A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEM **BIO** CHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Genetically encoded biotin analogs: Incorporation and application in bacterial and mammalian cells

Authors: Adrian Hohl, Yonatan G Mideksa, Ram Karan, Anastassja Akal, Malvina Vogler, Michael Groll, Magnus Rueping, Kathrin Lang, Matthias J Feige, and Jörg Eppinger

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: *ChemBioChem* 10.1002/cbic.201900015

Link to VoR: <http://dx.doi.org/10.1002/cbic.201900015>

WILEY-VCH

www.chembiochem.org

A Journal of

Genetically encoded biotin analogs: Incorporation and application in bacterial and mammalian cells

Adrian Hohl^{[a],[b]}, Yonatan G. Mideksa^{[b],[c]}, Ram Karan^[a], Anastassja Akal^{[a],[b]}, Malvina Vogler^[a], Michael Groll^[b], Magnus Rueping^[a], Kathrin Lang^{[b],[c]}, Matthias J. Feige^{[b],[c]}, Jörg Eppinger^[a]

Abstract: The biotin-streptavidin interaction is among the strongest known in nature. Here, we report the site-directed incorporation of biotin and 2-iminobiotin bearing non-canonical amino acids (ncAA) into proteins. 2-Iminobiotin lysine was employed for protein purification based on the pH-dependent dissociation constant to streptavidin. Using the high affinity binding of biotin lysine, we could specifically isolate the bacterial protein RecA and analyze its interaction partners. Furthermore, we successfully transferred our biotinylation approach to mammalian cells. The stringent control over the biotinylation site and the tunable affinity between ncAAs and streptavidin by the different biotin-analogs makes this approach an attractive tool for protein interaction studies, protein immobilization and the generation of well-defined protein drug conjugates.

Biotinylation of proteins is widely exploited for biological applications and offers advantages over immunoaffinity methodologies.^[1] The high affinity between biotin and streptavidin (dissociation constant $K_d = 10^{-14}$ M) permits stringent washing conditions to minimize non-specific interactions.^[2] Thus, different approaches towards site-specific protein biotinylation have been developed in earlier studies: Sortase A-mediated biotinylation has been used to label proteins at the N- and C-termini.^[3] Furthermore, biotinylated amino acids have been charged *in vitro* onto t-RNAs via chemical approaches and have been incorporated into proteins in a cell-free system or in cells.^[4] *In vivo* biotinylation can be achieved either by the *E. coli* biotin ligase BirA or orthologues.^[5] However, the biotin ligase-mediated biotinylation requires a recognition peptide fused to the protein of interest.^[6] This tag may interfere with the structure or function of the protein under investigation. Furthermore, while these approaches allow efficient isolation of labeled proteins, the subsequent release is problematic due to the strong interaction of biotin to streptavidin. To overcome these limitations, an engineered streptavidin mutant with a reduced affinity to biotin has been applied for purification of biotinylated proteins.^[7] Nowadays, a broad range of commercial biotinylation reagents is available for *in vitro* protein labeling. These reagents include biotin analogs with different affinities to streptavidin according to the specific demands of an experiment. For instance, 2-

iminobiotin binds tightly to streptavidin at basic pH but the interaction drops considerably at pH 4, thus allowing protein elution.^[8] If the target protein is affected by this pH change, the biotin precursor desthiobiotin provides an alternative, as it binds with high specificity but less tightly to streptavidin ($K_d = 10^{-11}$ M). Hence, protein elution under physiological conditions can be achieved through the addition of biotin.^[9] Despite the versatility of commercial biotinylation reagents, applications often suffer from poor control over stoichiometry and site-specificity.

Here, we report the genetic incorporation of biotin lysine and 2-iminobiotin lysine in bacterial cells by an engineered pyrrolysyl-tRNA synthetase (PylRS), termed BioRS. Using the pH-dependent affinity between 2-iminobiotin and streptavidin, we demonstrate the purification of the ancestral β -subunit protein (Anbu) from *E. coli* cell lysate. We also employ our *in vivo* biotinylation protocol to isolate the bacterial DNA repair protein RecA and to identify possible interaction partners in *E. coli*. Moreover, we demonstrate the compatibility of our approach with mammalian cells.

