DOI: 10.1002/anie.200805420

Pyrrolysine

A Pyrrolysine Analogue for Protein Click Chemistry**

*Tomasz Fekner, Xin Li, Marianne M. Lee, and Michael K. Chan**

The discovery of pyrrolysine (**1**, Figure 1), the 22nd genetically encoded amino acid,[1,2] and its subsequent incorporation into recombinant proteins in *E. coli*[3] have laid the foundation for the future development of novel biotechnologies and tactics in protein research. In principle, incorporation of the genes encoding the pyrrolysine tRNA synthase (PylS) and its cognate tRNA (PylT) could enable both bacteria[4,5] and eukaryotes[6] to use the UAG codon for the production of proteins containing pyrrolysine (**1**) and its analogues. Herein, we describe the synthesis of a new pyrrolysine congener that can be incorporated into recombinant protein and selectively modified with azide-containing dyes by copper(I) click chemistry. We demonstrate the usefulness of this technology for monitoring conformational changes of calmodulin (CaM) by Förster resonance energy transfer (FRET) measurements.

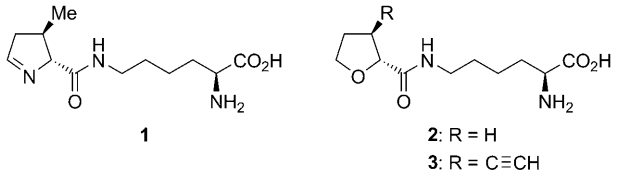
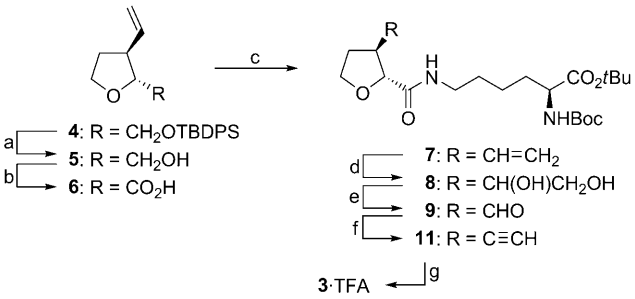


Figure 1. Pyrrolysine (**1**) and its THF analogues **2** and **3**.

To harness the unique UAG-codon–pyrrolysine system for biochemical applications, we decided to search for functionalized pyrrolysine surrogates that would be both more stable and amenable for further post-translational modification (i.e., tagging). Our primary focus centered on derivatives of the THF-containing amino acid **2** because of the steric and electronic similarity between the THF ring and the 3,4-dihydro-2*H*-pyrrole ring present in **1**. Recently, in collaboration with the Krzycki group, we demonstrated that *Methanosarcina barkeri* PylS could charge PylT with the parent pyrrolysine analogue **2**.[7] We subsequently postulated that this framework could be used as a starting point for the design of other functionalizable pyrrolysine analogues.

Herein, we report on the synthesis of the pyrrolysine analogue **3**, bearing a terminal alkyne functionality, that not only reads through the UAG codon in *E. coli* overexpressing the *Methanosarcina mazei* PylS and PylT, but also allows for site-specific modification of the resulting protein through the copper(I)-catalyzed azide–alkyne cycloaddition (CuAAC) reaction.[8] The taggable pyrrolysine analogue **3** was prepared from the known tetrahydrofuran **4**[9] by a seven-step sequence (Scheme 1). Thus, desilylation of **4** with TBAF/AcOH in THF[10] or AcCl in MeOH[11] gave alcohol **5** in 76% or 89% yield, respectively. Its subsequent oxidation with H_5IO_6 in the presence of a catalytic amount of CrO_3[12] provided acid **6** that, in turn, was coupled (BOP/NMM) with Boc-Lys-O*t*Bu to give amide **7** in 68% yield over the two steps. NMO/OsO_4-mediated dihydroxylation of **7** furnished an approximately 1:1 mixture of the corresponding diastereomeric diols **8** (93%) that were oxidatively cleaved with $NaIO_4$ to give aldehyde **9** in 94% yield. Subsequent Ohira–Bestmann alkynylation[13,14] with diazoketophosphonate **10**[15] in the presence of K_2CO_3 led to the protected target **11** (70%). Its final treatment with neat TFA cleanly freed the desired pyrrolysine analogue **3**.



