Protein Chemistry

DOI: 10.1002/ange.201309396

Click-Tag and Amine-Tag: Chemical Tag Approaches for Efficient Protein Labeling In Vitro and on Live Cells using the Naturally Split *Npu* DnaE Intein**

*Vivien Schütz and Henning D. Mootz**

Abstract: *Protein labeling with synthetic moieties remains in many cases a technically challenging or unresolved task. Two new and simple concepts are presented. In both approaches, a very short tag of only a few amino acids is prepared with the desired chemical modification and, in a second step, it is transferred to the protein of interest by protein trans-splicing. For the amine-tag, a recombinant intein fragment free of lysine residues was generated such that the amine group of the N terminus could be selectively modified with regular amine-reactive reagents. Thus, standard bioconjugation procedures without any chemical synthesis could be applied without modification of lysines in the protein of interest. For the click-tag, protein trans-splicing was combined with unnatural amino acid mutagenesis and subsequent bioorthogonal side chain modification, as demonstrated for click chemistry using p-azidophenylalanine. By the two-step strategy, exposure of the protein of interest to the copper catalyst was avoided.*

The selective covalent modification of proteins with synthetic moieties is a long-standing goal in basic research as well as in biotechnology and biomedicine. Applications include the introduction of biophysical probes to study protein function or protein localization in live cells and the preparation of protein diagnostics and therapeutics, for example.[1] Classical bioconjugation approaches are often limited by the presence of more than one of the suitable residues, for example, cysteine and lysine, in the sample.[1a] The incorporation of unnatural amino acids with bioorthogonal reactivity[2] and in vitro ligations, such as native chemical ligation,[3] can circumvent this problem, but may have limitations in terms of expression and ligation conditions. The use of genetically encoded fusion tags[4] to achieve selectivity has become very popular since the initial report of the cysteine-rich FLASH-tag.[4a] However, the current systems require either self-modifying enzymes of considerable size,[1b] for example SNAP/CLIP-tag,[4b] HaloTag,[4c] DHFR-tag,[4d] and BL-tag,[4e] or an extra enzyme that complicates the conjugation procedure,[1b] for example tags for sortase,[4f,g] biotin-ligase,[4h] lipoic acid ligase,[4i] and 4′-phosphopantetheine-transferase.[4j]

The transfer of a very short tag by split inteins can be a powerful alternative to the aforementioned techniques (Figure 1).[5] Protein *trans*-splicing (PTS) occurs in a self-processive fashion, that is, no additional factors are required, and with a concomitant removal of the intein fragments.

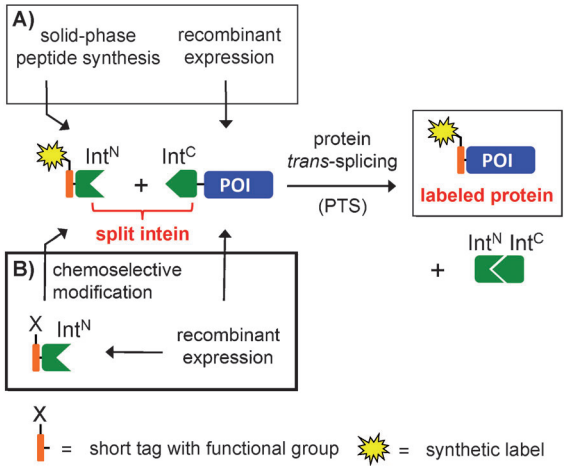

Figure 1. Protein labeling by PTS. Strategy A illustrates semi-synthetic PTS requiring peptide synthesis and split inteins with short fragments. Strategy B involves the chemical tag approaches described in this work in combination with the naturally split *Npu* DnaE intein (Int^N = 102 aa; Int^C = 36 aa). In this study, X represents the N-terminal amine group or the azide moiety of an unnatural amino acid. In principle, also the C-terminal end of the POI could be targeted by exchanging the intein N- and C-terminal fragments (Int^N and Int^C), except for the amine-tag approach.