First, we designed a mutant PylRS library to engineer a PylRS/tRNA^{CUA} pair for the site-specific incorporation of the biotinylated ncAAs **1-3** (Figure 1A). Our recently developed polyspecific PylRS variant (Y271A, L274V, C313V, M315Y, Y349F, V370R) served as a template.^[10] Positions M241, L372, W376 were selected for mutagenesis due to their close proximity to the biotin residue (Figure 1B) and were mutated according to Table S1-2. Positions 241 and 372 were mainly restricted to small hydrophobic amino acids to introduce additional space in the pocket and polar amino acids that may improve biotin binding through hydrogen-bonding.

Figure 1. (A) Chemical structures of biocytin and ncAAs **1, 2** and **3**. (B) Homology model of BioRS in complex with **2**. BioRS specific mutations are labeled. L372, M241, and W376 are highlighted in bold. (C) Incorporation test of biocytin **B** and ncAAs **1, 2** and **3** (1 mM) into mTFP1-128TAG by

- [a] Dr. A. Hohl, Dr. R. Karan, Dr. A. Akal, Dr. M. Vogler, Prof. Dr. M. Rueping, Prof. Dr. J. Eppinger
Division of Physical Sciences and Engineering
KAUST Catalysis Center (KCC)
King Abdullah University of Science and Technology
Thuwal, Kingdom of Saudi Arabia
E-mail: adrian-hohl@gmx.de, jorg.eppinger@gmail.com
- [b] Dr. A. Hohl, Y. G. Mideksa, Dr. A. Akal, Prof. Dr. M. Groll, Prof. Dr. K. Lang, Prof. Dr. M. J. Feige
Center for Integrated Protein Science at the Department of Chemistry, Technical University of Munich
Lichtenbergstr. 4, 85748 Garching, Germany
- [c] Y. G. Mideksa, Prof. Dr. K. Lang, Prof. Dr. M. J. Feige
Institute for Advanced Study, Technical University of Munich
Lichtenbergstr. 2a, 85748 Garching, Germany

Supporting information for this article is given via a link at the end of the document.

BioRS (top: Coomassie-stained SDS-PAGE; center: anti-His immunoblot; bottom: Alexa Fluor 532 streptavidin conjugate). (D) Left: Coomassie-stained SDS-PAGE of cell lysates with mTFP1 wt or mTFP1-128TAG containing **2** or **3**. Center: cell lysates after treatment with streptavidin coated dynabeads. Right: heat induced elution from the washed dynabeads.

The PyIRS library was co-expressed with the corresponding tRNA_{CUA} and the fluorescent protein mTFP1 bearing an amber stop codon at the permissive position 128 (mTFP1-128TAG). Cells displaying a fluorescence signal of mTFP1 in the presence of **2** were analyzed by fluorescence activated cell sorting (FACS). This procedure allowed us to identify BioRS (M241A, Y271A, L274V, C313V, M315Y, Y349F, V370R), showing the highest fluorescence signal among all sorted PyIRS variants (Figure S1). A homology model of BioRS predicted an enlarged binding pocket that accommodates the biotin residue (Figure 1B). Remarkably, SDS-PAGE analysis and anti-His immunoblot indicated aminoacylation and site-specific incorporation into mTFP1 of the biotinylated ncAA **2** but also of the shorter analog **1** and 2-iminobiotin lysine **3** by BioRS (Figure 1C). The ncAA **1-3** can be detected by streptavidin-dye-conjugates, substituting for immunoblots. The absence of a fluorescent signal for wt mTFP1 highlights the specific streptavidin/biotin interaction (Figure 1C). The binding affinities of biotin analogs towards streptavidin were determined (Figure S2). mTFP1_{ncAA1-3} was purified (Figure S3), and ESI-TOF confirmed the successful ncAA incorporation (Figure S4).

The high affinity of biotin to streptavidin enables protein isolation directly from cell lysate. We demonstrated this application for mTFP1 wt and mTFP1 with incorporated **2** or **3**. The proteins were separately expressed, and the cell lysates were diluted according to the fluorescence signal (Figure 1D). Samples were cleared from free ncAAs through filtering and incubated with streptavidin-coated dynabeads. mTFP1_{ncAA2,3} bonded to the beads and could be isolated by heating to 95 °C in Laemmli buffer, while no binding of mTFP1 wt was observed (Figure 1D).

