Cell PRESS

Chemistry & Biology
Brief Communication

Genetically Directing ε-N, N-Dimethyl-L-Lysine in Recombinant Histones

Duy P. Nguyen,[1] Maria M. Garcia Alai,[1] Satpal Virdee,[1] and Jason W. Chin[1,*]
[1]Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 0QH, UK
*Correspondence: chin@mrc-lmb.cam.ac.uk
DOI 10.1016/j.chembiol.2010.07.013

SUMMARY

A molecular understanding of the biological phenomena orchestrated by lysine N^{ε}-methylation is impeded by the challenge of producing site-specifically and quantitatively methylated histones. Here, we report a general method that combines genetic code expansion and chemoselective reactions, for the quantitative, site-specific installation of dimethyllysine in recombinant histones. We demonstrate the utility of our method by preparing H3K9me2 and show that this modified histone is specifically recognized by heterochromatin protein 1 beta. Extensions of the strategy reported here will allow a range of chemoselective reactions (which have been used for residue-selective, but not site-selective protein modification) to be leveraged for site-specific protein modification.

INTRODUCTION

The N^{ε}-modification status of specific lysine residues on histone proteins in chromatin regulates genomic imprinting, X chromosome inactivation, transcriptional regulation, heterochromatin formation, and DNA repair and may define epigenetic status (Köhler and Villar, 2008; Martin and Zhang, 2005; Spivakov and Fisher, 2007). Lysine residues in histones are subject to reversible acetylation, ubiquitination, and mono-, di-, and trimethylation. The methylation status of lysine residues is controlled by methylases and demethylases and the state and site of methylation may dictate biological fate in ways that are beginning to be deciphered (Strahl and Allis, 2000).

A molecular understanding of the biological changes that may be orchestrated by lysine N^{ε}-methylation is impeded by the challenge of producing site-specifically and quantitatively methylated histones. Methyltransferases have been used to methylate histones, but it is challenging to control the site, or extent or degree of methylation using these enzymes in vitro (Martino et al., 2009). Furthermore, this approach requires the availability of specific methyltransferase enzymes, many of which have not yet been identified. Native chemical ligation has been used to construct histones with modified N-terminal tails (Shogren-Knaak et al., 2003, 2006), and the role of H2B ubiquitination in the histone core has recently been investigated via an extension of this approach using multiple ligations (McGinty et al., 2008, 2009). Thioether analogs of methylated lysines (Guo et al., 2008; Simon et al., 2007), in which the γ methylene unit of lysine is replaced by a sulfur atom, can be installed in a protein. But, while these analogs are simple to install they are 0.3 Å longer than the native amino acid, decrease the pK_a of the ammonium proton by 1.1 unit, and have more degrees of freedom, which may lead to altered specificity or affinity in binding interactions (Gellman, 1991; Gloss and Kirsch, 1995). The differences between the analogs and the natural modification make them potentially problematic for discovering unknown properties of the natural system or explaining known phenomena in molecular detail.

We recently reported general methods that allow two important histone modifications, N^{ε}-acetyl-L-lysine, and N^{ε}-methyl-L-lysine, to be site-specifically incorporated in recombinant proteins (Neumann et al., 2008; Nguyen et al., 2009). We evolved an orthogonal pyrrolysyl-tRNA synthetase/$tRNA_{CUA}$ pair to incorporate N^{ε}-acetyl-L-lysine in response to the amber codon (Neumann et al., 2008, 2009), while to install N^{ε}-methyl-L-lysine we genetically directed the incorporation of N^{ε}-*tert*-butyloxycarbonyl-N^{ε}-methyl-L-lysine using an orthogonal pyrrolysyl-tRNA synthetase/$tRNA_{CUA}$ pair and subsequently removed the *tert*-butyloxycarbonyl group under mild conditions to reveal N^{ε}-methyl-L-lysine in the protein (Nguyen et al., 2009). We have begun to use these methods to provide a molecular understanding of the roles of acetylation and monomethylation on histone and nonhistone proteins in regulating cell and organism function (Lammers et al., 2010; Neumann et al., 2009; Zhao et al., 2010).

