Post-Translational Modifications

DOI: 10.1002/anie.201203349

Site-Specific Incorporation of ε-*N*-Crotonyllysine into Histones**

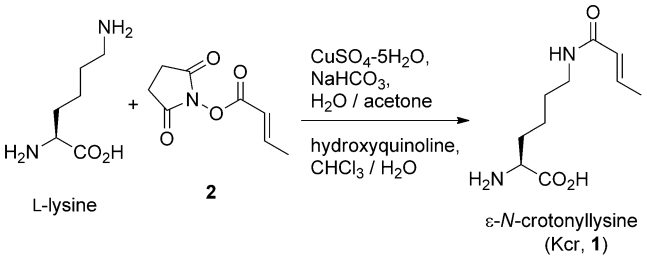
*Chan Hyuk Kim, Mingchao Kang, Hak Joong Kim, Abhishek Chatterjee, and Peter G. Schultz**

Histones, the major protein components of chromatin, are extensively altered by post-translational modifications (PTMs). Deciphering distinctive patterns of these PTMs (namely the "histone code") is key to understanding the epigenetic regulation of diverse biological processes.[1] For example, lysine acetylation is one of the most common PTMs found in histones, and its regulation has been demonstrated to be closely related to DNA replication/repair, transcriptional regulation of genes, and the status of chromatin assembly.[2] Misregulation of chromatin PTM is linked to various human diseases such as cancer,[3] central nervous system (CNS) disorders,[4] and autoimmune diseases.[5] Therefore, the enzymes responsible for lysine acetyl modifications, including histone acetyltransferases (HATs) and histone deacetylases (HDACs), are regarded as important drug targets.[6]

Besides acetylation, other known lysine modifications in histones include methylation, ribosylation, ubiqutination, and sumoylation.[7] Recently, Tan and co-workers described the identification of a new type of histone PTM, lysine crotonylation, by using a mass spectrometry-based proteomics approach.[8] This novel PTM is largely associated with transcriptionally active chromatin, and principally interacts with promoter and enhancer regions. Despite their structural similarity, lysine crotonylation substantially differs from lysine acetylation in genomic distribution and regulation. For example, the overexpression of two different histone acetyltransferases (HATs), CPB and p300, had no effect on cellular histone crotonylation levels, and histone lysine deacetylases (HDACs) 1, 2, 3, and 6 showed weak or no lysine decrotonylation activity.[8] These observations suggest that the two PTMs are likely regulated by separate sets of modifying enzymes and, hence, may possess distinctive cellular functions. The ability to generate histones with well-defined crotonyl modifications at specific lysine residues would facilitate a better understanding of the physiological roles and regulation of this novel PTM.[9]

A number of different approaches have been developed to prepare histones with defined PTMs.[10] Attempts to directly modify histones by enzymatic reactions[11] are limited either by the promiscuous nature of histone-modifying enzymes or by incomplete in vitro reactions that result in a heterogeneous mixture.[12] Semi-synthetic approaches using native chemical ligation (NCL)[13] and expressed protein ligation (EPL)[14] have been used successfully to generate site-specifically modified histones.[15] Additional chemical methods include the introduction of PTM analogues by using cysteine-based chemistries,[16] or the genetic incorporation and subsequent modification of reactive nonnatural amino acids.[17] These chemical approaches allow the selective, facile preparation of proteins with chemically defined PTMs in vitro. A complementary approach involves the genetic incorporation of post-translationally modified amino acids into histones directly in living cells by using the endogenous translational machinery. Herein, we report the preparation of site-specifically crotonyl-modified histones by means of an orthogonal aminoacyl-tRNA synthetase and tRNA pair, which was evolved to site-specifically incorporate ε-*N*-crotonyllysine (**1**) into proteins in response to amber codons in *Escherichia coli*. Furthermore, the evolved pair was used to overexpress a homogeneously crotonylated protein in mammalian cells.