Further advantages of PTS include the high reaction rate, the orthogonality to cellular chemistry, the low concentrations of the intein fragments, and the possibility to fully genetically encode them.[5] Naturally split inteins are rare, but they are considered to be superior over artificially split inteins, in particular in terms of solubility, association kinetics, and fragment affinity.[5] Previously, two distinct strategies were reported to exploit PTS for protein labeling. Figure 1 illustrates N-terminal protein labeling, but both strategies can also be used to modify the complementary intein fragment and achieve C-terminal labeling. First, in so-called semi-synthetic PTS,[6] one of the split intein fragments is prepared by solid-phase peptide synthesis (SPPS) and thereby incorporation of synthetic groups into the tag can be achieved (strategy A in Figure 1). However, for naturally occurring

[*] Dr. V. Schütz, Prof. Dr. H. D. Mootz
Institute of Biochemistry, University of Muenster
Wilhelm-Klemm-Strasse 2, 48149 Münster (Germany)
E-mail: Henning.Mootz@uni-muenster.de

[**] We thank Patrick Helmer for technical assistance, Tulika Dhar for discussions, and Peter G. Schultz for providing plasmids. Funding from the DFG (SPP1623, MO1073/5-1; and the Cells-in-Motion excellence cluster, EXC1003) is gratefully acknowledged.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201309396.

split inteins, only the C-terminal fragments (Int^C) are within the reasonable reach of SPPS, although they exhibit already considerable sizes of about 35 aa (Int^N fragments are ca. 100 aa),[6a,7] thus restricting labeling to the C terminus of the protein of interest (POI). Shorter fragments were reported for artificially split inteins useful for N- and C-terminal labeling,[6b–f] but these are generally less robust in the PTS reaction, show lower affinity interactions and lower solubility compared to the naturally split inteins. The second strategy, also referred to as Cys-tag labeling (strategy B in Figure 1),[8] exploits recombinantly expressed intein fragments devoid of cysteines to allow for selective bioconjugation at a unique cysteine in the short tag to be transferred to the POI by PTS. By this two-step strategy, the plethora of commercially available thiol-reactive reagents can be applied to proteins containing additional cysteine residues, and no synthetic operations are required. Cys-tag labeling was established with artificially split inteins for C- and N-terminal labeling.[8] However, N-terminal labeling was very inefficient because of the poorly behaved *Psp*-GBD Pol intein.[8c] Only very few naturally split inteins with cysteine-less fragments and that are suitable for handling at ambient temperatures are known, and these are also limited to potential C-terminal labeling.[9] Successful mutation of the catalytic Cys1 to serine in the Int^N fragment of naturally split inteins has not yet been reported, thus precluding this interesting combination for N-terminal labeling by the Cys-tag approach.

The aim of this study was to combine a chemical tag strategy for N-terminal protein labeling with the naturally split *Npu* DnaE intein, which is well-characterized for its superior performance with regard to rates, yields, and fragment affinity in the low nanomolar range.[10] However, also this intein has an essential cysteine at the N-terminal splice junction, that could not be mutated to serine;[10b] therefore the Cys-tag approach could not be applied. We therefore devised two new approaches, termed amine-tag and click-tag approaches.

For the amine-tag approach, all lysine residues (Lys20, 24, 70, and 73) in the 102 aa Int^N fragment of the *Npu* DnaE intein (hereafter only referred to as Int^N fragment) were mutated to arginine to allow for regioselective labeling of the amino terminus in a tag-$Int^N(\Delta K)$-H_6 construct with amine-reactive probes. The impact of the mutations on PTS activity was first evaluated with fused globular proteins using model constructs gpD-Int^N-H_6 (constructs **1a** and **1b**; gpD = phage coat protein) and GFP-Int^N-H_6 (**2a** and **2b**) with either wild-type Int^N or $Int^N(\Delta K)$ fragments, respectively (see the Supporting Information, Table S1 for all of the constructs). These were incubated with Int^C-GFP (**3**) and Int^C-Trx (**4**; Trx = thioredoxin). Lysine removal led to only about 3- to 4-fold lower rates and slightly reduced product yields of 55–75% after 24 h, depending on the protein (Figure 2A;

