FEBS openbio

FEBS PRESS
science publishing by scientists

METHOD

Genetically encoding thioacetyl-lysine as a non-deacetylatable analog of lysine acetylation in *Escherichia coli*

Sumana Venkat[1,*], Dharma Theja Nannapaneni[1,*], Caroline Gregory[2], Qinglei Gan[1], Matt McIntosh[1] and Chenguang Fan[1]

1 Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR, USA
2 Department of Biological Sciences, University of Arkansas, Fayetteville, AR, USA

Keywords
deacetylation; genetic code expansion; non-canonical amino acid; protein acetylation; synthetic biology; thioacetyl-lysine

Correspondence
M. McIntosh, Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR 72701, USA
Fax: +1 479 575 4049
Tel: +1 479 575 4692
E-mail: mcintosh@uark.edu
and
C. Fan, Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR 72701, USA
Fax: +1 479 575 4049
Tel: +1 479 575 4653
E-mail: cf021@uark.edu

*These authors contributed equally to this work

(Received 11 September 2017, accepted 21 September 2017)

doi:10.1002/2211-5463.12320

Reversible lysine acetylation is one of the most widely distributed post-translational modifications; it is involved in a variety of biological processes and can be found in all three domains of life. Acetyltransferases and deacetylases work coordinately to control levels of protein acetylation. In this work, we applied the genetic code expansion strategy to site-specifically incorporate N^{ε}-thioacetyl-L-lysine (TAcK) as an analog of N^{ε}-acetyl-L-lysine (AcK) into green fluorescent protein and malate dehydrogenase in *Escherichia coli*. We showed that TAcK could serve as an ideal functional mimic for AcK. It could also resist the bacterial sirtuin-type deacetylase CobB. Thus, genetic incorporation of TAcK as a non-deacetylatable analog of AcK into proteins will facilitate *in vivo* studies of protein acetylation.

The reversible acetylation of lysine residues in proteins has been recognized as one of the most widely distributed post-translational modifications in both eukaryotes and prokaryotes [1–5]. It plays key roles in regulating a wide range of important biological processes including gene transcription, stress response, apoptosis, cellular differentiation and metabolism [6–13]. Accumulating evidence has shown that protein acetylation is tightly associated with many human diseases such as cancers, cardiovascular diseases, diabetes

Abbreviations
AcK, N^{ε}-acetyl-L-lysine; AcKRS, acetyllysyl-tRNA synthetase; MDH, malate dehydrogenase; MS, mass spectrometry; ncAA, non-canonical amino acid; PylRS, pyrrolysyl-tRNA synthetase; sfGFP, superfolder green fluorescent protein; SOC, Super Optimal broth with Catabolite repression; TAcK, N^{ε}-thioacetyl-L-lysine; TAcKRS, thioacetyllysyl-tRNA synthetase.

 © 2017 The Authors. Published by FEBS Press and John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

and neurodegenerative disorders [14–20]. Studies on protein acetylation have not only deepened our knowledge of its functions and mechanisms but also guided therapeutic strategies for acetylation-associated diseases [21–24].

Lysine acetylation can take place with or without acetyltransferases, while the removal of the acetyl group is mediated by deacetylases [25–33]. Due to the existence of acetyltransferases and deacetylases as well as non-enzymatic acetylation in cells, one challenge for studying lysine acetylation is to synthesize homogeneously acetylated proteins at specific sites. To overcome this problem, the genetic code expansion strategy was applied to utilize an orthogonal pair of an evolved pyrrolysyl-tRNA synthetase (PylRS) variant and its cognate tRNA from *Methanosarcinaceae* species to cotranslationally incorporate N^{ε}-acetyl-L-lysine (AcK) in response to a stop codon at the desired position in target proteins [34–36]. This approach has already proved to be a powerful tool to study protein acetylation [37–45].

As a non-hydrolyzable analog of AcK, N^{ε}-thioacetyl-L-lysine (TAcK) has been used as an inhibitor for a number of deacetylases in the form of peptide or non-peptide substrates [46–52]. Recently, Söll's group utilized the flexizyme-mediated tRNA aminoacylation and cell-free translation systems to incorporate TAcK into histone H3, which was shown to be resistant to SIRT1 deacetylase [53]. However, this approach can only be used *in vitro*. In this work, we established an *in vivo* system to site-specifically incorporate TAcK into proteins for broader applications by using living cells.

Materials and methods

Thioacetyl-lysine synthesis

The synthesis of TAcK followed a previous protocol with slight modifications (Fig. 1) [53].

Step 1

A 5% (w/v) aqueous solution of Na_2CO_3 (25 mL) was added dropwise to a stirred suspension of N^{α}-Boc-L-lysine (3.0 g, 12 mmol) in ethanol (25 mL) at 0 °C. Ethyl dithioacetate (1.5 mL, 13.2 mmol) was then added dropwise at 0 °C. After the addition was complete, the reaction mixture was stirred at room temperature for 4 h before the addition of a 50% (v/v) solution of ethanol in water (3 mL). The ethanol was removed under reduced pressure and the residue was acidified with 6 M HCl to pH ~ 1–2 and extracted with CH_2Cl_2. The combined organic extracts were washed with saturated NaCl solution, dried over anhydrous Na_2SO_4, filtered and evaporated to dryness, affording an oily residue. The product was isolated as a white solid (3.0 g, 68%) after silica gel column chromatography (60 : 40 ethyl acetate/hexanes). Spectral data of the compound match with the previous report [53]: ^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.61 (s, 1H), 4.16 (m, 1H), 3.49–3.55 (m, 2H), 2.46 (s, 3H), 1.62–1.78 (m, 6H), 1.37 (s, 9H)

