Electronic Supplementary Information (ESI)

A Bioorthogonal Chemical Reporter for the Detection and Identification of Protein Lactylation

Yanan Sun,[†,#] Yanchi Chen,[†,#] and Tao Peng*[,†,‡]

[†] State Key Laboratory of Chemical Oncogenomics, Guangdong Provincial Key Laboratory of Chemical Genomics, School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, Shenzhen 518055, China
[‡] Shenzhen Bay Laboratory, Shenzhen 518132, China

[#] Y. S. and Y. C. contributed to this work equally.

*Correspondence: tpeng@pku.edu.cn

Table of Contents

Experimental procedures

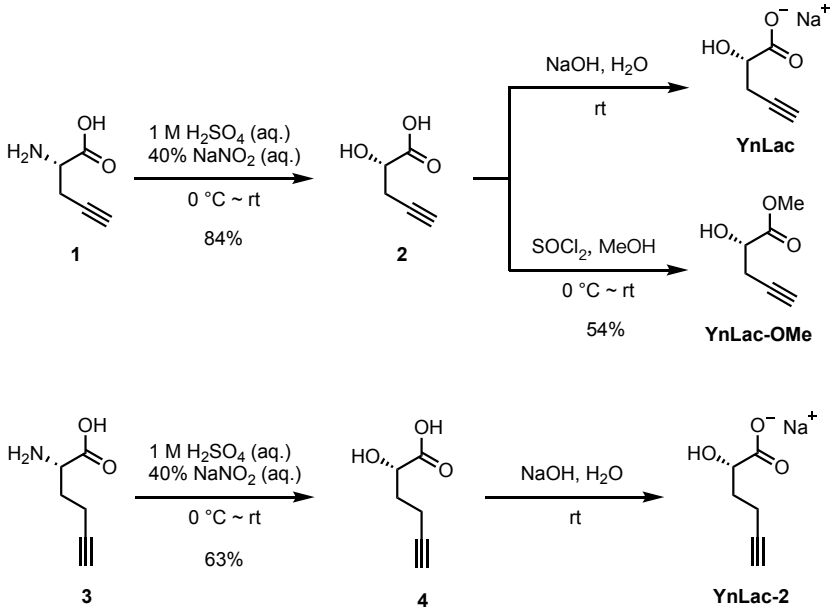
General methods and materials

Unless otherwise noted, chemicals and solvents were obtained from Sigma-Aldrich or Bidepharm and were used directly as received without further purification. Chemical reactions were performed in oven-dried flasks under a N_2 atmosphere when necessary. TLC was conducted on silica gel 60 GF254 glass plates (Qingdao Haiyang Chemical Co., Ltd), and spots were visualized by staining with iodine (I_2), potassium permanganate ($KMnO_4$), or phosphomolybdic acid (PMA). Flash column chromatography was performed with silica gel (230-400 mesh, reagent grade) from Qingdao Haiyang Chemical Co., Ltd. ^1H and ^{13}C spectra were recorded in $CDCl_3$ or CD_3OD at room temperature on Bruker Avance NMR Spectrometers operating at 300 or 400 MHz for ^1H. Chemical shifts are reported in δ ppm, and coupling constants (*J* values) are reported in Hz. ^1H NMR chemical shifts are calibrated using tetramethylsilane (TMS, δ = 0.00 ppm) in $CDCl_3$ as the internal standard or with the residual solvent peaks of CD_3OD (δ = 3.31 ppm). ^{13}C NMR chemical shifts are calibrated with the residual solvent peaks of $CDCl_3$ (δ = 77.16 ppm) or CD_3OD (δ = 49.00 ppm). High resolution ESI mass spectra were recorded with a Q Exactive Focus (ThermoFisher) mass spectrometer.

Azido-rhodamine and azido-biotin were synthesized in the lab according to literature.[1] Azido-DADPS-biotin was purchased from Click Chemistry Tools. PfuUltra High-Fidelity DNA polymerase was obtained from Agilent Technologies; restriction enzymes and dNTPs were obtained from New England Biolabs; oligonucleotide primers and gene fragments were synthesized by Tsingke. Plasmid DNA isolation was carried out with the Plasmid Mini Kit (Omega). Polyethylenimine (PEI) was purchased from Polysciences. Protease inhibitor cocktail was purchased from Roche (cOmplete ULTRA mini Tablets, EDTA-free). In-gel fluorescence and western blotting analyses were recorded on a Chemidoc MP imaging system (Biorad). Confocal fluorescence imaging was performed with a Nikon A1R confocal fluorescence microscope.

