

Dual Genetic Encoding of Acetyl-lysine and Non-deacetyltable Thioacetyl-lysine Mediated by Flexizyme

Hai Xiong, Noah M. Reynolds⁺, Chenguang Fan⁺, Markus Englert, Denton Hoyer, Scott J. Miller, and Dieter Söll*

Abstract: Acetylation of lysine residues is an important post-translational protein modification. Lysine acetylation in histones and its crosstalk with other post-translational modifications in histone and non-histone proteins are crucial to DNA replication, DNA repair, and transcriptional regulation. We incorporated acetyl-lysine (AcK) and the non-hydrolyzable thioacetyl-lysine (ThioAcK) into full-length proteins *in vitro*, mediated by flexizyme. ThioAcK and AcK were site-specifically incorporated at different lysine positions into human histone H3, either individually or in pairs. We demonstrate that the thioacetyl group in histone H3 could not be removed by the histone deacetylase sirtuin type 1. This method provides a powerful tool to study protein acetylation and its role in crosstalk between post-translational modifications.

Post-translational modifications (PTMs) of the N-terminal “tails” of histones serve as epigenetic indicators of chromatin structure associated with genome regulation, DNA repair, and transcriptional silencing.^[1] Lysine acetylation of histones is correlated with dynamic transitions in chromatin structure and function, transcriptional activation, and DNA replication.^[1a,c] Crosstalk between lysine acetylation of histones and other PTMs is crucial in chromatin-based control and shaping inheritable epigenetic programs.^[1f,2] As histone acetyltransferases (HATs) and histone deacetylases (HDACs) modify several lysine residues and multiple enzymes target common sites in histones, it is thought that function and regulation of histone acetylation is obtained through the cumulative effects of HATs and HDACs.^[3] Studies on the role of PTMs use non-hydrolysable analogues, which are novel tools in the analysis of enzymes that modulate modification *in vivo*. N^ε-thioacetyl-lysine is a single-atom substituted acetyl-lysine derivative with the most similar chemical properties to AcK among the known analogues.^[4] To date, solid-phase peptide synthesis was

used to incorporate ThioAcK into short-peptide-derived p53 and α -tubulin. Such peptides were effective human sirtuin inhibitors and served as multifunctional probes owing to the high stability of ThioAcK.^[4,5] However, ThioAcK has not been incorporated into a full-length protein, limiting its use in studies of histone modification.

Incorporation of non-canonical amino acids (ncAAs) into peptides and proteins was achieved by various methods, including protein semi-synthesis combined with chemical ligation, and by genetic-code expansion using orthogonal aminoacyl-tRNA synthetase (aaRS)-tRNA pairs.^[6] A combination of two pairs, pyrrolysyl-tRNA synthetase-tRNA^{Pyl} and *M. jannaschii* tyrosyl-tRNA synthetase-tRNA^{Tyr} allowed the suppression of UAG and UGA codons simultaneously, while suppression of quadruplet codons was also utilized to incorporate different ncAAs into a single protein.^[7]

Acylation of tRNAs with ncAAs can be performed by orthogonal aaRS-tRNA pairs or through chemical aminoacylation of a dinucleotide and subsequent ligation to the tRNA body; however, both require multistep experiments involving substantial labor and materials.^[8] An attractive, general method to form aminoacyl-tRNA was developed by Suga and co-workers,^[9] a catalytic RNA (Flexizyme) recognizes and aminoacylates tRNA using an amino acid cyanomethyl ester with an aromatic side-chain as the amino acid donor. Further improvement led to dinitro-flexizyme (dFx), which utilizes amino acid 3,5-dinitrobenzyl esters (DBE) as the aminoacyl-donor without dependence on the amino acid side-chain.^[9a] Flexizyme was used to produce N-terminal peptides of histone H3 containing methyl-lysine and acetyl-lysine residues; however, to date this has not been reported for full-length histone proteins.^[9b]

