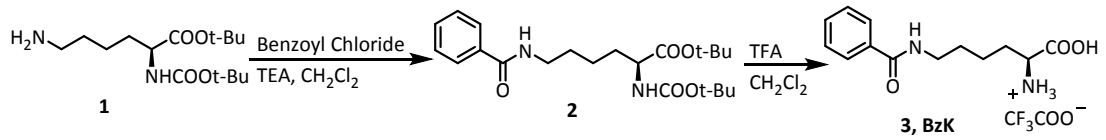
Genetically Encoding *ε-N*-Benzoyllysine in Proteins

Supporting Information

Materials and Methods

Isopropyl-β-D-thiogalactoside (IPTG) was purchased from Chemsynlab, and L-arabinose was purchased from Meryer. PrimeSTAR HS (Premix) DNA Polymerase was obtained from Takara. Gibson assembly master mix and dNTPs were obtained from New England Biolabs. All oligonucleotide primers were synthesized by Tsingke. Plasmid DNA preparation was carried out with the Plasmid Miniprep Kit from Zymo research and CWBIOTECH. Ni-NTA beads were purchased from GenScript, protease inhibitor was purchased from Roche, and the BCA kit was purchased from Coolaber. Antibiotics were obtained from SolarBio. *E. coli* ΔcobB (Kan^R) strain was obtained from CGSC (#9039). HEK293T cell line was purchased from National Collection of Authenticated Cell Cultures (Shanghai, China). Fetal bovine serum (FBS) was purchased from BioInd (04-001-1A). An ATPIO XO-900D sonic disruptor was used. AcetylLysine was purchased from Bide Pharmatech Ltd, SAHA from Cayman, TM from TargetMol, EX-527 and SirReal2 from Beyotime. Optical density was measured with Thermo Nanodrop OneC. Absorbance and emission spectra were measured on plater reader Varioskan Lux. Protein mass spectrometry was carried out on Agilent 6545 LC/ESI-QTOF. Data for 1H are reported as follows: chemical shift (ppm), and multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet). Data for 13C NMR are reported as ppm.

1. Chemical synthesis of BzK



(1) **Synthesis of *N'*-benzoyl-L-lysine tert-butyl ester (2).** To a solution of *N*-Boc-L-lysine tert-butyl ester (1.23 g, 4.07 mmol) in 15 ml CH_2Cl_2 at 0 ℃ was added triethylamine (848 µl, 6.1 mmol). Then benzoyl chloride (709 µl, 6.1 mmol) was added dropwise over 5 min, and the reaction mixture was stirred for 1 h at 0 ℃. The solvent was washed with saturated aqueous $NaHCO_3$ solution twice, and the organic layer was dried with anhydrous Na_2SO_4 and evaporated with a rotary evaporator. The crude product was purified by column chromatography on silica gel, eluting with 3:1 v/v petroleum ether/ethyl acetate to afford product **2** as a white solid (1.18 g, 71%). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ7.80-7.72 (m, 2H); 7.47 (t, J = 7.3 Hz, 1H); 7.40 (dd, J = 8.3, 6.6 Hz, 2H); 6.40 (s, 1H); 5.10 (d, J = 8.2 Hz, 1H); 4.15 (q, J = 7.2 Hz, 1H); 3.44 (q, J = 6.6 Hz, 2H); 1.84-1.57 (m, 4H); 1.40 (s, 20H). **^{13}C-NMR** (101 MHz, $CDCl_3$): δ171.87; 167.63; 155.53; 134.71 , 131.28; 128.47; 126.91; 81.87; 79.63; 53.72; 39.71; 32.74; 29.06; 28.30; 27.99; 22.63.

(2) **Synthesis of BzK.** To a solution of *N'*-benzoyl-L-lysine tert-butyl ester 2 (1.18 g, 2.90 mmol) in 22 ml CH_2Cl_2 was added 11 ml trifluoroacetic acid, and the solution was stirred for 5 h at room temperature. The solvent was removed with a rotary evaporator. The residue was washed with

diethyl ether, and the obtained precipitate was dried under high vacuum to afford the final product as brown oil in a quantitative yield. **^{1}H-NMR** (400 MHz, D_2O+NaOH): δ7.70-7.63 (m, 2H); 7.59-7.50 (m, 1H); 7.46 (dd, *J* = 8.3, 6.9 Hz, 2H); 3.32 (t, *J* = 7.0 Hz, 2H); 3.19 (t, *J* = 6.4 Hz, 1H); 1.35 (p, *J* = 7.8 Hz, 2H). **^{13}C-NMR** (101 MHz, D_2O+NaOH): δ183.63; 170.77; 133.80 , 131.92; 128.70; 126.94; 55.92; 39.80; 34.38; 28.37; 22.54. **HRMS–ESI** (m/z): $[M+H]^+$ calc. for $C_{13}H_{18}N_2O_3$, 251.1317; found, 251.1397.