Synthesis of L-lactate chemical reporters

Scheme S1. Synthesis of YnLac, YnLac-OMe, and YnLac-2



Synthesis of (S)-2-hydroxypent-4-ynoic acid (compound 2)

To a solution of *L*-propargylglycine hydrochloride (**1**) (0.50 g, 3.35 mmol) in 1 M H_2SO_4 (25 mL) was added 40% aqueous sodium nitrite (2.9 mL, 16.75 mmol) at −5 °C. The resulting reaction mixture was stirred at 0 °C for 2 h and then at room temperature for another 3 h. The reaction mixture was extracted with EtOAc (3 x 25 mL), and the organic phase was dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The residue was purified by flash chromatography (eluent: EtOAc/MeOH = 4/1) to give compound **2** as a yellow oil (320 mg, 84% yield). ^{1}H NMR (300 MHz, CD_3OD) δ 4.25 (t, *J* = 5.5 Hz, 1H), 2.70 – 2.53 (m, 2H), 2.31 (t, *J* = 2.6 Hz, 1H). ^{13}C NMR (75 MHz, CD_3OD) δ 175.98, 80.43, 71.73, 70.28, 25.36. HRMS (ESI) calcd for $C_5H_6O_3$ $[M-H]^-$ 113.02442, found 113.02319.

Synthesis of methyl (S)-2-hydroxypent-4-ynoate (YnLac-OMe)

To a solution of compound **2** (200 mg, 1.75 mmol) in MeOH (5 mL) was added $SOCl_2$ (0.5 mL) at 0 °C. The reaction mixture was stirred at room temperature for another 8 h and then concentrated under reduced pressure. The residue was purified by flash chromatography (eluent: petroleum ether/EtOAc = 2/1) to afford YnLac-OMe as a yellow oil (120 mg, 54% yield). ^{1}H NMR (300 MHz, $CDCl_3$) δ 4.32 (d, *J* = 4.6 Hz, 1H), 3.79 (s, 3H), 3.23 (br, 1H), 2.76 – 2.58 (m, 2H), 2.05 (t, *J* = 2.6 Hz, 1H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 173.53, 78.64, 71.45, 68.85, 52.93, 24.87. HRMS (ESI) calcd for $C_6H_8O_3Na$ $[M+Na]^+$ 151.03657, found 151.03665.

Synthesis of (S)-2-hydroxyhex-5-ynoic acid (compound 4)

To a solution of (*S*)-2-aminohex-5-ynoic acid hydrochloride (**3**) (163 mg, 1.0 mmol) in 1 M H_2SO_4 (8 mL) was added 40% aqueous sodium nitrite (0.87 mL, 5 mmol) at −5 °C. The resulting reaction mixture was stirred at 0 °C for 2 h and then at room temperature for another 3 h. The reaction mixture was extracted with EtOAc (3 x 25 mL), and the organic phase was dried over Na_2SO_4 and evaporated under reduced pressure. The residue was purified by flash chromatography (eluent: EtOAc/MeOH = 4/1) to give compound **4** as a yellow oil (80 mg, 63% yield). ^{1}H NMR (300 MHz, CD_3OD) δ 4.23 (dd, *J* = 8.8, 3.9 Hz, 1H), 2.41 – 2.28 (m, 2H), 2.25 (t, *J* = 2.6 Hz, 1H), 2.04 – 1.92 (m, 1H), 1.85 – 1.73 (m, 1H). ^{13}C NMR (75 MHz, CD_3OD) δ 177.53, 83.88, 70.09, 69.96, 34.46, 15.21. HRMS (ESI) calcd for $C_6H_7O_3$ $[M-H]^-$ 127.03897, found 127.03899.

Preparation of sodium (S)-2-hydroxypent-4-ynoate (YnLac) and sodium (S)-2-hydroxyhex-5-ynoate (YnLac-2)

(*S*)-2-Hydroxypent-4-ynoic acid (compound **2**) or (*S*)-2-hydroxyhex-5-ynoic acid (compound **4**) was dissolved in ddH_2O. Aqueous NaOH solution (1 equiv., ~0.1 M in ddH_2O) was then added dropwise. The mixture was filtered through 0.45 μm membrane, frozen in liquid nitrogen, and lyophilized to provide YnLac or YnLac-2 as white powders.