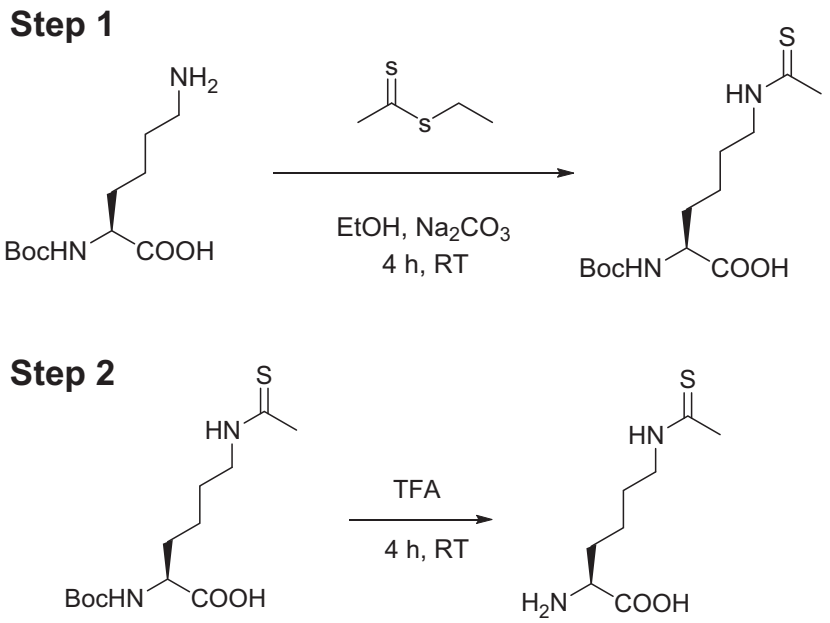

Fig. 1. The Scheme for TAcK synthesis.

(Fig. S1); ^{13}C-NMR (100 MHz, $CDCl_3$) δ: 200.4, 174.5, 156.0, 79.6, 52.9, 46.1, 33.7, 32.7, 28.3, 27.0, 22.8 (Fig. S2).

Step 2

N^{α}-Boc-N^{ε}-thioacetyl-L-lysine (2.75 g, 9 mmol) was dissolved in dry CH_2Cl_2 (25 mL), and trifluoroacetic acid (15 mL, 0.2 mol, 22 eq) was added dropwise to the reaction mixture over 20 min. The reaction mixture was stirred at room temperature for 4 h. The biphasic reaction mixture was separated, and the organic layer was washed with CH_2Cl_2 several times. The product was evaporated to dryness affording a white solid powder (1.65 g, 90%). Spectral data of the compound: ^{1}H-NMR (400 MHz, D_2O) δ: 3.96 (m, 1H), 3.51 (m, 2H), 2.39 (s, 3H), 1.78–1.97 (m, 2H), 1.58–1.75 (m, 2H); 1.24–1.51 (m, 2H) (Fig. S3); ^{13}C-NMR (100 MHz, D_2O) δ: 202.5, 172.6, 52.8, 45.6, 32.2, 29.3, 26.3, 21.5 (Fig. S4); ESI-FTMS calcd for $C_8H_{16}N_2O_4S$ (M + H) – 205.1005, m/z found 205.1000 (Fig. S5).

General molecular biology

Chemicals in this study were purchased from Sigma-Aldrich (St Louis, MO, USA). *E. coli* TOP10 cells (Thermo Fisher Scientific, Waltham, MA, USA) were used for cloning and expression. Cloning experiments were performed by PCR and the DNA Assembly Kit (New England Biolabs, Ipswich, MA, USA). Point mutations were made by the QuikChange II mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA).

Western blotting

Purified proteins were fractionated by SDS/PAGE and transferred onto a PVDF membrane. The membrane was incubated at room temperature with gentle shaking in TTBS (Tris Buffered Saline, with Tween 20, pH 8.0) and 5% BSA blocking buffer for 60 min. Horseradish peroxidase-conjugated acetylated-lysine (Ac-K^2-100) rabbit antibody (Cell Signaling Technology, Beverly, MA, USA) was diluted 1 : 1000 and soaked the membrane overnight at 4 °C. The membrane was prepared for detection using Pierce™ ECL western blotting substrates (Thermo Fisher Scientific).

The superfolder GFP readthrough assay

The assay follows a previous protocol [54]. Strains harboring the superfolder GFP (sfGFP) reporter gene in the pBAD plasmid and the genes of AcKRS and TAcKRS variants and tRNA^{Pyl} in the pTech plasmid were inoculated into 2 mL LB medium. The overnight culture was diluted with fresh LB medium to an attenuance of 0.15 at 600 nm, supplemented with non-canonical amino acids (ncAAs), 100 μg·mL^{-1} ampicillin and 50 μg·mL^{-1} chloramphenicol; 1 mM arabinose was added to induce the expression of sfGFP. Two hundred microliters of culture of each strain was transferred to a 96-well plate. Cells were shaken for 12 h at 37 °C, with monitoring of fluorescence intensity (excitation 485 nm, emission 528 nm, bandwidths 20 nm) and attenuance (D_{600}) by the microplate-reader.

AcKRS variant library construction and selection for TAcK-specific variants

For variant library construction, three residues (F271, F313 and W382) of AcKRS were randomly mutagenized with three primers (AcKRS271-QF: GCACCGAACCTGNNNAATTAT-GCCCGTAAACTG; AcKRS313-QF: CACCATGCTGAACTTTNNNCAAATGGGCTCG; and AcKRS382-QF: GGTATTGACAAACCGNNNATCGGCGCG, NNN is completely randomization of corresponding positions) by the QuikChange multisite-directed mutagenesis kit (Agilent Technologies). Selections were performed as described before [35]. For the first round positive selection, 50 ng of pBK-AcKRS (Km^r) library plasmid was introduced into 50 μL of *E. coli* TOP10 (10^8 cells) with the positive selection plasmid pCAT-pylT (Tet^r) that has a mutant *cat* gene with an amber stop codon at position 112 and tRNA^{Pyl}. The transformants were recovered in 1 mL SOC (Super Optimal broth with Catabolite repression) at 37 °C for 2 h, and then cultivated in 100 mL LB-TK (containing 10 μg·mL^{-1} tetracycline and 25 μg·mL^{-1} kanamycin) overnight at 37 °C. Of the overnight culture, 100 μL was inoculated into 5 mL fresh LB-TK-TAcK (LB-TK containing 2 mM TAcK). After growing at 37 °C for 4 h, 200 μL culture was plated on LB-TK-TAcK-Cm (LB-TK-TAcK containing 350 μg·mL^{-1} chloramphenicol) plates. Here we used 350 μg·mL^{-1} chloramphenicol because cells harboring the original AcKRS and tRNA^{Pyl} as well as the mutant *cat* gene could survival at 300 but not 350 μg·mL^{-1} chloramphenicol in the same condition (Fig. S6). The plates were incubated at 37 °C for 48 h, and all the colonies growing on the plates were scraped and resuspended in 10 mL LB-TK-TAcK. After additional incubation for 4 h at 37 °C, total plasmids were extracted, and the pBK-AcKRS library plasmids were isolated by agarose gel electrophoresis and purified by the gel purification kit.