In contrast to biotin, the analog **3** enables reversible binding to streptavidin. To investigate the feasibility of pH-dependent protein purification through ncAA **3**, we chose the ancestral β -subunit protein (Anbu) as a model protein. Anbu is found in cyanobacteria, and its topology adopts the fold of Ntn-hydrolases.^[11] Thus, it is suggested that Anbu represents the common origin of proteasome subunits. Based on the Anbu crystal structure (PDB: 5NYP),^[12] the amber codon was placed at the solvent-exposed position Asp71. After expression of Anbu-71TAG in the presence of **3** and the BioRS/tRNA_{CUA} pair, the cell lysate was equilibrated with basic buffer (pH 11) prior to purification. As expected 2-iminobiotin binds strongly to streptavidin at basic pH, hence protein containing **3** could be isolated from cell-lysates using streptavidin-coupled dynabeads. Upon reduction to pH 4, the streptavidin 2-iminobiotin complex dissociated and Anbu detached from the dynabeads. Analysis of the eluted sample by SDS-PAGE revealed a band at 30 kDa that was excised and further analyzed by MS/MS (Figure 2A). A Mascot search identified Anbu with a sequence coverage of

about 95% and the incorporation of ncAA **3** was confirmed at position 71 (Figure S6).

Protein-biotinylation is a typical approach for pull-down analysis. For a proof-of-concept we focused on the bacterial protein RecA that plays an essential role in the repair of DNA double-strand breaks in *E. coli*. Additionally, RecA interacts with several proteins to initiate the SOS response and assembles into spiral filaments.^[13] This functional diversity makes RecA sensitive to structural modifications normally introduced to allow affinity-purification of transient complexes. The small size of biotinylated ncAAs compared to conventional tags increases the probability to preserve the natural activity and interactions of the labeled protein. Hence, ncAA **2** was site-specifically inserted at the exposed position 36 of RecA by the BioRS/tRNA_{CUA} system to investigate transient interaction partners in *E. coli* (PDB-ID: 4TWZ).^[14] Cells were treated with H₂O₂, heat, UV, Fe²⁺ or Zeocin after the expression of RecA_{ncAA2} to initiate DNA breaks and to trigger the assembly of RecA filaments.^[15] The protein of interest was pulled down from the cell lysate through streptavidin-coupled dynabeads. Proteins were separated by SDS-PAGE, bands were analyzed by MS/MS, and the site-specific incorporation of **2** into RecA was confirmed by a tryptic digest (Figure S7, S8). A Mascot search revealed 21 possible interaction partners of RecA, which were clustered in an interaction network using the software STRING (Figure 2B).^[16]

Figure 2. (A) SDS-PAGE analysis of single-step purification of Anbu-71TAG. Cell lysate containing Anbu with incorporated **3** (L), cell lysate after treatment with dynabeads (TL), three washing steps of dynabeads (W1, W2, and W3) and protein elution (E). (B) RecA-36TAG and its identified interaction partners. Proteins are clustered in an interaction network using the software

STRING^[16]. The confidence of interaction is indicated by the thickness of the lines. Proteins are classified as ribonucleotide processing proteins (yellow), proteases and chaperones (green), ribosomal proteins (blue) or membrane proteins (red).

The identified candidates could be classified according to their biological function as ribonucleotide processing proteins, proteases and chaperones, ribosomal or membrane proteins which all orchestrate the SOS response in the cell.^[13b] Some of these interaction partners have previously been identified in pull-down experiments.^[17] In particular, RecA showed significant interaction with membrane proteins such as FtsZ, AtpA, AtpD or MreB, which might enhance the arrest of cell division to accomplish the DNA repair properly.^[18]

