



2015

Genetic encoding of the post-translational modification 2-hydroxyisobutyryl-lysine

William A. Knight
knightwa@vcu.edu

Thomas A. Cropp
tacropp@vcu.edu

Follow this and additional works at: http://scholarscompass.vcu.edu/chem_pubs

 Part of the [Chemistry Commons](#)

© Royal Society of Chemistry 2015

Downloaded from

http://scholarscompass.vcu.edu/chem_pubs/16

This Article is brought to you for free and open access by the Dept. of Chemistry at VCU Scholars Compass. It has been accepted for inclusion in Chemistry Publications by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.



Cite this: *Org. Biomol. Chem.*, 2015, **13**, 6479

Received 10th February 2015,
Accepted 7th May 2015

DOI: 10.1039/c5ob00283d

www.rsc.org/obc

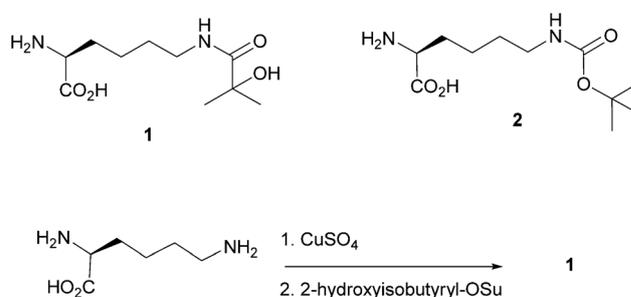
Genetic encoding of the post-translational modification 2-hydroxyisobutyryl-lysine†

William A. Knight and T. Ashton Cropp*

We report the synthesis and genetic encoding of a recently discovered post-translational modification, 2-hydroxyisobutyryl-lysine, to the genetic code of *E. coli*. The production of homogeneous proteins containing this amino acid will facilitate the study of modification in full-length proteins.

The number and complexity of identified lysine post-translational modifications (PTMs) continue to grow. Historically the most commonly observed protein modifications are located at lysine side chains, such as methylation or acetylation. In these examples the charge state of a lysine side chain is altered, which in turn can give rise to changes in interactions with DNA and other biomolecules.^{1,2} This epigenetic control of gene expression is vital, and malfunction of these systems can be a hallmark of disease.³

In addition to acetylation and methylation, it has been reported that lysine residues can be modified by malonylation,⁴ propionylation and butyrylation,⁵ succinylation,⁶ and crotonylation.⁷ These modifications are derived from intracellular acyl-CoA metabolites and provide a wide spectrum of epigenetic control of gene expression. The extent to which these modifications are actively added and removed by enzymes is a current interest in deciphering the “histone code”. Recently proteomics profiling revealed a new lysine modification that was identified as 2-hydroxyisobutyryl-lysine (K_{hib}) (1, Scheme 1).⁸ This modification appears to be conserved throughout evolution, appearing in human, mouse, *Drosophila*, and yeast cells.⁸ Intriguingly, unlike other histone marks which occur mainly at the N-terminus of histone proteins, K_{hib} is seen distributed throughout the globular protein structure, and includes many lysine residues that have not been previously reported to be modified. Moreover, in comparison with other lysine PTMs, K_{hib} is quite large and has the



Scheme 1 Synthesis of K_{hib} .

unique potential to serve as a hydrogen bond donor, raising the possibility of significant structural changes and alternative intermolecular interactions mediated by this modification. All of these point to an important role in chromatin function.

Critical to the study of PTMs is the ability to produce pure protein samples containing homogeneous compositions of a modified residue. Lysine modifications in particular may offer the ability to generate a designer chromatin⁹ to study the effects of PTMs either individually or in concert. Modified amino acids can be added to proteins using a combination of peptide synthesis and expressed protein ligation.¹⁰ This approach has the ability to precisely add PTMs, but it is technically challenging and cannot be scaled easily. As an alternative, some amino acids containing PTMs have also been added to the genetic codes of both prokaryotes and eukaryotes, allowing biosynthetic production of target proteins containing site-specific modifications. Indeed most lysine modifications have been incorporated using the pyrrolysine translational machinery.^{11–13} These *in vivo* production methods are very adaptable to biochemical laboratories. Moreover, biosynthetic production of proteins containing PTMs opens the door to more sophisticated experiments such as phage display,¹⁴ incorporation of isotopic labels,^{15,16} and a wide variety of *in vivo* experiments. Towards these goals we describe the synthesis and addition of 2-hydroxyisobutyryl-lysine to the genetic code of *E. coli*.