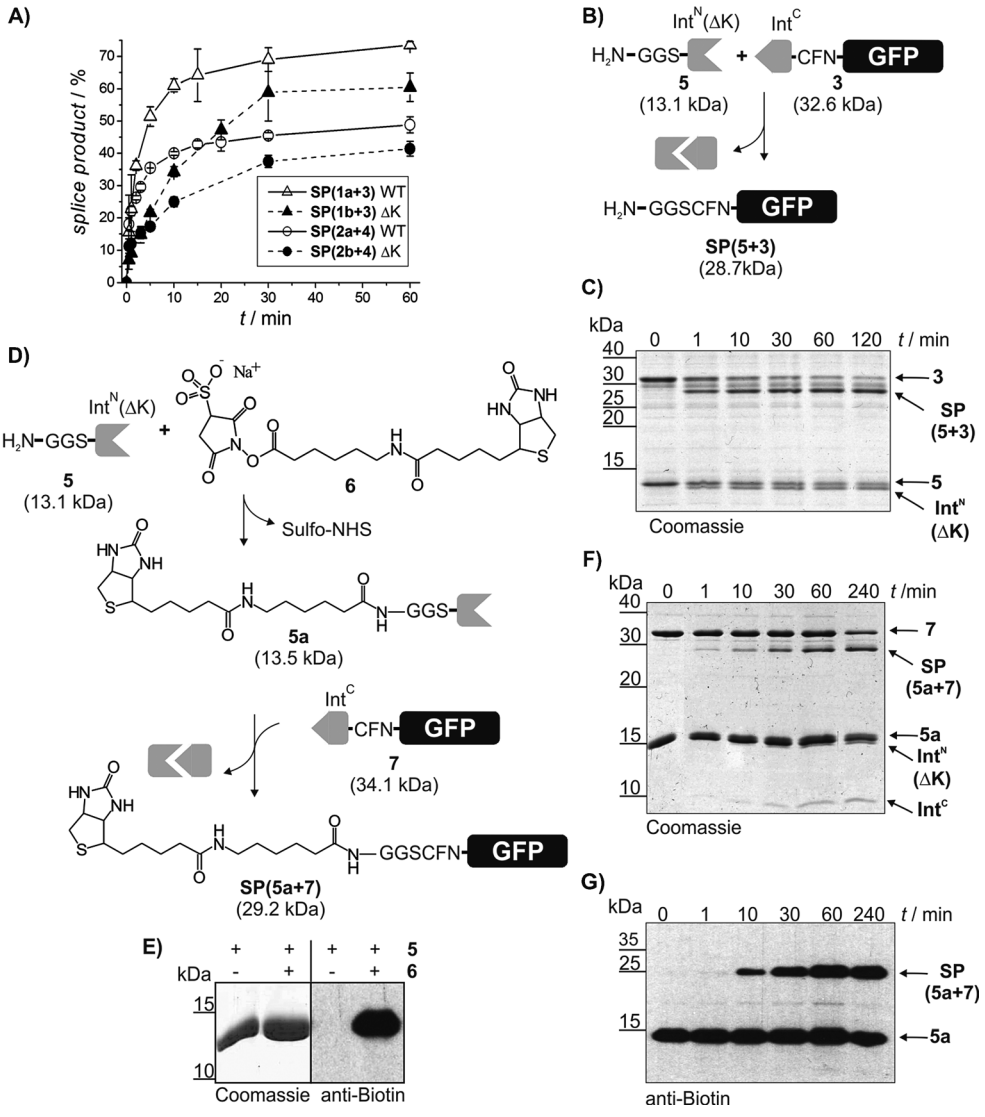

Figure 2. Amine-tag approach for protein labeling in vitro. Shown are SDS-PAGE gels. A) Comparison of PTS kinetics. B), C) PTS with a short N-terminal tag (NH_2-GGS-). The three flanking aa on either side of the intein are indicated. D) Tag modification with amine-directed active ester reagent and subsequent PTS. E) Protein **5** analyzed before and after the acylation reaction shown in (D). F), G) Monitoring of the PTS reaction shown in (D).

Supporting Information, Figure S1, Table S2). Next, a short tag of only four amino acids (MGGS) was placed in front of the Int^N. Analysis by mass spectrometry revealed that the starting methionine was virtually completely removed (Supporting Information, Figure S2), as expected,[11] leaving GGS-$Int^N(\Delta K)$-H_6 (**5**) with a tag of only three residues. To our knowledge, only one other intein has been reported to splice with such a short N-terminal tag.[12] Robust splicing activity was retained (Figure 2B,C). The N-terminal amino group was then modified using sulfo-NHS-biotin (**6**), a typical amine-reactive reagent, to give biotin-GGS-$Int^N(\Delta K)$-H_6 (**5a**) (Figure 2D,E). Importantly, upon mixing of **5a** (15 μM) with a StreptagII-Int^C-GFP-H_6 construct (**7**; 15 μM) the labeled tag was transferred onto the model protein GFP to give labeled splice product biotin-GFP (**SP(5a+7)**) (Figure 2F,G). Together, these results demonstrate the feasibility of the amine-tag concept.