2. Screening *Mb*PylRS variants to genetically incorporate BzK.

PylRS variants (table S1) in pUltra vector were transformed into *E. coli* DH10B containing pET22b-T5-sfGFP-Y151TAG. Three single colonies were picked and inoculated into 1 mL LB broth. After overnight growth at 37 °C, the culture was diluted 50 folds with LB and supplemented with or without 2 mM BzK. When the OD reached 0.8, 1 mM IPTG was added to the culture, and the culture was grown for an additional 10 h at 37 °C. *E. coli* was pelleted and re-suspended in PBS. The absorption and fluorescence spectra were measured with a microplate spectrophotometer, and the fluorescence was normalized by fluorescence/OD600.

3. Comparing the efficiency of PylT and opt-PylT in the genetic incorporation of BzK

pUltra vector vectoring harboring BzKRS and PylT or opt-PylT was transformed into electrocompetent *E. coli* DH10B containing pET22b-T5-sfGFP-Y151TAG. Three single colonies were picked and inoculated into 4 mL 2YT. After overnight growth at 37 °C, the culture was diluted 100-fold by 2YT. When OD600 reached 0.5, the culture was supplemented with or without 2 mM BzK. After 30 min growth, 1 mM IPTG was added. Then, the culture was grown for an additional 8 h at 37 °C. Cells were pelleted and re-suspended in PBS. The absorption and fluorescence intensity were measured with a microplate spectrophotometer , and the fluorescence was normalized by fluorescence/OD600.

4. Purification procedure of sfGFP-Y151BzK.

Electrocompetent *E. coli* ΔcobB was co-transformed with pUltra-opt-PylT-BzKRS and pET22b-T5-sfGFP-Y151TAG. A single colony was picked and inoculated into 4 mL 2YT media (50 μg/mL Spectinomycin, 100 μg/mL Ampicillin, 25 μg/mL Kanamycin) at 37 °C overnight. The saturated culture was diluted 100-fold (500 μL to 50 mL) with 2YT and grown at 37 °C. When the OD reached 0.6, 2 mM BzK was added and the culture was grown for another 30 min before induction with 1 mM IPTG. Cells were grown at 37 °C and harvested after 8 h. The cells were lysed with a sonic disruptor in the presence of protease inhibitors. Insoluble protein and cell debris were removed by centrifugation. The supernatants was purified by Ni^{2+} affinity chromatography (according to the manufacture's protocol), and finally buffer was exchanged with an Amicon centrifugal filter. The protein concentration was determined by BCA assay.

5. Procedure for the purification of H3 with site-specifically incorporated BzK and the DNA sequence used in this work for H3 expression.

Amber stop codon was introduced into H3 at K18 or K27 by site-directed mutagenesis. The plasmid

pET22b-H3-K18TAG or pET22b-H3-K27TAG was co-transformed into electrocompetent *E.coli* BL21(DE3) along with pUltra-opt-PylT-BzKRS. Single colonies were picked and inoculated in 2YT media (100 μg/mL Ampicillin, 50 μg/mL Spectinomycin). After overnight growth at 37 °C, the culture was diluted 100-fold into 2YT, and then grown at 37 °C. When OD600 reached 0.5, the culture was supplemented with 2 mM BzK. After another 30 min, 1 mM IPTG was added for induction. Cells were harvested after an additional 12 h growth at 37 °C, and sonicated in lysis buffer (50 mM Tris, 500 mM NaCl, 0.1% Triton X-100, pH 8.0). Inclusion bodies were collected and dissolved in LE buffer (100 mM Na_2HPO_4, 10 mM Tris, 8 M urea, pH 8.0). The solution was incubated with Ni-NTA beads, and then washed with a washing buffer (100 mM Na_2HPO_4, 10 mM Tris, 20 mM Imidazole, 8 M urea, pH 8.0). Finally, the protein was collected from Ni-NTA beads with an elution buffer (100 mM Na_2HPO_4, 10 mM Tris, 500 mM imidazole, 8 M urea, pH 8.0). The protein concentration was determined with BCA assay.

The DNA sequence of H3 used in this work (sequence 5'-3'):

ATGGCGCGTACCAAACAGACCGCGCGTAAAAGCACCGGTGGCAAAGCGCCGCGT**AAA**CAGCTGGCGAC
CAAAGCGGCGCGT**AAA**AGCGCGCCGGCGACCGGCGGCGTTAAAAAACCGCACCGTTACCGTCCGGGCAC
CGTTGCGCTGCGTGAAATCCGTCGTTACCAGAAAAGCACCGAACTGCTGATCCGTAAACTGCCGTTCCAG
CGTCTGGTTCGTGAAATCGCGCAGGATTTCAAAACCGATCTGCGTTTCCAGAGCAGCGCGGTTATGGCGC
TGCAGGAAGCGTGCGAAGCGTACCTGGTTGGCCTGTTCGAAGATACCAACCTGTGCGCGATCCACGCGA
AACGTGTTACCATCATGCCGAAAGATATCCAGCTGGCGCGTCGTATCCGTGGCGAACGTGCGCACCACCA
CCACCACCACTAA