For the negative selection, 10 ng pBK-AcKRS library plasmids from the positive selection were transformed into 50 μL *E. coli* TOP10 (10^8 cells) with negative selection plasmid pAraCB2-pylT (Cm^r) that has tRNA^{Pyl} and a mutant *ccdB* gene with two amber stop codons at positions 13 and 44. The transformants were recovered in 2 mL SOC at 37 °C for 2 h, and then 50 μL was plated on LB-CKara (containing 50 μg·mL^{-1} chloramphenicol, 25 μg·mL^{-1} kanamycin and 0.2% arabinose) plates. After incubation at 37 °C overnight, all the colonies were harvested, and pBK-AcKRS library plasmids were isolated as described above for the next positive selection. One nanogram of pBK-AcKRS library plasmids from the negative selection were

introduced into 50 μL TOP10 (10^8 cells) with the positive selection plasmid pCAT-pylT. The transformants were recovered in 1 mL SOC at 37 °C for 2 h. It was transferred and cultivated in 50 mL LB-TK overnight at 37 °C. Of the overnight culture, 100 μL was inoculated into 5 mL fresh LB-TK-TAcK. To obtain individual positive colonies at this stage, the cultures were diluted with fresh LB-TK-TAcK and plated on LB-TK-TAcK-Cm plates. After 48 h of incubation, 69 colonies were selected, and each clone was inoculated into 1 mL LB-TK and incubated overnight at 37 °C. Of the overnight culture, 100 μL was diluted with fresh 1 mL LB-TK with or without 2 mM TAcK. After growing at 37 °C for 3 h, cell cultures were spotted on an LB-TK-TAcK-Cm or LB-TK-Cm plate, separately. Finally, 17 clones were found to grow only on LB-TK-TAcK-Cm plate. The pBK-TAcKRS plasmid isolated and extracted from each clone was sent for DNA sequencing.

Protein expression and purification

The procedure has small modifications from previous protocols [54,55]. The genes of target proteins were cloned into the pBAD plasmid with a C-terminal His_6-tag, and transformed into Top10 cells together with the pTech plasmid harboring genes of $tRNA^{Pyl}$ and TAcKRS for expression. The expression strain was grown on 400 mL LB medium supplemented with 100 $\mu g \cdot mL^{-1}$ ampicillin and 50 $\mu g \cdot mL^{-1}$ chloramphenicol at 37 °C to an attenuance of 0.6–0.8 at 600 nm, and protein expression was induced by the addition of 1 mM arabinose and supplemented with 5 mM TAcK. Cells were incubated at 30 °C for an additional 8 h, and harvested by centrifugation at 5000 ***g*** for 10 min at 4 °C. The cell paste was suspended in 15 mL of lysis buffer (50 mM Tris pH 7.5, 300 mM NaCl, 20 mM imidazole) with the protease inhibitor cocktail (Roche, Basel, Switzerland), and broken by sonication. The crude extract was centrifuged at 20 000 ***g*** for 25 min at 4 °C. The soluble fraction was filtered with a 0.45 μm filter and loaded onto a column containing 1 mL of Ni-NTA resin (Qiagen, Hilden, Germany) previously equilibrated with 20 mL lysis buffer. The column was washed with 20 mL wash buffer (50 mM Tris pH 7.5, 300 mM NaCl, 50 mM imidazole), and eluted with 2 mL elution buffer (50 mM Tris pH 7.5, 300 mM NaCl, 150 mM imidazole). The elution fraction was desalted with desalting buffer (50 mM Tris pH 7.5, 20 mM NaCl) with a PD-10 column (GE Healthcare Life Sciences, Pittsburgh, PA, USA).

LC-MS/MS analyses

The procedure has small modifications from the previous protocol [55]. The purified proteins were trypsin digested by a standard in-gel digestion protocol, and analyzed by LC-MS/MS on an LTQ Orbitrap XL (Thermo Fisher Scientific) equipped with a nanoACQUITY UPLC system (Waters, Milford, MA, USA). A Symmetry C18 trap column (180 μm × 20 mm; Waters) and a nanoACQUITY UPLC column (1.7 μm, 100 μm × 250 mm, 35 °C) were used for peptide separation. Trapping was done at 15 $\mu L \cdot min^{-1}$, 99% buffer A (0.1% formic acid) for 1 min. Peptide separation was performed at 300 $nL \cdot min^{-1}$ with buffer A and buffer B (CH_3CN containing 0.1% formic acid). The linear gradient was from 5% buffer B to 50% buffer B at 50 min, and to 85% B at 51 min. MS data were acquired in the Orbitrap with one microscan, and a maximum inject time of 900 ms followed by data-dependent MS/MS acquisitions in the ion trap (through collision-induced dissociation). The MASCOT search algorithm was used to search for the appropriate non-canonical substitution (Matrix Science, Boston, MA, USA).