To test, whether **2** and **3** can also be genetically incorporated into proteins in mammalian cells, the BioRS/tRNA_{CUA} pair was co-expressed with a superfolder GFP (sfGFP) bearing an amber codon at position 150 and a C-terminal His-tag (sfGFP-150TAG-His) in HEK-293T cells in the presence or absence of the ncAAs. As a positive control, the established ncAA Nε-(tert-butoxycarbonyl)-L-lysine (BockK) was incorporated into sfGFP-150TAG-His by wt PylRS.^[19] Fluorescence microscopy revealed amber codon suppression in the presence of BockK, **2** and **3**. The lack of a fluorescence signal in the absence of any ncAA indicated the orthogonality and specificity of the BioRS/tRNA_{CUA} suppression system in mammalian cells (Figure 3A). The site-specific incorporation of BockK, **2** and **3** into sfGFP-150TAG-His was additionally confirmed by MS/MS analysis (Figure S9). The incorporation efficiencies of the ncAAs were compared by an anti-His immunoblot on the cell lysates (Figure 3B, center). Streptavidin-coupled dynabeads were used to separate ncAA-labeled proteins from the cell lysates. The proteins were released from the dynabeads at 95 °C, and an anti-His immunoblot of the eluted proteins indicated that only sfGFP bearing **2** or **3** was isolated (Figure 3B, bottom).

Figure 3. (A) Fluorescence micrographs of HEK-293T cells expressing the BioRS/tRNA_{CUA} pair and sfGFP-150TAG-His in the presence or absence of **2** and **3**. BockK was incorporated by wt PylRS/tRNA_{CUA}. (B) Immunoblot of sfGFP-150TAG-His6 in cell lysate (center) and after pull-down with streptavidin-coated dynabeads (bottom) using an anti-His antibody and an anti-Hsc70 antibody as a loading control (top). Dashed line: lane not relevant

for the figure. EV (empty vector) indicates a vector containing 4 pPylT copies but no amber-containing reporter gene.

In conclusion, the genetically encoded biotinylation provides a homogeneous expression product, which is a major advantage over chemical methods. The site-directed incorporation of lysine-conjugated biotin derivatives can potentially simplify protein analyses, protein immobilization and the production of protein drug conjugates, which require a stringent control over the site of modification and stoichiometry.^[20]

Acknowledgements

We thank Marie-Theres Vielberg (TUM) for providing AnbuTAG71. This study was supported by the King Abdullah University of Science and Technology (KAUST). This work was performed in the framework of the German Research Foundation (DFG) Sonderforschungsbereich 1035, projects A02, B10, and B11. YGM gratefully acknowledges funding via a DAAD Ph.D. scholarship. KL and MJF are Rudolf Mößbauer Tenure Track Professors and as such gratefully acknowledge funding through the Marie Curie COFUND program and the Technical University of Munich Institute for Advanced Study, funded by the German Excellence Initiative and the European Union Seventh Framework Program under Grant Agreement 291763.

Keywords: amber suppression • unnatural amino acids • site-specific incorporation • pyrrolysyl-tRNA synthetase • biotinylated amino acid