Department of Chemistry, Virginia Commonwealth University, 1001 W. Main Street, Richmond, Virginia 23284, USA. E-mail: tacropp@vcu.edu

† Electronic supplementary information (ESI) available: Full procedures for chemical synthesis and protein expression and purification. See DOI: 10.1039/c5ob00283d

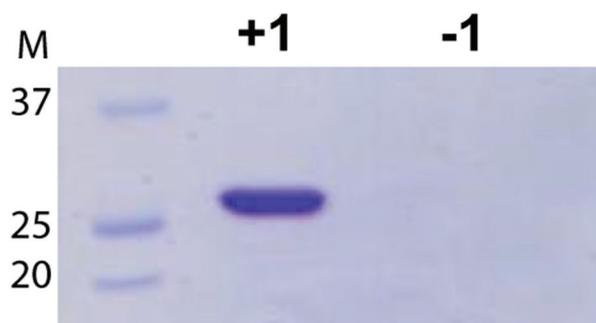


Fig. 1 Production of sfGFP containing **1** at position 151. No protein is produced in the absence of added unnatural amino acid.

The synthesis of K_{hib} was performed from lysine by first preparing a copper complex, which allows for selective acylation of the ϵ -nitrogen (Scheme 1). This complex was then treated with 2-hydroxyisobutyryl-*O*-succinimide ester and the copper was removed by a chelating agent to generate the final product. This synthesis is simple and can provide gram-scale quantities of amino acid for protein expression studies.

We decided to test whether K_{hib} is a substrate for several variants of the pyrrolysyl-tRNA synthetase (PylRS) derived from either *Methanosarcina barkeri* (Mb) or *Methanosarcina mazei* (Mm) (see the ESI†). These included the wild-type enzymes and variants that have been shown to have relaxed substrate specificity towards other, larger unnatural amino acids. The screen utilized an expression plasmid for superfolder green fluorescent protein (sfGFP) containing an amber stop codon, TAG, in place of the codon for Y151. The plasmid also contains the gene encoding the Mm-pyrrolysyl tRNA (pylT). Incorporation of unnatural amino acid leads to the production of a full-length protein and a corresponding increase in cellular fluorescence. Using a plate-based screening assay we first examined fluorescence in the presence and absence of K_{hib} for any observable differences. As a positive control, we also used N^{ϵ} -(*tert*-butyloxycarbonyl)-*L*-lysine (BocLys, (**2**), Scheme 1), which is a known substrate for PylRS. Among the five variants we screened the most observable fluorescence difference was obtained using wild-type Mm PylRS. While the observable fluorescence was weak in comparison with BocLys, it did show clear differences when compared to controls (see the ESI,

Fig. S2†). Variants with larger active sites did not appear to accept K_{hib} as the substrate.

We chose to perform medium-scale expression of sfGFP in the presence and absence of 5 mM K_{hib} and purified the resulting His-tagged proteins using Ni^{2+} affinity chromatography. As shown in Fig. 1, we observed robust protein expression ($\sim 10 \text{ mg L}^{-1}$) only in the presence of K_{hib} , indicating that this amino acid can serve as a substrate, without further evolution of PylRS. No protein is seen in the absence of K_{hib} , verifying that endogenous amino acids are not substrates for Mm PylRS. In order to verify the position and identity of the mutation, a gel slice of the produced protein was excised and subjected to in-gel tryptic digest.¹⁷ Upon examining the tryptic fragments by LC/MS/MS, the spectrum of the expected fragment was trapped in a +2 charge state (Fig. 2). Fragment masses of this ion are consistent with site-specific incorporation of K_{hib} at the correct position, in place of Y151. No masses that correspond to the same fragment containing other natural amino acids at position 151 were seen. In addition to tryptic peptide analysis, the protein samples were analyzed by ESI-MS on intact protein which also confirms incorporation of the amino acid (see the ESI, Fig. S3†). Interestingly, we did not observe any masses corresponding to a lysine residue at position 151 (or a resulting tryptic fragment), which would be indicative of active deacylation of K_{hib} in *E. coli*. Removal of other lysine PTMs has been previously observed and ascribed to bacterial sirtuins¹³¹², and can be prevented by the use of a nicotinamide enzyme inhibitor. It is possible that K_{hib} residues are not a substrate for these enzymes at all or when in the context of this mutation position in sfGFP.