For the click-tag approach, we incorporated the unnatural amino acid *p*-azidophenylalanine (AzF) into the short tag by using the suppression technology in *E. coli*.[13] An amber stop codon was inserted into the expression plasmid to encode for G(AzF)GS-Int^N-H_6 (**8**) with the wild-type Int^N fragment. AzF was added to the growth medium during expression. The purified protein **8** retained splicing activity (Supporting Information, Figures S3,S4). Chemical modification of the tag was achieved by the copper(I)-catalyzed alkyne–azide cycloaddition (CuAAC)[14] using $CuSO_4$, TBTA, and TCEP with dansyl-alkyne **9** (Figure 3A,B). Excess conjugation reagents were removed by dialysis in the presence of EDTA. Figure 3C shows that the resulting construct G(AzF-dansyl)GS-Int^N-H_6 (**8a**) was active in PTS, as demonstrated by the transfer of the modified tag to the model protein GFP, upon incubation with the complementary Int^C-construct **7**, with excellent rate and nearly quantitative conversion. One advantage of the click-tag approach over direct incorporation of the unnatural amino acid into the POI is that exposure of the POI to the potentially harmful CuAAC reaction conditions, in particular the presence of copper ions and the addition of metal chelators to remove them, is circumvented. Thereby, the scope of the CuAAC reaction could be extended, for example to thiol-rich or metal-containing proteins and to proteins in live cells (see below). Note that it would also be conceivable to first perform the PTS reaction and then carry out the click reaction.

Finally, we tested the applicability of the amine-tag and the click-tag approaches to protein labeling on a cell surface (Figure 4A).[15] Importantly, the high affinity of the two intein

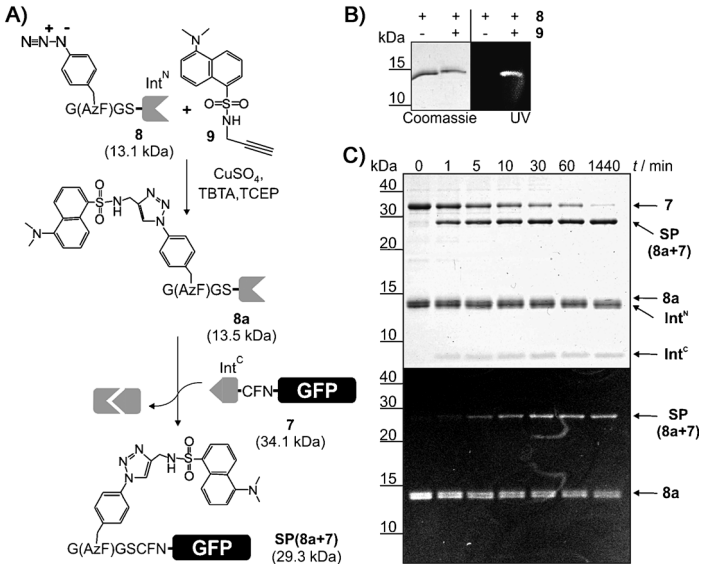

Figure 3. Click-tag approach for protein labeling in vitro. A) Tag modification with click chemistry and subsequent PTS. B) Protein **8** analyzed before and after the click reaction shown in (A). C) PTS-mediated tag-transfer to GFP as the model POI monitored by SDS-PAGE using Coomassie staining (top) and UV illumination (bottom).