Scheme 1. Synthesis of taggable pyrrolysine analogue **3**: a) TBAF, AcOH, THF, room temperature, 14 h, 76% or AcCl (25 mol%), MeOH, room temperature, 10 h, 89%; b) H_5IO_6, CrO_3 (4 mol%), MeCN, H_2O, 0°C, 90 min; c) Boc-Lys-O*t*Bu, BOP, NMM, CH_2Cl_2, room temperature, 48 h, 68% from **5**; d) OsO_4 (2 mol%), NMO, Me_2CO, H_2O, room temperature, 48 h, 93%; e) $NaIO_4$, THF, H_2O, room temperature, 30 min, 94%; f) $AcC(N_2)P(O)(OMe)_2$ (**10**), K_2CO_3, MeOH, room temperature, 3 h, 70%; g) TFA, room temperature, 1 h, ca. 100%. TBDPS: *tert*-butyldiphenylsilyl, TBAF: tetrabutylammonium fluoride, BOP: (benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate, NMM: *N*-methylmorpholine, NMO: *N*-methylmorpholine *N*-oxide; Boc: *tert*-butoxycarbonyl; Ac: acetyl; TFA: trifluoroacetic acid.

To test if **3** could be recognized by a pyrrolysyl-tRNA synthetase and incorporated into recombinant protein in vivo,

[*] Dr. T. Fekner,[+] X. Li,[+] M. M. Lee, Prof. M. K. Chan
Departments of Chemistry and Biochemistry, The Ohio State University
484 West 12th Avenue, Columbus, OH 43210 (USA)
Fax: (+1) 614-292-6773
E-mail: chan@chemistry.ohio-state.edu
Homepage: http://www.chemistry.ohio-state.edu/~chan/

[+] These authors contributed equally to this work.

[**] This research was supported by a grant from the US National Institutes of Health (GM061796) and an American Heart Association Great Rivers Affiliate Predoctoral Fellowship (0815449D) to X.L. We thank Liwen Zhang (CCIC, OSU) for protein mass-spectrometry analysis.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200805420.

E. coli (strain BL21(DE3)) harboring a plasmid containing the genes for *M. mazei* PylS and PylT[3,5] and *Rattus norvegicus* calmodulin-Thr34Pyl (CaM-Thr34Pyl) were grown in liquid Luria–Bertani medium supplemented with **3** (Figure 2a). The amber codon readthrough efficiency, as measured by the production of the full-length CaM, increased as the concentration of **3** present in the growth medium was raised. At [**3**] = 5 mM, approximately 15% of the total protein in the cell was the full-length CaM based on the sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) band quantification. In addition, mass-spectrometry studies (see Supporting Information) verified the incorporation of **3** at the expected residue location. This result, in turn, unequivocally demonstrated that the unnatural amino acid **3** could be incorporated specifically into a protein by means of an internal amber (UAG) codon.

Next, we tested whether CaM Thr34Pyl could be tagged by a CuAAC click process with a fluorophore containing the azido group (Figure 2b).[16] For this purpose, we selected azidocoumarin **12** owing to its considerable fluorescence increase following the cycloaddition reaction.[17] The results (Figure 2c,d) revealed that only the CaM with the engineered amber codon was labeled. The wild-type CaM did not show any detectable fluorescence, although it did give a dark band in the coomassie-stained gel. The other cellular proteins, including the highly expressed PylS, remained unlabeled.

FRET is a powerful tool for the study of protein dynamics.[18,19] A typical FRET experiment involves specific incorporation of a fluorophoric donor–acceptor (DA) pair into a protein and the measurement of the distance-dependent energy transfer efficiency between the two dyes. In reported CaM FRET studies, a double cysteine mutant was used.[20] However, the dominant forms found in the post-incorporation mixture were unlabeled CaM, CaM–AA, and CaM–DD, while the requisite CaM–DA constituted only a small fraction. As a potential improvement to this approach, we prepared the CaM Thr34Pyl/Thr110Cys double mutant. Consequently, the reactivity of azides towards **3** and that of maleimides towards cysteine, combined with their mutual orthogonality, made it possible to label specifically Pyl34 of the mutant CaM with azidocoumarin **12** (D) and Cys110 with Alexa Fluor 488 C5-maleimide (A). Thus the doubly labeled CaM mutant could be prepared in high yield.

CaM undergoes a significant structural change in the presence of Ca^{II} ions and different regulatory peptides.[21] One such peptide, M13, is derived from the CaM-binding domain of skeletal muscle myosin light-chain kinase. In the presence of Ca^{II} ions and the M13 peptide, CaM adopts a compact conformation, whereas in their absence, it is extended. Given that the FRET donor and acceptor are located on different domains, the FRET efficiency between these two dyes could potentially be used to monitor conformational states of the protein. Figure 3 shows the recorded fluorescence emission spectra of the doubly labeled CaM under 390 nm excitation. The different conformation of CaM in the presence and absence of Ca^{II} ions and M13 peptide is reflected in the change of the relative donor and acceptor peak heights.

