathbf{LF_{45}}$-MES was isolated after HPLC purification (isolated yield 50%). Next, H-$\mathbf{LF_{45}}$-MES

(13) Yang, R.; Pasunooti, K. K.; Li, F.; Liu, X.-W.; Liu, C.-F. *Chem. Commun.* **2010**, *46*, 7199.

was reacted with peptide **6** under normal ligation conditions. 1.5 mg of H-**LF$_{45}$**-MES and 1.5 mg of **6** were dissolved in 200 μL of NCL buffer (6 M Gdn•HCl, 0.2 M phosphate, 20 mM TCEP, and 0.2 M MESNa, pH 8.0). The reaction was completed within 6.5 h (Supporting Information, Figure S8). 1.4 mg of ligation product was isolated (isolated yield: 57%).

For the second option, direct ligation between H-**LF$_{45}$**-BMEA and **6** was performed under microwave irradiation at pH 5.0. We first used MESNa as the thiol additive and found that when 1.5 to 2 equiv of **6** relative to H-**LF$_{45}$**-BMEA were used, the ligation yield was modest. After 5 h of irradiation, all H-**LF$_{45}$**-BMEA was transformed to H-**LF$_{45}$**-MES. But only 30−40% of the thioester reacted with **6** to form the ligation product. Continuing irradiation did not increase the yield (data not shown). The ligation reaction was more efficient when a larger excess (about 3 equiv) of **6** was used. After microwave irradiation for 10 h, more than 70% of H-**LF$_{45}$**-MES was converted to the ligation product (Supporting Information, Figure S9). Based on the above observations, we realized that MESNa might not be a good thiol additive for BMEA-mediated ligation when very long peptide segments were involved. It appeared that, although BMEA to thioester conversion was not a problem, the resultant MES thioester would not react efficiently with the cysteinyl peptide at the weakly acidic pH. This is not unexpected considering that, at pH 5, the alkyl thiol of an N-Cys is not appreciably deprotonated for efficient reaction with an MES thioester which is not very reactive itself because its negatively charged sulfonate moiety reduces the electrophilicity of the thioester carbonyl. We reasoned that methyl mercaptoacetate might be a good thiol additive as the presence of an electron-withdrawing ester group near the thiol would make the resultant thioester much more reactive toward a Cys-peptide even at weakly acidic pH. To our delight, this alkyl thiol was indeed an excellent thiol additive in our reaction system. So, when 1 mg of H-**LF$_{45}$**-BMEA (1.9 mM) and 1.4 mg of **6** (4 mM) were mixed in 100 μL of buffer (6 M Gdn•HCl, 0.2 M NaOAc, 2% v/v methyl mercaptoacetate and 40 mM TCEP, pH 5.0), the ligation reaction was completed after microwave irradiation for 7 h (Figure 2C). Almost all of the H-**LF$_{45}$**-BMEA was consumed and ligated with peptide **6**. Compared to MESNa, the use of methyl mercaptoacetate as the thiol additive gave a much more efficient ligation reaction.

To generate the native sequence of ubiquitin, free radical mediated desulfurization[14] was performed on the full-length ubiquitin synthesized by our N-to-C sequential ligation approach, which would convert the two cysteine residues at the ligation junctions to Ala residues. Desulfurization was performed with VA-044 as a free radical initiator[14] and glutathione as the hydride source.[15] The desulfurization process was monitored with analytical HPLC and ESI-MS and was completed within 8 h (Supporting Information, Figure S10). The end product was purified by C18 semipreparative HPLC. 1.1 mg of the final product was obtained from 1.5 mg of the starting material (73% isolated yield). Refolding of the lyophilized material was performed through dialysis to give the native folded protein as confirmed by CD analysis (Supporting Information, Figure S11).

In summary, we have demonstrated the utility of a new N-to-C sequential ligation strategy which is based on the use of BMEA as a latent thioester. We have shown that the BMEA moiety remains stable under the conventional conditions of NCL. During the review process of this manuscript, a similar sequential ligation scheme from the Melnyk lab was published in which the BMEA moiety was blocked as a preformed disulfide during the first NCL step.[16] Our work shows that it is unnecessary to protect BMEA during the NCL step because of the very large differences in reaction conditions and kinetics between a conventional thioester and BMEA. For the key BMEA mediated ligation step, we have explored methods of ligation with an isolated thioester or direct ligation via an in situ generated thioester for which a new alkyl thiol, methyl mercaptoacetate, has been found to be an effective additive. In principle, this N-to-C sequential ligation can be performed for infinite steps if the BMEA is converted to an isolated thioester at every intermediate step. The BMEA ligation methodology works well not only with small peptides but also with large peptide segments. With this and the fact that peptide BMEA segments can be readily synthesized by Fmoc chemistry, we believe that our newly developed N-to-C sequential ligation strategy has great value in protein synthesis.

Acknowledgment. We thank A*Star of Singapore (BMRC 08/1/22/19/588) and Nanyang Technological University for financial support.

Supporting Information Available. Experimental procedures, HPLC and MS data. This material is available free of charge via the Internet at http://pubs.acs.org.

(14) Wan, Q.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 9248.

(15) Haase, C.; Rohde, H.; Seitz, O. *Angew. Chem., Int. Ed.* **2008**, *47*, 6807.

(16) Ollivier, N.; Vicogne, J.; Vallin, A.; Drobecq, H.; Desmet, R.; El Mahdi, O.; Leclercq, B.; Goormachtigh, G.; Fafeur, V.; Melnyk, O. DOI: 10.1002/anie.201105837.