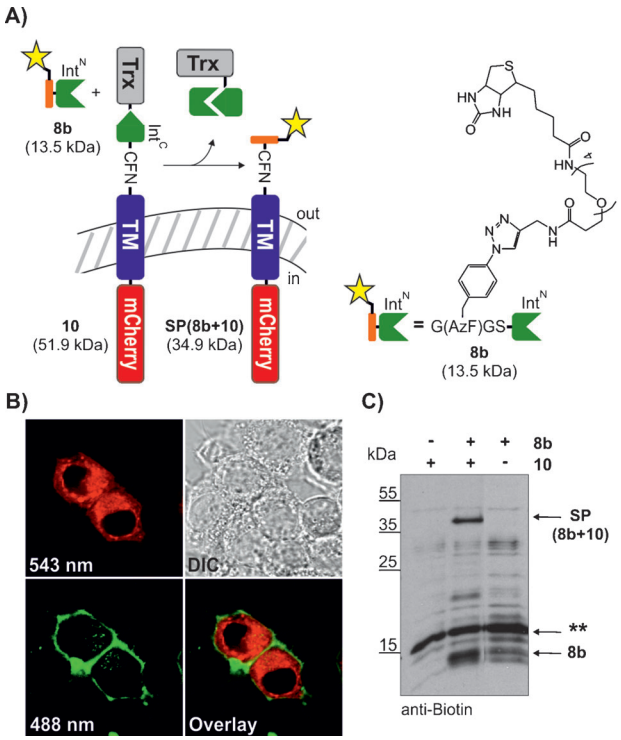

Figure 4. Chemical labeling on living cells. A) The PTS reaction. B) Confocal microscopy images of selected transfected and untransfected cells. After the PTS reaction (60 min), cells were incubated with streptavidin-Alexa488. Note that only transfected cells give rise to staining with the synthetic fluorophore underlining the selectivity of PTS. C) Western blot analysis of the PTS reaction using total cell samples. (TM = transmembrane domain, Trx = thioredoxin).

fragments and the high reaction rate of this intein can translate into low reactant concentrations and short labeling times,[15] conditions that are beneficial for live cell microscopy to study dynamic processes. Short tags and synthetic probes are of particular interest for these kind of investigations.[1b] However, our previous work with the *Npu* DnaE intein has only shown the extracellular, N-terminal labeling of membrane-anchored proteins with the large recombinant fluorescent protein GFP as a marker.[15] We used the mouse neuro 2a (N2a) cell line for transient expression and the transmembrane domain (TM) of the platelet-derived growth-factor receptor (PDGF-R) to localize a model protein in the plasma

membrane (HA-Trx-Int^C-Myc-TM-mCherry (**10**); HA = hemagglutinin tag, Myc = Myc-epitope). mCherry served to monitor protein localization in the red channel. Trx improved the levels of correctly localized protein,[15] probably through stabilization of the largely unfolded Int^C fragment[16] on the pathway through the endoplasmatic reticulum. While in previous work Trx remained in the splice product,[15] here we improved the arrangement by fusing it N-terminally to the Int^C fragment to effect its removal during the PTS reaction (Figure 4A). Thus, virtually traceless labeling of the membrane protein was enabled (Supporting Information, Figure S5). 24 h after transfection, the click-tag protein biotin-tag-Int^N-H$_6$ (**8b**) was added to the growth medium (at a final concentration of 5 μM) for 1 h at 37 °C, followed by washing of the cells. Figure 4B shows that efficient labeling at the plasma membrane of transfected cells occurred. Untransfected cells were not stained indicating that excess **8b** could be efficiently removed by washing. Western blot analysis confirmed the formation of the desired splice product biotin-tag-Myc-TM-mCherry (**SP**(**8b**+**10**); Figure 4C). Comparable results were obtained using fluorescein-labeled amine-tag protein **5b** (Supporting Information, Figure S6), indicating that the lysine-free Int^N(ΔK) fragment, although being somewhat compromised compared to Int^N as shown above, still exhibited sufficient efficiency. Together, these results show that a transmembrane protein could be modified with the short chemical tags.

In conclusion, we present two new approaches for the selective chemical modification of purified proteins and proteins on live cells using very short, pre-labeled chemical tags that are transferred by a split intein. The tags are not restricted to a particular sequence and could also be part of the POI sequence. Both approaches require only standard molecular biology techniques and a minimum of steps with handling of chemicals such as usually associated with classical bioconjugation procedures. With the amine-tag approach, the entire portfolio of amine-reactive compounds can be used without the traditional problem of also labeling internal lysine side chains of the POI. The click-tag approach provides even more freedom in the bioconjugation strategy as it could also be devised with other unnatural amino acids and bioorthogonal reactions or for C-terminal protein labeling. Both approaches could be easily extended for selective dual labeling of a single protein, for example by incorporating an additional unnatural amino acid in the POI sequence of the Int^C-POI construct. Thus, they should provide powerful alternatives to address specific problems in protein chemical modification labeling in vitro and in vivo.