Malate dehydrogenase activity assay

Malate dehydrogenase (MDH) activity assays were performed by following the instructions for the EnzyChrom™ Malate Dehydrogenase Assay Kit (BioAssay Systems,

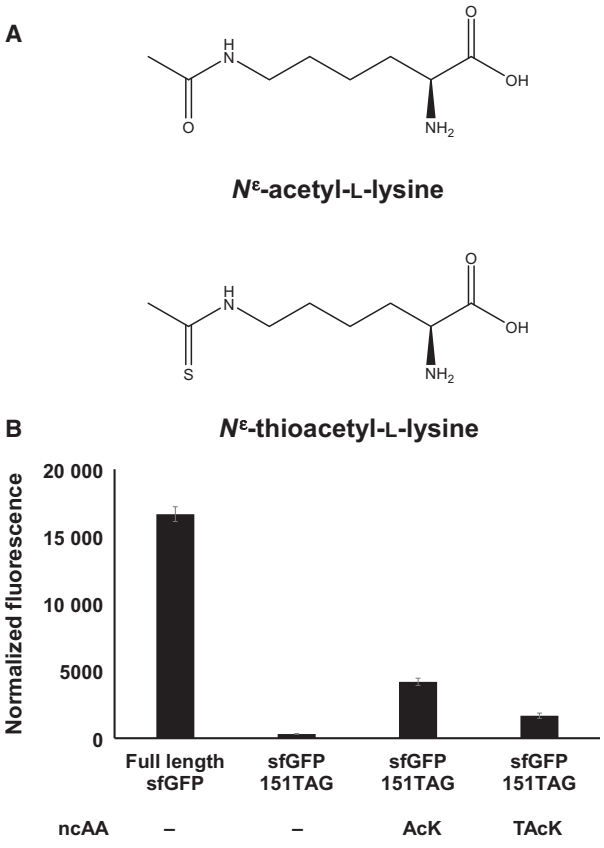

Fig. 2. Recognition of TAcK by AcKRS. (A) The structures of AcK and TAcK. (B) The readthrough of the TAG codon at position 151 in sfGFP with AcK- or TAcK-charged $tRNA^{Pyl}$ generated by AcKRS; 5 mM AcK or TAcK was used. Normalized fluorescence intensities were calculated from absolute fluorescence intensities read at 12 h normalized by corresponding cell densities. Mean values and standard errors were calculated from three replicates.

Hayward, CA, USA). This non-radioactive, colorimetric MDH assay is based on the reduction of the tetrazolium salt 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) in an NADH-coupled enzymatic reaction to a reduced form of MTT that exhibits an absorption maximum at 565 nm. The increase in absorbance at 565 nm is proportional to the enzyme activity.

CobB-mediated deacetylation

The reaction was performed in buffer containing 40 mM HEPES (pH 7.0), 6 mM $MgCl_2$, 1.0 mM NAD^+, 1 mM DTT and 10% glycerol. Ten micrograms of MDH variants, 10 μg CobB and reaction buffer were incubated at 37 °C in a total volume of 100 μL. The treated proteins were used directly for western blotting.

Table 1. Sequence comparison of TAcKRSs and AcKRS.

WT PylRS	L266	L270	Y271	L274	C313	W382
AcKRS (ref. [37])	M	I	F	A	F	W
TAcKRS-1 (this work)	M	I	L	A	C	W
TAcKRS-2 (this work)	M	I	L	A	M	W
TAcKRS-3 (this work)	M	I	N	A	I	W

Results and Discussion

Recognition of TAcK by acetyllysyl-tRNA synthetase

Due to structural similarity between AcK and TAcK (Fig. 2A) as well as the substrate flexibility of PylRS-derived acetyllysyl-tRNA synthetase (AcKRS) [56,57],