Unfortunately, simple extensions of the methods we previously reported do not allow the site-specific installation of dimethyl-lysine in recombinant histones. The pyrrolysyl-tRNA synthetase does not accept ε-N,N-dimethyl-L-lysine as a substrate, and our efforts to evolve a pyrrolysyl-tRNA synthetase for the direct genetic encoding of ε-N,N-dimethyl-L-lysine, essentially as previously described for acetyl-lysine, did not yield specific enzymes (data not shown). Moreover, we cannot extend the method used for installing N^{ε}-methyl-L-lysine to installing ε-N,N-dimethyl-L-lysine because the required quaternary amine-N^{ε}-*tert*-butyloxycarbonyl-N^{ε}-dimethyl-L-lysine is not stable.

It is well established that the amino groups of proteins can be efficiently methylated by reductive alkylation (Means and Feeney, 1968; Rayment, 1997), but it is challenging to site-specifically alkylate one lysine in a histone protein that contains numerous chemically identical lysine residues (for example, histone H3 contains 13 lysines). We realized that, by using a strategy involving a combination of genetic code expansion and chemoselective chemistry, it might be possible to tracelessly differentiate one lysine in a protein from all the others

 ©2010 Elsevier Ltd All rights reserved

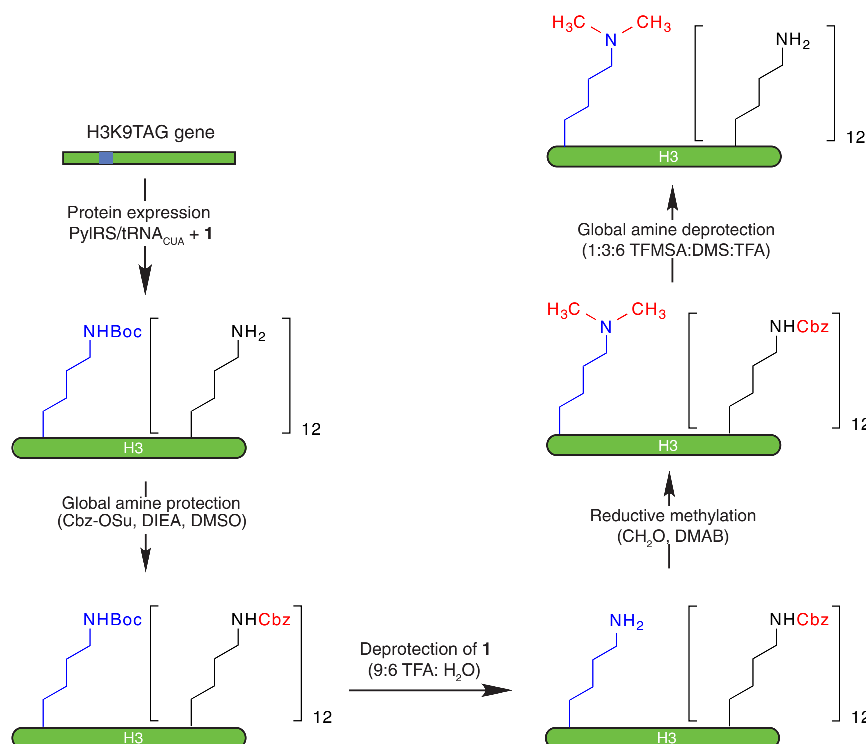

Figure 1. Strategy for Genetically Directing the Site-Specific Incorporation of Dimethyl-Lysine in Recombinant Histones
1 is N^{ε}-*tert*-butyloxycarbonyl-L-lysine, shown in blue. Cbz-OSu is N-(benzyloxycarbonyl)succinimide which reacts with the N^{ε}-amine of lysine in proteins to give N^{ε}-(benzyloxycarbonyl)-L-lysine. TFA is trifluoroacetic acid, DIEA is N,N-diisopropylethylamine, DMSO is dimethyl sulfoxide, HOSu is N-(hydroxy)succinimide, DMAB is dimethylamine borane, TFMSA is trifluoromethanesulfonic acid, and DMS is dimethyl sulfide. The N-terminal amine and its Cbz derivative are not shown for clarity.

and to direct the quantitative site-specific dimethylation of a chosen lysine residue.

RESULTS AND DISCUSSION

We reasoned that site-specific methylation of a lysine residue in a protein might be achieved by (1) replacing the codon for the lysine to be methylated with the amber codon, (2) genetically encoding a protected version of lysine in response to the amber codon, (3) quantitatively protecting the remaining amino groups in the protein under mild conditions, (4) specifically and quantitatively deprotecting the amino acid incorporated in response to the amber codon, (5) specifically and quantitatively methylating the amino group of the genetically encoded, deprotected amino acid, and (6) effecting the removal of all the remaining protecting groups under mild conditions (Figure 1).