ε-*N*-crotonyllysine (**1**, Kcr) was synthesized from unprotected lysine in two steps in > 80 % overall yield. The α-amino and carboxylic acid groups of lysine were transiently protected by chelation with copper(II) ion in basic aqueous solution,[18] and the ε-amino group was subsequently reacted with *N*-hydroxysuccinimide (NHS) ester of (*E*)-but-2-enoic (crotonic) acid (**2**) (Scheme 1). Orthogonal *Methanosarcina*



Scheme 1. Synthesis of ε-*N*-corotonyllysine (Kcr, **1**).

barkeri tRNA/pyrrolysyl-tRNA synthetase (Mb-PylRS) pairs have been used to encode a variety of lysine analogues, including lysine PTMs such as acetyl lysine[19] and methyl lysine.[17d] In order to evolve an Mb-PylRS that can efficiently charge Kcr to Pyl-tRNA, we used an Mb-PylRS library[20] randomized at residues Leu270, Tyr271, Leu274, and Cys313

[*] Dr. C. H. Kim,[+] M. Kang,[+] Dr. H. J. Kim,[#] Dr. A. Chatterjee, Prof. P. G. Schultz
Department of Chemistry, The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
E-mail: schultz@scripps.edu
[#] Current address: Department of Chemistry, Korea University Seoul 136-701 (Korea)
[+] These authors contributed equally to this work.

[**] This work was supported by the National Institutes of Health grant R01 GM062159 and the Skaggs Institute for Chemical Biology (P.G.S.). We thank Jun Y. Axup, Josh Payette and Virginia Seely for manuscript preparation. This is paper no 21779 at The Scripps Research Institute.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201203349.

(with a fixed Tyr349Phe mutation)[21] to NNK (N = A, T, G, or C, K = T or G) obtaining a diversity of 1×10^6 mutants (Supporting Information, Figure S1). A series of negative and positive selections were then performed in *E. coli* strain DH10B as previously described.[17b,22] In brief, in the negative selection, the Mb-PylRS library was introduced into DH10B cells harboring a plasmid containing the toxic barnase gene with amber mutations at two permissive sites (Gln2TAG, Asp44TAG), and cells were grown in the absence of Kcr to remove clones containing mutant synthetases that can recognize endogenous amino acids. The positive selection was carried out by subsequently transforming the library from the negative selection into DH10B cells containing a plasmid encoding chloramphenicol acetyltransferase with a Asp112TAG mutation. Cells were grown in the presence of 2 mM Kcr and 70 $\mu g\,mL^{-1}$ chloramphenicol to select clones containing mutant synthetases that can recognize the non-natural amino acid. After two alternating negative and positive rounds of selection, single colonies of Mb-PylRS mutants were obtained. The sequences of mutant synthetases from 20 colonies all converged to a unique clone: Kcr-RS (Leu270, Tyr271, Leu274Ala, Cys313Phe, Tyr349Phe), and this Kcr-RS was used for further studies.

To determine the fidelity and efficiency of Kcr incorporation into proteins, we expressed a mutant green fluorescent protein (GFP) using the Kcr-RS/Pyl-tRNA pair. In short, codon Asp149 in a C-terminal His-tagged GFP gene was mutated to TAG and was inserted into pLei vector containing a single copy of Pyl-tRNA. pBK-Kcr, which encodes the Kcr-RS gene was co-transformed into *E. coli* DH10B. Cells were grown in Luria–Bertani (LB) medium supplemented with and without 5 mM Kcr; 20 mM nicotinamide (NAM)[19] was added after 1 mM isopropyl-β-D-thiogalactopyranosid (IPTG) induction at OD_{600} 0.6. The protein was expressed overnight at 37 °C, followed by Ni-NTA (nickel nitrilotriacetic acid) affinity purification. SDS-PAGE analysis showed that full-length GFP was produced in good yield in the presence of 5 mM Kcr (ca. 10 $mg\,L^{-1}$), and not observed in the absence of Kcr (Figure 1 A). The purified mutant GFP was also characterized by electrospray ionization mass spectrometry (ESI-MS); the observed molecular mass (27786) confirmed that Kcr was site-specifically substituted for Asp149 (Figure 1 B).

