ChemComm

RSC Publishing

COMMUNICATION

View Article Online
View Journal | View Issue

Synthesis of ε-*N*-propionyl-, ε-*N*-butyryl-, and ε-*N*-crotonyl-lysine containing histone H3 using the pyrrolysine system†

Cite this: *Chem. Commun.*, 2013, **49**, 379

Received 29th October 2012,
Accepted 14th November 2012

DOI: 10.1039/c2cc37836a

www.rsc.org/chemcomm

Michael J. Gattner, Milan Vrabel and Thomas Carell*

Recently new lysine modifications were detected in histones and other proteins. Using the pyrrolysine amber suppression system we genetically inserted three of the new amino acids ε-*N*-propionyl-, ε-*N*-butyryl-, and ε-*N*-crotonyl-lysine site specifically into histone H3. The lysine at position 9 (H3 K9), which is known to be highly modified in chromatin, was replaced by these unnatural amino acids.

The histone code is based on the post-translational modification of critical amino acid residues in different histones. Among these, lysine acetylations and methylations are the most abundant and affect the transcriptional status of the genes associated with the corresponding histones.[1–3] The role of these modifications is sequence dependent. Typically acetylation is associated with transcriptionally active genes, while methylation induces transcriptional silencing.[4,5] Recently Zhao and co-workers discovered a number of new modified amino acids in histones.[6–8] Some of these were also detected in proteins other than histones, raising the possibility that they are of more widespread importance.[9–11] The new post-translational modifications (PTMs) are acylated derivatives of lysine at the ε-amino position. The acylation partners are propionic acid, butyric acid, malonic acid, succinic acid, crotonic acid or fatty acids. A common characteristic of these compounds is that they are key metabolic intermediates that typically exist as CoA activated thioester species in the cell.[12] It is currently not clear how these newly discovered PTMs are biosynthetically established within the histones and we do not fully understand if and how they influence genetic processes.

So far no specific deacylases for ε-*N*-propionyl- or ε-*N*-butyryl-lysine have been identified. However Sirt5, a member of NAD-dependent sirtuins is able to specifically deacylate ε-*N*-malonyl- and ε-*N*-succinyllysine.[11] Recently it was discovered that HDAC3 exhibits decrotonylase activity *in vitro*[13] and it was found that lysine crotonylation activates genes even in a globally repressive environment.[7,14]

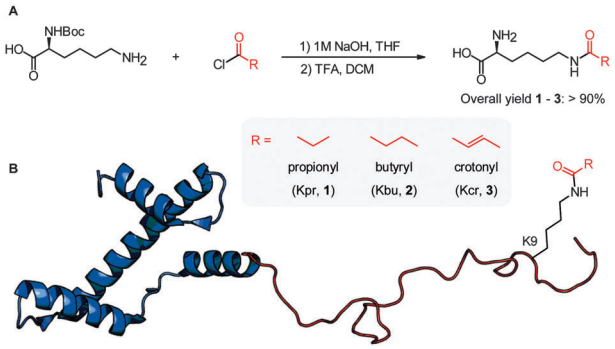
In order to enable investigation of the new lysine derivatives in histones it is essential to generate histones that contain these amino acids site specifically.[15] In this direction semi-synthetic chemical ligation based methods[16,17] were utilized and chemical methods were employed to generate acetyllysine (Kac) and methylated-lysines or derivatives thereof.[18–21] Chin and co-workers reported the introduction of Kac into H3 K56[22] using the pyrrolysine system. This system was also employed for the synthesis of monomethyl- and dimethyl-lysine containing histones.[23,24] Schultz *et al.* recently described the synthesis of histone H2B containing ε-*N*-crotonyllysine at position K11 using an evolved pyrrolysyl-tRNA synthetase from *Methanosarcina barkeri*.[25]

Here we show that the pyrrolysine system can be used to insert ε-*N*-propionyl-(Kpr, **1**), ε-*N*-butyryl-(Kbu, **2**), and ε-*N*-crotonyl-lysine (Kcr, **3**) into histones at critical positions such as H3 K9 using the wild type pyrrolysyl-tRNA synthetase. The synthesis of the three modified amino acids **1–3** is shown in Scheme 1. All three amino acids were obtained using the depicted procedure in multi-gram quantities starting from commercially available materials without the need for chromatographic purification (for synthetic details see ESI†).