Conclusions

In conclusion, we have demonstrated that the pyrrolysyl-tRNA synthetase (PylRS) has a suitably relaxed substrate specificity to accommodate 2-hydroxyisobutyryl-lysine (K_{hib}). This, coupled with a quick and high-yielding synthesis of this important amino acid makes this technology accessible to anyone working in the field of protein chemistry. We anticipate that the ability to produce homogeneous samples of proteins containing K_{hib} will facilitate the study of the histone code and chromatin function.

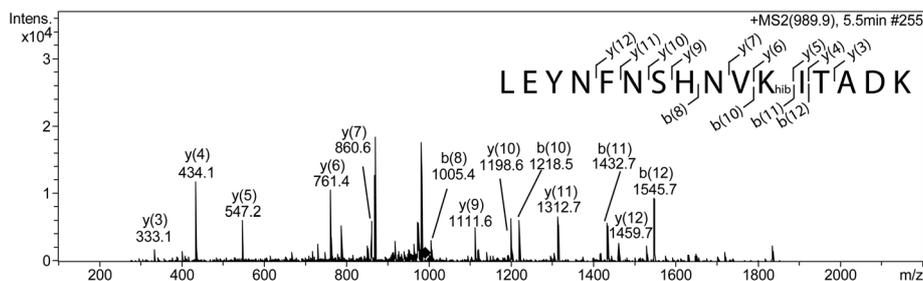


Fig. 2 MS/MS spectrum of tryptic fragment of sfGFP bearing K_{hib} at position 151.

Acknowledgements

This work was funded in part by NIH/NIGMS (GM084396).

Notes and references

- 1 P. Chi, C. D. Allis and G. G. Wang, *Nat. Rev. Cancer*, 2010, **10**, 457–469.
- 2 C. T. Walsh, S. Garneau-Tsodikova and G. J. Gatto Jr., *Angew. Chem., Int. Ed.*, 2005, **44**, 7342–7372.
- 3 S. Timmermann, H. Lehrmann, A. Poleskaya and A. Harel-Bellan, *Cell Mol. Life Sci.*, 2001, **58**, 728–736.
- 4 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.
- 5 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.
- 6 J. Park, Y. Chen, D. X. Tishkoff, C. Peng, M. Tan, L. Dai, Z. Xie, Y. Zhang, B. M. Zwaans, M. E. Skinner, D. B. Lombard and Y. Zhao, *Mol. Cell*, 2013, **50**, 919–930.
- 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 L. Dai, C. Peng, E. Montellier, Z. Lu, Y. Chen, H. Ishii, A. Debernardi, T. Buchou, S. Rousseaux, F. Jin, B. R. Sabari, Z. Deng, C. D. Allis, B. Ren, S. Khochbin and Y. Zhao, *Nat. Chem. Biol.*, 2014, **10**, 365–370.
- 9 B. Fierz and T. W. Muir, *Nat. Chem. Biol.*, 2012, **8**, 417–427.
- 10 T. W. Muir, D. Sondhi and P. A. Cole, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 6705–6710.
- 11 B. J. Wilkins, L. E. Hahn, S. Heitmuller, H. Frauendorf, O. Valerius, G. H. Braus and H. Neumann, *ACS Chem. Biol.*, 2015, **10**, 939–944.
- 12 H. Neumann, S. Y. Peak-Chew and J. W. Chin, *Nat. Chem. Biol.*, 2008, **4**, 232–234.
- 13 M. J. Gattner, M. Vrabel and T. Carell, *Chem. Commun.*, 2013, **49**, 379–381.
- 14 F. Tian, M. L. Tsao and P. G. Schultz, *J. Am. Chem. Soc.*, 2004, **126**, 15962–15963.
- 15 C. Castaneda, J. Liu, A. Chaturvedi, U. Nowicka, T. A. Cropp and D. Fushman, *J. Am. Chem. Soc.*, 2011, **133**, 17855–17868.
- 16 C. A. Castaneda, J. Liu, T. R. Kashyap, R. K. Singh, D. Fushman and T. A. Cropp, *Chem. Commun.*, 2011, **47**, 2026–2028.
- 17 A. Shevchenko, H. Tomas, J. Havlis, J. V. Olsen and M. Mann, *Nat. Protocols*, 2006, **1**, 2856–2860.