To investigate the utility of this approach, we aimed to install dimethyl-lysine at position 9 of histone H3, which is a physiologically important methylation site (Strahl and Allis, 2000; Fischle et al., 2003; Kim et al., 2006). N^{ε}-*tert*-butyloxycarbonyl-L-lysine (**1**) (see Figure S1 available online) is a good substrate for the *Methanosarcina barkeri* pyrrolysyl-tRNA synthetase/$tRNA_{CUA}$ pair (*Mb*PylRS/*Mb*$tRNA_{CUA}$) and we have recently shown that **1** can be incorporated into recombinant histones in good yield (Nguyen et al., 2009). Moreover we have specifically removed the *tert*-butyloxycarbonyl group from N^{ε}-*tert*-butyloxycarbonyl-N^{ε}-methyl-L-lysine in recombinant histones under mild conditions (Nguyen et al., 2009). We prepared His_6-H3K9-**1** (N-terminally hexahistidine-tagged histone H3 containing **1** at position 9), with a yield of 13 mg per liter of culture, by expressing *his₆-H3K9TAG* in the presence of the *Mb*PylRS/*Mb*$tRNA_{CUA}$ pair. His_6-H3K9-**1** was purified by Ni-NTA chromatography and the N-terminal His_6 tag was removed by TEV protease to reveal H3K9-**1**. ESI-MS analysis of the purified H3K9-**1** confirms the incorporation of **1** into histone H3 (Figure 2A).

To protect the N-terminal amine and 12 epsilon amino groups of lysine residues in H3K9-**1**, we treated the purified protein with *N*-(benzyloxycarbonyloxy)succinimide (Cbz-OSu) in basic DMSO/6 M Gdn-Cl (9:1) to obtain the globally Cbz-protected sample H3K9-**1**-(Cbz_{13}). The Cbz protection was complete after 1 hr at room temperature, as demonstrated by ESI-MS of the Cbz-OSu-treated sample (Figure 2B). We revealed the free amino group of lysine 9 by treating H3K9-**1**-(Cbz_{13}) with 9:6 trifluoroacetic acid (TFA)/water at 4°C. The ESI-MS spectrum of the TFA-treated sample indicates that TFA treatment quantitatively removed the Boc group from **1**, while leaving the Cbz groups intact (Figure 2C). We methylated the epsilon amino group of lysine 9 in H3K9-(Cbz_{13}) by reductive alkylation using formaldehyde and a dimethylamine borane complex (Means and Feeney, 1968; Rayment, 1997). After 4 hr, ESI-MS analysis of the alkylated sample demonstrates that the epsilon amino group of lysine 9 is quantitatively dimethylated to yield H3K9me2-(Cbz_{13}) (Figure 2D). To remove the Cbz groups, we used a cocktail of trifluoromethanesulfonic acid (TFMSA)/trifluoroacetic acid (TFA)/dimethylsulfide (DMS) (1:3:6) for 2 hr at 0°C. Two milligrams of the deprotected protein (H3K9me2) was recovered using the material from a 1 liter expression.

H3K9me2 prepared by our method was recognized by an antibody against H3K9me2 which does not recognize H3K9 and H3K9me1 (Figure 3). Similarly H3K9me2 was not recognized by an anti-H3K9me1 antibody. These experiments confirm the presence of dimethyl-lysine at position 9 in H3. The ESI-MS spectra of the deprotected protein revealed that greater than 90% of the sample was indeed H3K9me2 and top-down MS/MS protein sequencing through the entire histone (Neumann et al., 2009) confirmed that dimethyl-lysine was exclusively incorporated at the genetically encoded site (Figures S2 and S3). Traces of incompletely deprotected protein were subsequently removed by HPLC purification (Figure 2E).

