Protein Labeling

DOI: 10.1002/anie.201109252

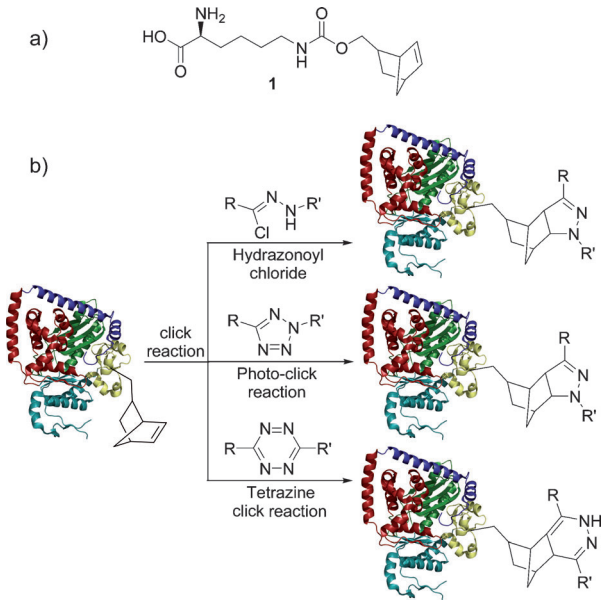
A Genetically Encoded Norbornene Amino Acid for the Mild and Selective Modification of Proteins in a Copper-Free Click Reaction**

*Emine Kaya, Milan Vrabel, Christian Deiml, Stefan Prill, Viviana S. Fluxa, and Thomas Carell**

Methods for the site-specific chemical modification of proteins are currently of immense importance for the synthesis of protein–hybrid compounds for pharmaceutical and diagnostic purposes.[1] Most of the methods rely on the reaction of free protein thiols with maleimides[2] or the reaction of lysine side chains with activated esters.[3] These methods provide only limited specificity, which is prompting researchers to develop alternative strategies that involve the incorporation of special unnatural amino acid into proteins to enable site-specific bioorthogonal functionalization.[4] Among the developed methods, the Cu^{I}-catalyzed reaction of a protein containing an alkyne amino acid with azides stands out as the most thoroughly investigated technology.[5,6] However, the need for Cu^{I} salts, which may harm the protein structure, limits the technology.[7] This fuels current interest to develop copper-free coupling reactions that are compatible with fragile protein structures.[8] Here we show that these requirments can be met with a specially encoded norbornene amino acid which reacts selectively with nitrile imines.

In order to insert a norbornene amino acid into a protein we used the amber suppression technique based on the pyrrolysyl tRNA/pyrrolysyl-tRNA synthetase ($tRNA^{Pyl}$/PylRS) pair from *Methanosarcina mazei*.[9] The main task of the project was to evolve the pyrrolysine synthetase so that it accepts the synthetic norbornene amino acid **1** (Scheme 1 a) for loading onto the pyrrolysyl-tRNA.

For the study we synthesized the norbornene-containing Pyl analogue **1** in seven steps from readily available starting materials (see the Supporting Information). To test to what extent **1** is accepted by wild-type (wt) PylRS, we used *E. coli* cells encoding the full $tRNA^{Pyl}$/PylRS pair and a modified yellow fluorescent protein (YFP) containing one in-frame TAG stop codon.[10] In this system the full-length and hence fluorescent YFP can only be produced when the corresponding Pyl analogue is accepted by the PylRS and successfully loaded onto the $tRNA^{Pyl}$ for subsequent incorporation into the protein at the amber stop codon site.



Scheme 1. a) Structure of norbornene Pyl analogue **1** and b) schematic representation of the click reactions. Top: A nitrile imine is generated by base-promoted HCl elimination from the hydrazonoyl chloride and then used in a cycloaddition reaction with the norbornene. Middle: Alternatively the nitrile imine is generated from a tetrazole in a photochemical reaction. Bottom: The norbornene modification can also react with tetrazines in a reversed-electron-demand Diels–Alder reaction. (Protein representations generated from PDB 3IN5.)