- [1] a) J. Kim, A. B. Cantor, S. H. Orkin, J. Wang, *Nat. Protoc.* **2009**, 4, 506-517; b) C. M. Dundas, D. Demonte, S. Park, *Appl. Microbiol. Biotechnol.* **2013**, 97, 9343-9353.
- [2] N. M. Green, *Methods Enzymol.* **1990**, 184, 51-67.
- [3] M. P. Madej, G. Coia, C. C. Williams, J. M. Caine, L. A. Pearce, R. Attwood, N. A. Bartone, O. Dolezal, R. M. Nisbet, S. D. Nuttall, T. E. Adams, *Biotechnol. Bioeng.* **2012**, 109, 1461-1470.
- [4] a) J. P. Gallivan, H. A. Lester, D. A. Dougherty, *Chem. Biol.* **1997**, 4, 739-749; b) T. V. Kurzchalia, M. Wiedmann, H. Breuer, W. Zimmermann, E. Bauschke, T. A. Rapoport, *Eur. J. Biochem.* **1988**, 172, 663-668.
- [5] J. E. Cronan, *J. Biol. Chem.* **1990**, 265, 10327-10333.
- [6] a) D. Beckett, E. Kovaleva, P. J. Schatz, *Protein Sci.* **1999**, 8, 921-929; b) I. Chen, M. Howarth, W. Lin, A. Y. Ting, *Nat. Methods* **2005**, 2, 99-104.
- [7] a) S. C. Wu, S. L. Wong, *J. Biol. Chem.* **2005**, 280, 23225-23231; b) V. J. O'Sullivan, I. Barrette-Ng, E. Hommema, G. T. Hermanson, M. Schofield, S. C. Wu, C. Honetschlaeger, K. K. S. Ng, S. L. Wong, *Plos One* **2012**, 7, e35203.
- [8] B. Fudem-Goldin, G. A. Orr, *Methods Enzymol.* **1990**, 184, 167-173.
- [9] J. D. Hirsch, L. Eslamizar, B. J. Filanoski, N. Malekzadeh, R. P. Haugland, J. M. Beechem, R. P. Haugland, *Anal. Biochem.* **2002**, 308, 343-357.
- [10] A. Hohl, Dissertation, Technische Universität München, **2017**.
- [11] a) R. E. Valas, P. E. Bourne, *J. Mol. Evol.* **2008**, 66, 494-504; b) A. C. D. Fuchs, V. Alva, L. Maldoner, R. Albrecht, M. D. Hartmann, J. Martin, *Structure* **2017**, 25, 834-845; c) A. Piasecka, H. Czapinska, M. T. Vielberg, R. H. Szczepanowski, R. Kiefersauer, S. Reed, M. Groll, M. Bochtler, *J. Mol. Biol.* **2018**, 430, 611-627.
- [12] M. T. Vielberg, V. C. Bauer, M. Groll, *J. Mol. Biol.* **2018**, 430, 628-640.
- [13] a) H. Y. Wu, C. H. Lu, H. W. Li, *Sci. Rep.* **2017**, 7, 11876; b) K. N. Kreuzer, *Cold Spring Harb. Perspect. Biol.* **2013**, 5, a012674.

- [14] T. Shinohara, S. Ikawa, W. Iwasaki, T. Hiraki, T. Hikima, T. Mikawa, N. Arai, N. Kamiya, T. Shibata, *Nucleic Acids Res.* **2015**, *43*, 973-986.
- [15] N. Renzette, S. J. Sandler, *Mol. Microbiol.* **2008**, *67*, 1347-1359.
- [16] D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K. P. Tsafou, M. Kuhn, P. Bork, L. J. Jensen, C. von Mering, *Nucleic Acids Res.* **2015**, *43*, D447-452.
- [17] a) M. Zhang, Y. Zhou, T. Li, H. L. Wang, F. Cheng, Y. F. Zhou, L. J. Bi, X. E. Zhang, *Arch. Biochem. Biophys.* **2012**, *517*, 98-103; b) G. Butland, J. M. Peregrin-Alvarez, J. Li, W. H. Yang, X. C. Yang, V. Canadien, A. Starostine, D. Richards, B. Beattie, N. Krogan, M. Davey, J. Parkinson, J. Greenblatt, A. Emili, *Nature* **2005**, *433*, 531-537; c) M. Arifuzzaman, M. Maeda, A. Itoh, K. Nishikata, C. Takita, R. Saito, T. Ara, K. Nakahigashi, H. C. Huang, A. Hirai, K. Tsuzuki, S. Nakamura, M. Altaf-Ul-Amin, T. Oshima, T. Baba, N. Yamamoto, T. Kawamura, T. Ioka-Nakamichi, M. Kitagawa, M. Tomita, S. Kanaya, C. Wada, H. Mori, *Genome Res.* **2006**, *16*, 686-691.
- [18] N. Garvey, A. C. St John, E. M. Witkin, *J. Bacteriol.* **1985**, *163*, 870-876.
- [19] D. P. Nguyen, H. Lusic, H. Neumann, P. B. Kapadnis, A. Deiters, J. W. Chin, *J. Am. Chem. Soc.* **2009**, *131*, 8720-8721.
- [20] P. Dennler, E. Fischer, R. Schibli, *Antibodies* **2015**, *4*, 197-224.

Entry for the Table of Contents

COMMUNICATION

Author(s), Corresponding Author(s)*

Page No. – Page No.

Title

Genetically encoded biotin and 2-iminobiotin lysine provides stringent control over the site of modification and stoichiometry. Well-defined biotinylated proteins can potentially simplify protein investigation, protein immobilization and the production of protein drug conjugates by the strong interaction to streptavidin.