Primers used to generate H3-K18TAG and H3-K27TAG

K18TAG:
F: CGCCGCGT**TAG**CAGCTGGCGACCAAAGCGG
R: GCCAGCTG**CTA**ACGCGGCGCTTTGCCA

K27TAG:
F: CGGCGCGT**TAG**AGCGCGCCGGCGAC
R: GGCGCGCT**CTA**ACGCGCCGCTTTGGTCG

6. Procedure for the genetic incorporation of BzK in HEK293T cells and western blot analysis.

HEK293T cells were grown in Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with 10% fetal bovine serum (FBS, BI) at 37 °C and 5% CO_2 in a humidified incubator. For cell transfection, about 2×10^5 HEK293T cells per well were seeded on glass coverslips (pretreated with TC, SolarBio) for 24-well plate 12 h prior to transfection. Cells were transiently co-transfected with 1 μg intended plasmids and 2 μL lipofectamine 3000 (Invitrogen) following the manufacturer's instructions at 70% ~ 90% confluency. After 4 h, the media was supplied with or without 2 mM BzK. Then, the cells were incubated for another 24 h at 37 °C with 5% CO_2. Finally, HEK293T cells

were fixed by 4% paraformaldehyde for 10 min, and washed with 10 mM PBS (pH 7.4) for three times. Confocal fluorescence imaging was performed using Olympus FV3000RS Confocal Laser Scanning Microscope. EGFP channel: 488nm excitation wavelength and 500 to 580 nm detector slit; mcherry channel: 594nm excitation wavelength and 600 to 700 nm detector slit.

For western blot analysis, about 10^6 HEK293T cells per well plated in a 6-well plate were transfected with 5 μg of intended plasmids and cultured in the presence of 2 mM BzK for 24 h. Then, Cells were rinsed with 0.6 mL of 10 mM PBS (pH 7.4), and lysed with cell lysis buffer [500 μL CelLytic™M reagent (SIGMA), 1×Protease Inhibitor Cocktail] for 15 min on ice. Suspensions were then centrifuged (4200 rpm, 30 min, 4 °C, Thermo Heraeus Multifuge X3R), and supernatants were collected. Protein in supernatant was mixed with 5×protein loading buffer and heated at 98 °C for 5 min. The target protein was separated using SDS-PAGE (15%) for 1 h at 180V and transferred onto polyvinylidene fluoride (PVDF) transfer membrane (SolarBio). Then, the PVDF membrane was blocked for 1 h at room temperature using blocking buffer (5% skim milk in 1×TBST). After blocking, the membrane was incubated with primary antibody [anti-His antibody (Trans,1:2000 diluted)] at 4 °C overnight and washed in 1×TBST, 10 min each time for three times. The secondary antibody [anti-Mouse IgG-HRP (Trans, 1:5000 diluted)] was used to incubate the membrane at room temperature for 1 h. Next, the membrane was washed three times again, 10 min for each time. Last, chemiluminescence imaging was taken using High-sig ECL western-blotting substrate (Tanon). Western blot images were recorded by an automatic chemiluminescence imaging system (Tanon).

7. Procedure for evaluating the effect of L-arabinose concentration on the fluorescence intensity in the detection of sirtuin activity.

SIIRT1-7 were cloned into pBAD15A vector by Gibson Assembly. *E.coli* ΔcobB strains were co-transformed with pET22b-T5-EGFP, pUltra-opt-PylT-BzKRS (or pUltra-PylT-AcKRS3) and pBAD15A-SIRT1-7. Three single colonies were picked and inoculated in 96-well plate (0.4 mL 2YT media with 100 μg/mL Ampicillin, 50 μg/mL Spectinomycin, 50 μg/mL Chloramphenicol, 25 μg/mL Kanamycin per well). After overnight growth at 37 °C, the culture was diluted 100-fold into 2YT supplemented with 2 mM BzK. About 3 h later, 1 mM IPTG and different concentrations of L-arabinose were added into the culture for induction. Cells were incubated for additional 8h at 37 °C. Cells were pelleted at 4200 rpm for 15 min, and then resuspended with 10 mM PBS (pH7.4). The absorption and fluorescence intensity were measured with a microplate spectrophotometer (Thermo Varioskan LUX), and the fluorescence was standardized by fluorescence intensity/OD600, and relative fluorescence was determined by dividing fluorescence intensity obtained in the absence of sirtuin.

Table S1. The mutation sites of *Mb*PylRS variants used in this study.

AcKRS2	L270I Y271L L274A C313F
AcKRS3	L266M L270I Y271F L274A C313F
hibKRS1	C313T
hibKRS2	C313S
ThzKRS	A267S C313V M315F D344G
AcrKRS	L270I L274A C313F

Figure S1. Fluorescence screening of *Mb*PylRS variants to incorporate BzK. Error bars represent

s.d. of three independent biological replicates from a single LB agar plate.

