Supplementary Information

A Genetically Encoded ^{19}F NMR Probe for Lysine Acetylation

Feng Zhang,‡[a,b,c] Qing Zhou,‡[b] Guiwen Yang,[a] Liguo An,*[a] Fahui Li*[b] and Jiangyun Wang*[b]

a. *Shandong Provincial Key Laboratory of Animal Resistance Biology, College of Life Science, Shandong Normal University, No. 88 East Wenhua Road, Jinan, 250014, (China).E-mail: an.liguo58@gmail.com*

b. *Laboratory of RNA Biology, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Chaoyang District, Beijing, 100101 (China).E-mail: lifahui@moon.ibp.ac.cn, jwang@ibp.ac.cn*

c. *Key Laboratory of Atherosclerosis in Universities of Shandong and Institute of Atherosclerosis, Taishan Medical University, Taian, Shandong, 271000 (China).*

‡ *These authors contributed equally to the work.*

Materials and General Procedures

N$^{\varepsilon}$-Trifluoroacetyl-L-lysine was purchased from J&K chemical. All other chemicals were purchased from Sigma-Aldrich and used without further purification. Primers were synthesized by Shanghai Sangon. Restriction enzymes and T4 DNA ligase were purchased from TaKaRa. PCR was carried out with the High Fidelity polymerase system provided by Fermentas. Protein purification was performed at AKTA UPC 900 FPLC system (GE healthcare). Cellular fluorescence imaging experiments were performed at an Olympus LSCMFV1000 confocal microscope. Protein mass spectra were run on a thermo LTQ-orbitrap at the IBP, CAS (Beijing, China).

Plasmids and cell lines used

Plasmid pBK-lib-jw1 encodes a library of *Methanosarcina Barkeri* pyrrolysyl tRNA synthetase (PylRS) mutants randomized at residues Leu266, Leu270, Tyr271, Leu274, and Cys313. Plasmid pREP(2)/YC encodes *Mb* $tRNA_{CUA}^{Pyl}$, the chloramphenicol acetyltransferase (CAT) gene with a TAG codon at residue 112, the GFP gene under control of the T7 promoter, and a Tet^{r} marker; plasmid pLWJ17B3 encodes *Mb* $tRNA_{CUA}^{Pyl}$ under the control of the *lpp* promoter and *rrnC* terminator, the barnase gene (with three amber codons at residues 2, 44 and 65) under the control of the *ara* promoter, and an Amp^{r} marker. Plasmid pBAD/JYAMB-4TAG encodes the mutant sperm whale myoglobin gene with an arabinose promoter and *rrnB* terminator, *Mb* $tRNA_{CUA}^{Pyl}$ with an *lpp* promoter and *rrnC* terminator, and a tetracycline resistance marker; *E. coli* strain GeneHog®-Fis, F- *mcr*A (*mrr-hsd*RMS-*mcr*BC) 80*lac*Z M15 *lac*X74 *rec*A1 *end*A1 *ara*D139 (*ara-leu*)7697 *gal*U*gal*K*rps*L(StrR) *nup*G, *fis*::Tn*7* (DHFR).

Genetic selection of the mutant synthetase specific for TFAcK

pBK-lib-jw1 consisting of 2×10^{9} PylRS independent clones was constructed using standard PCR methods. *E. coli* DH10B harboring the pREP(2)/YC plasmid was used as the host strain for the positive selection. Cells were transformed with the pBK-lib-jw1 library, recovered in SOC for 1 h, washed twice with LB media before plating on LB-agar plates supplemented with kanamycin, chloramphenicol, tetracycline and TFAcK at 50 mg/mL, 60 mg/mL, 15 mg/mL and 1 mM, respectively. Plates were incubated at 37 °C for 30 hours, surviving cells were scraped, and plasmid DNA was extracted and purified by gel electrophoresis. The pBK lib-jw1 DNA was then transformed into electro-competent cells harboring the negative selection plasmid pLWJ17B3, recovered for 1 h in SOC and then plated on LB-agar plates containing 0.2% arabinose, 50 mg/mL ampicillin and 50 mg/mL kanamycin. The plates were then incubated at 37 °C for 8-12 h, and pBK-lib-jw1 DNA from the surviving clones was extracted as described