To demonstrate the protein prepared using our approach is functional in biochemical assays, we performed coimmunoprecipitations of heterochromatin protein 1 (HP1), a chromodomain-containing protein, which is known to specifically bind to

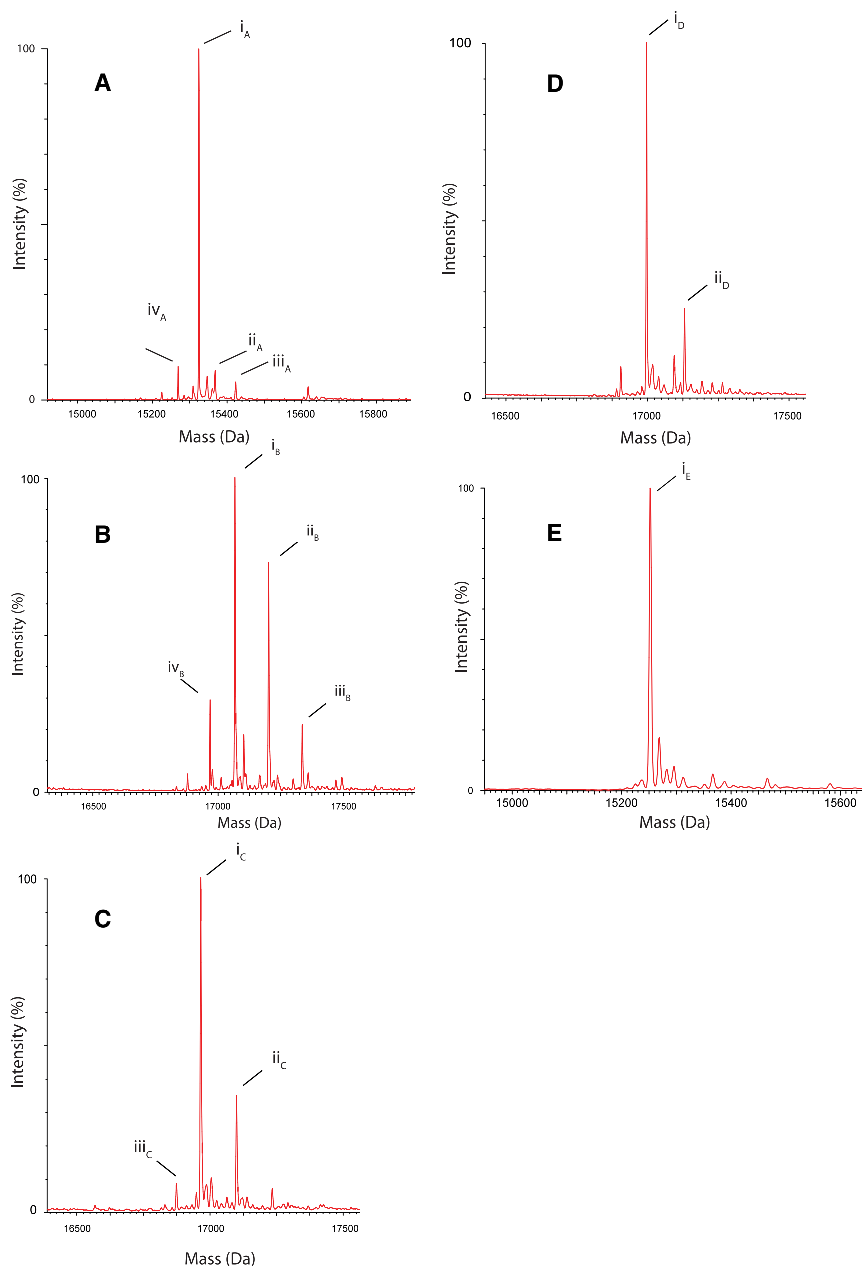

Figure 2. Synthesis and Characterization of Histone H3K9me2

(A) ESI-MS analysis of the purified H3K9-**1** produced by *Mb*PylRS/$tRNA_{CUA}$ pair revealed a mass of 15,325.5 ± 1.0 Da (peak i_A, expected mass 15,325 Da). The disodium and phosphate adducts, differing by 44 and 98 Da, respectively, are also detected in the spectrum (peak ii_A and iii_A). These counter ions are often found associated with positively charged proteins (Chowdhury et al., 1990). The peak of 15,269.5 Da (peak iv_A) may correspond to the loss of *t*-butyl from the *tert*-butyloxycarbonyl group (−56 Da) (Raju et al., 2010).

(B) All the free amines in H3K9-**1** are protected with Cbz. ESI-MS analysis of H3bocK9(Cbz_{13}) revealed a mass of 17,068.0 ± 1.0 Da (peak i_B, expected mass 17,067 Da). Two peaks each differing by 134 Da are seen in the spectrum (peak ii_B and iii_B, and these may result from the reversible protection of the imidazole functionality in histidine residues) (Sieber and Riniker, 1987).The peak of mass 16,968.5 Da (peak iv_B) corresponds to loss of Boc group (−100 Da) during electrospray ionization process (Raju et al., 2010).