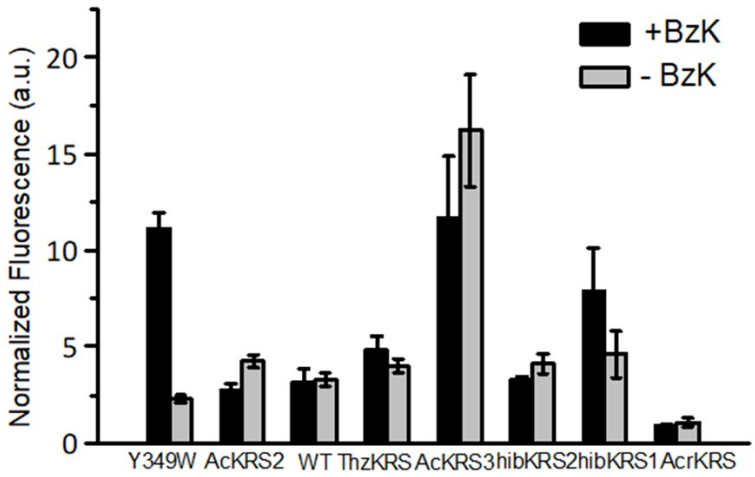


Figure S2. A) Western blot analysis of expressed EGFP-Y39BzK in HEK293T cells. The control was

wild type sfGFP. B) ESI-QTOF mass spectrometry of EGFP-Y39BzK purified from HEK293T cells.

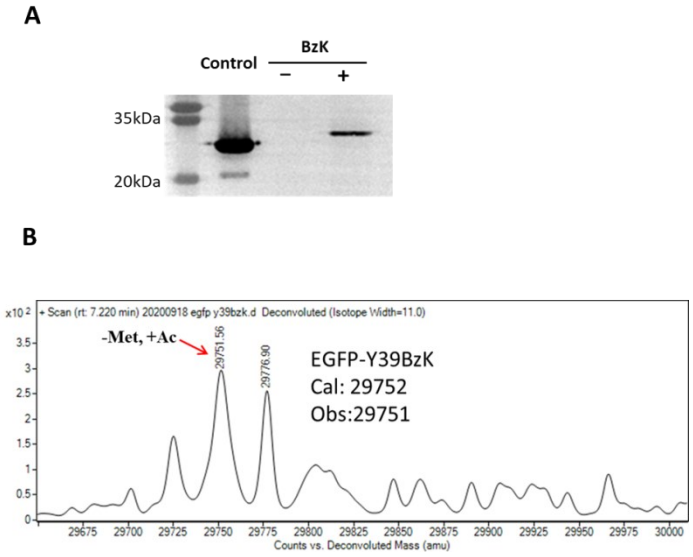


Figure S3. Mass spectra of H3 variants purified from *E. coli* ΔcobB.

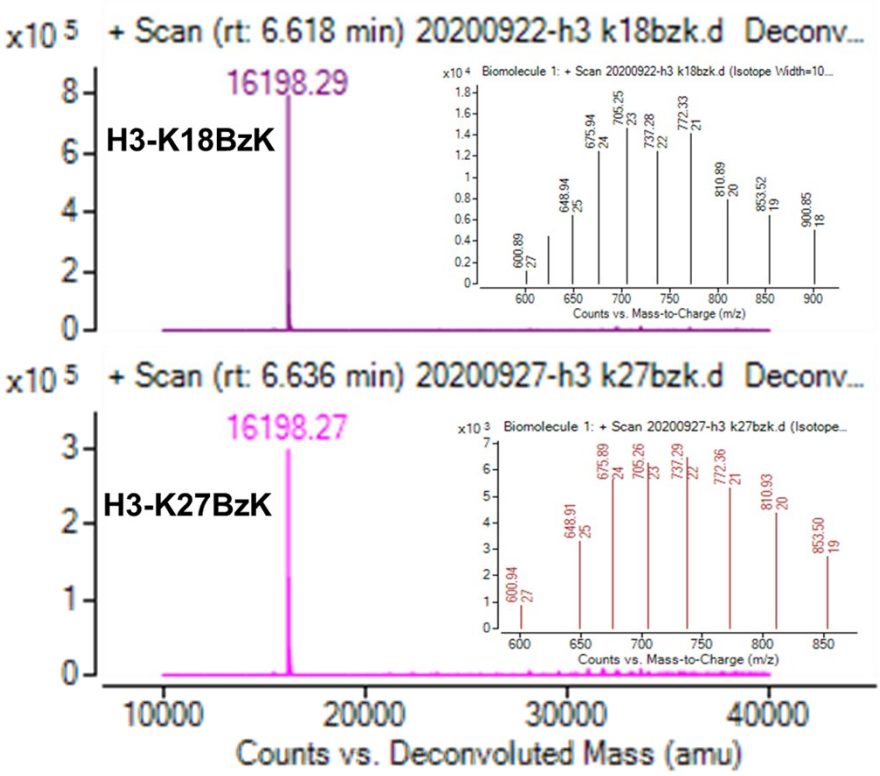


Figure S4. The effect of SIRT2 expression level on the fluorescence intensity of EGFP-K85BzK.

