BIOCHEMISTRY
including biophysical chemistry & molecular biology

Rapid Report

pubs.acs.org/biochemistry

A Facile Method to Synthesize Histones with Posttranslational Modification Mimics

Zhiyong U. Wang,[#] Yane-Shih Wang,[#] Pei-Jing Pai, William K. Russell, David H. Russell, and Wenshe R. Liu*

Department of Chemistry, Texas A&M University, College Station, Texas 77843, United States

Supporting Information

ABSTRACT: Using an evolved pyrrolysyl-tRNA synthetase-tRNA^{Pyl} pair, a *Se*-alkylselenocysteine was genetically incorporated into histone H3 with a high protein expression yield. Quantitative oxidative elimination of *Se*-alkylselenocysteine followed by Michael addition reactions with various thiol nucleophiles generated biologically active mimics of H3 with posttranslational modifications including lysine methylation, lysine acetylation, and serine phosphorylation.

Epigenetic changes are crucial for the development and differentiation of the various cell types in an organism and typically involve postreplicational modifications to DNA[1] and posttranslational modifications to proteins that are closely associated with DNA.[2] Among posttranslational modifications of proteins, those on histones are probably the most diversified.[3] Modifications of histones are central to the regulation of chromatin dynamics and regulate biological processes such as replication, repair, transcription, and genome stability. The importance and complexity of histone modifications has led to the coining of the term "the histone code". In order to decipher the histone code, recently several methods have been introduced to synthesize histones with posttranslational modifications. Native chemical ligation and its derivative, expressed protein ligation, have been combined together with solid phase peptide synthesis to synthesize acetylated and ubiquitinated histones.[4,5] Although elegant, the synthesis generally takes lengthy steps and also yields only small quantity of histones. Another method to synthesize histones with posttranslational modifications is to site-specifically incorporate modified amino acids directly into histones at amber mutation sites during protein translation using evolved pyrrolysyl-tRNA synthetase (PylRS)-tRNA^{Pyl} pairs.[6−8] High purity histones can be obtained in large quantities this way, but each time a different PylRS-tRNA^{Pyl} pair needs to be evolved for a specific modification. In a more diversified approach, dehydroalanine (Dha) precursors can be genetically installed at designated sites followed by reactions with thiol-containing molecules to install histones with different posttranslational mimics.[9,10] Using the genetic incorporation of phenylselenocysteine (PheSec) followed by oxidative elimination to generate Dha that then underwent Michael additions with a series of cysteamine derivatives, Schultz and co-workers showed that histones with acetylated and methylated lysine mimics could be synthesized.[9] We have been primarily focusing on the Dha-based approach for its easy access to histones with multiple types of modifications, among which N^{ϵ},N^{ϵ}-dimethyllysine and N^{ϵ},N^{ϵ},N^{ϵ}-trimethyllysine have not been directly installed into histones through the genetic NAA incorporation approach. Along this line, Schultz's method of oxidatively eliminating PheSec in a histone protein to install Dha is straightforward. However the genetic incorporation of PheSec has relatively low efficiency. For wild type histone H3, its expression in *Escherichia coli* can reach to 300 mg/L in LB medium. But the expression of H3 incorporated with PheSec could only reach to 15 mg/L in LB medium under the same conditions. The low solubility of PheSec (up to 1 mM)[9] at the physiological pH also excludes the possibility of improving its incorporation level by increasing its concentration. For these reasons, we have been searching alternative methods to incorporate Dha into histones for the site-specific installation of posttranslational mimics on them.