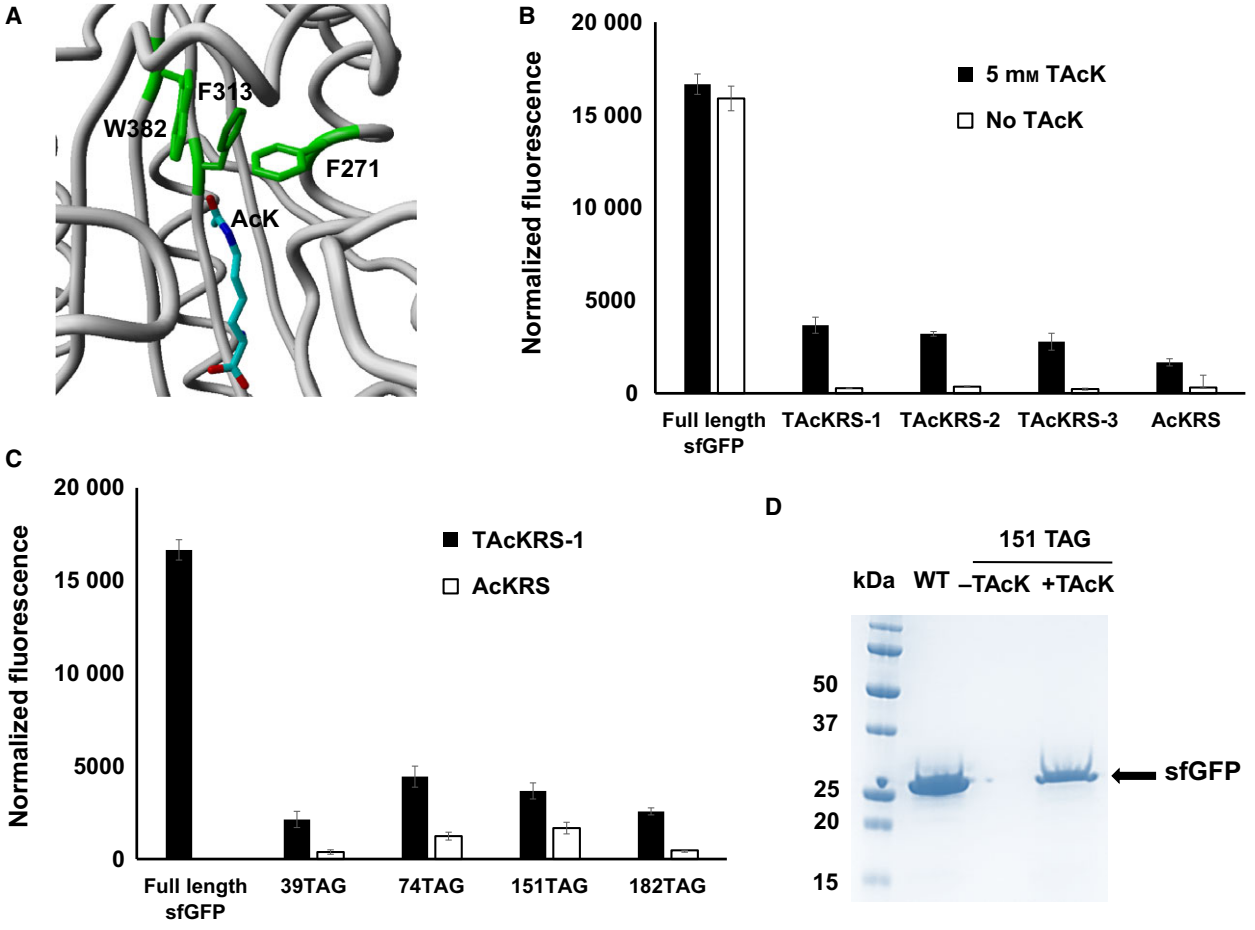

Fig. 3. TAcKRS engineering. (A) The active site of AcKRS bound to AcK (PDB ID: 4q6 g). Residues F271, F313 and W382 were mutated in the complete randomization library. (B) The sfGFP readthrough assay for TAcKRS variants. Normalized fluorescence intensities were calculated from absolute fluorescence intensities read at 12 h normalized by corresponding cell densities. Mean values and standard errors were calculated from three replicates. (C) Comparison of TAcKRS-1 and AcKRS efficiencies for TAcKRS incorporation at different positions in sfGFP; 5 mM TAcK was used in assays. The background of normalized fluorescence read from media without TAcK was subtracted for each reading. Mean values and standard errors were calculated from three replicates. (D) The Coomassie blue-stained SDS/PAGE gel of purified full-length sfGFP and its TAcK-containing variant. The same volumes of elution fractions were loaded on the SDS/PAGE gel.

we firstly tested whether the original AcKRS could incorporate TAcK into proteins. Here we used the superfolder green fluorescent protein (sfGFP) as a reporter. The result showed that the suppression of the TAG codon at the permissive position 151 of sfGFP was only about 10% with the TAcK-charged $tRNA^{Pyl}$ that was generated by AcKRS (Fig. 2B). Thus we further engineered AcKRS for higher TAcK incorporation efficiency.

Selection of TAcK-specific aminoacyl-tRNA synthetase variants

Based on the co-crystal structure of the AcKRS and AcK (Fig. 3A), three residues of AcKRS (F271, F313 and W382) that are proximate to the oxygen atom of the acetyl-group in AcK (replaced by a sulfur atom in TAcK) were selected for optimizing TAcK binding [56]. A library of AcKRS variants with complete randomization of these three residues was generated and subjected to a series of chloramphenicol resistance-based positive selections and a toxin CcdB-based negative selection. After the second positive selection, we obtained three unique variants: TAcKRS-1 (F271L and F313C), TAcKRS-2 (F271L and F313M) and TAcKRS-3 (F271N and F313I) (Table 1). The phenylalanine residue at position 271 or 313 was replaced by amino acids with smaller side chains to make a larger volume for the sulfur atom in TAcK. The tryptophan residue at position 382 was not changed in all the three variants, consistent with previous studies that have indicated that this tryptophan residue is important for the recognition of lysine analogs [35,58,59].

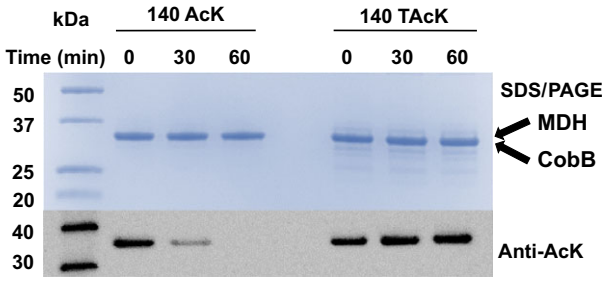

Fig. 5. Resistance of TAcK-containing MDH against CobB. The Coomassie blue-stained SDS/PAGE gel and western blotting of CobB treatment of AcK- and TAcK-containing MDH variants. The same amounts of proteins were loaded on the gels. CobB and MDH have similar molecular masses and were overlapped in the SDS/PAGE gel. Loaded samples were obtained at the start, 30 min and 60 min from deacetylation reaction mixtures.

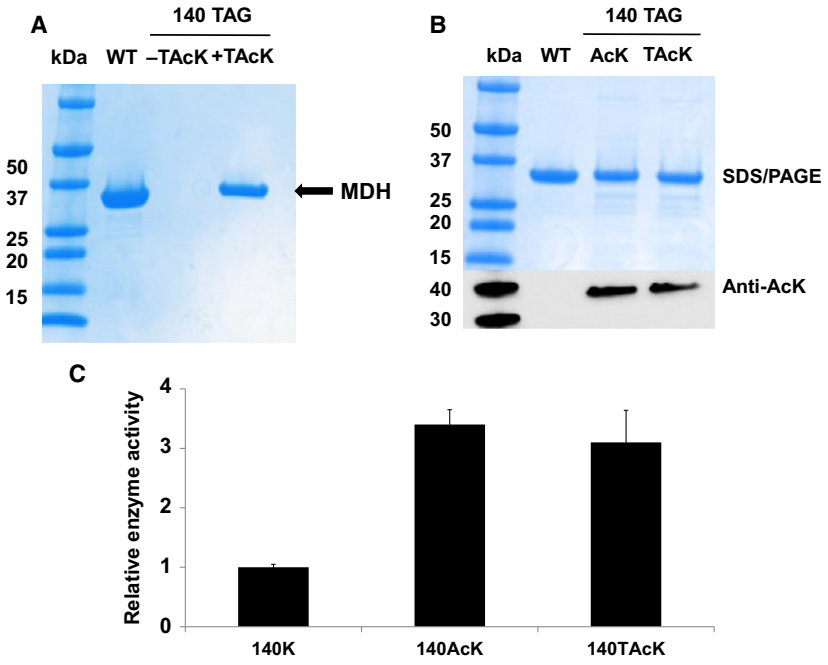

Fig. 4. Characterization of TAcK-containing MDH. (A) The Coomassie blue-stained SDS/PAGE gel of purified full-length MDH and its TAcK-containing variant. The same volumes of elution fractions were loaded on the SDS/PAGE gel. (B) The Coomassie blue-stained SDS/PAGE gel and western blotting of purified full-length MDH and its AcK- and TAcK-containing variants. The same amount of proteins were loaded on the gels. (C) The enzyme activities of MDH and its variants. Mean values and standard errors were calculated from three replicates. The enzyme activity of wild-type MDH was set as 1.