To generate full-length H3 histones containing the non-deacetyltable ThioAcK, we used dFx to acylate tRNA with the 3,5-dinitrobenzyl esters of N^ε-acetyl-lysine (AcK-DBE, **3**), N^ε-thioacetyl-lysine (ThioAcK-DBE, **4**), and N^ε-selenoacetyl lysine (SeAcK-DBE, **5**; Figure 1). SeAcK was synthesized to show the effectiveness and general usefulness of our methods as AcK, ThioAcK, and SeAcK contain oxygen, sulfur, and selenium, respectively; all of them belong to the same group of elements in the periodic table. The novel compounds ThioAcK-DBE and SeAcK-DBE were synthesized and further characterized by ¹H NMR, ¹³C NMR spectroscopy, and HPLC-MS (Supporting Information). As controls for dFx aminoacylation, Val-DBE (**1**), Tyr-DBE (**2**), and AcK-DBE (**3**) were synthesized as described elsewhere with minor changes.^[9a]

The flexizyme and its derivatives recognize the 3' end of tRNA, using the RCCA-3' on the tRNA, where R is G or A at

[*] Dr. H. Xiong, Dr. N. M. Reynolds,^[+] Dr. C. Fan,^[+] Dr. M. Englert, Prof. D. Söll
Department of Molecular Biophysics and Biochemistry
Yale University
Whitney Avenue 266, New Haven, CT 06511 (USA)
E-mail: dieter.soll@yale.edu

Dr. D. Hoyer
Yale Center for Molecular Discovery, Yale University
West Haven, CT 06516 (USA)
Prof. S. J. Miller, Prof. D. Söll
Department of Chemistry, Yale University
225 Prospect Street, New Haven, CT 06511 (USA)

[+] These authors contributed equally to this work.

Supporting information for this article can be found under <http://dx.doi.org/10.1002/anie.201511750>.

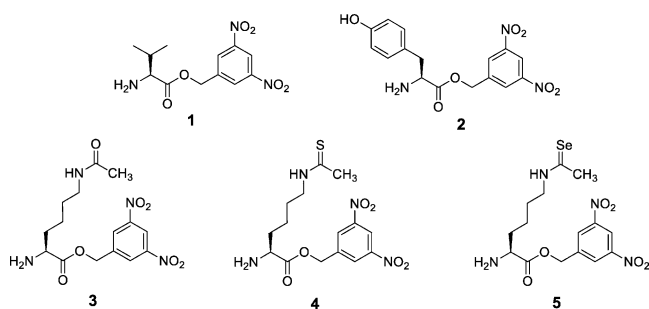


Figure 1. Structures of valine 3,5-dinitrobenzyl ester (Val-DBE, **1**), tyrosine 3,5-dinitrobenzyl ester (Tyr-DBE, **2**), acetyl-lysine 3,5-dinitrobenzyl ester (AcK-DBE, **3**), thioacetyl-lysine 3,5-dinitrobenzyl ester (ThioAcK-DBE, **4**), and selenoacetyl-lysine 3,5-dinitrobenzyl ester (SeAcK-DBE, **5**).

position 73. To match the 3'-end of dFx with ACCA-3' in the dFx-based acylation system with our tRNA, we designed a stable U73A-tRNA^{Sep}_{CUA} derived from tRNA^{Sep}_{CUA} (UCCA-3' end) by mutating the nucleotide in position 73 (Figure S1 in the Supporting Information). We tested aminoacylation efficiency of tRNA^{Sep}_{CUA} or the U73A-tRNA^{Sep}_{CUA} variant with compounds **1** and **2** with the dFx system and all combinations showed charging in the 36–45% range (Figure S2). As previous investigations using the flexizyme system had utilized tRNA^{Asn}, tRNA^{MetE}, and Mj-tRNA^{Tyr}_{CUA}, we tested the aminoacylation efficiency of the dFx system using these tRNAs as controls to ensure aminoacylation activity with our synthesized amino acid benzyl esters (Figure S2).

To investigate the effect of altering anticodons on dFx activity, the U73A-tRNA^{Sep}_{CUA} anticodon was mutated to match UAA and UGA as well as quadruplet codons in coding sequences. All combinations of tRNA^{Sep} mutants and amino acids resulted in significant charging, with efficiencies in the 26–38% range (Figure S2). Using the PURExpress (NEB) in vitro translation system lacking release factor 1 (PURE ΔRFI), we tested the ability of dFx aminoacylated U73A-tRNA^{Sep}_{CUA} to direct incorporation of AcK, ThioAcK, and SeAcK into proteins by UAG suppression. Suppression of UAG at position 151 of superfolder-GFP (GFP-Y151TAG) resulted in incorporation efficiencies of 15% for Val, 32% for Tyr, 41% for AcK, 37% for ThioAcK, and 26% for SeAcK compared to wild-type sfGFP when measured by fluorescence intensity (Figure S3–S5 and Table S1–S2).