We previously showed that an evolved *Methanosarcina mazei* PylRS, mKRS1 together with tRNA^{Pyl} allows the specific incorporation of N^{ϵ}-Cbz-lysine[11] (**1** in Figure 1) at an amber

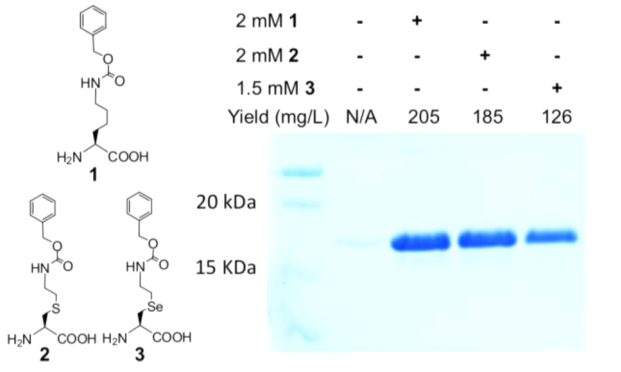


Figure 1. The site-specific incorporation of **1**, **2**, and **3** at the K9 position of H3. Proteins were expressed in BL21 cells transformed with pEVOL-mKRS1-pylT and pET-H3K9TAG in LB medium supplemented with 2 mM **1**, 2 mM **2**, or 1.5 mM **3**.

mutation site of a protein in *E. coli*.[7] This mutant enzyme also shows high substrate promiscuity and recognizes **2** and **3** shown in Figure 1. When growing BL21 cells transformed with two plasmids pEVOL-mKRS1-pylT and pET-H3K9TAG that

Received: April 25, 2012
Revised: June 12, 2012
Published: June 14, 2012

carry genes coding $tRNA^{Pyl}$, mKRS1 with an additional Y384F mutation and *Xenopus laevis* H3 with an amber mutation at its K9 position in LB medium supplemented with **1**, **2**, or **3**, overexpression of H3 was observed. The expression levels for all three conditions are above 100 mg/L. Given that the expression level of wild type H3 at the same condition was 300 mg/L, these expression levels represented more than 40% of that of wild type H3 and are significantly better than the incorporation level of PheSec. When either **2** or **3** was not provided in the medium, only a trace amount of H3 could be observed. As shown in Table 1, the purified proteins all showed

Table 1. Theoretical and Detected Molecular Weights of Various H3 Proteins[a]

proteins[b]	theoretic MW_{avg} (Da)	detected MW_{avg} (Da)
H3K9-**2**	16423[c], 16465[d]	16423[c], 16466[d]
H3K9-**3**	16470[c], 16512[d]	16470[c], 16513[d]
H3K9Dha	16228[e], 16244[f] 16270[g], 16286[h]	16229[e], 16244[f] 16287[h]
H3K9AcsK	16347[e], 16363[f] 16389[g], 16405[h]	16363[f]
H3K9m^1sK	16319[e], 16335[f] 16361[g], 16377[h]	16318[e], 16334[f] 16361[g], 16377[h]
H3K9m^2sK	16333[e], 16349[f] 16375[g], 16391[h]	16332[e], 16348[f] 16391[h]
H3K9m^3sK	16348[e], 16364[f] 16390[g], 16406[h]	16348[e], 16364[f]
H3K9pC	16342[e], 16358[f] 16384[g], 16400[h]	16341[e], 16357[f]

[a]ESI-MS and FT-ICR-MS spectra of these proteins are shown as SI Figures 4−10 and 15−18. [b]Please see main text for the definitions of these abbreviations. [c]A full-length protein without M1. [d]A full-length protein without M1 but containing the N-terminal acetylation. [e]A full-length protein without M1 but containing one oxidized methionine residue. [f]A full-length protein without M1 but containing two oxidized methionine residues. [g]A full-length protein without M1 but containing one oxidized methionine residues and the N-terminal acetylation. [h]A full-length protein without M1 but containing two oxidized methionine residues and the N-terminal acetylation.