Relative fluorescence is defined as fluorescence intensity divided by fluorescence obtained without SIRT. Error bars represent s.d. of three independent biological replicates from a single LB agar plate.

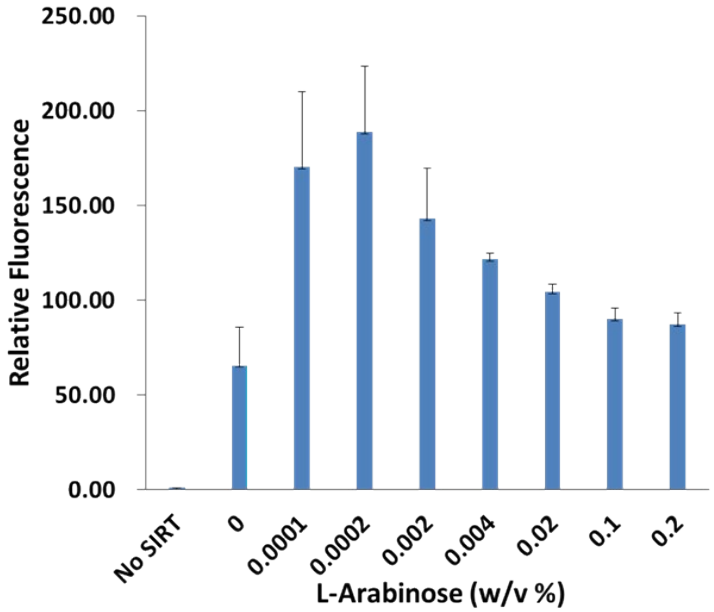


Figure S5. The expression level of SIRT1 (A) and SIRT2 (B) can significantly affect the fluorescence

intensity of EGFP-K85AcK co-expressed in *E. coli* ΔcobB. The sirtuin expression level is controlled by L-arabinose. Relative fluorescence is defined as fluorescence intensity divided by fluorescence obtained without SIRT. Error bars represent s.d. of three independent biological replicates from a single LB agar plate.