Plasmids and cloning

Full-length cDNAs encoding proteins of interest (PARP1, HMGB1, NOLC1, NCL, NUCKS1, PCNP, and DNAJC8) were purchased from WZ Biosciences and cloned into the pCMV-HA vector (Clontech). Site-directed mutagenesis was performed with QuikChange II Site-Directed Mutagenesis Kit (Agilent) with primers designed by Agilent Primer Design Program.

Cell culture and transfection

HEK293T, HeLa, MCF-7, RAW264.7, MDA-MB-231, HepG2, and A549 cells were obtained from ATCC. PARP1 knockout HEK293T cell line was generated by Abcam (cat#ab266598) and verified in the lab. Cells were grown in DMEM (Dulbecco's modified Eagle's medium; Corning, cat#10-013-CVR) supplemented with 10% FBS (fetal bovine serum; Corning, cat#35-076-CV) at 37 °C in a humidified incubator with an atmosphere of 5% CO_2. For transfection, cells were grown on cell culture dishes or plates to 70% confluence and transfected with indicated plasmids using PEI (Polysciences) at a ~2.5:1 ratio of transfection reagent/DNA in Opti-MEM media (ThermoFisher) for about 18-24 h.

Metabolic labeling in mammalian cells with L-lactate chemical reporters

L-lactate chemical reporters (YnLac and YnLac-2) were dissolved in PBS to make the 2 M stock solutions. YnLac-OMe was dissolved in PBS to make the 0.2 M stock solution. For metabolic labeling of cellular proteins, cells were incubated with the chemical reporters at desired concentrations in DMEM supplemented with 10% FBS for indicated time periods at 37 °C. Generally, YnLac was incubated with cells at 20 mM concentration for 8 h, unless otherwise indicated. For cellular competition experiments, sodium L-lactate (Sigma, cat#71718) was dissolved in PBS to make the 20 M stock solution. Cells were pre-treated with sodium L-lactate at indicated concentrations for 4 h, and then co-incubated with the chemical reporters and sodium L-lactate at indicated concentrations in DMEM supplemented with 10% FBS for 8 h. For inhibition of p300-dependent protein labeling, HEK293T cells were either metabolically labeled with 10 mM YnLac for 6 h or pre-treated with 60 μM curcumin (MedChemExpress, cat#HY-N0005; 10 mM stock solution in DMSO) for 6 h then labeled with 10 mM YnLac for additional 6 h. For inhibition of histone deacetylases (HDACs) or sirtuins, HEK293T cells were labeled with 20 mM YnLac for 8 h and then treated with 0.4 μM trichostatin A (TSA; Cell Signaling Technology, cat#9950S; 50 mM stock solution in DMSO) or 10 mM nicotinamide (NAM; Sigma, cat#72340; 50 mM stock solution in ddH_2O) for different time periods. Probe-labeled cells were harvested, washed with cold PBS, and flash-frozen in liquid nitrogen before stored at −80 °C.

Preparation of cell lysates and core histones

For click reactions with azido-rhodamine in whole cell lysates, cell pellets were lysed in SDS lysis buffer (1% SDS, 150 mM NaCl, 50 mM HEPES, pH 7.4, supplemented with benzonase) with vigorous vortexing. The resulting lysate was centrifuged at 12,000g for 20 min to remove cellular debris.

For separation of the nuclear and cytoplasmic fractions,[2] cell pellets were lysed in ice-chilled Brij lysis buffer (1% Brij 97, 150 mM NaCl, 50 mM triethanolamine, pH 7.4, supplemented with EDTA-free protease inhibitor cocktail) with brief vortexing. The resulting cell lysates were centrifuged at 5,000g for 5 min at 4 °C and the supernatant was collected as the cytoplasmic fractions containing membrane and cytosolic proteins. The nuclear pellets were gently washed with the above lysis buffer without Brij 97 and further lysed with SDS lysis buffer (4% SDS, 150 mM NaCl, 50 mM triethanolamine, pH 7.4, supplemented with benzonase) as the nuclear fractions containing nuclear proteins. Both fractions were centrifuged at 12,000g for 20 min at room temperature to remove debris.