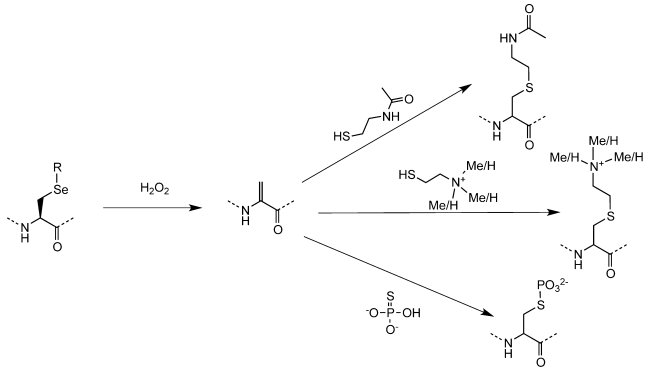
the expected molecular weight (MW) when analyzed by electrospray ionization mass spectrometry (ESI-MS). To independently confirm the incorporation of **3**, **3** was also genetically incorporated at the Q204 position of GFP_{UV}. The purified protein was digested by Asp-N protease and the target fragment N-DNHYLSTXSALSK-C (X denotes **3**) was analyzed and confirmed by tandem mass spectrometry (SI Figure 2). Importantly, the isotope distribution pattern observed for the **3**-containing peptide fragments matches the theoretical selenium isotope distributions (SI Figure 3), further confirming the incorporation of **3**.

It was previously demonstrated by Davis and co-workers that an *S*-alkylcysteine in a protein is susceptible to oxidative elimination to generate Dha in the presence of *O*-mesitylenesulfonylhydroxylamine (MSH).[10] To convert **2** in H3K9-**2** to Dha under forcing conditions, H3K9-**2** in 8 M urea, pH 8.0 was reacted with 1000 equiv of MSH for 0.5 h at room temperature, but no conversion was observed from the ESI-MS analysis of the final reaction product. When the reaction time was prolonged, unexpected products were observed, indicating that MSH reacted with other amino acids in the protein. Since H3K9-**2** is fully denatured, the structure rigidity is not a factor that may influence the conversion.[12] At this stage, we suspect the somewhat bulky long side chain in **2** may inhibit attack from MSH.

To convert **3** in H3K9-**3** to Dha, we tried two oxidative reagents, sodium periodate and H_2O_2. Although it was reported that sodium periodate can effectively convert PheSec in peptides to Dha, it apparently did not work on H3K9-**3**.[13] Only adducts with single and multiple oxygen atom(s) added to the protein were detected. On the other hand, oxidative elimination of H3K9-**3** using H_2O_2 proceeded smoothly. In the presence of 100 equiv of H_2O_2, **3** in H3K9-**3** was efficiently converted to Dha in 1 h at room temperature. ESI-MS analysis of the final product also indicated the oxidation of the two methionine residues in H3K9-**3** to methionine sulfoxide (SI Figure 9). The detected molecular mass (16244 and 16287 Da) agreed well with the calculated mass of H3 with Dha incorporated at its K9 position (H3K9Dha) that also has two oxidized methionine residues (16244 and 16286 Da). There is also one peak at 12229 Da that matched the mass of H3K9Dha with only one oxidized methionine residue (calculated mass: 16228 Da). Similar methionine oxidations were observed with the H_2O_2 oxidation of PheSec-containing H3 and could be circumvented by replacing methionine residues with leucine residues.[9]

With an effective strategy to install the Dha precursor in H3 at hand, we then explored conditions to synthesize histone mimics with various posttranslational modifications (Scheme 1). To generate the acetylated histone mimic, H3K9Dha was

Scheme 1. Oxidative Elimination Reactions to Generate Dha that Undergoes Michael Addition Reactions To Form Different Posttranslational Modification Mimics