The amber suppression technology was used extensively for the incorporation of various unnatural amino acids into proteins.[26] Recently we reported the development of an *E. coli* based assay that allows examination of how efficiently the pyrrolysine system from *Methanosarcina mazei* works.[27] *E. coli* cells containing a plasmid with an YFP reporter gene bearing an amber stop codon at position K114 were generated. The plasmid also contains the genes of the pyrrolysine tRNA (*pyl*T) and the corresponding synthetase (*pyl*S) or of *pyl*S mutants, respectively, that were previously shown to improve the amber suppression efficiency of bulkier pyrrolysine analogues.[28,29] Upon addition of an unnatural amino acid to the expression medium that is accepted by the tRNA–pyrrolysyl-tRNA synthetase pair ($tRNA^{Pyl}$–PylRS), full-length fluorescent YFP is produced.

Center for Integrated Protein Science at the Department of Chemistry, Ludwig-Maximilians-Universität, Munich, Butenandtstraße 5-13, D-81377, Munich, Germany. E-mail: Thomas.Carell@cup.uni-muenchen.de; Fax: +49 (0)89 2180 77756; Tel: +49 (0)89 2180 77750

† Electronic supplementary information (ESI) available: Synthesis of compounds **1–3**, preparation of expression vectors, protein expression and purification procedures, HPLC-MS/MS and HPLC-MS data, western blot analysis. See DOI: 10.1039/c2cc37836a



Scheme 1 (A) Synthesis of the three amino acids ε-*N*-propionyl-(Kpr, **1**), ε-*N*-butyryl-(Kbu, **2**), and ε-*N*-crotonyl-lysine (Kcr, **3**): (1) N_α-Boc-Lys-OH, propionyl-, butyryl or crotonyl chloride, 1 M NaOH/THF = 1/1, 0 °C to r.t., overnight; (2) DCM/TFA = 5/1, r.t., 1.5 h. (B) Depiction of histone H3 and of the position K9 in the tail where the modifications were incorporated (protein representation generated from PDB 1KX5).

Testing the three amino acids Kpr, Kbu and Kcr with this reporter system, the best results were obtained in the presence of wild type PylRS (see ESI†).

To further determine the efficiency of the incorporation we transformed *E. coli* cells with a plasmid containing the gene of thioredoxin A (trxA) with a C-terminal Strep-Tag and an amber codon at position N56 as a model protein. A second plasmid carrying the genes *pyl*T and *pyl*S was transformed. The cells were grown in the presence of 5 mM Kcr and the protein expression was induced by the addition of 1 mM IPTG at an OD_{600} of 0.6. After an overnight expression at 30 °C the cells were lysed and purified by Strep-Tag affinity chromatography. We obtained 2 mg of the Kcr containing TrxA from 1 L cell culture. The purified intact protein was analyzed by mass spectrometry. The observed protein molecular weight of 13 014 Da (calculated: 13 016 Da) confirmed the presence of the Kcr moiety in the protein. In addition tryptic digest was performed followed by HPLC-MS/MS analysis, which further showed that the modification was in the desired position within the protein structure (see ESI†).

We next investigated if the system is able to provide histone proteins containing one of the three modified lysine amino acids. To this end we used a first plasmid containing *pyl*S and three copies of *pyl*T under the control of a constitutive glutamine promoter. A second plasmid was generated containing an N-terminally His-tagged H3 gene with the amber stop codon at position K9 under the control of an IPTG-inducible T7 promoter. Both plasmids were transformed into *E. coli* BL21(DE3). We individually added the three amino acids Kpr, Kbu, and Kcr to the medium at concentrations of 5 mM for Kbu and Kcr and 10 mM for Kpr. The *E. coli* cells were grown until an OD_{600} of 0.6 and then gene expression was induced by the addition of 1 mM IPTG. After 16 h at 30 °C we harvested and lysed the cells and isolated the H3 proteins under denaturing conditions (7 M Urea) using Ni–NTA column chromatography. The purified histones containing the modifications in position 9 were refolded using standard procedures (see ESI†). In all three cases we isolated 0.5–1.0 mg of modified histone H3 per L of cell culture (for comparison we obtained 5 mg of unmodified histone H3).