In summary, we have designed and synthesized a stable, taggable pyrrolysine analogue **3** and demonstrated its ability to be incorporated into recombinant protein through the UAG codon. By utilizing its unique CuAAC reactivity, in combination with a known cysteine tagging method, we were able to generate a CaM labeled with two distinct fluorophores

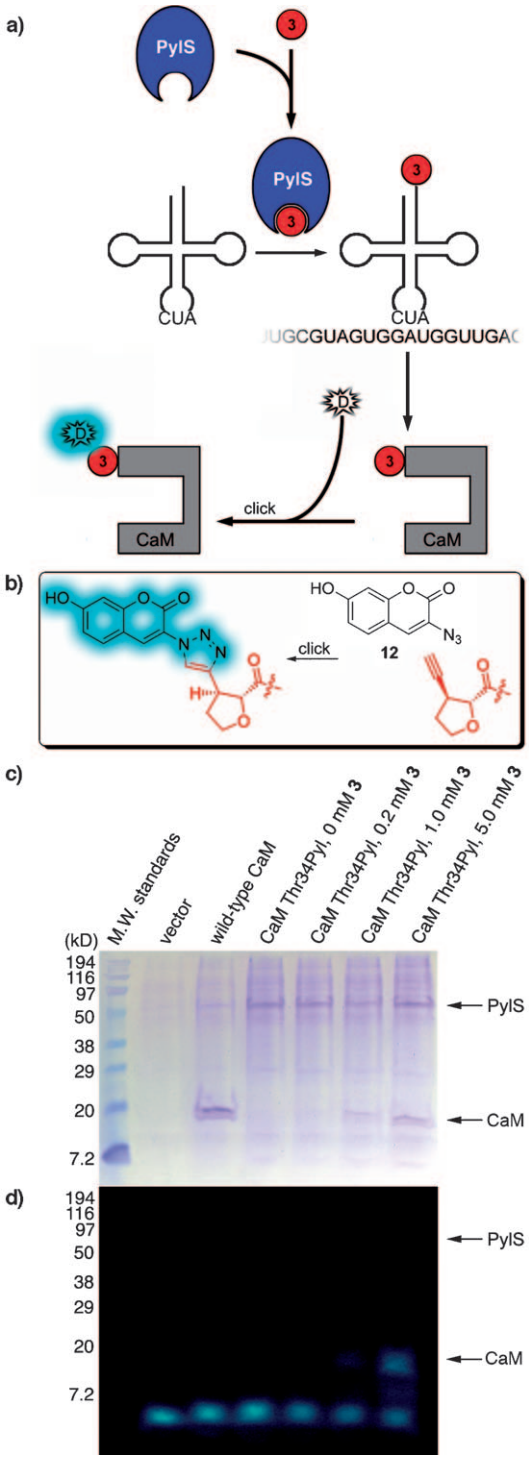


Figure 2. Translational incorporation of **3** into CaM and subsequent post-translational labeling with azidocoumarin **12** (a) through a CuAAC reaction (b). The process was monitored by SDS-PAGE gels with the protein bands visualized by both coomassie blue staining (c) and UV excitation at 312 nm (d) of the fluorescent dye-labeled proteins. The lower fluorescent band corresponds to residual dye.

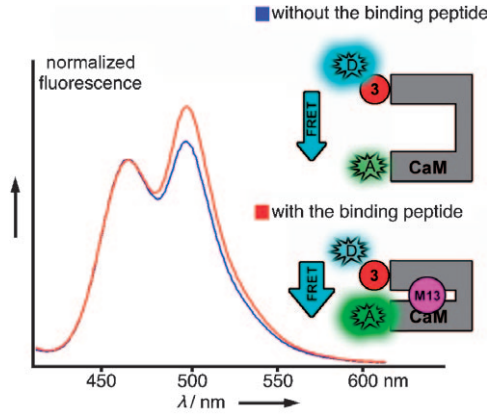


Figure 3. Tracking CaM conformational changes with FRET spectroscopy. Spectra normalized to donor fluorescence. See text for details.

that allowed us to monitor the conformational changes of this protein by FRET spectroscopy. Notably, this orthogonal double labeling strategy should be applicable to the study of many other protein systems.