above. The library was then carried through a subsequent round of positive selection, followed by a negative selection and a final round of positive selection (with chloramphenicol at 70 mg/mL). At this stage, 96 individual clones were selected and suspended in 0.2 mL of LB in a 96-well plate, and replica-spotted on two sets of LB plates. One set of LB-agar plates was supplemented with tetracycline (15 mg/mL), kanamycin (50 mg/mL) and chloramphenicol at concentrations of 60, 80, 100 and 120 mg/mL with 1 mM TFAcK. The other set of plates were identical but did not contain TFAcK, and the chloramphenicol concentrations used were 0, 20, 40 and 60 mg/mL. After 30 h incubation at 37 °C, one clone was found to survive at 100 mg/mL chloramphenicol in the presence of 1 mM TFAcK, but only at 20 mg/mL chloramphenicol in the absence TFAcK.

Mutant myoglobin expression

To express mutant myoglobin protein, plasmid pBADJYA MB-Ser4 TAG His was cotransformed with pBK-TFAcKRS into TOP10 *E. coli* competent cells. Cells were amplified in LB media (5 mL) supplemented with kanamycin (50 μg/mL) and tetracycline (15 μg/mL). A starter culture (1 mL) was used to inoculate 100 mL of liquid LB supplemented with appropriate antibiotics and corresponding unnatural amino acid (2mM). Cells were then grown at 37°C to OD600 (optical density at 600 nm) of 0.5, and protein expression was induced by the addition of 0.2% arabinose. After 12 h of growth at 37°C, cells were harvested by centrifugation. Mutant myoglobin was then purified by Ni-NTA affinity chromatography and size exclusion chromatography to achieve homogeneity.

Construction of pEVOL-TFAcKRS, pCMV-TFAcKRS and pET22b-p53 DBD vectors

TFAcKRS gene was amplified by primers P1 and P2. The PCR product was digested with restriction enzymes BglII and SalI and ligated into pEVOL vector, affording pEVOL-1. TFAcKRS gene was again amplified by primers P3 and P4. The PCR product was digested with restriction enzymes PstI and NdeI and ligated into pEVOL-1 vector, affording pEVOL-2. pylT was amplified with primers P5 and P6. The PCR product was digested with restriction enzymes ApaLI and XhoI and ligated into pEVOL-2vector, affording pEVOL-TFAcKRS.

The pCMV-NBK-1 vector was amplified with the primers P7 and P8. The TFAcKRS gene was amplified with the primers P9 and primer P10. Both the PCR products were digested with BglII and Xho1, and ligated to create pCMV-TFAcKRS vector.

p53 DBD (94-312aa) were amplified from full length of p53 gene by primers P11 and P12. The PCR products were digested by Nde1 and Xho1 and ligated into pET22b vector to create pET22b-p53 DBD. And then, p53 DBD were amplified by primers P13 and P14 using the Quikchange mutagenesis kit and generated pET22b-p53 K164TAG

All of the constructs and mutations were verified by DNA sequencing.

Primers:

P1: GGAAGATCTATGGATAAAAAACCGCTGGATG

P2: ACGCGTCGACTTACAGGTTCGTGCTAATG

P3: GTCCAGTCATATGGATAAAAAACCGCTGGATGTG

P4: ATCAGACTCGAGCAGGTTCGTGCTAATGCCGTTATA

P5: GTGCACGGCTAACTAAGCGGCCTGC

P6: CTCGAGCATGCAAAAAAGCCTGCTC

P7: GGAAGATCTGGTGGCAAGCTTCCGTGCAGTTG

P8: GACTCTCGAGACCCGCTGATCAGCCTCGACTGTG

P9: GGAAGATCTATGGATAAAAAACCGCTGGATG

P10: GACTCTCGAGTTACAGGTTCGTGCTAATG

p11: GGGAATTCCATATGTCATCTTCTGTCCCTTCCCAGAAAAC

p12: GACTCTCGAGGGTGTTGTTGGGCAGTGCTCGCTTAG

p13: CATCTACTAGCAGTCACAGCACATGACGGAG

p14: CATGTGCTGTGACTGCTAGTAGATGGCCATG

Circular dichroism measurements

CD measurements were performed with Chirascan Plus in the UV region (200–260 nm), using a protein concentration of 0.2 mg/mL in buffer (20 mM Tris, 150 mM NaCl, pH7.4). Melting temperatures were measured with 222nm UV in the range from 20 to 90℃, at a rate of 1℃/min.

The expression and purification of p53 K164AcK

For the expression of p53 mutants, pEVOL-TFAcKRS was cotransformed with pET22b-p53 K164TAG into BL21 (DE3). A single colony was grown overnight at 37 ℃ in 4 mL of LB medium. The transformed cells were induced with 1 mM IPTG and 0.02 % L-arabinose at OD600nm of 1.2 in the presence of 2 mM TFAcK or AcK. After growing for 13 hours at 25 ℃, the cells were harvested and resuspended in lysis buffer (20 mM Na_2HPO4, 300 mM NaCl, pH 7.5, 100 $ZnCl_2$, 10 mM imidazole) and lysed by sonication. The lysate after sonication was loaded onto a Ni-NTA column (Histrap 5 mL, GE healthcare). The column was washed with 5 ml washing buffer twice (20 mM Na_2HPO4, 300 mM NaCl, pH 7.5, 20 mM imidazole) and then eluted with elution buffer (20 mM Na_2HPO4, 300 mM NaCl, pH 7.5, 500 mM imidazole). The eluted proteins were purified using Superdex G75 column and stored in stock buffer (20 mM Tris-HCl, 100 mM NaCl, pH 7.5).

The expression and purification of SIRT1 and SIRT2

For the expression of SIRT1 and SIRT2 protein, pET28a-SIRT1 (89-747aa) or pET22b-SIRT2 were transformed into BL21 (DE3), respectively. A single colony was grown overnight at 37 ℃ in 4 mL of LB medium. The transformed cells were induced with 0.2 mM IPTG at OD600nm of 0.6. After growing for 16 hours at 18℃, the cells were harvested and resuspended in lysis buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 10 mM imidazole) and lysed by sonication. The lysate after sonication was loaded onto a Ni-NTA column (Histrap 5 mL, GE healthcare). The column was washed with 5 ml washing buffer twice (50 mM Tris, pH 7.5, 150 mM NaCl, 20 mM imidazole) and then eluted with elution buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 500 mM imidazole). The eluted proteins were purified using Superdex G75 column and stored in stock buffer (20 mM Tris-HCl, 100 mM NaCl, pH 7.5).

P53 binding DNA

The single strand DNAs were synthesized by Shanghai Sangon, annealed at 95°C for 5 minutes and gradually cooling to room temperature for one hour. The DNA sequence is as follows: GGAACATGTTCC.

***In vitro* deacetylation assay**

SIRT1 or SIRT2 were incubated with 10 uM p53 mutants, 1 mM NAD in 20 mM Tris-HCl (pH 7.5), 150 mM NaCl buffer containing 1 mM DTT in 60uL reactions for 2 h at 37 ℃. The reaction mixtures were subjected to SDS-PAGE and detected by Western blot.