In order to prove the presence of the unnatural lysine amino acids and in particular to investigate potential deacylation processes during protein expression and purification we performed HPLC-MS/MS studies. When we analyzed the tryptic digests of the H3 proteins by HPLC-MS/MS, peptides containing K9 (or the corresponding modifications) were not detected due to the position of this amino acid in a lysine rich sequence giving rise to polar and in particular too short peptides for subsequent analysis. We hence modified a method first described by Hunt *et al.*[30] The H3 proteins containing Kbu and Kcr were reacted with propionic acid anhydride (**4**) directly in the gel. For H3 proteins containing the expected Kpr we used butyric acid anhydride (**5**) instead. This chemical treatment blocks all free lysine ε-amino groups leading to longer and more hydrophobic tryptic peptides (*e.g.* KSTGGKAPR instead of STGGK for the K9 containing peptide). After in-gel tryptic digest we repeated the chemical treatment to cap also the N-terminal amino groups of the generated peptides. The peptides were then eluted from the gel and analyzed by HPLC-MS/MS. The obtained data are depicted in Fig. 1 (for MS/MS data see ESI†). For wild type histone H3, the positions K9 and K14 are either propionyl- (K9Kpr K14Kpr, as shown in Fig. 1) or butyryl-capped. The MS analyses of the corresponding modified histone H3 peptides gave the correct mass spectra for K9Kpr K14Kbu, K9Kbu K14Kpr and K9Kcr K14pr, respectively, showing the correct incorporation of the three lysine derivatives. We noted that the MS-analyzed peptides were contaminated in all three cases with

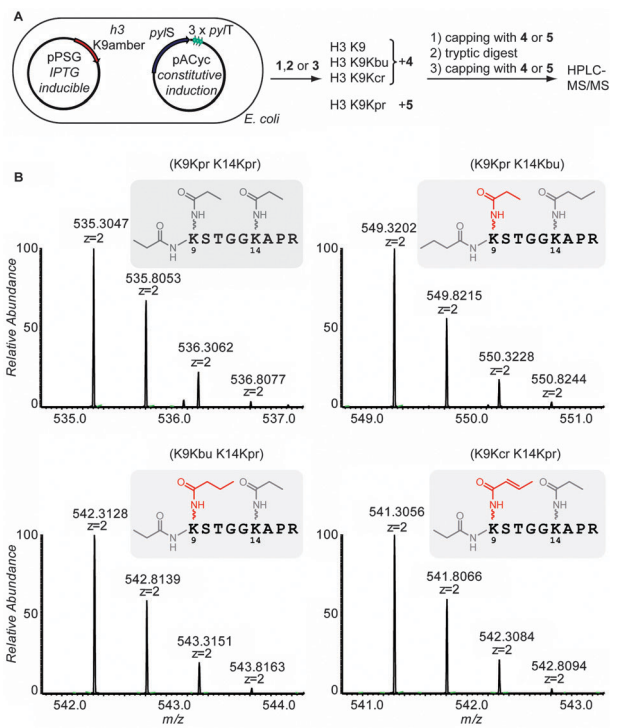


Fig. 1 (A) Schematic overview of the workflow: expression of the four histone H3 variants followed by in-gel capping of the purified proteins with propionic anhydride (**4**) or butyric anhydride (**5**), tryptic digestion, repeated in-gel capping and HPLC-MS/MS analysis. (B) MS spectra and illustrations of the capped K9 containing tryptic peptides (MS/MS spectra are shown in ESI†). The capping modifications are shown in grey while the incorporated PTMs are shown in red.