For extraction of core histones, a standard acid-extraction protocol was adapted.[3] The nuclear pellets were resuspended in 0.4 N H_2SO_4 and shaked overnight on a rotator at 4 °C.

The nuclear debris was pelleted by centrifugation at 16000g for 10 min at 4 °C. The supernatant containing core histones was collected and then precipitated with MeOH at −80 °C overnight. Precipitated histone proteins were centrifuged at 16000g for 10 min at 4 °C and washed twice with ice-cold MeOH. Protein pellets were air-dried at room temperature and resuspended in ddH_2O.

Protein concentrations were generally determined by the BCA assay (Pierce). Cell lysates were normalized with lysis buffer to equal protein concentrations.

CuAAC click reaction and in-gel fluorescence analysis

Cell lysates (100 μg) were diluted with the corresponding lysis buffers to 89 μL, and then reacted with 11 μL freshly prepared click reaction cocktail containing azido-rhodamine (2 μL, 10 mM stock solution in DMSO), tris-(2-carboxyethyl)phosphine hydrochloride (TCEP, 2 μL, 50 mM freshly prepared stock solution in ddH_2O), tris-[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA, 5 μL, 10 mM stock solution in DMSO/t-butanol), and $CuSO_4{\cdot}5H_2O$ (2 μL, 50 mM freshly prepared stock solution in ddH_2O) for 2 h at room temperature in the dark. The click reactions were terminated by addition of ice-cold methanol (500 μL), placed at −20 °C overnight, and then centrifuged at 20,000g for 15 min at 4 °C to precipitate the proteins. The supernatants were discarded and protein pellets were washed with ice-cold methanol twice and air-dried. The resulting protein pellets were resuspended with 35 μL of SDS lysis buffer (4% SDS, 150 mM NaCl, 50 mM triethanolamine, pH 7.4), and diluted with 12.5 μL 4X SDS-loading buffer (40% glycerol, 200 mM Tris-HCl pH 6.8, 8% SDS, 0.4% bromophenol blue) and 2.5 μL 2-mercaptoethanol. The resulting samples were heated for 5 min at 95 °C before loaded onto 4−20%, unless otherwise noted, ExpressPlus™ PAGE gels (Genscript) for SDS-PAGE separation. Generally, 20 μg of protein per gel lane is loaded for in-gel fluorescence visualization. For in-gel fluorescence, gels were scanned on a ChemiDoc MP Imager (Bio-Rad) with the rhodamine filter. After in-gel fluorescence scanning, gels were stained with Coomassie Brilliant Blue staining reagent.

Fluorescence imaging

HeLa cells were cultured on sterilized coverslips and incubated with YnLac (20 mM, 2 M stock solution in PBS) in culture medium. The same volume of PBS was used as the vehicle control. After 12 h of labeling at 37°C, the cells were washed once with warmed PBS, fixed with 4% formaldehyde in PBS for 10 min at room temperature, and then washed twice with ice-cold PBS. Cells were permeabilized with 0.3% Triton X-100 in PBS for 10 min at room temperature, blocked with 3% BSA in PBS for 30min at room temperature, and washed with PBS (3 x 5 min with gentle agitation). The cells were then treated with a freshly premixed click reaction cocktail (50 μM azido-rhodamine, 1 mM TCEP, 100 μM TBTA, and 1 mM $CuSO_4{\cdot}5H_2O$) in PBST (0.1% Tween-20 in PBS) for 1 h at room temperature. After gentle washes three times with 1% Tween-20 in PBS, cells were then stained with Hoechst 33342 (Beyotime, cat#C1029) and imaged on a Nikon A1R confocal fluorescence microscope. For Hoechst channel, the 405 nm laser was used as the excitation, and emission was collected between 425 nm to 475 nm. For rhodamine channel, the 561 nm laser was used as the excitation, and emission was collected between 570 nm to 620 nm.