Next, we used the sfGFP readthrough assay mentioned above to evaluate the TAcK incorporation efficiencies of these three variants. The best variant is TAcKRS-1 (Fig. 3B). It is known that the ncAA incorporation efficiency depends on incorporation positions [60], and thus we tested the TAcK incorporation at different positions in sfGFP. The results showed that the engineered TAcKRS-1 could increase TAcK incorporation up to 6-fold from the original AcKRS (Fig. 3C). To lower the near cognate suppression of the TAG codon with canonical amino acids, TOP10 cells were used as the expression strain [61]. We purified the TAcK-containing sfGFP with a yield of 32 $mg{\cdot}L^{-1}$ culture (the yield of wild-type sfGFP was 149 $mg{\cdot}L^{-1}$ culture in the same growth condition) (Fig. 3D), and the TAcK incorporation at the position 151 of sfGFP was confirmed with mass spectrometry (MS) (Fig. S7). The MS results also did not show any canonical amino acid incorporation at this position.

Function of TAcK as a mimic of AcK

To validate our system in functional proteins, we chose malate dehydrogenase (MDH), which plays a crucial role in the tricarboxylic acid cycle and glyoxylate bypass, as the target to characterize the replacement of lysine acetylation with thioacetylation. Our previous study showed that the acetylation of lysine residue 140 in MDH could increase the enzyme activity by 3.4-fold [44,54], and thus we mutated the corresponding position of K140 in the MDH gene to a TAG stop codon and expressed the MDH variant with the TAcK incorporation system. The yield of TAcK-containing MDH was 9 $mg{\cdot}L^{-1}$ culture, while the yield of wide-type MDH was 32 $mg{\cdot}L^{-1}$ culture in the same growth conditions (Fig. 4A). The incorporation of TAcK in MDH was confirmed by MS (Fig. S8). Also, MS analysis did not show any canonical amino acid incorporation at position 140 in MDH. Western blotting demonstrated that TAcK could also be detected by the anti-AcK antibody with a similar intensity of AcK detection (Fig. 4B). The enzyme assay showed that the MDH-140TAcK variant could increase the enzyme activity by 3-fold, similar to that of the MDH-140AcK variant (Fig. 4C). These results indicated that TAcK could serve as an ideal functional mimic of AcK.

Deacetylation of TAcK by deacetylase

To test resistance of TAcK-containing proteins against deacetylases, we chose the CobB protein, which is a bacterial sirtuin deacetylase with a wide range of substrates as a representative [62,63]. We have shown that acetylation of K140 in MDH was sensitive to CobB [44]. Since TAcK and AcK had a similar intensity against the anti-AcK antibody (Fig. 3B), we performed western blotting to determine the deacetylation effect. After incubating the TAcK- or AcK-containing MDH variant at position K140 with CobB individually, there was significant loss of acetylation of AcK-containing MDH, while TAcK-containing MDH could resist CobB deacetylase (Fig. 5).

Conclusions

In summary, we established an *in vivo* genetic incorporation system of TAcK, which is both a good functional mimic and a non-deacetylatable analog of AcK, and this will facilitate studies of protein acetylation in living bacterial cells. Furthermore, the pair of PylRS–$tRNA^{Pyl}$ and its derivatives are orthogonal in eukaryotic cells, and previous studies have showed that AcKRS could incorporate AcK into protein by using mammalian cells as the hosts [36], so our approach could be further extended to living eukaryotic cells for a broader range of medical and pharmaceutical applications.

Acknowledgements

This work was supported by a grant from National Institutes of Health (National Institute of Allergy and Infectious Diseases; AI119813 to CF), a start-up fund from University of Arkansas to CF and awards from Arkansas Biosciences Institute to CF and MM.

Author contributions

MM and CF designed the experiments. SV, DTN, CG, QG and CF performed the experiments. SV, DTN, MM and CF analyzed the data and wrote the paper.

References

1 Kouzarides T (2000) Acetylation: a regulatory modification to rival phosphorylation? *EMBO J* **19**, 1176–1179.
2 Kim GW and Yang XJ (2011) Comprehensive lysine acetylomes emerging from bacteria to humans. *Trends Biochem Sci* **36**, 211–220.
3 Cohen T and Yao TP (2004) AcK-knowledge reversible acetylation. *Sci STKE* **2004**, pe42.
4 Escalante-Semerena JC (2010) Nε-acetylation control conserved in all three life domains. *Microbe* **5**, 340–344.