Interestingly, when the mutant GFP is expressed in the absence of NAM, ESI-MS analysis shows two peaks: one small peak corresponds to the GFP with Kcr incorporated, and the second large peak has a mass 68 Da lower (Figure S2), which corresponds to a GFP mutant bearing a lysine at Asp149. Previous studies have shown that CobB, a sirtuin family nicotinamide adenine dinucleotide (NAD)-dependent enzyme, can deacetylate acetyllysine (Kac)-containing proteins in *E. coli*; the enzyme can be inhibited by high concentration (ca. 20 mM) of NAM.[19] We speculate that CobB may also possess decrotonylation activity, and when we expressed the mutant GFP in the presence of 20 mM NAM, indeed a single MS peak of mutant GFP with intact Kcr was obtained (Figure S2).

We next expressed a site-specifically crotonylated human histone subunit 2B (H2B) in *E. coli*. Multiple lysine crotonylation sites have been identified in H2B.[8] We choose to

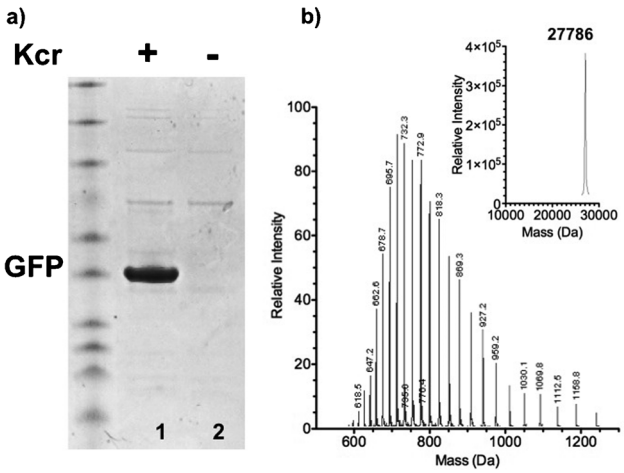


Figure 1. A) A mutant Mb-Pyl Kcr-RS/Pyl-tRNA pair selectively and efficiently incorporates Kcr in response to the TAG codon: GFP-Asp149TAG was expressed with or without 5 mM Kcr (lane 1 and lane 2, respectively) in the presence of an Mb-Pyl Kcr-RS/Pyl-tRNA pair. GFP mutant was purified by Ni-NTA affinity chromatography, analyzed by SDS-PAGE and stained with Coomassie Blue. B) ESI-MS analysis of purified mutant GFP; the peak corresponds to full-length GFP-Asp149Kcr-His_6 (expected: 27789, observed: 27786).

substitute Kcr for K11, since the N-terminal domain of histone is known to play an essential role in modulating nucleosome structures.[23] Furthermore, Lys11 residue is a unique crotonylation site while other residues (K5, K12, and K15) can be acetylated or crotonylated. The human H2B gene with an N-terminal His-tag and Lys11TAG mutation (His_6-H2B-K11TAG) was inserted into pCDF vector[24] containing a single copy of Pyl-tRNA. pBK-Kcr, which encodes the Kcr-RS gene was co-transformed into *E. coli* BL21(DE3) strains. Cells were grown in LB medium supplemented with and without 5 mM Kcr; 20 mM NAM was added after induction with 1 mM IPTG at OD_{600} 0.8. After overnight expression at 37 °C, cells were harvested and lysed with BugBuster (EMD), and protein was purified with Ni-NTA agarose beads (Qiagen) under denaturing conditions. SDS-PAGE analysis showed that full-length H2B was produced in good yield in the presence of 5 mM Kcr (ca. 2 $mg\,L^{-1}$) while full-length His_6-H2B was not observed in the absence of Kcr (Figure 2 A). Western blot analysis was carried out to confirm the incorporation of Kcr using an anti-corotonyllysine polyclonal antibody (PTM BioLab). Both His_6-H2B-wt and His_6-H2B-K11Kcr revealed a band with the anti-His_6 antibody, but only the His_6-H2B-K11Kcr mutant was detected with the anti-crotonyllysine antibody (Figure 2 A). In addition, the purified mutant H2B was also characterized by ESI-MS; the observed mass (16071) confirmed that Kcr was site-specifically substituted for Lys11 (Figure 2 B).