Western blot procedure

To test the anti-AcK antibody recognition to TFAcK mutants, the samples were boiled in SDS loading buffer and separated by 10% SDS-PAGE, followed by transferring the proteins from the gel onto PVDF membrane under the condition of 200 mA for 40 min. The membranes were placed into blocking buffer (20 mM Tris-HCl, pH7.5, 150 mM NaCl, 0.05％ Tween 20，5% BSA) and incubated for 2 hours at room temperature. The membranes were immunobloted with anti-AcK monoclonal antibody (sc-32268, Santa Cruz Biotechnology, diluted in blocking buffer (dilution rate=1:2000)) or anti His monoclonal antibody (sc-8036, Santa Cruz Biotechnology, diluted in blocking buffer (dilution rate=1:2000)) at room temperature for 1 hour, respectively. After the incubation with primary antibody, the membranes were washed three times by washing buffer (20 mM Tris-HCl, pH7.5, 150 mM NaCl, 0.05％ Tween 20) for 10 min each time. And then the membranes were incubated with secondary antibody (goat anti mouse HRP antibody, Jackson ImmunoResearch Laboratories, diluted in blocking buffer (dilution rate=1:10000)) at room temperature for 1 hour. After washing for three times, the membranes were subjected to enhanced

chemiluminescence (ECL) reaction.

The incorporation of TFAcK in mammalian cells

293T cells were allowed to grow 50-60% confluence on a 20-mm culture dish in DMEM medium supplemented with 10% FBS. 1.5 g pCMV-TFAcKRS and 0.5 g pSWAN-EGFP 37 TAG were cotransfected into 293T cells by using Lipofectamine 2000 Reagent (Invitrogen), according to the manufacturer's instructions. The cells were cultured in the presence or absence of 1 mM TFAcK for 24 hours. Fluorescence images were recorded by confocal microscope in EGFP channel and differential interference contrast (DIC)-channel. All image acquisitions were performed under identical conditions.

^{19}F solid state NMR experiments

The p53 protein, p53-DNA complex (the mole ratio of p53: DNA=1:4) and p53-SIRT2 complex (the mole ratio of p53: SIRT2=1:3) were subjected to ^{19}F solid state NMR. NMR experiments were performed on a Bruker Avance WB 400MHz spectrometer using a 2.5 mm broad-band triple-resonance H-F-X MAS probe. All ^{19}F NMR spectra were acquired with 4 s delays at 300 K, accumulated 30000 scans for 1D ^{19}F spectrum. The sample spinning rate was 30 KHz controlled by a Bruker pneumatic MAS unit. One dimensional ^{19}F spectra was acquired with one pulse program with 90 pulse width of 3.3 us. The spectra width (SW) was 498.03 ppm and offset was -97.3 ppm. ^{19}F chemical shifts were referenced to fluorobenzene (-113.5 ppm). The ^{19}F solid NMR data were processed using an exponential window function with 10 Hz line-broadening. The data were analyzed and plotted using Topspin 3.1.

TFAcKRS:

ATGGATAAAAAACCGCTGGATGTGCTGATTAGCGCGACCGGCCTGTGGATGAGCCGTACCGGCACCCTGCATAAAATCAAACAT
CATGAAGTGAGCCGCAGCAAAATCTATATTGAAATGGCGTGCGGCGATCATCTGGTGGTGAACAACAGCCGTAGCTGCCGTACC
GCGCGTGCGTTTCGTCATCATAAATACCGCAAAACCTGCAAACGTTGCCGTGTGAGCGGTGAAGATATCAACAACTTTCTGACC
CGTAGCACCGAAAGCAAAAACAGCGTGAAAGTGCGTGTGGTGAGCGCGCCGAAAGTGAAAAAAGCGATGCCGAAAAGCGTG
AGCCGTGCGCCGAAACCGCTGGAAAATAGCGTGAGCGCGAAAGCGAGCACCAACACCAGCCGTAGCGTTCCGAGCCCGGCG
AAAAGCACCCCGAACAGCAGCGTTCCGGCGTCTGCGCCGGCACCGAGCCTGACCCGCAGCCAGCTGGATCGTGTGGAAGCG
CTGCTGTCTCCGGAAGATAAAATTAGCCTGAACATGGCGAAACCGTTTCGTGAACTGGAACCGGAACTGGTGACCCGTCGTAA
AAACGATTTTCAGCGCCTGTATACCAACGATCGTGAAGATTATCTGGGCAAACTGGAACGTGATATCACCAAATTTTTTGTGGAT
CGCGGCTTTCTGGAAATTAAAAGCCCGATTCTGATTCCGGCGGAATATGTGGAACGTATGGGCATTAACAACGACACCGAACTG
AGCAAACAAATTTTCCGCGTGGATAAAAACCTGTGCCTGCGTCCGATGATGGCCCCGACCATTTATAACTATGCGCGTAAACTGG
ATCGTATTCTGCCGGGTCCGATCAAAATTTTTGAAGTGGGCCCGTGCTATCGCAAAGAAAGCGATGGCAAAGAACACCTGGAA
GAATTCACCATGGTTAACTTTTTTCAAATGGGCAGCGGCTGCACCCGTGAAAACCTGGAAGCGCTGATCAAAGAATTCCTGGAT
TATCTGGAAATCGACTTCGAAATTGTGGGCGATAGCTGCATGGTGTATGGCGATACCCTGGATATTATGCATGGCGATCTGGAAC
TGAGCAGCGCGGTGGTGGGTCCGGTTAGCCTGGATCGTGAATGGGGCATTGATAAACCGTGGATTGGCGCGGGTTTTGGCCT
GGAACGTCTGCTGAAAGTGATGCATGGCTTCAAAAACATTAAACGTGCGAGCCGTAGCGAAAGCTACTATAACGGCATTAGCA
CGAACCTGTAA

P53 K164TAG:

ATGTCATCTTCTGTCCCTTCCCAGAAAACCTACCAGGGCAGCTACGGTTTCCGTCTGGGCTTCTTGCATTCTGGGACAGCCAAGT
CTGTGACTTGCACGTACTCCCCTGCCCTCAACAAGATGTTTTGCCAACTGGCCAAGACCTGCCCTGTGCAGCTGTGGGTTGATT
CCACACCCCCGCCCGGCACCCGCGTCCGCGCCATGGCCATCTAC**TAG**CAGTCACAGCACATGACGGAGGTTGTGAGGCGCTGC
CCCCACCATGAGCGCTGCTCAGATAGCGATGGTCTGGCCCCTCCTCAGCATCTTATCCGAGTGGAAGGAAATTTGCGTGTGGAG
TATTTGGATGACAGAAACACTTTTCGACATAGTGTGGTGGTGCCCTATGAGCCGCCTGAGGTTGGCTCTGACTGTACCACCATCC
ACTACAACTACATGTGTAACAGTTCCTGCATGGGCGGCATGAACCGGAGGCCCATCCTCACCATCATCACACTGGAAGACTCCA
GTGGTAATCTACTGGGACGGAACAGCTTTGAGGTGCGTGTTTGTGCCTGTCCTGGGAGAGACCGGCGCACAGAGGAAGAGA