5 Soppa J (2010) Protein acetylation in archaea, bacteria, and eukaryotes. *Archaea* **2010**, 820681.
6 Kouzarides T (2007) Chromatin modifications and their function. *Cell* **128**, 693–705.
7 Shahbazian MD and Grunstein M (2007) Functions of site-specific histone acetylation and deacetylation. *Annu Rev Biochem* **76**, 75–100.
8 Close P, Creppe C, Gillard M, Ladang A, Chapelle JP, Nguyen L and Chariot A (2010) The emerging role of lysine acetylation of non-nuclear proteins. *Cell Mol Life Sci* **67**, 1255–1264.
9 Scott I (2012) Regulation of cellular homoeostasis by reversible lysine acetylation. *Essays Biochem* **52**, 13–22.
10 Banreti A, Sass M and Graba Y (2013) The emerging role of acetylation in the regulation of autophagy. *Autophagy* **9**, 819–829.
11 Choudhary C, Weinert BT, Nishida Y, Verdin E and Mann M (2014) The growing landscape of lysine acetylation links metabolism and cell signalling. *Nat Rev Mol Cell Biol* **15**, 536–550.
12 Koprinarova M, Schnekenburger M and Diederich M (2016) Role of histone acetylation in cell cycle regulation. *Curr Top Med Chem* **16**, 732–744.
13 Menzies KJ, Zhang H, Katsyuba E and Auwerx J (2016) Protein acetylation in metabolism - metabolites and cofactors. *Nat Rev Endocrinol* **12**, 43–60.
14 McKinsey TA and Olson EN (2004) Cardiac histone acetylation–therapeutic opportunities abound. *Trends Genet* **20**, 206–213.
15 Batta K, Das C, Gadad S, Shandilya J and Kundu TK (2007) Reversible acetylation of non histone proteins: role in cellular function and disease. *Subcell Biochem* **41**, 193–212.
16 Glozak MA and Seto E (2007) Histone deacetylases and cancer. *Oncogene* **26**, 5420–5432.
17 Iyer A, Fairlie DP and Brown L (2012) Lysine acetylation in obesity, diabetes and metabolic disease. *Immunol Cell Biol* **90**, 39–46.
18 Al-Haddad R, Karnib N, Assaad RA, Bilen Y, Emmanuel N, Ghanem A, Younes J, Zibara V, Stephan JS and Sleiman SF (2016) Epigenetic changes in diabetes. *Neurosci Lett* **625**, 64–69.
19 Bonnaud EM, Suberbielle E and Malnou CE (2016) Histone acetylation in neuronal (dys)function. *Biomol Concepts* **7**, 103–116.
20 Gong F, Chiu LY and Miller KM (2016) Acetylation reader proteins: linking acetylation signaling to genome maintenance and cancer. *PLoS Genet* **12**, e1006272.
21 Guarente L (2006) Sirtuins as potential targets for metabolic syndrome. *Nature* **444**, 868–874.
22 Gan L (2007) Therapeutic potential of sirtuin-activating compounds in Alzheimer's disease. *Drug News Perspect* **20**, 233–239.
23 Marchion D and Munster P (2007) Development of histone deacetylase inhibitors for cancer treatment. *Expert Rev Anticancer Ther* **7**, 583–598.
24 Kaypee S, Sudarshan D, Shanmugam MK, Mukherjee D, Sethi G and Kundu TK (2016) Aberrant lysine acetylation in tumorigenesis: implications in the development of therapeutics. *Pharmacol Ther* **162**, 98–119.
25 Marmorstein R and Roth SY (2001) Histone acetyltransferases: function, structure, and catalysis. *Curr Opin Genet Dev* **11**, 155–161.
26 Thiagalingam S, Cheng KH, Lee HJ, Mineva N, Thiagalingam A and Ponte JF (2003) Histone deacetylases: unique players in shaping the epigenetic histone code. *Ann N Y Acad Sci* **983**, 84–100.
27 Hodawadekar SC and Marmorstein R (2007) Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design. *Oncogene* **26**, 5528–5540.
28 Lee KK and Workman JL (2007) Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol* **8**, 284–295.
29 Yang XJ and Seto E (2007) HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* **26**, 5310–5318.
30 Wagner GR and Payne RM (2013) Widespread and enzyme-independent Nepsilon-acetylation and Nepsilon-succinylation of proteins in the chemical conditions of the mitochondrial matrix. *J Biol Chem* **288**, 29036–29045.
31 Seto E and Yoshida M (2014) Erasers of histone acetylation: the histone deacetylase enzymes. *Cold Spring Harb Perspect Biol* **6**, a018713.
32 Baeza J, Smallegan MJ and Denu JM (2015) Site-specific reactivity of nonenzymatic lysine acetylation. *ACS Chem Biol* **10**, 122–128.
33 Simic Z, Weiwad M, Schierhorn A, Steegborn C and Schutkowski M (2015) The ε-amino-amino group of protein lysine residues is highly susceptible to nonenzymatic acylation by several physiological acyl-CoA thioesters. *ChemBioChem* **16**, 2337–2347.
34 Neumann H, Peak-Chew SY and Chin JW (2008) Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. *Nat Chem Biol* **4**, 232–234.
35 Umehara T, Kim J, Lee S, Guo LT, Soll D and Park HS (2012) N-acetyl lysyl-tRNA synthetases evolved by a CcdB-based selection possess N-acetyl lysine specificity *in vitro* and *in vivo*. *FEBS Lett* **586**, 729–733.
36 Mukai T, Kobayashi T, Hino N, Yanagisawa T, Sakamoto K and Yokoyama S (2008) Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases. *Biochem Biophys Res Commun* **371**, 818–822.
37 Neumann H, Hancock SM, Buning R, Routh A, Chapman L, Somers J, Owen-Hughes T, Van Noort J,

Rhodes D and Chin JW (2009) A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. *Mol Cell* **36**, 153–163.

38 Lammers M, Neumann H, Chin JW and James LC (2010) Acetylation regulates cyclophilin A catalysis, immunosuppression and HIV isomerization. *Nat Chem Biol* **6**, 331–337.

39 Thao S, Chen CS, Zhu H and Escalante-Semerena JC (2010) Nepsilon-lysine acetylation of a bacterial transcription factor inhibits its DNA-binding activity. *PLoS ONE* **5**, e15123.

40 Arbely E, Natan E, Brandt T, Allen MD, Veprintsev DB, Robinson CV, Chin JW, Joerger AC and Fersht AR (2011) Acetylation of lysine 120 of p53 endows DNA-binding specificity at effective physiological salt concentration. *Proc Natl Acad Sci USA* **108**, 8251–8256.

41 Oppikofer M, Kueng S, Martino F, Soeroes S, Hancock SM, Chin JW, Fischle W and Gasser SM (2011) A dual role of H4K16 acetylation in the establishment of yeast silent chromatin. *EMBO J* **30**, 2610–2621.