It also has been shown that the Mb-PylRS/Pyl-tRNA pair is orthogonal in mammalian cells.[20] We therefore next demonstrated the feasibility of shuttling the Kcr-RS/Pyl-tRNA system from *E. coli* to mammalian cells for expressing homogeneous crotonylated proteins in the latter cells. A plasmid pCMV-Kcr containing Kcr-RS and Pyl-tRNA under the CMV and U6 promoters, respectively, was transiently co-

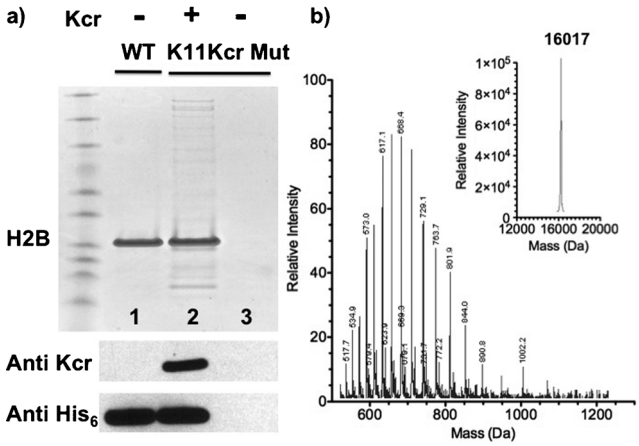


Figure 2. A) Selective incorporation of Kcr into recombinant human histone H2B: Wild type H2B was expressed (lane 1); the His_6-H2B K11 Kcr mutant was expressed with or without 5 mM Kcr (lane 2 and 3, respectively). Proteins were purified by Ni-NTA affinity chromatography under denaturing conditions and analyzed by SDS-PAGE. The gel was stained with Coomassie Blue (upper). Western blot analysis was carried out with anti-crotonyllysine (middle) and anti-His_6 antibodies (lower) to confirm the specific incorporation of Kcr. B) ESI-MS analysis reveals a mass corresponding to His_6-H2B K11Kcr with the N-terminal Met removed (expected: 16020, observed: 16017).

transfected with a plasmid pEGFP-Tyr39TAG containing a C-terminal His-tagged enhanced green fluorescent protein (EGFP) gene under the CMV promoter into HEK293T cells. Cells were grown in DMEM (Dulbecco's modified Eagle's medium) media supplemented with 10% FBS (fetal bovine serum) in the presence or absence of 1 mM Kcr for 48 h. Analysis of the cells by fluorescence microscopy revealed bright green fluorescence only in the presence of 1 mM Kcr (Figure 3A). After confirming Kcr-dependent TAG suppression, we also purified EGFP proteins and confirmed

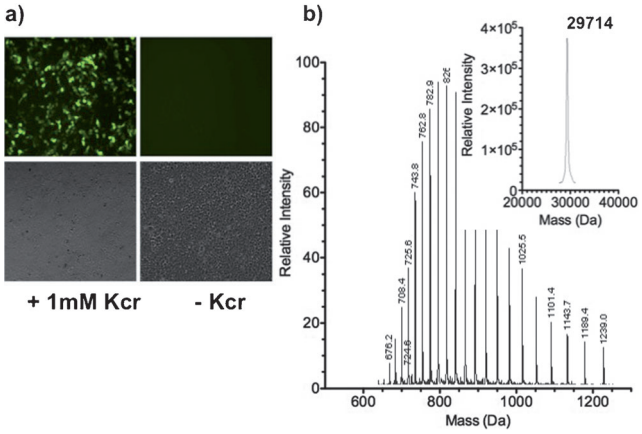


Figure 3. A) Selective incorporation of Kcr into EGFP-Y39TAG with the Mb-Pyl Kcr-RS/Pyl-tRNA pair in HEK293T cells; green fluorescent (top) and bright field (bottom) images were obtained to confirm the expression of full-length EGFP in the presence of 1 mM Kcr (left). The fluorescent signal was not observed in the absence of Kcr (right). B) ESI-MS analysis of the purified protein revealed a single mass corresponding to EGFP-Y39Kcr (expected: 29713, observed: 29714).