(C) The Boc group of H3K9-**1**-(Cbz_{13}) is quantitatively deprotected with 60% TFA. ESI-MS analysis of TFA-treated H3K9-**1**-(Cbz_{13}) revealed a mass of 16,965.0 ± 1.0 Da (peak i_C, expected mass 16,967 Da). The minor peak of mass 17,100.0 Da (peak ii_C) corresponds to the overprotected polypeptide H3K9(Cbz_{14}). The peak of mass 16,875.0 Da (peak iii_C) may correspond to loss of benzyl group (−90 Da) during electrospray ionization process.

(D) Quantitative dimethylation of histone H3 lysine 9 was performed by reductive alkylation of H3K9 (Cbz_{13}) with formaldehyde and dimethylamine borane complex. ESI-MS analysis of methylated sample revealed a mass of 16,996.5 ± 1.0 Da (peak i_D, expected mass 16,995 Da) corresponding to the addition of two methyl groups (+28 Da). The minor peak of mass 17,131.0 Da (peak ii_D) corresponds to the overprotected polypeptide H3K9me2(Cbz_{14}).

(E) ESI-MS analysis of HPLC-purified H3K9me2 revealed a peak of 15,253.0 ± 1.0 Da (peak i_E, expected mass 15253 Da) (Figure S3).

See also Figures S1 and S2.

short peptides based on histone H3 tail containing mono-, di-, or trimethylated K9 (Fischle et al., 2003; Kim et al., 2006). HP1 immunoprecipitation of full-length H3, H3K9me1, and H3K9me2, synthesized by our approach, allows us to demonstrate that full-length H3K9me2, but not unmethylated H3, binds specifically to HP1 (Figure 3B).

In conclusion, we have demonstrated a general method for the site-specific installation of dimethyl-lysine in recombinant histones. Extensions of the strategy reported here will allow a range of chemoselective reactions (Hermanson, 1996) to be leveraged for site-specific protein modification. Indeed, we recently demonstrated that a similar approach can be used for the first synthesis of atypical ubiquitin chains (Virdee et al., 2010). We are currently using this approach in combination with methods for installing other modifications into histones to understand how combinations of posttranslational modifications program cellular outcomes and epigenetic inheritance.

SIGNIFICANCE

A molecular understanding of the biological phenomena orchestrated by lysine N^{ε}-methylation is impeded by the challenge of producing site-specifically and quantitatively methylated histones. Here, we demonstrate a general method for the quantitative, site-specific installation of dimethyl-lysine in recombinant histones. Extensions of the

 ©2010 Elsevier Ltd All rights reserved

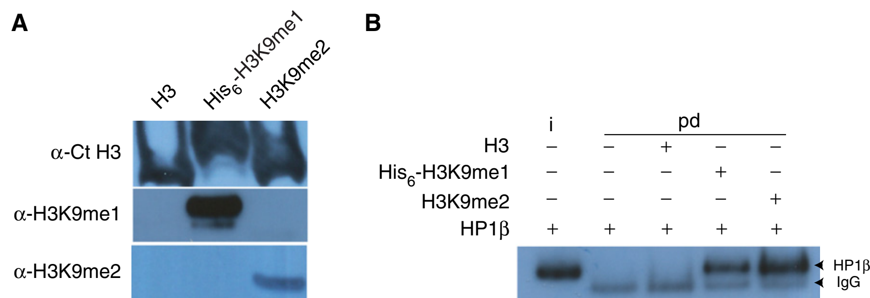

Figure 3. Characterization of H3K9me2 and Its Binding to HP1

(A) H3 detection with an antibody to the C terminus of H3, H3K9me1, and H3K9me2-specific antibodies.

(B) Interaction between HP1β and methylated H3K9. The immunoprecipitation of HP1β is described in the Experimental Procedures. Anti C terminus H3 was used to immunoprecipitate H3, H3K9me1 or H3K9me2 in the presence of HP1β. HP1β recognition was probed by immunoblotting with an anti-HP1 antibody. i, 2% of total HP1β, pd, "pull-down" assay.