The so-prepared *E. coli* cells were grown in a medium containing 5 mM **1**. In addition to the wild-type PylRS we also tested a PylRS mutant (Y384F) previously used by Yanagisawa and co-workers.[4f] These initial experiments provided a just faint fluorescence when the mutant PylRS(Y384F) was used. No flurescence and hence no full-length YFP was generated in the presence of wild-type PylRS. In order to increase the PylRS activity we evolved the protein using iterative saturation mutagenesis (ISM) developed by Reetz and co-workers.[11]

Based on the co-crystal structure of PylRS in complex with adenylated pyrrolysine (PDB 2Q7H) we selected five residues in the substrate-binding pocket of wt-PylRS for the experiments. After transformation of the plasmid-based PylRS library into *E. coli*, single colonies were grown in liquid cultures supplemented with **1**. The incorporation was monitored by means of the YFP fluorescence intensity of the cells. The most efficient PylRS variants were then sequenced and used in the next round of the saturation mutagenesis. In

[*] Dipl.-Ing. E. Kaya,[+] Dr. M. Vrabel,[+] M. Sc. C. Deiml, M. Sc. S. Prill, Dr. V. S. Fluxa, Prof. Dr. T. Carell
Department Chemie, Ludwig-Maximilians-Universität
Butenandtstrasse 5–13, 81377 Munich (Germany)
E-mail: thomas.carell@cup.uni-muenchen.de
Homepage: http://www.carellgroup.de

[+] These authors contributed equally to this work.

[**] We thank Prof. Dr. Manfred T. Reetz and Dr. Despina Bougioukou for their assistance with the site-directed mutagenesis experiments. We thank Bayer Pharma AG, the Excellence Cluster CiPS^{M}, and the Volkswagen Foundation for generous support. M.V. thanks the LMU Munich for a postdoctoral fellowship. V.S.F. thanks the SNSF for a postdoctoral fellowship.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201109252.

total, three cycles of ISM were performed with full saturation of all five positions. The saturation mutagenesis provided a norb-PylRS enzyme with the following mutations: Tyr384Phe, Tyr306Gly, and Ile405Arg. This protein accepted the amino acid **1** by a factor four better than the Tyr384Phe mutant protein.

To investigate whether the norbornene unit can be selectively functionalized within the protein, we created *E. coli* cells containing two plasmids. One plasmid encoded three copies of *M. mazei* $tRNA^{Pyl}$ and one copy of the norb-PylRS. The second plasmid encoded for a truncated and C-terminally His_6-tagged human Polymerase κ (*h*Polκ), which contained an in-frame amber codon at position 163. We chose *h*Polκ for the experiment because this protein is an important and fragile member of human translesion polymerases.[12] A method for the site-selective modification of this protein is highly desirable in order to study the dynamics of the DNA repair mechanism. The corresponding *E. coli* cells were grown in a standard medium in the presence of 2 mM **1**. After a three-step purification procedure (Ni-NTA affinity chromatography, heparin affinity chromatography, and Superdex 200 gel filtration) about 2 mg of the norbornene-containing *h*Polκ protein (norb-*h*Polκ) was typically isolated from 1 L cell culture. The site-specific incorporation of the norbornene amino acid **1** at position 163 of *h*Polκ was confirmed by high-resolution mass spectrometry (see the Supporting Information). No peptide fragment with a lysine at this position (as a result of a possible carbamate linker hydrolysis) was detected by MS, indicating high content of the norbornene modification.

For the following functionalization we tested various cycloaddition reactions including reactions with nitrile oxides.[13] However, in all cases reaction was also observed with the protein itself. More successful were experiments with nitrile imines, which are weak electrophiles known to react efficiently with strained alkenes. To investigate this reaction in more detail, we prepared the hydrazonoyl chloride **2** and generated the corresponding nitrile imine by a base-promoted HCl elimination, as reported originally by Huisgen and co-workers.[14] We discovered that the elimination of HCl proceeds already with sufficient yield in aqueous buffer at a moderate pH of 7.4, perfect for proteins. Since nitrile imines can also be generated in a photochemical reaction from tetrazoles, as pioneered by Lin and co-workers,[8h,15] we also prepared tetrazole **3** to investigate this possibility. Finally, norbornenes were shown to react also with tetrazines in a reversed-electron-demand Diels–Alder reaction.[8d,e] We explored this possibility as an alternative to the nitrile imine chemistry with compound **4**. For the protein functionalization, we added the reagents **2**, **3**, and **4** to a buffered solution of *h*Polκ (and to wild-type *h*Polκ in a control reaction) at room temperature for 1 hour in 10 mM Tris-HCl buffer, pH 7.4. In the case of the tetrazole **3**, the reaction was also irradiated with a handheld UV lamp at $\lambda = 302$ nm for 10 min to generate the nitrile imine (see Scheme 1). These are the conditions that were developed by Lin and co-workers to functionalize the maleimide-based norbornene-modified lysozyme.[8g] The in-gel fluorescence analysis of the SDS-PAGE clearly showed that the norbornene functionalization proceeded successfully in all cases (Figure 1; lanes 2, 4, and 6), while no reaction was observed with the wild-type *h*Polκ (Figure 1; lanes 1, 3, and 5). This proves the specificity of the reaction.