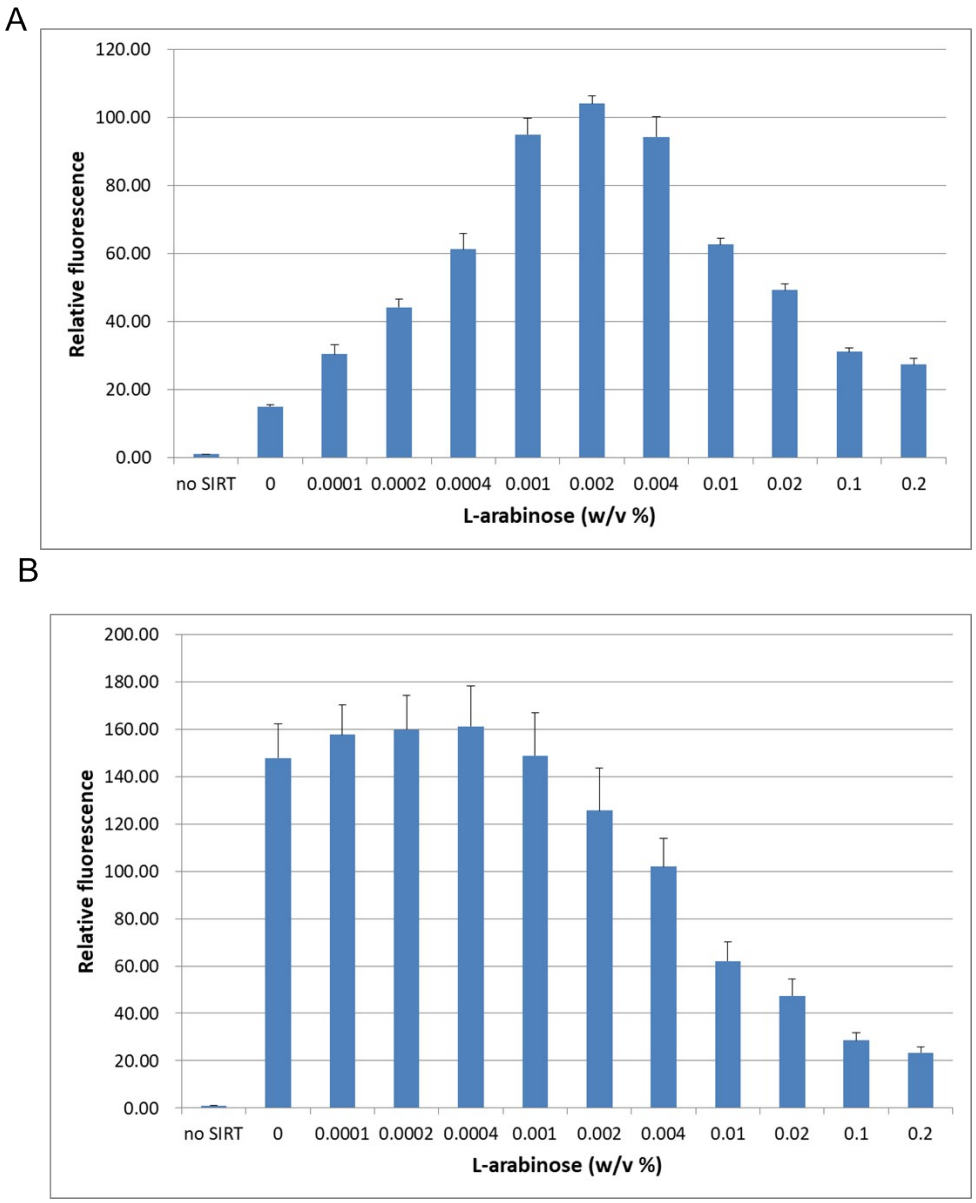


Figure S6. An evaluation of the deacetylase activity of recombinantly expressed sirtuins. The

fluorescence was normalized by fluorescence/OD600. Error bars represent s.d. of three independent biological replicates from a single LB agar plate.