Experimental Section

Comprehensive experimental procedures are given in the Supporting Information. Chemical tag conjugation: Protein **5** was purified by Ni-NTA chromatography under denaturing conditions (8 M urea), dialyzed against ddH_2O, flash-frozen in liquid N_2, and lyophilized. 1 mg mL^{-1} protein was solved in borate buffer (pH 8.5) and 24-fold molar excess of NHS-fluorescein in DMF or 5-fold molar excess Sulfo-NHS-LC-Biotin in PBS buffer were added at 4 °C. The reaction was quenched by adding 50 mM NH_4Cl and the protein was purified and dialyzed against PBS buffer (136 mM NaCl, 2.7 mM KCl, 8.2 mM Na_2HPO_4 1.5 mM KH_2PO_4; pH 7.4). Protein **8** (20–100 μM in PBS buffer) was reacted under CuAAC conditions by adding one of the alkyne reagents (100 μM in DMF), 1 mM TCEP, and a mixture of TBTA (100 μM) in DMSO/*t*BuOH (1:4) and 1 mM $CuSO_4$ in H_2O. After 30 min at room temperature, the reaction was quenched with EDTA (10 mM). The protein was purified and dialyzed against PBS buffer. Protein *trans*-splicing: In vitro reactions with labeled chemical tags were carried out at 25 °C in PBS buffer with 2 mM DTT using purified proteins (10–20 μM). Reactions on cells were performed for 1 h at 37 °C in serum-free DMEM without DTT 24 h after transient transfection using proteins **8b** or **5b** (5 μM). Cells were washed three times with PBS buffer before further analysis.[17]

Received: October 28, 2013
Revised: December 16, 2013
Published online: March 11, 2014

Keywords: bioconjugation · bioorthogonality · fusion tags · lysine modification · unnatural amino acids

[1] a) G. T. Hermanson, *Bioconjugate Techniques*, Academic Press, New York, **2008**; b) M. J. Hinner, K. Johnsson, *Curr. Opin. Biotechnol.* **2010**, *21*, 766.
[2] C. C. Liu, P. G. Schultz, *Annu. Rev. Biochem.* **2010**, *79*, 413.
[3] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776; b) T. W. Muir, *Annu. Rev. Biochem.* **2003**, *72*, 249.
[4] a) B. A. Griffin, S. R. Adams, R. Y. Tsien, *Science* **1998**, *281*, 269; b) A. Gautier, A. Juillerat, C. Heinis, I. R. Correa, Jr., M. Kindermann, F. Beaufils, K. Johnsson, *Chem. Biol.* **2008**, *15*, 128; c) G. V. Los, L. P. Encell, M. G. McDougall, D. D. Hartzell, N. Karassina, C. Zimprich, M. G. Wood, R. Learish, R. F. Ohana, M. Urh, D. Simpson, J. Mendez, K. Zimmerman, P. Otto, G. Vidugiris, J. Zhu, A. Darzins, D. H. Klaubert, R. F. Bulleit, K. V. Wood, *ACS Chem. Biol.* **2008**, *3*, 373; d) Z. Chen, C. Jing, S. S. Gallagher, M. P. Sheetz, V. W. Cornish, *J. Am. Chem. Soc.* **2012**, *134*, 13692; e) S. Mizukami, S. Watanabe, Y. Hori, K. Kikuchi, *J. Am. Chem. Soc.* **2009**, *131*, 5016; f) H. Mao, S. A. Hart, A. Schink, B. A. Pollok, *J. Am. Chem. Soc.* **2004**, *126*, 2670; g) M. W. Popp, J. M. Antos, G. M. Grotenbreg, E. Spooner, H. L. Ploegh, *Nat. Chem. Biol.* **2007**, *3*, 707; h) I. Chen, M. Howarth, W. Lin, A. Y. Ting, *Nat. Methods* **2005**, *2*, 99; i) C. Uttamapinant, M. I. Sanchez, D. S. Liu, J. Z. Yao, A. Y. Ting, *Nat. Protoc.* **2013**, *8*, 1620; j) J. Yin, P. D. Straight, S. M. McLoughlin, Z. Zhou, A. J. Lin, D. E. Golan, N. L. Kelleher, R. Kolter, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15815.
[5] a) H. D. Mootz, *ChemBioChem* **2009**, *10*, 2579; b) N. H. Shah, T. W. Muir, *Isr. J. Chem.* **2011**, *51*, 854.
[6] a) I. Giriat, T. W. Muir, *J. Am. Chem. Soc.* **2003**, *125*, 7180; b) C. Ludwig, M. Pfeiff, U. Linne, H. D. Mootz, *Angew. Chem.* **2006**, *118*, 5343; *Angew. Chem. Int. Ed.* **2006**, *45*, 5218; c) J. H. Appleby, K. Zhou, G. Volkmann, X. Q. Liu, *J. Biol. Chem.* **2009**, *284*, 6194; d) A. S. Aranko, S. Zuger, E. Buchinger, H. Iwai, *PLoS ONE* **2009**, *4*, e5185; e) J. H. Appleby-Tagoe, I. V. Thiel, Y. Wang, Y. Wang, H. D. Mootz, X. Q. Liu, *J. Biol. Chem.* **2011**, *286*, 34440; f) Y. T. Lee, T. H. Su, W. C. Lo, P. C. Lyu, S. C. Sue, *PLoS One* **2012**, *7*, e43820.
[7] a) N. K. Chu, D. Olschewski, R. Seidel, K. F. Winklhofer, J. Tatzelt, M. Engelhard, C. F. Becker, *J. Pept. Sci.* **2010**, *16*, 582; b) R. Borra, D. Dong, A. Y. Elnagar, G. A. Woldemariam, J. A. Camarero, *J. Am. Chem. Soc.* **2012**, *134*, 6344.
[8] a) T. Kurpiers, H. D. Mootz, *Angew. Chem.* **2007**, *119*, 5327; *Angew. Chem. Int. Ed.* **2007**, *46*, 5234; b) T. Kurpiers, H. D. Mootz, *ChemBioChem* **2008**, *9*, 2317; c) S. Brenzel, M. Cebi, P. Reiss, U. Koert, H. D. Mootz, *ChemBioChem* **2009**, *10*, 983.
[9] a) B. Dassa, N. London, B. L. Stoddard, O. Schueler-Furman, S. Pietrokovski, *Nucleic Acids Res.* **2009**, *37*, 2560; b) P. Carvajal-