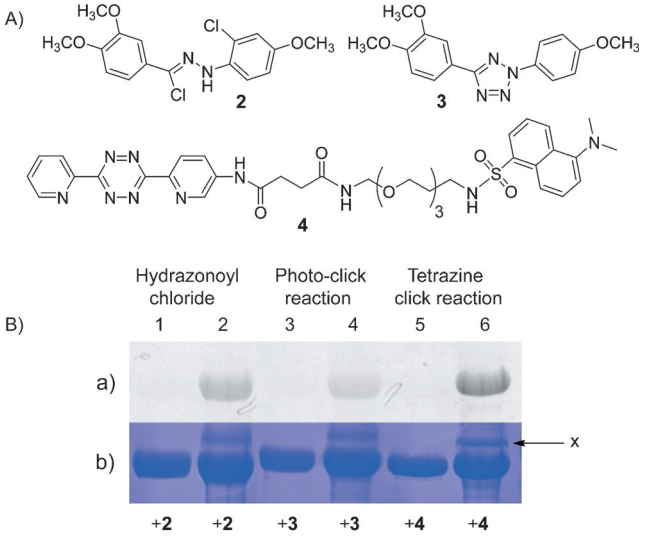


Figure 1. A) Structures of compounds **2**–**4**. B) SDS-PAGE of norbornene-containing *h*Polκ after fluorescent labeling with compounds **2**–**4**: a) Imaged gel showing the fluorescent click products; b) the same gel after Coomassie staining. Lanes 1, 3, and 5 show control experiments with wild-type *h*Polκ; lanes 2, 4, and 6 show the reaction with the norbornene-containing *h*Polκ. X = protein impurity.

The exact position of the chemical modification was determined using high-resolution HPLC-ESI-MS (see Figure 2 and the Supporting Information) after tryptic digest of the functionalized *h*Polκ. In the mass spectrum, the expected peptides for *h*Polκ were found with an excellent

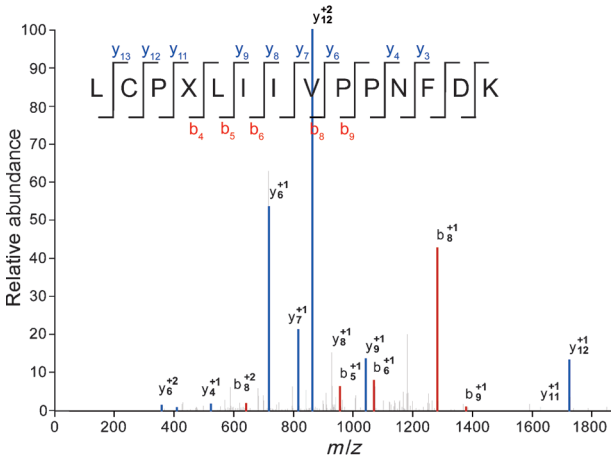


Figure 2. MS/MS spectrum of the peptide LCPXLIIVPPNFDK derived from tryptic digest of the modified *h*Polκ after click reaction with hydrazonoyl chloride **2** (X^1 is the position of the hydrazonoyl chloride modified norbornene).

sequence coverage of 80%. We also identified the expected hydrazonoyl chloride modified peptide (LCP**X**[1]LIIVPPNFDK, with **X**[1] as the modification site). For this peptide a mass of *m*/*z* 2031.10 was measured, which is in excellent agreement with the calculated value (2031.06). This peptide was further fragmented in the mass spectrometer (MS sequencing) to identify the critical b- and y-ions (Figure 2). The detected ions prove that the functionalization occurred exclusively at the desired position.