Received: November 6, 2008
Published online: January 20, 2009

Keywords: amino acids · calmodulin · click chemistry · FRET · pyrrolysine

[1] G. Srinivasan, C. M. James, J. A. Krzycki, *Science* **2002**, *296*, 1459 – 1462.
[2] B. Hao, W. Gong, T. K. Ferguson, C. M. James, J. A. Krzycki, M. K. Chan, *Science* **2002**, *296*, 1462 – 1466.
[3] S. K. Blight, R. C. Larue, A. Mahapatra, D. G. Longstaff, E. Chang, G. Zhao, P. T. Kang, K. B. Green-Church, M. K. Chan, J. A. Krzycki, *Nature* **2004**, *431*, 333 – 335.
[4] D. G. Longstaff, R. C. Larue, J. E. Faust, A. Mahapatra, L. Zhang, K. B. Green-Church, J. A. Krzycki, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1021 – 1026.
[5] O. Namy, Y. Zhou, S. Gundllapalli, C. R. Polycarpo, A. Denise, J. P. Rousset, D. Söll, A. Ambrogelly, *FEBS Lett.* **2007**, *581*, 5282 – 5288.
[6] T. Mukai, T. Kobayashi, N. Hino, T. Yanagisawa, K. Sakamoto, S. Yokoyama, *Biochem. Biophys. Res. Commun.* **2008**, *371*, 818 – 822.
[7] W.-T. Li, A. Mahapatra, D. G. Longstaff, J. Bechtel, G. Zhao, P. T. Kang, M. K. Chan, J. A. Krzycki, *J. Mol. Biol.* **2009**, DOI: 10.1016/j.jmb.2008.11.032.
[8] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem.* **2001**, *113*, 2056 – 2075; *Angew. Chem. Int. Ed.* **2001**, *40*, 2004 – 2021.
[9] S. Hanessian, M. Brassard, *Tetrahedron* **2004**, *60*, 7621 – 7628.
[10] F. Kocieński, *Protecting Groups*, 3rd ed., Thieme, Stuttgard, **2004**, p. 208.
[11] A. T. Khan, E. Mondal, *Synlett* **2003**, 694 – 698.
[12] M. Zhao, J. Li, Z. Song, R. Desmond, D. M. Tschaen, E. J. J. Grabowski, P. J. Reider, *Tetrahedron Lett.* **1998**, *39*, 5323 – 5326.
[13] S. Ohira, *Synth. Commun.* **1989**, *19*, 561 – 564.
[14] G. J. Roth, B. Liepold, S. G. Muller, H. J. Bestmann, *Synthesis* **2004**, 59 – 62.
[15] P. Callant, L. D'Haenens, M. Vandewalle, *Synth. Commun.* **1984**, *14*, 155 – 161.
[16] For leading references on protein labeling, see: a) J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang, P. G. Schulz, *J. Am. Chem. Soc.* **2002**, *124*, 9026 – 9027; b) N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046 – 15047; c) B. P. Duckworth, Z. Zhang, A. Hosokawa, M. D. Distefano, *ChemBioChem*, **2007**, *8*, 98 – 105; d) T. H. Yoo, D. A. Tirrell, *Angew. Chem.* **2007**, *119*, 5436 – 5439; *Angew. Chem. Int. Ed.* **2007**, *46*, 5340 – 5343; e) I. S. Carrico, B. L. Carlson, C. R. Bertozzi, *Nat. Chem. Biol.* **2007**, *3*, 321 – 322.
[17] a) K. Sivakumar, F. Xie, B. M. Cash, S. Long, H. N. Barnhill, Q. Wang, *Org. Lett.* **2004**, *6*, 4603 – 4606; b) Z. Zhou, C. J. Fahrni, *J. Am. Chem. Soc.* **2004**, *126*, 8862 – 8863; c) K. E. Beatty, F. Xie, Q. Wang, D. A. Tirrell, *J. Am. Chem. Soc.* **2005**, *127*, 14150 – 14151.
[18] P. R. Selvin, *Nat. Struct. Biol.* **2000**, *7*, 730 – 734.
[19] S. Weiss, *Nat. Struct. Biol.* **2000**, *7*, 724 – 729.
[20] M. W. Allen, R. J. Urbauer, A. Zaidi, T. D. Williams, J. L. Urbauer, C. K. Johnson, *Anal. Biochem.* **2004**, *325*, 273 – 284.
[21] M. Ikura, G. M. Clore, A. M. Gronenborn, G. Zhu, C. B. Klee, A. Bax, *Science* **1992**, *256*, 632 – 638.