See also Figure S2.

strategy reported here will allow a range of chemoselective reactions (which have been used for residue selective, but not site-selective protein modification) to be leveraged for site-specific protein modification. We are currently using our approach in combination with methods for installing other modifications into histones to understand how combinations of posttranslational modifications program cellular outcomes and epigenetic inheritance.

EXPERIMENTAL PROCEDURES

Information on protein expression and purification, mass spectrometric analysis of recombinant histones, and immunoprecipitation of HP1β can be found in the Supplemental Experimental Procedures.

Global Protection of Histone H3K9-1

Lyophilized histone H3K9-**1** (12 mg, 810 nmol) was dissolved in 120 μl 6 M Gdn-Cl and diluted with DMSO to 1.25 ml. DIEA (37 μl, 210 μmol) and Cbz-Osu (53 mg, 21 μmol) were then added to a vigorously stirred solution. The solution was further stirred for 1 hr at room temperature. The protected polypeptide was then precipitated with cold 1:1 MeOH/acetone (10 ml). The precipitate was collected by centrifugation, washed with cold 1:1 MeOH/acetone, and air-dried.

Deprotection of Boc Protecting Group

The globally protected protein, obtained from 12 mg of H3K9-**1**, was dissolved in cold 9:6 TFA/H_2O and incubated at 4°C for 1 hr to deprotect the genetically encoded Boc group. The protein was precipitated and washed with ice-cold ether. The aqueous and ether layers were removed, and the precipitate was allowed to air-dry.

Dimethylation of Histone H3K9-$(Cbz)_{13}$

The protein was dissolved in 2 ml of 6 M Gdn-Cl, 100 mM phosphate buffer (pH 7.5) by sonication. Dimethylation of histone H3K9-$(Cbz)_{13}$ was carried out using JBS methylation kit (Jena Bioscience). After 4 hr of reaction, the protein was precipitated with cold 1:1 MeOH/acetone (10 ml). The precipitate was collected by centrifugation, washed with cold 1:1 MeOH/acetone, and air-dried. The protein was then dissolved in an ice-cold cocktail (2 ml) of 7:11:2 DMS/TFA/TfOH and further stirred for 2 hr on ice to deprotect Cbz groups. The protein was precipitated with cold ether supplemented with 1% pyridine. The precipitate was collected by centrifugation and washed twice with cold ether. The dried precipitate was dissolved in unfolding buffer (8 M urea, 100 mM Na_2HPO_4 [pH 7.4], 500 mM NaCl). The protein was dialyzed in the same buffer overnight at room temperature and subsequently extensively dialyzed into 1 mM DTT to obtain histone H3K9me2. The overall yield of recovered H3K9me2 was approximately 2 mg (~17%). The protein was further purified by RP-HPLC using a gradient of 5%–75% Buffer B over 30 min at a flow rate of 5 ml/min. Buffer A = 0.1%TFA in H_2O, Buffer B = 10% Buffer A in acetonitrile.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and two figures and can be found with this article online at doi:10.1016/j.chembiol.2010.07.013.

ACKNOWLEDGMENTS

We are grateful to Rachel Overington for assistance (R.O. was funded by a summer studentship from Trinity College, Cambridge) and Sew Peak-Chew (MRC-LMB) for acquiring the top-down mass spectra. This work was supported by the Medical Research Council, UK, and grants from the European Research Council & the Human Frontiers of Science Program to J.W.C.

Received: June 14, 2010
Revised: July 16, 2010
Accepted: July 20, 2010
Published: October 28, 2010

REFERENCES

Chowdhury, S.K., Katta, V., Beavis, R.C., and Chait, B.T. (1990). Origin and removal of adducts (molecular mass = 98 u) attached to peptide and protein ions in electrospray ionization mass spectra. J. Am. Soc. Mass *1*, 382–388.

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev. *17*, 1870–1881.

Gellman, S.H. (1991). On the role of methionine residues in the sequence-independent recognition of nonpolar protein surfaces. Biochemistry *30*, 6633–6636.

Gloss, L.M., and Kirsch, J.F. (1995). Decreasing the basicity of the active site base, Lys-258, of Escherichia coli aspartate aminotransferase by replacement with gamma-thialysine. Biochemistry *34*, 3990–3998.

Guo, J., Wang, J., Lee, J.S., and Schultz, P.G. (2008). Site-specific incorporation of methyl- and acetyl-lysine analogues into recombinant proteins. Angew. Chem. Int. Ed. Engl. *47*, 6399–6401.