In order to determine the functionalization yield we reacted the norbornene amino acid **1** with the nitrile imine generated from **2**. We then measured the fluorescence of the resulting click product at different concentrations. The data were used to generate a calibration curve that was found to be linear (see the Supporting Information). We next synthesized the decapeptide AFD**X**KDKPAA containing the norbornene amino acid **1** at position **X**. This peptide was reacted in different molar ratios with the nitrile imine generated from **2.** Using the calibration curve we determined quantitative conversion of the peptide with 10 equiv of **2** after a reaction time of about 1 h at room temperature. We also measured MALDI-Tof mass spectra and observed no residual unreacted peptide after the functionalization reaction (see the Supporting Information).

In order to quantify the efficiency of the click reaction on the protein we reacted 240 pmol of the norbornene-containing *h*Polκ with increasing amounts of the nitrile imine generated from **2** and observed full conversion with 100 equiv of **2** (Figure 3). Using the calibration curve we determined a modification yield of 98%. Most important is the observation that this high modification yield is reached after only 1 h at room temperature. Finally, we wanted to prove that the nitrile imine reaction leaves the fragile *h*Polκ protein in the active state. Therefore, we compared the click-modified *h*Polκ and the wild-type *h*Polκ in a primer extension activity assay. The non-photochemical nitrile imine reaction and the reversed-electron-demand Diels–Alder reaction provided fully active protein (Figure S4 in the Supporting Information). Only the protein obtained after the photo-induced click reaction had compromised activity, showing that the UV light needed to create the nitrile imine may harm the protein.[16]

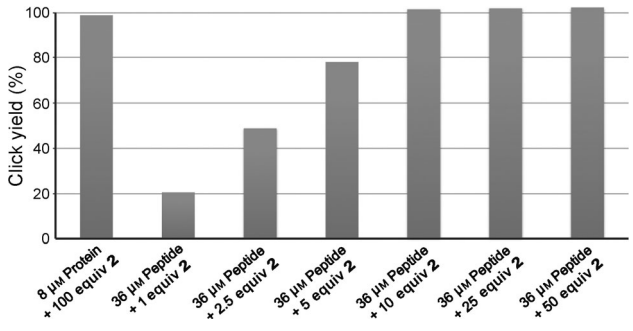


Figure 3. Quantification of the click efficiency based on the fluorescence emission of the click products. norb-Polκ was reacted with 100 equiv of the nitrile imine derived from **2,** and the norbornene-containing peptide AFDXKDKPAA with 1, 2.5, 5, 10, 25, and 50 equiv of the nitrile imine derived from **2**.

In summary, we have shown that the generation of nitrile imines from the corresponding hydrazonoyl chlorides by direct HCl elimination under ambient conditions (pH 7.4) followed by reaction with a norbornene-containing protein gives a fully active, modified protein. The click reaction with the norbornene-containing protein proceeds with almost quantitative yield in just 1 h at room temperature. The ability to insert norbornene amino acids into proteins using the pyrrolysine system and the application of a special evolved synthetase now enables the reliable and efficient chemical modification of even sensitive proteins.[17]

Received: December 30, 2011
Published online: March 21, 2012

Keywords: copper-free click chemistry · nitrile imines · norbornene amino acid · protein modification · pyrrolysine