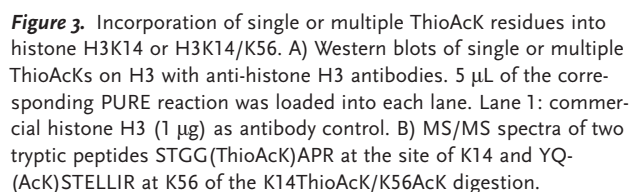
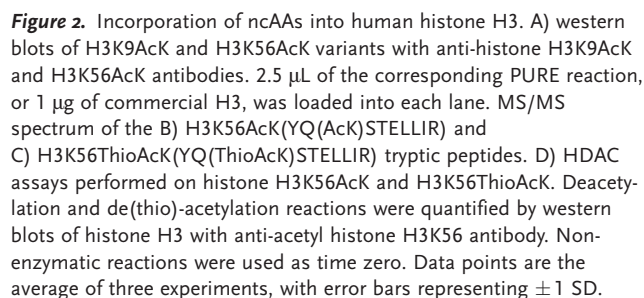
To study the incorporation of AcK and ThioAcK modifications in histone proteins, we directed AcK, ThioAcK, and SeAcK to positions K9 and K56 in the human histone H3 protein by UAG suppression with dFx aminoacylated U73A-tRNA^{Sep}_{CUA} in the PURE ΔRFI system. Detection of incorporation by site-specific anti-acetyl histone H3K9 and H3K56 antibodies showed high efficiency of ThioAcK incorporation at K9 and K56 (Figure 2A). Further experiments showed that AcK and ThioAcK have similar affinities for AcK-specific antibodies (Figure S14). Incorporation of AcK and ThioAcK at position K56 was confirmed by MS/MS (Figure 2B,C). However, we were unable to detect peptides from H3 that contained AcK or ThioAcK at position 9. Position 9 is in a lysine-rich sequence that gives rise to short,

polar peptides after trypsin digestion; such peptides would not be detectable by MS.^[6b] Furthermore, we did not detect peptides containing SeAcK, but found that the corresponding SeAcK modifications were chemically changed into AcK. While we found good incorporation of ThioAcK in response to UAG and UGA codons, incorporation of ThioAcK at position K56 in response to the UAA codon was unsuccessful (Figure 2A). This is in agreement with incorporation of ThioAcK at other positions, which shows incorporation at UAA codons to be mRNA-context dependent and weak overall.

Peptides prepared by SPPS that contain ThioAcK have been used as a mimetic of p53 in a HDAC assay, inhibiting human SIRT1-catalyzed deacetylation;^[4,10] however, few HDAC studies have been successfully performed directly on full-length proteins containing ncAAs owing to the difficulty in synthesis of full-length proteins. Furthermore, although chemically modified peptides have been used as mimetic histones, so far only weak interactions have been detected between PTM and PTM-reading domains.^[11] Having incorporated both AcK and ThioAcK into the H3 protein, we sought to examine the ability of HDACs to act on AcK residues in full-length proteins. We chose human SIRT1, a class III enzyme that catalyzes the NAD⁺-dependent conversion of acetyl-lysine residues. We compared the activity of SIRT1 on full-length H3 containing the ThioAcK at position K56 (K56ThioAcK) and with that of H3 containing AcK at position K56 (K56AcK). Analysis of a 90-minute time course at 37°C showed that treatment with SIRT1 resulted in a 70% loss in AcK at position K56, while SIRT1 treatment of K56ThioAcK containing H3 resulted in insignificant loss in histone K56ThioAcK (Figure 2D). These results demonstrate that the new K56ThioAcK-modified H3 at position K56 afforded H3 the ability to resist SIRT1 deacetylation.