ATCTCCGCAAGAAAGGGGAGCCTCACCACGAGCTGCCCCCAGGGAGCACTAAGCGAGCACTGCCCAACAACACCCTCGAGCA
CCACCACCACCACCACTGA

Red: amber stop codon for TFAcK incorporation

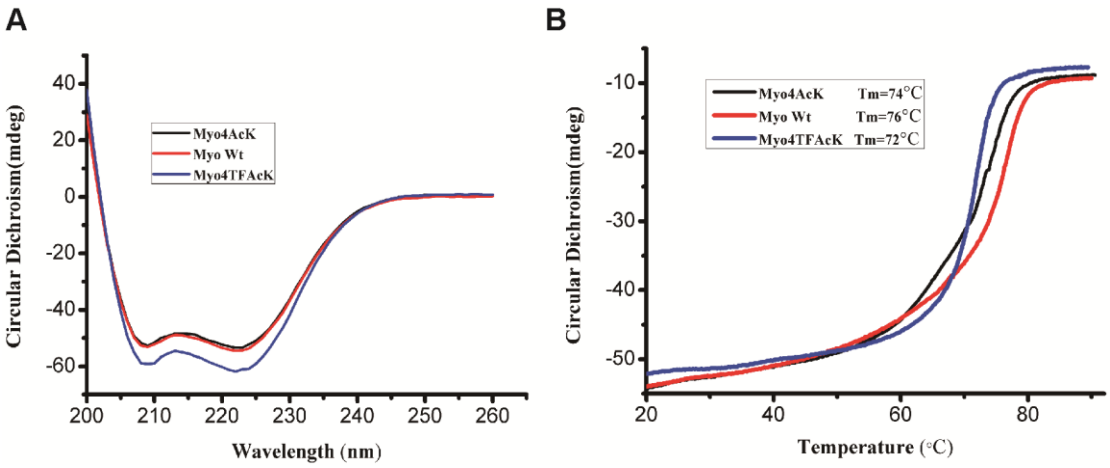

Fig. S1 Circular dichroism spectroscopy of myoglobin wt and mutants. A) Circular dichroism spectroscopy of myoglobin wt and mutants was performed in the UV region (200–260 nm), B) Melting temperatures of myoglobin wt and mutants were measured with 222nm UV in the range from 20 to 90℃.

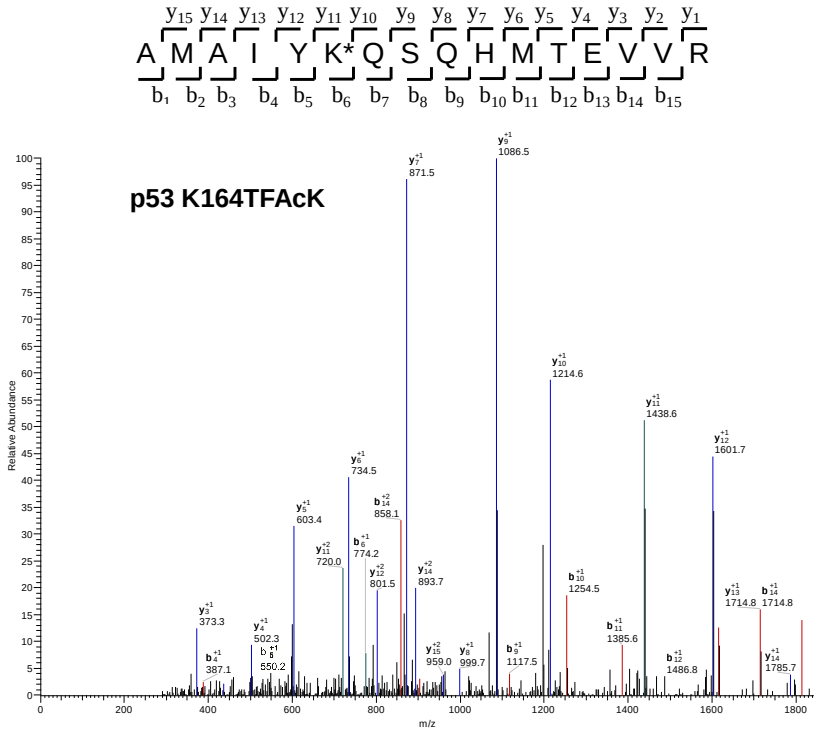

Fig. S2 MS-MS spectrum of the p53 K164TFAcK peptide fragment containing TFAcK (denoted K*) in the 164th position. The TFAcK mass of 242.2 Da was determined from the y11 and y10 ions (mass = y11-y10), which closely matches the calculated mass of 242.2 Da.