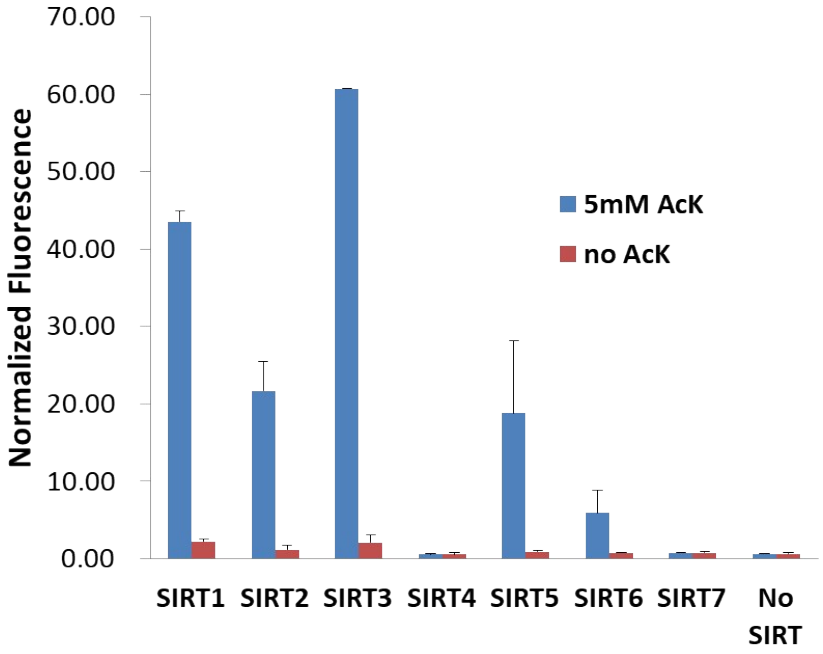
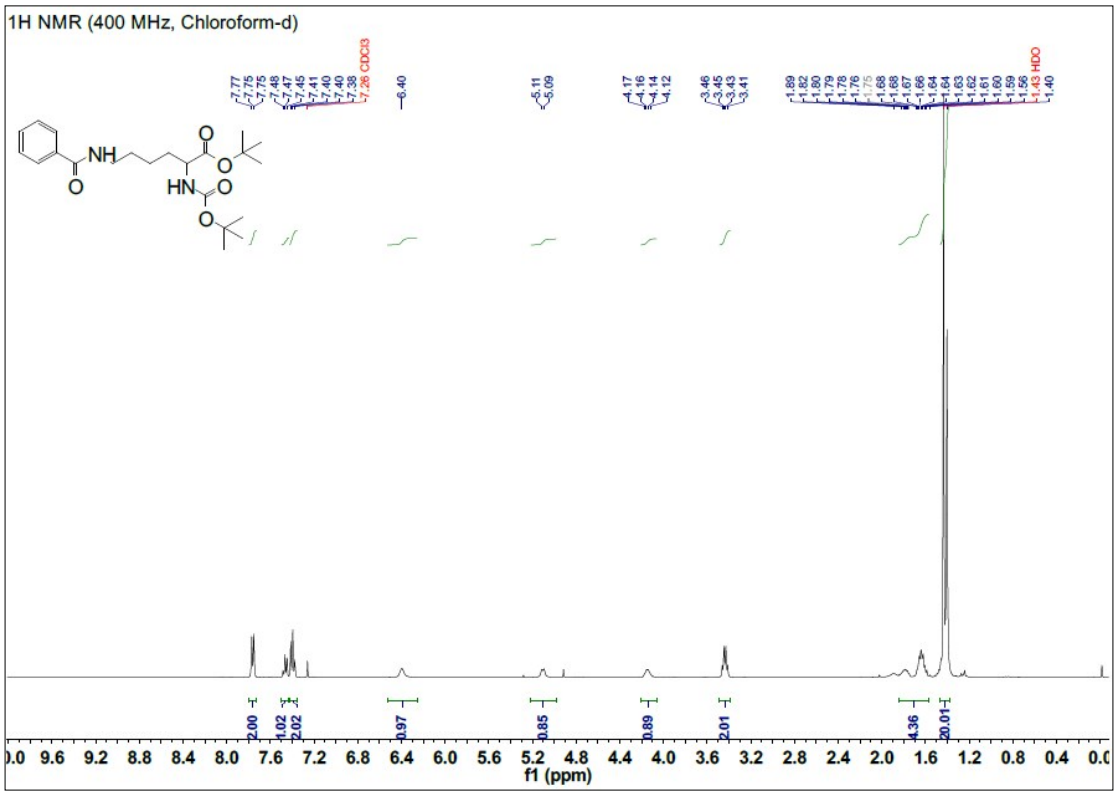
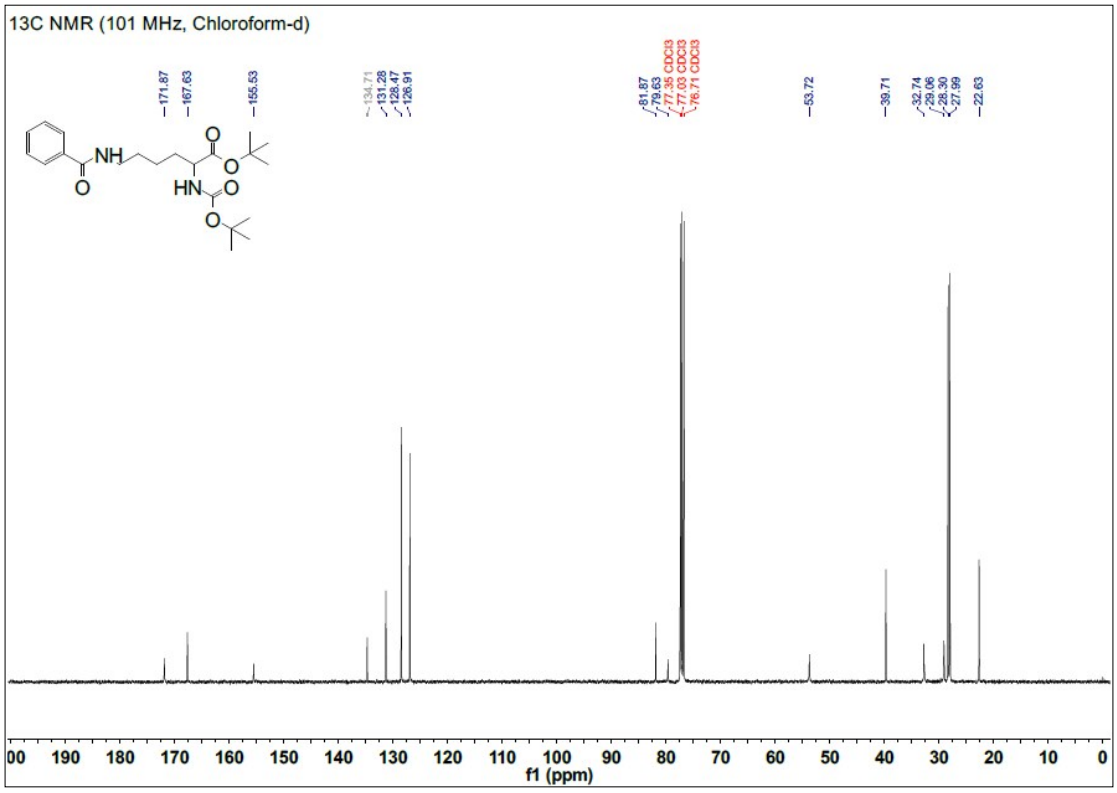
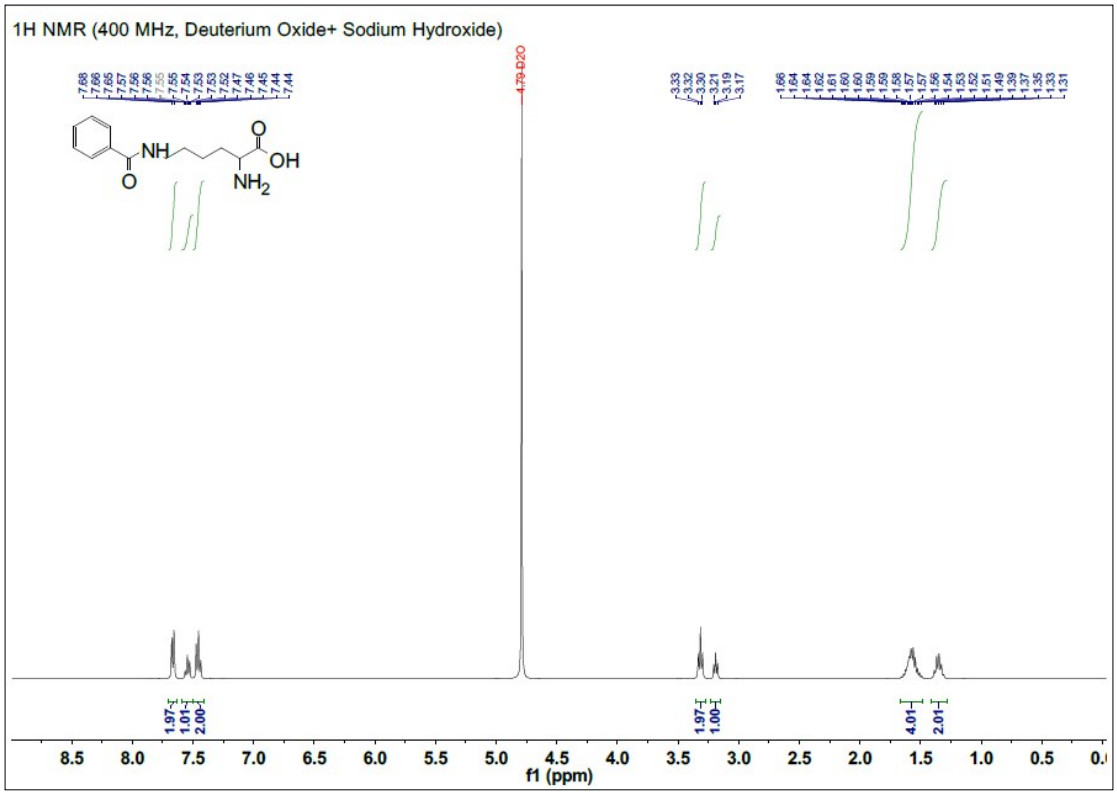
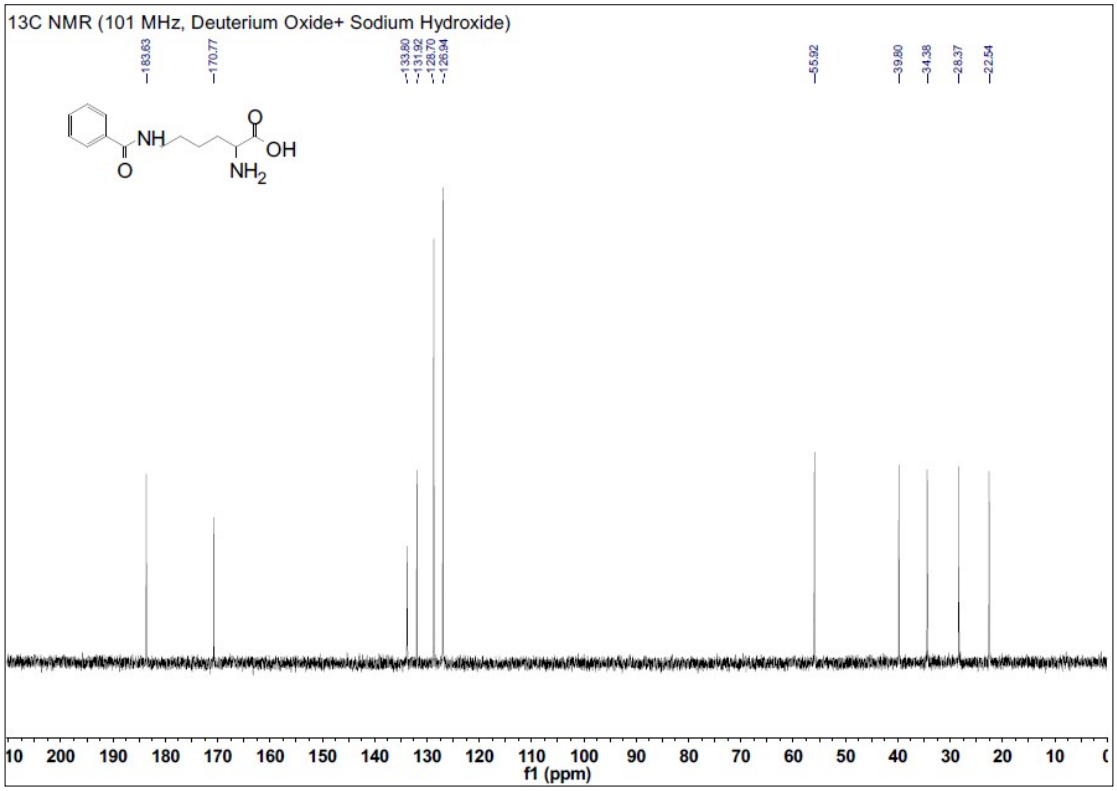


Figure S7. NMR spectra.


1H NMR (400 MHz, Chloroform-d)
f1 (ppm)


13C NMR (101 MHz, Chloroform-d)
f1 (ppm)


1H NMR (400 MHz, Deuterium Oxide+ Sodium Hydroxide)
f1 (ppm)


13C NMR (101 MHz, Deuterium Oxide+ Sodium Hydroxide)
f1 (ppm)