Pull-down and immunoprecipitation

For pull-down experiments, HEK293T cells were transfected with plasmids expressing HA-tagged proteins of interest for 18-24 h, metabolically labeled with YnLac (20 mM) for another 8 h, and harvested. Cell lysates were prepared in SDS lysis buffer (1% SDS, 150 mM NaCl, 50 mM HEPES, pH 7.4, supplemented with benzonase) with brief vortexing. Cell lysates (2~3 mg protein) were incubated with freshly prepared click reaction cocktail containing 100 μM azido-biotin,[1] 1 mM TCEP, 100 μM TBTA, and 1 mM $CuSO_4 \cdot 5H_2O$ at room temperature for 2 h in the dark. After protein precipitation and resuspension as described above, the biotinylated proteins were incubated with streptavidin agarose beads (ThermoFisher) at room temperature on a rotator for 2 h. The beads were washed six times with 1 mL of 1% SDS lysis buffer and the proteins were eluted with SDS-PAGE sample loading buffer (~30 μL) containing 70% SDS buffer (4% SDS, 150 mM NaCl, 50 mM triethanolamine, pH 7.4), 25% 4X SDS-loading buffer, and 5% 2-mercaptoethanol after heating at 95 °C for 5 min. The supernatant (20 μL per gel lane) was separated by SDS-PAGE for Western blotting analysis.

For immunoprecipitation experiments, HEK293T cells were transfected with plasmids expressing HA-tagged proteins of interest for 18-24 h and harvested. The cells were lysed with RIPA buffer (1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 150 mM NaCl, 50 mM Tris, pH 7.4, supplemented with EDTA-free protease inhibitor cocktail) with vigorous vortexing. The resulting cell lysates were centrifuged at 12,000g for 15 min at 4 °C to remove cellular debris. Protein concentrations were determined by the BCA assay (Pierce). Equal amounts of cell lysates (~1 mg) were incubated with anti-HA agarose beads (ThermoFisher) on a rotator at 4 °C overnight. The beads were washed six times with 1 mL of chilled 1% Triton X-100 buffer and then resuspended in SDS-PAGE sample loading buffer (~30 μL) with heating at 95 °C for 5 min. The supernatant (20 μL per gel lane) was separated by SDS-PAGE for Western blotting analysis.

Detection of PARP1 auto-ADP-ribosylation

PARP1 knockout HEK293T cells were transfected with plasmids expressing PARP1 or mutants for 18-24 h and incubated with sodium L-lactate (10 mM) for another 8 h. The cells were stimulated with 2 mM H_2O_2 for 10 min and harvested. Cell lysates were prepared in SDS lysis buffer (4% SDS, 150 mM NaCl, 50 mM HEPES, pH 7.4) supplemented with 5 μM olaparib (MedChemExpress, 10 mM stock solution in DMSO), 5 μM PDD00017273 (MedChemExpress, 10 mM stock solution in DMSO), Roche protease inhibitor cocktail, and benzonase with brief vortexing and heating at 95 °C for 5 min. For Western blotting analysis on whole cell lysates, the protein concentrations were determined by the BCA assay (Pierce). The lysates (20 μL per gel lane) was separated by SDS-PAGE. For immunoprecipitation, the whole cell lysates were diluted with 40 volumes of 1% Triton X-100 buffer (SDS level lower than 0.1%). The resulting cell lysates were centrifuged at 12,000g for 15 min at 4 °C to remove cellular debris. Protein concentrations were determined by the BCA assay (Pierce). HA-tagged PARP1 proteins were immunoprecipitated from ~1 mg of cell lysates using anti-HA agarose beads (ThermoFisher). After overnight incubation on a rotator at 4 °C, the beads were washed six times with 1 mL of chilled 1% Triton X-100 buffer. The proteins were eluted with SDS-PAGE sample loading buffer (~30 μL) after heating at 95 °C for 5 min. The supernatant (20 μL per gel lane) was separated by SDS-PAGE for Western blotting analysis using the antibody against ADP-ribosylation.

Western blotting

Gels were transferred to nitrocellulose membranes using Bio-Rad Trans-Blot Turbo Transfer System (25 V, 30 min). The membranes were blocked with PBST (0.05% Tween-20 in PBS) containing 5% nonfat milk for 30 min at room temperature and then incubated with primary antibodies at 4 °C overnight. Membranes were washed with PBST three times, incubated with appropriate secondary antibodies, and developed using Bio-Rad Clarity Western ECL substrate. Membranes were imaged with a ChemiDoc MP Imager (Bio-Rad). Anti-HA-HRP conjugate (3F10, 1:1000 dilution) was purchased from Roche for anti-HA blots. Pan anti-lactyllysine (PTM-1401, 1:1000 dilution) was purchased from PTM Biolabs for anti-Kla blots. Anti-PARP1 (9532S, 1:1000 dilution) was purchased from Cell Signaling Technology. Anti-ADP-ribosylation (MABE1016, 1:1000 dilution) was purchased from Sigma. Goat anti-rabbit- and anti-mouse HRP secondary antibodies (1:10000 dilution) were purchased from Jackson ImmunoResearch Laboratories.