42 Elsasser SJ, Huang H, Lewis PW, Chin JW, Allis CD and Patel DJ (2012) DAXX envelops a histone H3.3-H4 dimer for H3.3-specific recognition. *Nature* **491**, 560–565.

43 Yu W, Dittenhafer-Reed KE and Denu JM (2012) SIRT3 protein deacetylates isocitrate dehydrogenase 2 (IDH2) and regulates mitochondrial redox status. *J Biol Chem* **287**, 14078–14086.

44 Venkat S, Gregory C, Sturges J, Gan Q and Fan C (2017) Studying the lysine acetylation of malate dehydrogenase. *J Mol Biol* **429**, 1396–1405.

45 Venkat S, Gregory C, Gan Q and Fan C (2017) Biochemical characterization of the lysine acetylation of tyrosyl-tRNA synthetase in *Escherichia coli*. *Chembiochem* **18**, 1928–1934.

46 Fatkins DG, Monnot AD and Zheng W (2006) Nepsilon-thioacetyl-lysine: a multi-facet functional probe for enzymatic protein lysine Nepsilon-deacetylation. *Bioorg Med Chem Lett* **16**, 3651–3656.

47 Smith BC and Denu JM (2007) Mechanism-based inhibition of Sir2 deacetylases by thioacetyl-lysine peptide. *Biochemistry* **46**, 14478–14486.

48 Fatkins DG and Zheng W (2008) Substituting N (epsilon)-thioacetyl-lysine for N(epsilon)-acetyl-lysine in peptide substrates as a general approach to inhibiting human NAD(+)-dependent protein deacetylases. *Int J Mol Sci* **9**, 1–11.

49 Kiviranta PH, Suuronen T, Wallén EA, Leppänen J, Tervonen J, Kyrylenko S, Salminen A, Poso A and Jarho EM (2009) N(epsilon)-thioacetyl-lysine-containing tri-, tetra-, and pentapeptides as SIRT1 and SIRT2 inhibitors. *J Med Chem* **52**, 2153–2156.

50 Suzuki T, Asaba T, Imai E, Tsumoto H, Nakagawa H and Miyata N (2009) Identification of a cell-active non-peptide sirtuin inhibitor containing N-thioacetyl lysine. *Bioorg Med Chem Lett* **19**, 5670–5672.

51 Chen B, Wang J, Huang Y and Zheng W (2015) Human SIRT3 tripeptidic inhibitors containing N (epsilon)-thioacetyl-lysine. *Bioorg Med Chem Lett* **25**, 3481–3487.

52 Huang Y, Liu J, Yan L and Zheng W (2016) Simple N (epsilon)-thioacetyl-lysine-containing cyclic peptides exhibiting highly potent sirtuin inhibition. *Bioorg Med Chem Lett* **26**, 1612–1617.

53 Xiong H, Reynolds NM, Fan C, Englert M, Hoyer D, Miller SJ and Söll D (2016) Dual genetic encoding of acetyl-lysine and non-deacetylatable thioacetyl-lysine mediated by flexizyme. *Angew Chem Int Ed Engl* **55**, 4083–4086.

54 Fan C, Xiong H, Reynolds NM and Soll D (2015) Rationally evolving tRNAPyl for efficient incorporation of noncanonical amino acids. *Nucleic Acids Res* **43**, e156.

55 Gan Q, Lehman BP, Bobik TA and Fan C (2016) Expanding the genetic code of Salmonella with non-canonical amino acids. *Sci Rep* **6**, 39920.

56 Guo LT, Wang YS, Nakamura A, Eiler D, Kavran JM, Wong M, Kiessling LL, Steitz TA, O'Donoghue P and Söll D (2014) Polyspecific pyrrolysyl-tRNA synthetases from directed evolution. *Proc Natl Acad Sci USA* **111**, 16724–16729.

57 Wan W, Tharp JM and Liu WR (2014) Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochim Biophys Acta* **1844**, 1059–1070.

58 Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K and Yokoyama S (2008) Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode N(epsilon)-(o-azidobenzyloxycarbonyl) lysine for site-specific protein modification. *Chem Biol* **15**, 1187–1197.

59 Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K and Yokoyama S (2008) Crystallographic studies on multiple conformational states of active-site loops in pyrrolysyl-tRNA synthetase. *J Mol Biol* **378**, 634–652.

60 Albayrak C and Swartz JR (2013) Cell-free co-production of an orthogonal transfer RNA activates efficient site-specific non-natural amino acid incorporation. *Nucleic Acids Res* **41**, 5949–5963.

61 O'Donoghue P, Prat L, Heinemann IU, Ling J, Odoi K, Liu WR and Soll D (2012) Near-cognate suppression of amber, opal and quadruplet codons competes with aminoacyl-tRNAPyl for genetic code expansion. *FEBS Lett* **586**, 3931–3937.

62 Zhao K, Chai X and Marmorstein R (2004) Structure and substrate binding properties of cobB, a Sir2

homolog protein deacetylase from *Escherichia coli*. *J Mol Biol* **337**, 731–741.

63 Starai VJ, Celic I, Cole RN, Boeke JD and Escalante-Semerena JC (2002) Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. *Science* **298**, 2390–2392.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Fig. S1. The ^{1}H-NMR spectrum of N^{α}-Boc-N^{ε}-thioacetyl-L-lysine.

Fig. S2. The ^{13}C-NMR spectrum of N^{α}-Boc-N^{ε}-thioacetyl-L-lysine.

Fig. S3. The ^{1}H-NMR spectrum of N^{ε}-thioacetyl-L-lysine.

Fig. S4. The ^{13}C-NMR spectrum of N^{ε}-thioacetyl-L-lysine.

Fig. S5. The ESI-FTMS spectrum of N^{ε}-thioacetyl-L-lysine.

Fig. S6. The growth of cells harboring the original AcKRS, tRNA^{Pyl} and a TAG-containing mutant *cat* gene on different concentrations of chloramphenicol (Cm) in the LB-TK-TAcK plates.

Fig. S7. LC-MS/MS analysis of sfGFP 151-TAcK.

Fig. S8. LC-MS/MS analysis of MDH 140-TAcK.