Protein Modification

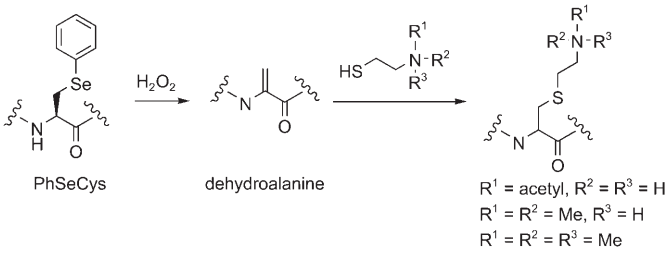
DOI: 10.1002/anie.200802336

Site-Specific Incorporation of Methyl- and Acetyl-Lysine Analogues into Recombinant Proteins**

*Jiantao Guo, Jiangyun Wang, Jong Seok Lee, and Peter G. Schultz**

The posttranslational modification (PTM) of lysine residues in proteins, either by acetylation or methylation, can modify the activities, biomolecular interactions, and lifetimes of proteins. For example, the modification of specific lysine side chains in histones can act epigenetically to modulate gene expression.[1-3] To better understand the role of these PTMs, both biosynthetic and semisynthetic strategies have been developed to produce homogeneous recombinant proteins with defined methylation and acetylation states at sites of interest. Unfortunately, the use of enzymatic methods, for example, histone methyltransferases (HMT) or histone acetyltransferases (HAT), is often complicated by the limited availability of the enzymes, insufficient site specificity, and the lack of control over degree of modification. Efficient semisynthetic methods have been developed to access modified proteins, but these are generally limited to modifications at N- or C-terminal residues.[4–6] Methods to chemically generate N^{ε}-methyl-lysine analogues have been reported by two different groups.[7–8] Both focused on the unique reactivity of cysteine, which can be directly alkylated[7] or first converted to dehydroalanine[8] and subsequently modified to generate N^{ε}-methyl-lysine analogues. Site-specific methylation using these methods is quite effective, but it is limited to proteins with only one reactive cysteine residue. And finally, it has recently been reported that N^{ε}-acetyl-lysine can be genetically introduced into proteins in *Escherichia coli* (*E. coli*) in response to amber nonsense mutations.[9]

To augment and extend these methods, we have exploited the unique reactivity of the unnatural amino acid phenylselenocysteine (PhSeCys) to chemically generate site-specific N^{ε}-methyl- and N^{ε}-acetyl-lysine analogues in proteins. A mutant *Methanococcus jannaschii* tyrosyl amber suppressor tRNA, Tyr *Mj*tRNA CUR/tyrosyl-tRNA synthetase (MjTyrRS) pair has been evolved to uniquely incorporate PhSeCys in response to the amber TAG codon in *E. coli*.[10] After being efficiently converted into dehydroalanine under mild conditions, Michael addition reactions with the corresponding thiols can be used to synthesize N^{ε}-methyl- and N^{ε}-acetyl-lysine analogues (Scheme 1). These N^{ε}-methyl-lysine analogues have been selectively incorporated into histone H3 and shown to function similarly to their natural counterparts.[7]



Scheme 1. The genetic incorporation of phenylselenocysteine (PhSeCys) with subsequent oxidative elimination; Michael addition of the corresponding thiols results in the formation of methyl- and acetyl-lysine analogues.