the site-specific incorporation of Kcr by ESI-MS (Figure 3B, expected: 29713, observed: 29714).

In conclusion, we have evolved an Mb-PylRS to site-specifically incorporate Kcr into proteins with high fidelity and efficiency both in *E. coli* and mammalian cells. The ability to produce histones with homogenous, site-specific Kcr modifications will be valuable in elucidating the biological role of this recently identified PTM.

Received: May 1, 2012
Published online: June 11, 2012

Keywords: crotonylations · histones · nonnatural amino acids · post-translational modifications

[1] a) E. I. Campos, D. Reinberg, *Annu. Rev. Genet.* **2009**, *43*, 559–599; b) M. Biel, V. Wascholowski, A. Giannis, *Angew. Chem.* **2005**, *117*, 3248–3280; *Angew. Chem. Int. Ed.* **2005**, *44*, 3186–3216.
[2] a) W. Dang, K. K. Steffen, R. Perry, J. A. Dorsey, F. B. Johnson, A. Shilatifard, M. Kaeberlein, B. K. Kennedy, S. L. Berger, *Nature* **2009**, *459*, 802–807; b) S. Zhao, W. Xu, W. Jiang, W. Yu, Y. Lin, T. Zhang, J. Yao, L. Zhou, Y. Zeng, H. Li, Y. Li, J. Shi, W. An, S. M. Hancock, F. He, L. Qin, J. Chin, P. Yang, X. Chen, Q. Lei, Y. Xiong, K.-L. Guan, *Science* **2010**, *327*, 1000–1004.
[3] P. Marks, R. A. Rifkind, V. M. Richon, R. Breslow, T. Miller, W. K. Kelly, *Nat. Rev. Cancer* **2001**, *1*, 194–202.
[4] A. G. Kazantsev, L. M. Thompson, *Nat. Rev. Drug Discovery* **2008**, *7*, 854–868.
[5] A. M. Grabiec, S. Krausz, W. de Jager, T. Burakowski, D. Groot, M. E. Sanders, B. J. Prakken, W. Maslinski, E. Eldering, P. P. Tak, K. A. Reedquist, *J. Immunol.* **2010**, *184*, 2718–2728.
[6] P. A. Cole, *Nat. Chem. Biol.* **2008**, *4*, 590–597.
[7] a) M. Yun, J. Wu, J. L. Workman, B. Li, *Cell Res.* **2011**, *21*, 564–578; b) S. Messner, M. O. Hottiger, *Trends Cell Biol.* **2011**, *21*, 534–542.
[8] M. Tan, H. Luo, S. Lee, F. Jin, J. S. Yang, E. Montellier, T. Buchou, Z. Cheng, S. Rousseaux, N. Rajagopal, Z. Lu, Z. Ye, Q. Zhu, J. Wysocka, Y. Ye, S. Khochbin, B. Ren, Y. Zhao, *Cell* **2011**, *146*, 1016–1028.
[9] C. A. Olsen, *Angew. Chem.* **2012**, *124*, 3817–3819; *Angew. Chem. Int. Ed.* **2012**, *51*, 3755–3756.
[10] A. Dhall, C. Chatterjee, *ACS Chem. Biol.* **2011**, *6*, 987–999.
[11] C. T. Walsh, S. Garneau-Tsodikova, G. J. Gatto, Jr., *Angew. Chem.* **2005**, *117*, 7508–7539; *Angew. Chem. Int. Ed.* **2005**, *44*, 7342–7372.
[12] P. J. J. Robinson, W. An, A. Routh, F. Martino, L. Chapman, R. G. Roeder, D. Rhodes, *J. Mol. Biol.* **2008**, *381*, 816–825.
[13] P. E. Dawson, S. B. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923–960.
[14] T. W. Muir, D. Sondhi, P. A. Cole, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6705–6710.
[15] a) R. K. McGinty, J. Kim, C. Chatterjee, R. G. Roeder, T. W. Muir, *Nature* **2008**, *453*, 812–816; b) M. Shogren-Knaak, H. Ishii, J. M. Sun, M. J. Pazin, J. R. Davie, C. L. Peterson, *Science* **2006**, *311*, 844–847.
[16] a) M. D. Simon, F. Chu, L. R. Racki, C. C. de La Cruz, A. L. Burlingame, B. Panning, G. J. Narlikar, K. M. Shokat, *Cell* **2007**, *128*, 1003–1012; b) R. Huang, M. A. Holbert, M. K. Tarrant, S. Curtet, D. R. Colquhoun, B. M. Dancy, B. C. Dancy, Y. Hwang, Y. Tang, K. Meeth, R. Marmorstein, R. N. Cole, S. Khochbin, P. A. Cole, *J. Am. Chem. Soc.* **2010**, *132*, 9986–9987; c) J. M. Chalker, L. Lercher, N. R. Rose, C. J. Schofield, B. G. Davis, *Angew. Chem.* **2012**, *124*, 1871–1875; *Angew. Chem. Int. Ed.*