Hermanson, G.T. (1996). Bioconjugate Techniques (San Diego: Academic Press).

Kim, J., Daniel, J., Espejo, A., Lake, A., Krishna, M., Xia, L., Zhang, Y., and Bedford, M.T. (2006). Tudor, MBT and chromo domains gauge the degree of lysine methylation. EMBO Rep. *7*, 397–403.

Köhler, C., and Villar, C.B. (2008). Programming of gene expression by Polycomb group proteins. Trends Cell Biol. *18*, 236–243.

Lammers, M., Neumann, H., Chin, J.W., and James, L.C. (2010). Acetylation regulates Cyclophilin A catalysis, immunosuppression and HIV isomerization. Nat. Chem. Biol. *6*, 331–337.

Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nat. Rev. Mol. Cell Biol. *6*, 838–849.

Martino, F., Kueng, S., Robinson, P., Tsai-Pflugfelder, M., van Leeuwen, F., Ziegler, M., Cubizolles, F., Cockell, M.M., Rhodes, D., and Gasser, S.M. (2009). Reconstitution of yeast silent chromatin: multiple contact sites and O-AADPR binding load SIR complexes onto nucleosomes in vitro. Mol. Cell *33*, 323–334.

McGinty, R.K., Kim, J., Chatterjee, C., Roeder, R.G., and Muir, T.W. (2008). Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. Nature *453*, 812–816.

McGinty, R.K., Köhn, M., Chatterjee, C., Chiang, K.P., Pratt, M.R., and Muir, T.W. (2009). Structure-activity analysis of semisynthetic nucleosomes: mechanistic insights into the stimulation of Dot1L by ubiquitylated histone H2B. ACS Chem. Biol. *4*, 958–968.

Means, G.E., and Feeney, R.E. (1968). Reductive alkylation of amino groups in proteins. Biochemistry *7*, 2192–2201.

Neumann, H., Peak-Chew, S.Y., and Chin, J.W. (2008). Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. Nat. Chem. Biol. *4*, 232–234.

Neumann, H., Hancock, S.M., Buning, R., Routh, A., Chapman, L., Somers, J., Owen-Hughes, T., van Noort, J., Rhodes, D., and Chin, J.W. (2009). A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. Mol. Cell *36*, 153–163.

Nguyen, D.P., Garcia Alai, M.M., Kapadnis, P.B., Neumann, H., and Chin, J.W. (2009). Genetically encoding N(epsilon)-methyl-L-lysine in recombinant histones. J. Am. Chem. Soc. *131*, 14194–14195.

Raju, G., Ramesh, V., Sharma, G.V.M., and Shoban Babu, B. (2010). Differentiation of Boc-protected α,δ-/δ,α- and β,δ-/δ,β-hybrid peptide positional isomers by electrospray ionization tandem mass spectrometry. J. Mass Spectrom. *45* (Spec.), 651–663.

Rayment, I. (1997). Reductive alkylation of lysine residues to alter crystallization properties of proteins. Methods Enzymol. *276*, 171–179.

Shogren-Knaak, M.A., Fry, C.J., and Peterson, C.L. (2003). A native peptide ligation strategy for deciphering nucleosomal histone modifications. J. Biol. Chem. *278*, 15744–15748.

Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science *311*, 844–847.

Sieber, P., and Riniker, B. (1987). Protection of histidine in peptide synthesis: a reassessment of the trityl group. Tet. Lett. *28*, 6031–6034.

Simon, M.D., Chu, F., Racki, L.R., de la Cruz, C.C., Burlingame, A.L., Panning, B., Narlikar, G.J., and Shokat, K.M. (2007). The site-specific installation of methyl-lysine analogs into recombinant histones. Cell *128*, 1003–1012.

Spivakov, M., and Fisher, A.G. (2007). Epigenetic signatures of stem-cell identity. Nat. Rev. Genet. *8*, 263–271.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41–45.

Virdee, S., Ye, Y., Nguyen, D.P., Komander, D., and Chin, J.W. (2010). Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OUT deubiquitinase. Nat. Chem. Biol. *6*, 750–757.

Zhao, S., Xu, W., Jiang, W., Yu, W., Lin, Y., Zhang, T., Yao, J., Zhou, L., Zeng, Y., Li, H., et al. (2010). Regulation of cellular metabolism by protein lysine acetylation. Science *327*, 1000–1004.

 ©2010 Elsevier Ltd All rights reserved