An amber codon was substituted for Lys 9 (H3 K9) in *Xenopus laevis* histone H3 containing a C-terminal His_6 tag. Protein expression was carried out in the presence of the selected synthetase (PhSeCysRS) and Tyr *Mj*tRNA CUR (on plasmid pSup-PhSeCysRS) in *E. coli* grown in lysogeny broth (LB) media supplemented with 1 mM PhSeCys. The yield of the mutant histone H3 was 15 mg L^{-1} after Ni–NTA affinity purification (NTA = nitrilotriacetic acid; Figure S2 in the Supporting Information). After being treated with 100 mM H_2O_2 in Tris buffer (50 mM, pH 8.5, Tris = $(HOCH_2)_3CNH_2$) at room temperature for 40 min, conversion of PhSeCys to dehydroalanine was determined by ESI-MS. The predicted mass of H3 containing dehydroalanine at position 9 is 16231 Da. However, the two major observed mass peaks are 16247 and 16262 Da (Figure S1 in the Supporting Information), which are 16 and 31 mass units higher than the calculated value. These two peaks were assigned to histone H3 in which Met 90, Met 120, or both are oxidized by H_2O_2. To subsequently convert dehydroalanine into the N^{ε}-acetyl-lysine analogue, the protein was allowed to react with 100 mM *N*-acetylcysteamine in pH 8.8, 50 mM Tris buffer at room temperature. ESI-MS showed complete disappearance of the dehydroalanine-containing histone protein after two hours. Two new peaks were detected at 16368 and 16384 Da, in close agreement with the calculated mass of 16366 and 16382 Da for the histone H3 mutant containing *S*-acetylethylamine-cysteine (N^{ε}-acetyl-lysine analogue) with single or/and double oxidation of the two methionine residues (Table S1 in

[*] J. Guo, J. Wang,[+] J. S. Lee, Prof. P. G. Schultz
Department of Chemistry and the Skaggs Institute for Chemical Biology, The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
Fax: (+1) 858-784-9440
E-mail: schultz@scripps.edu

[+] Current address: National Key Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101 (China)

[**] We thank Bill Webb for help in protein mass spectrometry, and we are grateful to the DOE (ER46051) and the Skaggs Institute for Chemical Biology for support.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200802336.

the Supporting Information). The total yield of lysine-modified protein was greater than 90% from the H3 containing PhSeCys.

To avoid the complication of methionine oxidation, the two methionine-encoding codons (ATGs) were replaced by CTGs to encode the structurally similar amino acid leucine in histone H3 (H3K9M90M120). After treatment with H_2O_2, only one major peak, 16195 Da (without the N-terminal methionine), was observed by ESI-MS (Figure 1; a small peak

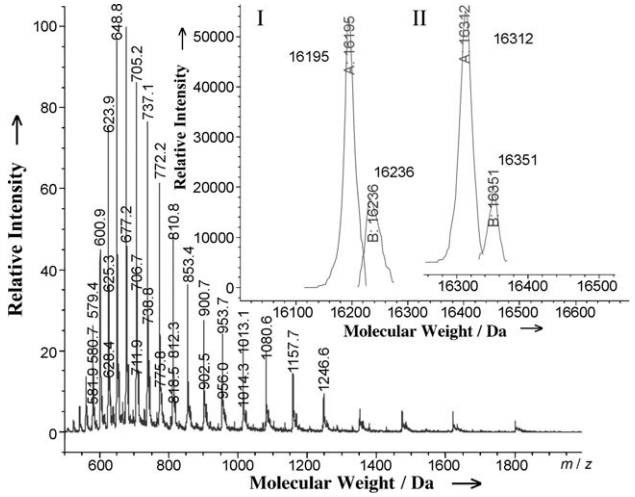


Figure 1. ESI-MS spectra of the H3K9→dehydroalanine and *S*-acetylethylamine-cysteine double methionine mutant. Inset I shows the deconvoluted spectrum of the dehydroalanine-containing histone H3 double methionine mutant: expected mass: 16195 (without the N-terminal methionine) and 16237 Da (without the N-terminal methionine and with N-terminal acetylation); found: 16195 and 16236 Da. Inset II shows the deconvoluted spectrum of the *S*-acetylethylamine-cysteine-containing histone H3 double methionine mutant: expected mass: 16314 (without the N-terminal methionine) and 16356 Da (without the N-terminal methionine and with N-terminal acetylation); found: 16312 and 16351 Da.