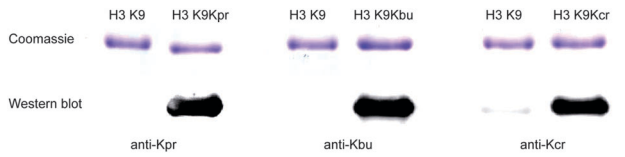


Fig. 2 Western blot analysis of the three synthetic histone variants H3 K9Kpr, H3 K9Kbu and H3 K9Kcr compared to unmodified histone H3 using the corresponding anti-Kpr, -Kbu and -Kcr antibodies.

peptides in which the ε-*N*-modified lysines were replaced by lysine. These peptides gave MS-signals for all-propionylated or all-buturylated peptides, showing that the acyl groups were missing at H3 K9 already before the peptides were analyzed by mass spectrometry. This discovery excludes the possibility that the acyl groups were cleaved during MS-measurement.

We can also exclude that unmodified lysine was inserted during amber suppression by potentially lysine-mischarged $tRNA^{Pyl}$. When we performed the protein expressions in medium lacking the unnatural amino acids, no YFP or histone H3 was produced, showing that successful amber suppression depends on the presence of these amino acid derivatives. We consequently speculate that *E. coli* contains one or more proteins (such as the bacterial sirtuin CobB) able to unspecifically deacylate Kpr- and Kbu-modified histone tails as previously observed for Kac[31] and Kcr.[25] In order to test this hypothesis we repeated the expression experiment in the presence of 20 mM nicotine amide (NAM), a known inhibitor of sirtuins.[32] After careful MS analysis we indeed found a strongly reduced amount of peptides containing natural lysine before the capping. We could reduce the deacylation of Kpr, Kbu and Kcr containing histones from about 20% to below 5% after addition of NAM.

To further verify the presence of the incorporated PTMs we performed a western blot analysis using polyclonal antibodies (PTM BioLabs) against the three amino acids Kpr, Kbu and Kcr. As shown in Fig. 2 the desired modifications were selectively detected by the corresponding antibodies while the unmodified histone H3 was not recognized. The antibodies show compromised selectivities in cross labeling experiments (see ESI†).

In summary, we show that by using the pyrrolysine system the newly discovered lysine derivatives ε-*N*-propionyl-, ε-*N*-butyryl-, and ε-*N*-crotonyl-lysine can be inserted into histones (here H3) at critical positions (here K9) using the wild type pyrrolysyl-tRNA synthetase. The observed deacylation in *E. coli* can be reduced by the addition of nicotine amide and thus homogenous proteins bearing these PTMs are now available. Histones containing defined lysine modifications will enable us to further study the biochemistry behind the acylation and deacylation processes in order to learn how the modification chemistry influences gene activity.

This work was supported by the Excellence Cluster CIPSM, the Volkswagen Foundation and the DFG SFB 1032. M.V. thanks the LMU Munich for postdoctoral fellowship.