[1] a) J. Kalia, R. T. Raines, *Curr. Org. Chem.* **2010**, *14*, 138–147; b) R. K. Lim, Q. Lin, *Chem. Commun.* **2010**, *46*, 1589–1600; c) C. C. Liu, P. G. Schultz, *Annu. Rev. Biochem.* **2010**, *79*, 413–444.
[2] J. M. Chalker, G. J. Bernardes, Y. A. Lin, B. G. Davis, *Chem. Asian J.* **2009**, *4*, 630–640.
[3] M. Morpurgo, E. A. Bayer, M. Wilchek, *J. Biochem. Biophys. Methods* **1999**, *38*, 17–28.
[4] a) T. Fekner, X. Li, M. M. Lee, M. K. Chan, *Angew. Chem.* **2009**, *121*, 1661–1663; *Angew. Chem. Int. Ed.* **2009**, *48*, 1633–1635; b) Z. Hao, Y. Song, S. Lin, M. Yang, Y. Liang, J. Wang, P. R. Chen, *Chem. Commun.* **2011**, *47*, 4502–4504; c) D. P. Nguyen, H. Lusic, H. Neumann, P. B. Kapadnis, A. Deiters, J. W. Chin, *J. Am. Chem. Soc.* **2009**, *131*, 8720–8721; d) T. Plass, S. Milles, C. Koehler, C. Schultz, E. A. Lemke, *Angew. Chem.* **2011**, *123*, 3964–3967; *Angew. Chem. Int. Ed.* **2011**, *50*, 3878–3881; e) L. Wang, Z. Zhang, A. Brock, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 56–61; f) T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto, S. Yokoyama, *Chem. Biol.* **2008**, *15*, 1187–1197.
[5] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2708–2711; *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599.
[6] C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057–3064.
[7] D. Soriano Del Amo, W. Wang, H. Jiang, C. Besanceney, A. C. Yan, M. Levy, Y. Liu, F. L. Marlow, P. Wu, *J. Am. Chem. Soc.* **2010**, *132*, 16893–16899.
[8] a) J. C. Jewett, E. M. Sletten, C. R. Bertozzi, *J. Am. Chem. Soc.* **2010**, *132*, 3688–3690; b) B. C. Sanders, F. Friscourt, P. A. Ledin, N. E. Mbua, S. Arumugam, J. Guo, T. J. Boltje, V. V. Popik, G. J. Boons, *J. Am. Chem. Soc.* **2011**, *133*, 949–957; c) S. S. van Berkel, A. T. Dirks, M. F. Debets, F. L. van Delft, J. J. Cornelissen, R. J. Nolte, F. P. Rutjes, *ChemBioChem* **2007**, *8*, 1504–1508; d) M. L. Blackman, M. Royzen, J. M. Fox, *J. Am. Chem. Soc.* **2008**, *130*, 13518–13519; e) N. K. Devaraj, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2008**, *19*, 2297–2299; f) J. Wang, W. Zhang, W. Song, Y. Wang, Z. Yu, J. Li, M. Wu, L. Wang, J. Zang, Q. Lin, *J. Am. Chem. Soc.* **2010**, *132*, 14812–14818; g) Z. Yu, R. K. Lim, Q. Lin, *Chem. Eur. J.* **2010**, *16*, 13325–13329; h) W. Song, Y. Wang, J. Qu, Q. Lin, *J. Am. Chem. Soc.* **2008**, *130*, 9654–9655.
[9] a) M. Ibba, D. Soll, *Curr. Biol.* **2002**, *12*, 464–466; b) S. M. Hancock, R. Uprety, A. Deiters, J. W. Chin, *J. Am. Chem. Soc.* **2010**, *132*, 14819–14824; c) C. Polycarpo, A. Ambrogelly, A. Berube, S. M. Winbush, J. A. McCloskey, P. F. Crain, J. L. Wood,

D. Soll, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12450 – 12454; d) C. R. Polycarpo, S. Herring, A. Berube, J. L. Wood, D. Soll, A. Ambrogelly, *FEBS Lett.* **2006**, *580*, 6695 – 6700.

[10] E. Kaya, K. Gutsmiedl, M. Vrabel, M. Müller, P. Thumbs, T. Carell, *ChemBioChem* **2009**, *10*, 2858 – 2861.

[11] M. T. Reetz, D. Kahakeaw, J. Sanchis, *Mol. Biosyst.* **2009**, *5*, 115 – 122.

[12] S. S. Lange, K. Takata, R. D. Wood, *Nat. Rev. Cancer* **2011**, *11*, 96 – 110.

[13] a) K. Gutsmiedl, D. Fazio, T. Carell, *Chem. Eur. J.* **2010**, *16*, 6877 – 6883; b) K. Gutsmiedl, C. T. Wirges, V. Ehmke, T. Carell, *Org. Lett.* **2009**, *11*, 2405 – 2408.

[14] R. Huisgen, M. Seidel, G. Wallbillich, H. Knupfer, *Tetrahedron* **1962**, *17*, 3 – 29.

[15] a) J. S. Clovis, A. Eckell, R. Huisgen, R. Sustmann, *Chem. Ber.* **1967**, *100*, 60 – 70; b) A. Padwa, S. Nahm, E. Sato, *J. Org. Chem.* **1978**, *43*, 1664 – 1671.

[16] R. L. Wells, A. Han, *Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med.* **1985**, *47*, 17 – 21.

[17] Note added in proof: During publication of this work a similar study was reported by: K. Lang et al., *Nature Chem.* **2012**, DOI: 10.1038/NCHEM.1250.