at 16236 Da corresponds to the protein without the N-terminal methionine and with N-terminal acetylation). For comparison, the predicted masses of the dehydroalanine-containing histone H3 double methionine mutant is 16195 Da (without the N-terminal methionine) and 16237 (without the N-terminal methionine and with N-terminal acetylation). Again, a nearly quantitative conversion from PhSeCys was observed. Subsequent reaction with *N*-acetylcysteamine resulted in a new major peak at 16312 Da (Figure 1), in close agreement with the calculated mass of 16314 Da for the histone H3 double methionine mutant containing *S*-*N*-acetylcysteamine (without the N-terminal methionine). Clean conversion to the N^{ε}-acetyl-lysine analogue can be achieved in less than two hours under the conditions described above. These reaction conditions were then used to generate N^{ε}-monomethyl-, dimethyl-, and trimethyl-lysine analogues. Close to quantitative installation of these N^{ε}-methyl-lysine analogues in both histone H3 and its double methionine mutant was obtained (Table S2 in the Supporting Information). Given the ease of synthesis of PhSeCys-containing proteins in *E. coli* and the straightforward, efficient conversion of PhSeCys to the lysine analogues, our dehydroalanine approach allows simple access to large quantities of proteins with methylated and acetylated lysine mimetics at genetically defined sites.

While it has been reported previously that Michael addition to dehydroalanine in small peptides lacks diastereoselectivity,[11] high selectivity is likely achieved with well-folded proteins owing to the local steric environments. To further demonstrate that chemically modified H3 functions in a similar manner to naturally methylated or acetylated H3, we examined the behavior of the N^{ε}-acetyl-lysine analogue in enzymatic reactions. An S28A mutant of H3 was generated with the N^{ε}-acetyl-lysine analogue installed at the K9 position (the S28A mutant removes the other Aurora B phosphorylation site in histone H3 and simplifies analysis). The protein was treated with the HDAC3/NCOR1 complex with subsequent phosphorylation at the S10 position using Aurora B kinase (H3S10ph, Figure 2a). As a control, the S28A mutant

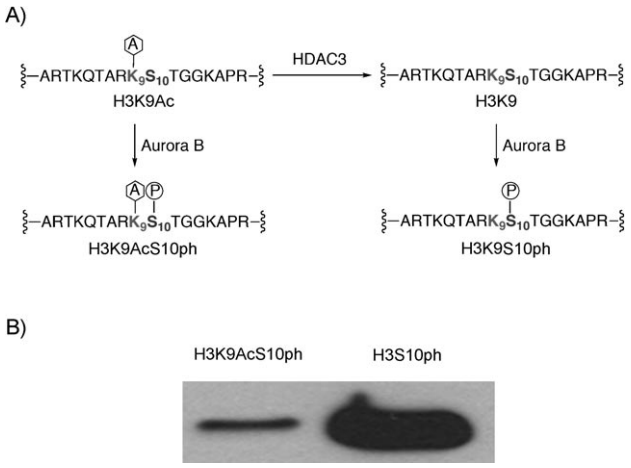


Figure 2. The H3S10 kinase Aurora B preferentially phosphorylates hypoacetylated H3. a) Histone deacetylase 3 (HDAC3) removes the acetyl group from lysine 9 of histone H3 (S28A) and facilitates phosphorylation of serine 10 by Aurora B kinase. b) Deacetylation of H3K9Ac by HDAC3 complex increased phosphorylation of H3S10 by Aurora B kinase. H3S10 phosphorylation was detected by anti-H3S10ph antibody.

H3 bearing the N^{ε}-acetyl-lysine analogue (H3K9Ac) was directly treated with Aurora B kinase. Western blot by anti-H3S10ph antibody showed significantly more phosphorylated H3 (Figure 2b) when the modified H3 was first treated with the HDAC3/NCOR1 complex. This result is consistent with previous reports that phosphorylation of H3S10 by Aurora B kinase prefers the deacetylated substrate.[12] The data indicate that the N^{ε}-acetyl-lysine analogue is a substrate of HDAC3 and that the resulting deacetylated H3 can be phosphorylated by Aurora B kinase. Moreover, the results also indicate that both chemically modified H3 with oxidized methionine and H3 with double methionine to leucine mutants (data not shown) function similarly to the naturally acetylated H3.

In conclusion, we have developed a new chemoselective method to site- and degree-specifically install modified lysine

analogues into recombinant proteins. All chemical modification reactions are carried out at room temperature under mild conditions. Since this method is based on the orthogonal reactivity of the unnatural amino acid PhSeCys, it can be used along with the two previously reported cysteine-modification strategies to install different lysine analogues simultaneously into proteins. This approach provides a facile method to obtain homogeneously methylated or acetylated histones, which can be readily used to investigate the biochemical mechanisms by which chromatin structure and function are modulated by lysine modifications.