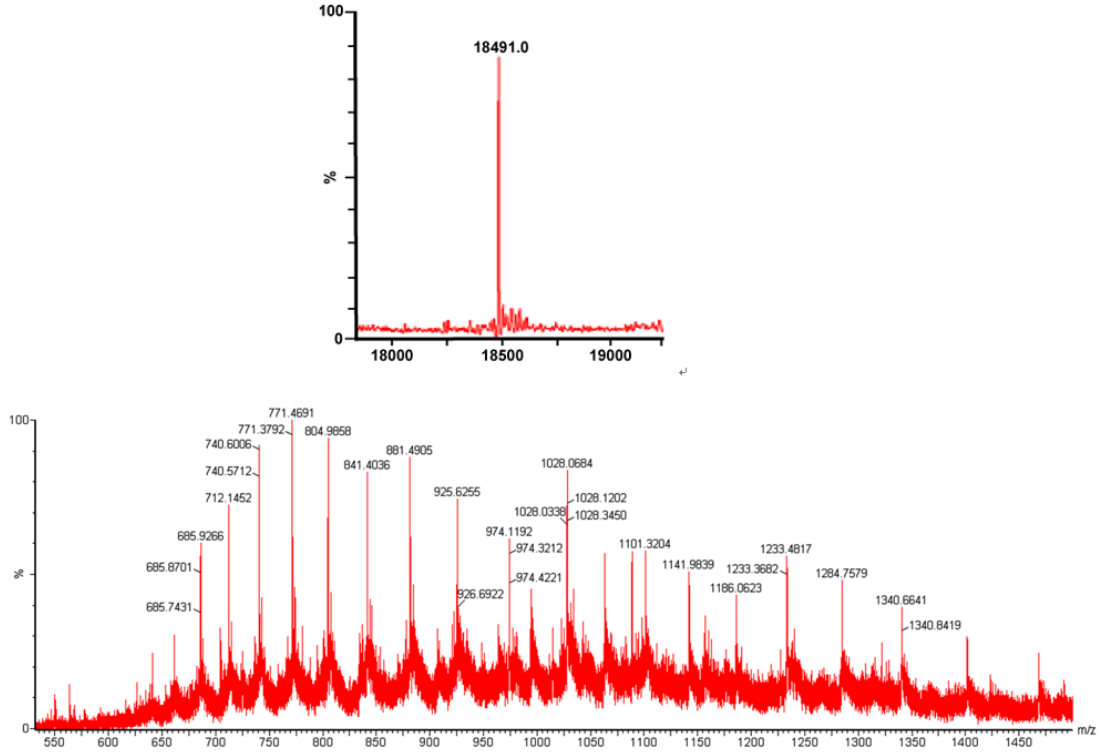

Fig. S3 ESI-MS analysis of the Ser4 TFAcK mutant myoglobin in the absence of nicotinamide.