2012, *51*, 1835 – 1839; d) F. Li, A. Allahverdi, R. Yang, G. B. Lua, X. Zhang, Y. Cao, N. Korolev, L. Nordenskiold, C. F. Liu, *Angew. Chem.* **2011**, *123*, 9785 – 9788; *Angew. Chem. Int. Ed.* **2011**, *50*, 9611 – 9614.

[17] a) J. Guo, J. Wang, J. S. Lee, P. G. Schultz, *Angew. Chem.* **2008**, *120*, 6499 – 6501; *Angew. Chem. Int. Ed.* **2008**, *47*, 6399 – 6401; b) H. W. Ai, J. W. Lee, P. G. Schultz, *Chem. Commun.* **2010**, *46*, 5506 – 5508; c) D. P. Nguyen, M. M. G. Alai, S. Virdee, J. W. Chin, *Chem. Biol.* **2010**, *17*, 1072 – 1076; d) D. P. Nguyen, M. M. Garcia Alai, P. B. Kapadnis, H. Neumann, J. W. Chin, *J. Am. Chem. Soc.* **2009**, *131*, 14194 – 14195.

[18] Z. Shen, Y. Zhang, Y. Chen, *Synth. Commun.* **2000**, *30*, 2525 – 2532.

[19] H. Neumann, S. Y. Peak-Chew, J. W. Chin, *Nat. Chem. Biol.* **2008**, *4*, 232 – 234.

[20] H. W. Ai, W. Shen, A. Sagi, P. R. Chen, P. G. Schultz, *ChemBiochem* **2011**, *12*, 1854 – 1857.

[21] T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto, S. Yokoyama, *Chem. Biol.* **2008**, *15*, 1187 – 1197.

[22] P. R. Chen, D. Groff, J. Guo, W. Ou, S. Cellitti, B. H. Geierstanger, P. G. Schultz, *Angew. Chem.* **2009**, *121*, 4112 – 4115; *Angew. Chem. Int. Ed.* **2009**, *48*, 4052 – 4055.

[23] M. A. Parra, D. Kerr, D. Fahy, D. J. Pouchnik, J. J. Wyrick, *Mol. Cell. Biol.* **2006**, *26*, 3842 – 3852.

[24] H. Neumann, S. M. Hancock, R. Buning, A. Routh, L. Chapman, J. Somers, T. Owen-Hughes, J. van Noort, D. Rhodes, J. W. Chin, *Mol. Cell* **2009**, *36*, 153 – 163.