Vallejos, R. Pallisse, H. D. Mootz, S. R. Schmidt, *J. Biol. Chem.* **2012**, *287*, 28686; c) J. J. Choi, K. H. Nam, B. Min, S. J. Kim, D. Soll, S. T. Kwon, *J. Mol. Biol.* **2006**, *356*, 1093.

[10] a) H. Iwai, S. Zuger, J. Jin, P. H. Tam, *FEBS Lett.* **2006**, *580*, 1853; b) J. Zettler, V. Schütz, H. D. Mootz, *FEBS Lett.* **2009**, *583*, 909; c) N. H. Shah, M. Vila-Perello, T. W. Muir, *Angew. Chem.* **2011**, *123*, 6641; *Angew. Chem. Int. Ed.* **2011**, *50*, 6511.

[11] W. T. Lowther, B. W. Matthews, *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.* **2000**, *1477*, 157.

[12] A. Wasmuth, C. Ludwig, H. D. Mootz, *Bioorg. Med. Chem.* **2013**, *21*, 3495.

[13] J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang, P. G. Schultz, *J. Am. Chem. Soc.* **2002**, *124*, 9026.

[14] a) C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057; b) V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2708; *Angew. Chem. Int. Ed.* **2002**, *41*, 2596.

[15] T. Dhar, H. D. Mootz, *Chem. Commun.* **2011**, *47*, 3063.

[16] Y. Zheng, Q. Wu, C. Wang, M. Q. Xu, Y. Liu, *Biosci. Rep.* **2012**, *32*, 433.

[17] Note added in proof (February, 26, 2014): While this paper was under review, the first atypical naturally split intein was reported, which exhibits a short N-terminal intein fragment of only 25 amino acids amenable to solid-phase peptide synthesis: I. V. Thiel, G. Volkmann, S. Pietrokovski, H. D. Mootz, *Angew. Chem.*, **2014**, *126*, 1330; *Angew. Chem. Int. Ed.*, **2014**, *53*, 1306.