Notes and references

1 E. I. Campos and D. Reinberg, *Annu. Rev. Genet.*, 2009, **43**, 559–599.
2 C. T. Walsh, S. Garneau-Tsodikova and G. J. Gatto Jr., *Angew. Chem., Int. Ed.*, 2005, **44**, 7342–7372.
3 M. Biel, V. Wascholowski and A. Giannis, *Angew. Chem., Int. Ed.*, 2005, **44**, 3186–3216.
4 B. M. Turner, *Cell*, 2002, **111**, 285–291.
5 S. Khorasanizadeh, *Cell*, 2004, **116**, 259–272.
6 Y. Chen, R. Sprung, Y. Tang, H. Ball, B. Sangras, S. C. Kim, J. R. Falck, J. Peng, W. Gu and Y. Zhao, *Mol. Cell. Proteomics*, 2007, **6**, 812–819.
7 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 and Y. Zhao, *Cell*, 2011, **146**, 1016–1028.
8 Z. Xie, J. Dai, L. Dai, M. Tan, Z. Cheng, Y. Wu, J. D. Boeke and Y. Zhao, *Mol. Cell. Proteomics*, 2012, **11**, 100–107.
9 J. Garrity, J. G. Gardner, W. Hawse, C. Wolberger and J. C. Escalante-Semerena, *J. Biol. Chem.*, 2007, **282**, 30239–30245.
10 Z. Zhang, M. Tan, Z. Xie, L. Dai, Y. Chen and Y. Zhao, *Nat. Chem. Biol.*, 2011, **7**, 58–63.
11 J. Du, Y. Zhou, X. Su, J. J. Yu, S. Khan, H. Jiang, J. Kim, J. Woo, J. H. Kim, B. H. Choi, B. He, W. Chen, S. Zhang, R. A. Cerione, J. Auwerx, Q. Hao and H. Lin, *Science*, 2011, **334**, 806–809.
12 H. Lin, X. Su and B. He, *ACS Chem. Biol.*, 2012, **7**, 947–960.
13 A. S. Madsen and C. A. Olsen, *Angew. Chem., Int. Ed.*, 2012, **51**, 9083–9087.
14 E. Montellier, S. Rousseaux, Y. Zhao and S. Khochbin, *Bioessays*, 2012, **34**, 187–193.
15 C. A. Olsen, *Angew. Chem., Int. Ed.*, 2012, **51**, 3755–3756.
16 P. E. Dawson and S. B. Kent, *Annu. Rev. Biochem.*, 2000, **69**, 923–960.
17 T. W. Muir, D. Sondhi and P. A. Cole, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 6705–6710.
18 R. K. McGinty, J. Kim, C. Chatterjee, R. G. Roeder and T. W. Muir, *Nature*, 2008, **453**, 812–816.
19 J. Guo, J. Wang, J. S. Lee and P. G. Schultz, *Angew. Chem., Int. Ed.*, 2008, **47**, 6399–6401.
20 M. D. Simon, F. Chu, L. R. Racki, C. C. de la Cruz, A. L. Burlingame, B. Panning, G. J. Narlikar and K. M. Shokat, *Cell*, 2007, **128**, 1003–1012.
21 J. M. Chalker, L. Lercher, N. R. Rose, C. J. Schofield and B. G. Davis, *Angew. Chem., Int. Ed.*, 2012, **51**, 1835–1839.
22 H. Neumann, S. M. Hancock, R. Buning, A. Routh, L. Chapman, J. Somers, T. Owen-Hughes, J. van Noort, D. Rhodes and J. W. Chin, *Mol. Cell*, 2009, **36**, 153–163.
23 D. P. Nguyen, M. M. Garcia Alai, S. Virdee and J. W. Chin, *Chem. Biol.*, 2010, **17**, 1072–1076.
24 D. P. Nguyen, M. M. Garcia Alai, P. B. Kapadnis, H. Neumann and J. W. Chin, *J. Am. Chem. Soc.*, 2009, **131**, 14194–14195.
25 C. H. Kim, M. Kang, H. J. Kim, A. Chatterjee and P. G. Schultz, *Angew. Chem., Int. Ed.*, 2012, **51**, 7246–7249.
26 M. G. Hoesl and N. Budisa, *Curr. Opin. Biotechnol.*, 2012, **23**, 751–757.
27 E. Kaya, K. Gutsmiedl, M. Vrabel, M. Muller, P. Thumbs and T. Carell, *ChemBioChem*, 2009, **10**, 2858–2861.
28 T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto and S. Yokoyama, *Chem. Biol.*, 2008, **15**, 1187–1197.
29 E. Kaya, M. Vrabel, C. Deiml, S. Prill, V. S. Fluxa and T. Carell, *Angew. Chem., Int. Ed.*, 2012, **51**, 4466–4469.
30 B. A. Garcia, S. Mollah, B. M. Ueberheide, S. A. Busby, T. L. Muratore, J. Shabanowitz and D. F. Hunt, *Nat. Protocols*, 2007, **2**, 933–938.
31 H. Neumann, S. Y. Peak-Chew and J. W. Chin, *Nat. Chem. Biol.*, 2008, **4**, 232–234.
32 B. Schwer, J. Bunkenborg, R. O. Verdin, J. S. Andersen and E. Verdin, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 10224–10229.