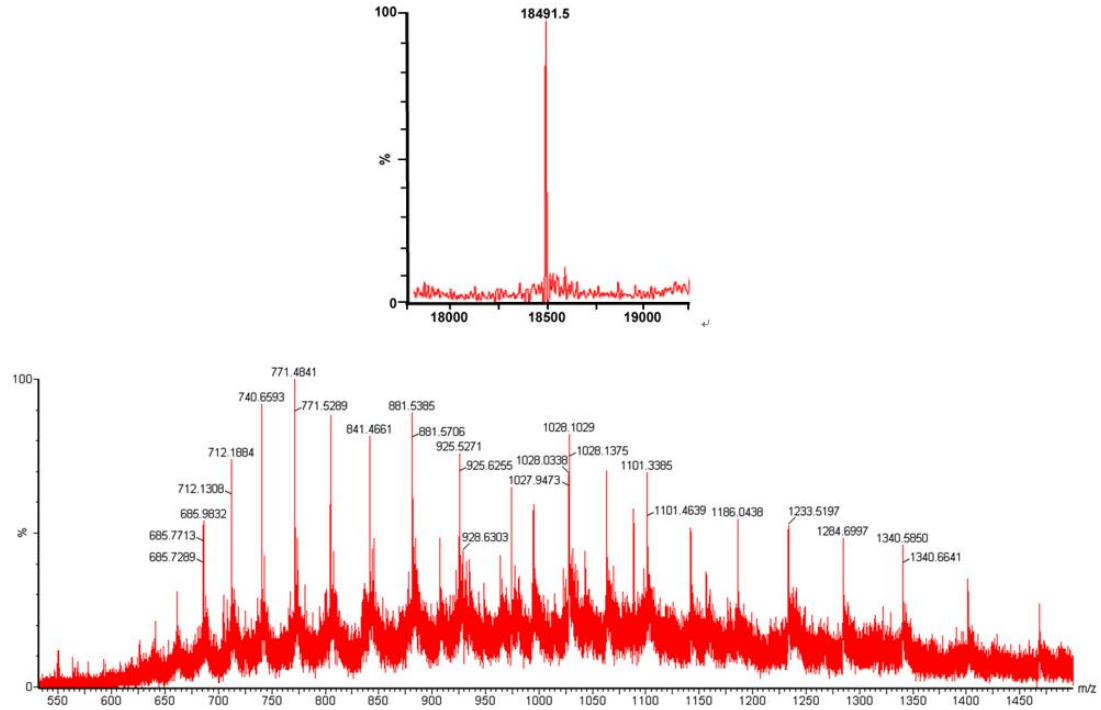

Fig. S4 ESI-MS analysis of the Ser4 TFAcK mutant myoglobin in the presence of nicotinamide.

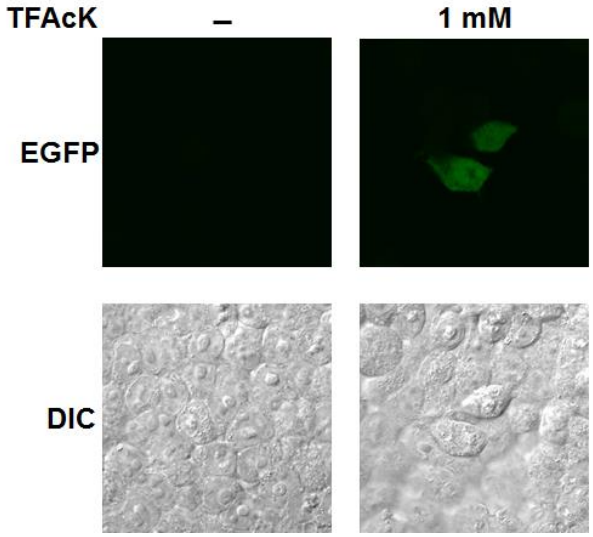

Fig. S5 Genetic incorporation of TFAcK into the 37th position of EGFP in 293T cells by using the TFAcKRS/*Mb*tRNA$_{Pyl}^{CUA}$ pair. Full length EGFP is produced only in the presence of 1 mM TFAcK, but not in its absence.

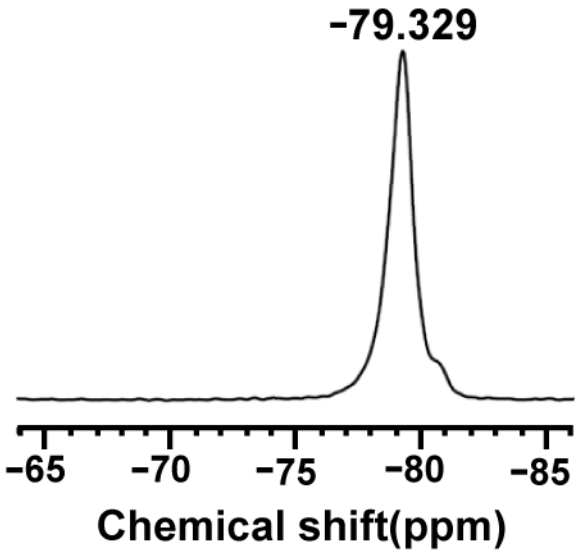

Fig. S6 ^{19}F solid-state NMR spectra of N$^{\varepsilon}$-Trifluoroacetyl-L-lysine. Chemical shift is referenced to fluorobenzene as an internal standard.