A comparison of the effects of directing ThioAcK insertion into H3 with differing stop codons and positions (K9, K14, K36, K56) was undertaken and western blots performed with a histone H3 antibody. ThioAcK was most efficiently incorporated in response to an UAG codon at position K14 and an UGA codon at K56 (Figure 3A, Figures S7 and S8). Further validation of AcK and ThioAcK incorporation was performed by western blotting with position-specific anti-AcK histone H3K14 and H3K56 antibodies (Figure 3A and Figure S9). To incorporate simultaneously AcKs or ThioAcKs into H3 we tested insertion of AcK and ThioAcK at positions K14/K56 and K36/K56. We found that incorporation of the same ncAA into multiple positions of a single protein was highest when using UGA codons as opposed to UAG codons (Figure 3 and Figure S7). Analysis by MS/MS showed the correct multiple incorporation of ThioAcK at K36/K56 or K14/K56 (Figure S12, S13 and Table S5, S6).

After successful incorporation of multiple AcKs and ThioAcKs into full-length H3, we attempted to simultaneously incorporate both AcK and ThioAcK. Because insertion efficiency at K14 was higher than that at K36 (Figure S7), we chose K14/K56 in histone H3 to incorporate AcK and ThioAcK simultaneously by a combination of UAG and UGA codon suppression (Figure 3A). We randomized stop



codons and positions for AcK/ThioAcK incorporation into full-length histone H3: ThioAcK with UGA at position K14 and AcK with UAG at position K56 (K14ThioAcK(UGA)/K56AcK(UGA)), AcK with UGA at position K14 and ThioAcK with UAG at position K56 (K14AcK(UGA)/K56ThioAcK(UGA)), ThioAcK with UAG at position K14 and AcK with UGA at position K56 (K14ThioAcK(UGA)/K56AcK(UGA)), AcK with UAG at position K14 and ThioAcK with UGA at position K56 (K14AcK(UGA)/K56ThioAcK(UGA)). Western blots showed observable signals in nearly all of the above samples (Figure 3A and Figures S8,S9). MS/MS analyses gave the correct mass spectra for two digested peptides containing, respectively, ThioAcK at K14 and AcK at the site of K56 of histone H3K14ThioAcK/K56AcK (Figure 3B).

The work presented herein is the first example of lysine modifications in full-length human histone H3 with combinations of ThioAcK and AcK. The use of these and similar histone variants should lead to a deeper understanding of the important processes of histone acetylation and deacetylation.^[4,9b,12]

Acknowledgements

We thank the members of the Söll lab for helpful comments. This work was supported grants from the National Institutes of Health (GM22854 to D.S.; GM68649 to S.J.M.; AI119813 to C.F.), and the Defense Advanced Research Projects Agency (contract N66001-12-C-4211) to D.S.

Keywords: flexizyme · histone · lysine acetylation · post-translational modifications · thioacetyl-lysine

How to cite: *Angew. Chem. Int. Ed.* **2016**, 55, 4083–4086
Angew. Chem. **2016**, 128, 4151–4154