Experimental Section

To express mutant *Xenopus laevis* histone H3, plasmid pET-H3K9TAG was cotransformed with pSup vector containing two copies of PhSeRS into *E. coli* BL21(DE3). Cells were amplified in LB media (5 mL) supplemented with ampicillin (100 $\mu g\,mL^{-1}$) and chloramphenicol (50 $\mu g\,mL^{-1}$). The starter culture (5 mL) was used to inoculate LB (100 mL) supplemented with appropriate antibiotics and grown at 37 °C. When cells reached OD600 0.6, PhSeCys was added to 1 mM and at the same time, protein expression was induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After another 4–6 h of growth at 37 °C, cells were harvested by centrifugation. The recombinant histone mutant was then purified by Ni–NTA affinity chromatography under denaturing conditions.

To install modified lysine analogues, 10 μM recombinant histone was treated with 100 mM H_2O_2 for 40 min at room temperature in pH 8.5, 50 mM Tris buffer. The protein sample was then dialyzed to remove H_2O_2 and phenylselenic acid. After oxidative elimination, 10 μM protein was allowed to react with 100 mM methylated or acetylated cysteamine in 50 mM Tris buffer, pH 8.8, at room temperature for two hours. The modified histone was then purified from excess reagents by dialysis.

Modified H3 (1 μg) was incubated with HDAC3/NCOR1 (6 $ng\,\mu l^{-1}$, Biomol) complex in HDAC assay buffer (50 mM Tris, pH 8.0, 137 mM NaCl, 2.7 mM KCl, 1 mM $MgCl_2$, and 0.1 $mg\,mL^{-1}$ bovine serum albumin) at 37 °C for 10 min. The reaction mixture was then subjected to diafiltration (3×) against kinase assay buffer (50 mM Tris, pH 7.5, 1 mM ethylene glycol tetraacetic acid, 38 mM $MgCl_2$, 250 μM ATP, and 15 mM dithiothreitol) and concentrated. The protein sample was subsequently incubated with Aurora B kinase (1 μg, Millipore) at 30 °C for 15 min. The reaction mixture was diluted with 2X sodium dodecyl sulfate (SDS) loading buffer and analyzed by western blot.

Received: May 19, 2008
Published online: July 15, 2008

Keywords: dehydroalanine · histones · lysine derivatives · protein modifications

[1] T. Jenuwein, C. D. Allis, *Science* **2001**, *293*, 1074.
[2] P. Trojer, D. Reinberg, *Cell* **2006**, *125*, 213.
[3] B. M. Turner, *Cell* **2002**, *111*, 285.
[4] S. He, D. Bauman, J. S. Davis, A. Loyola, K. Nishioka, J. L. Gronlund, D. Reinberg, F. Meng, N. Kelleher, D. G. McCafferty, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12033.
[5] M. A. Shogren-Knaak, C. J. Fry, C. L. Peterson, *J. Biol. Chem.* **2003**, *278*, 15744.
[6] M. A. Shogren-Knaak, C. L. Peterson, *Methods Enzymol.* **2004**, *375*, 62.
[7] 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.
[8] G. J. L. Bernardes, J. M. Chalker, J. C. Errey, B. G. Davis *J. Am. Chem. Soc.* **2008**, *130*, 5052.
[9] H. Neumann, S. Y. Peak-Chew, J. W. Chin, *Nat. Chem. Biol.* **2008**, *4*, 232.
[10] J. Wang, S. M. Schiller, P. G. Schultz, *Angew. Chem.* **2007**, *119*, 6973; *Angew. Chem. Int. Ed.* **2007**, *46*, 6849.
[11] Y. Zhu, W. A. van der Donk, *Org. Lett.* **2001**, *3*, 1189.
[12] Y. Li, G. D. Kao, B. A. Garcia, J. Shabanowitz, D. F. Hunt, J. Qin, C. Phelan, M. A. Lazar1, *Genes Dev.* **2006**, *20*, 2566.