incubated together with 1500 equiv of *N*-acetylcysteamine at 25 °C for 1 h. ESI-MS analysis of the final product (H3K9AcsK) indicated the quantitative conversion of Dha to N^{ϵ}-acetylthialysine (AcsK) (Table 1). Top-down tandem MS analysis of the final product also confirmed the site-specific installation of AcsK at the K9 position (SI Figures 11−12).[14,15] This top-down tandem MS analysis also confirmed an additional oxidation at M120 (SI Figures 13−14). We performed similar reactions to convert Dha in H3K9Dha to N^{ϵ}-methylthialysine (m^1sK), N^{ϵ},N^{ϵ}-dimethylthialysine (m^2sK), and N^{ϵ},N^{ϵ},N^{ϵ}-trimethylthialysine (m^3sK) by reacting H3K9Dha with the corresponding cysteamine derivatives to afford H3K9m^1sK, H3K9m^2sK, and H3K9m^3sK. The final products were analyzed by Fourier transform ion cyclotron resonance (FT-ICR) MS. As shown in Table 1, the detected mass agreed well with the calculated mass. Finally a phosphocysteine (pC)-containing histone analogue was made by reacting H3K9Dha with 1000 equiv of thiophosphate for 1 h at room temperature. FT-ICR-MS analysis of the final product H3K9pC indicated a quantitative conversion. This chemically installed pC was stable at the physiological pH. Storing the final product in a 4 °C fridge for a week did not cause any significant degradation.

dx.doi.org/10.1021/bi300535a | *Biochemistry* 2012, 51, 5232−5234

To demonstrate that m^{1}sK, m^{2}sK, and m3sK can closely mimic their naturally counterparts, N^ϵ-methyllysine, N^ϵ,N^ϵ-dimethyllysine, and N^ϵ,N^ϵ,N^ϵ-trimethyllysine, H3K9m^{1}sK, H3K9m^{2}sK, and H3K9m^{3}sK were immunoprecipitated by heterochromatin protein 1, HP-1β that specifically binds to H3 with mono-, di-, or trimethylated lysine at K9 but not wild type H3. The immunoprecipitates were then probed by anti-H3 antibody that specifically recognizes the C-terminal domain of H3. All three proteins were pulled out by HP-1β and detected by anti-H3 antibody (Figure 2). As a control, wild type H3 was

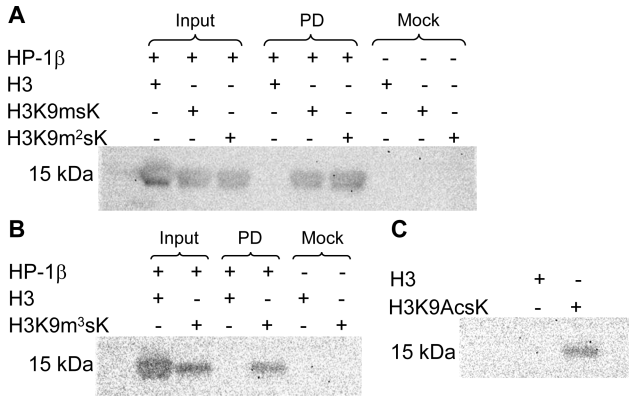


Figure 2. (A) Immunoprecipitation assay of wild type H3, H3K9m^{1}sK, and H3K9m^{2}sK using HP-1β. Each protein was incubated with HP-1β and anti-HP-1β antibody and pulled down by immobilized protein A resins that bind the Fc region of anti-HP-1β. PD represents pull down. (B) Immunoprecipitation assay of wild type H3 and H3K9m^{3}sK using HP-1β. (C) Western blot analysis of wild type H3 and H3K9AcsK probed by anti-H3K9Ac antibody.

not immunoprecipitated by HP-1β at all. To demonstrate N^ϵ-acetylthialysine closely mimics N^ϵ-acetyllysine, H3K9AcsK was probed by anti-H3K9Ac antibody that was raised against H3 with N^ϵ-acetyllysine at K9 in a Western blot analysis. The intense detected band for H3K9AcsK proved the specific interaction between H3K9AcsK and anti-H3K9Ac antibody (Figure 2C). A similar assay for wild type H3 did not give any detectable signal. All these experiments clearly demonstrated that our histone mimics with various degrees of methylation and acetylation are biologically active analogues of their natural counterparts and thus are practically useful.