- [1] a) T. Tsubota, C. E. Berndsen, J. A. Erkmann, C. L. Smith, L. Yang, M. A. Freitas, J. M. Denu, P. D. Kaufman, *Mol. Cell* **2007**, 25, 703–712; b) C. C. Chen, J. J. Carson, J. Feser, B. Tamburini, S. Zabaronick, J. Linger, J. K. Tyler, *Cell* **2008**, 134, 231–243; c) C. Choudhary, C. Kumar, F. Gnäd, M. L. Nielsen, M. Rehman, T. C. Walther, J. V. Olsen, M. Mann, *Science* **2009**, 325, 834–840; d) H. S. Park, M. J. Hohn, T. Umehara, L. T. Guo, E. M. Osborne, J. Benner, C. J. Noren, J. Rinehart, D. Söll, *Science* **2011**, 333, 1151–1154; e) M. R. Duan, M. J. Smerdon, *J. Biol. Chem.* **2014**, 289, 8353–8363; f) X. J. Yang, E. Seto, *Mol. Cell* **2008**, 31, 449–461.
- [2] J. S. Lee, E. Smith, A. Shilatifard, *Cell* **2010**, 142, 682–685.
- [3] a) K. K. Lee, J. L. Workman, *Nat. Rev. Mol. Cell Biol.* **2007**, 8, 284–295; b) C. Dhalluin, J. E. Carlson, L. Zeng, C. He, A. K. Aggarwal, M. M. Zhou, *Nature* **1999**, 399, 491–496.
- [4] D. G. Fatkins, A. D. Monnot, W. Zheng, *Bioorg. Med. Chem. Lett.* **2006**, 16, 3651–3656.
- [5] a) P. Mellini, T. Kokkola, T. Suuronen, H. S. Salo, L. Tolvanen, A. Mai, M. Lahtela-Kakkonen, E. M. Jarho, *J. Med. Chem.* **2013**, 56, 6681–6695; b) P. H. Kiviranta, T. Suuronen, E. A. A. Wallen, J. Leppanen, J. Tervonen, S. Kyrilenko, A. Salminen, A. Poso, E. M. Jarho, *J. Med. Chem.* **2009**, 52, 2153–2156.
- [6] a) C. C. Liu, P. G. Schultz, *Annu. Rev. Biochem.* **2010**, 79, 413–444; b) M. J. Gattner, M. Vrabel, T. Carell, *Chem. Commun.* **2013**, 49, 379–381; c) Z. Zhang, M. Tan, Z. Xie, L. Dai, Y. Chen, Y. Zhao, *Nat. Chem. Biol.* **2011**, 7, 58–63; d) C. H. Kim, M. Kang, H. J. Kim, A. Chatterjee, P. G. Schultz, *Angew. Chem. Int. Ed.* **2012**, 51, 7246–7249; *Angew. Chem.* **2012**, 124, 7358–7361; e) W. Wan, J. M. Tharp, W. R. Liu, *Biochim. Biophys. Acta Proteins Proteomics* **2014**, 1844, 1059–1070.
- [7] a) A. Sachdeva, K. Wang, T. Elliott, J. W. Chin, *J. Am. Chem. Soc.* **2014**, 136, 7785–7788; b) K. Wang, A. Sachdeva, D. J. Cox, N. M. Wilf, K. Lang, S. Wallace, R. A. Mehl, J. W. Chin, *Nat. Chem.* **2014**, 6, 393–403; c) W. Wan, Y. Huang, Z. Wang, W. K. Russell, P. J. Pai, D. H. Russell, W. R. Liu, *Angew. Chem. Int. Ed.* **2010**, 49, 3211–3214; *Angew. Chem.* **2010**, 122, 3279–3282; d) H. Neumann, K. Wang, L. Davis, M. Garcia-Alai, J. W. Chin, *Nature* **2010**, 464, 441–444.
- [8] a) S. M. Hecht, B. L. Alford, Y. Kuroda, S. Kitano, *J. Biol. Chem.* **1978**, 253, 4517–4520; b) L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, 292, 498–500.
- [9] a) H. Murakami, A. Ohta, H. Ashigai, H. Suga, *Nat. Methods* **2006**, 3, 357–359; b) T. J. Kang, S. Yuzawa, H. Suga, *Chem. Biol.* **2008**, 15, 1166–1174; c) Y. Goto, T. Katoh, H. Suga, *Nat. Protoc.* **2011**, 6, 779–790; d) Y. Bessho, D. R. Hodgson, H. Suga, *Nat. Biotechnol.* **2002**, 20, 723–728; e) N. Terasaka, G. Hayashi, T. Katoh, H. Suga, *Nat. Chem. Biol.* **2014**, 10, 555–557.
- [10] a) B. M. Hirsch, W. Zheng, *Mol. Biosyst.* **2011**, 7, 16–28; b) H. Vaziri, S. K. Dessain, E. N. Eagon, S. I. Imai, R. A. Frye, T. K. Pandita, L. Guarente, R. A. Weinberg, *Cell* **2001**, 107, 149–159.
- [11] S. D. Taverna, H. Li, A. J. Ruthenburg, C. D. Allis, D. J. Patel, *Nat. Struct. Mol. Biol.* **2007**, 14, 1025–1040.
- [12] P. Gut, E. Verdin, *Nature* **2013**, 502, 489–498.

Received: December 28, 2015

Published online: February 23, 2016