In conclusion, we have developed a general method to synthesize biologically active histone mimics with posttranslational modifications including acetylation, methylation, and phosphorylation in large quantities and high purity. Compound **3** could be readily synthesized in multigram scale and genetically incorporated into representative histone H3 with protein expression yield up to 126 mg/L. The high expression yield reaches more than 40% of that of the natural histones and is among the highest yields that have been achieved so far with mutated histones.[9] Oxidative conversion of **3** to dehydroalanine and subsequent Michael addition reactions with readily available thiol nucleophiles could be conveniently done under mild nondenaturing conditions in near quantitative yield for each step to install many desired posttranslational modifications in histones. The nature of the amber mutation incorporation approach makes it possible to specifically install posttranslational modifications at any desired position of a target histone of interest. We anticipate the application of our method in the histone biology research area will facilitate addressing fundamental questions such as how posttranslational modifications affect the chromatin structure, how these modifications regulate the interactions between chromatin and other non-chromatin regulatory proteins, and how modifications on chromatin cross regulate each other. One limitation with our approach is the lack of stereocontrol on the newly formed unnatural amino acid mimic from Michael addition of Dha by thiol nucleophiles. We however still deem the strategy to have considerable utility given its methodological simplicity and versatility. In addition, reports have also proven that similarly modified proteins from Dha are indeed biologically active.[9,10]

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Synthesis of **2** and **3**, plasmid constructions, protein expression, ESI-MS and FT-ICR-MS analysis of proteins, and supplementary figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: wliu@chem.tamu.edu. Telephone: 979-845-1746.

Author Contributions

#These authors contributed equally.

Funding

This work was supported in part by the National Institute of Health (Grant 1R01CA161158 to W.R.L.), the Welch Foundation (Grant A-1715 to W.R.L.), and National Science Foundation (Grant DBI0821700 to D.H.R.).

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Jones, P. A., and Takai, D. (2001) *Science 293*, 1068−1070.
(2) Turner, B. M. (2002) *Cell 111*, 285−291.
(3) Kouzarides, T. (2007) *Cell 128*, 693−705.
(4) Shogren-Knaak, M., Ishii, H., Sun, J. M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006) *Science 311*, 844−847.
(5) McGinty, R. K., Kim, J., Chatterjee, C., Roeder, R. G., and Muir, T. W. (2008) *Nature 453*, 812−816.
(6) Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2008) *Nat. Chem. Biol. 4*, 232−234.
(7) Wang, Y. S., Wu, B., Wang, Z., Huang, Y., Wan, W., Russell, W. K., Pai, P. J., Moe, Y. N., Russell, D. H., and Liu, W. R. (2010) *Mol. Biosyst. 6*, 1557−1560.
(8) Groff, D., Chen, P. R., Peters, F. B., and Schultz, P. G. (2010) *ChemBiochem 11*, 1066−1068.
(9) Guo, J., Wang, J., Lee, J. S., and Schultz, P. G. (2008) *Angew. Chem., Int. Ed. 47*, 6399−6401.
(10) Bernardes, G. J., Chalker, J. M., Errey, J. C., and Davis, B. G. (2008) *J. Am. Chem. Soc. 130*, 5052−5053.
(11) Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., and Yokoyama, S. (2008) *Chem. Biol. 15*, 1187−1197.
(12) Chalker, J. M., Gunnoo, S. B., Boutureira, O., Gerstberger, S. C., Fernandez-Gonzalez, M., Bernardes, G. J. L., Griffin, L., Hailu, H., Schofield, C. J., and Davis, B. G. (2011) *Chem. Sci. 2*, 1666−1676.
(13) Zhu, Y., and van der Donk, W. A. (2001) *Org. Lett. 3*, 1189−1192.
(14) Reid, G. E., and McLuckey, S. A. (2002) *J. Mass. Spectrom. 37*, 663−675.
(15) Zinnel, N. F., Pai, P. J., and Russell, D. H. (2012) *Anal. Chem. 84*, 3390−3397.
(16) Chalker, J. M., Lercher, L., Rose, N. R., Schofield, C. J., and Davis, B. G. (2012) *Angew. Chem. Int. Ed. 51*, 1835−1839.