Chemoproteomic profiling of lactylated proteins

HEK293T cells were treated with 20 mM YnLac for 8 h. The same concentration of L-lactate was used as the vehicle control. For both YnLac labeling and L-lactate control treatment, four biological replicates were performed in the proteomics analysis. Cells were harvested and lysed to separate the nuclear and cytoplasmic fractions as described above. Cell lysates from both fractions were centrifuged at 12,000g for 20 min to remove cellular debris. Protein concentrations were determined by the BCA assay (Pierce). The cell lysates (5 mg protein) were then clicked with acid cleavable azido-DADPS-biotin (Click Chemistry Tools, cat#1330) in the presence of TCEP, TBTA, and $CuSO_4$ as described above. Methanol-precipitated and washed protein pellets were again resuspended in 4% SDS buffer. Protein concentrations were determined and equal amounts of each protein sample were diluted 1/4 by volume with 50 mM triethanolamine buffer to 1% SDS (~1.6 mg/mL protein). Then prewashed streptavidin agarose beads (~50 µL slurry; ThermoFisher) were added to each sample. The protein and beads mixtures were incubated at room temperature on a nutating mixer for 1.5 h. The beads were then washed 4~6 times with 1% SDS (in PBS, pH 8) with rotation for 5 min and transferred into spin-columns (ThermoFisher). The beads were then washed with 5 M urea (50 mM Tris-HCl, pH 8), followed by washes with PBS (pH 7.4) and 100 mM ammonium bicarbonate (ABC) buffer. The beads were then transferred into new 1.5 mL centrifuge tubes and incubated with 10 mM DTT for 1 h, followed by treatment with 20 mM iodoacetamide for another 1 h in the dark. After that, the beads were washed with 100 mM ABC buffer and digested with 0.5 µg of trypsin (Promega) in ABC buffer (100 µL) at 37 °C overnight. The beads were again washed with ABC buffer and ddH_2O. Finally, the beads were resuspended with 5% formic acid in ddH_2O (100 µL) to cleave the DADPS linker. The elution was repeated twice and the supernatants were pooled and dried by SpeedVac for LC-MS/MS analysis.

LC-MS/MS analysis was performed with a Q Exactive HF-X Hybrid Quadrupole-Orbitrap mass spectrometer (ThermoFisher) coupled to an EASY-nLC 1200 system (ThermoFisher). The peptide samples were loaded onto an Acclaim PepMap RSLC C18 reverse-phase column (75 µm x 25 cm, nanoViper, C18, 2 µm, 100 Å; ThermoFisher). A 120-minute gradient increasing from 90% buffer A (water with 0.1% formic acid) and 10% buffer B (80% acetonitrile in water with 0.1% formic acid) to 45% buffer B in 30 minutes and to 95% buffer B in 70 minutes, and then keeping at 95% buffer B for 20 minutes was used at 0.3 µL/min for HPLC separation. The mass spectrometer was calibrated using Tune instrument

control software. Spray voltage was set to 2.1 kV, heated capillary at 320 °C, and funnel RF level at 40. The mass spectrometer was configured for data-dependent acquisition mode using the full MS/DD–MS/MS setup. All data were acquired in profile mode with positive polarity. Full MS resolution was set to 60,000 at m/z 200 and full MS automatic gain control (AGC) target was 3 x 10^6 with an maximum IT of 50 ms. Mass range was set to 350–1800 m/z. MS2 resolution was set to 15,000 at m/z 200, AGC target value for MS2 at 1 x 10^5, and intensity threshold at 2.2 x 10^4. Isolation width was set at 1.6 m/z. A fixed first mass of 100 m/z was used. Normalized collision energy was set at 28%. Peptide match was set to preferred, and isotope exclusion was on. Precursor ions with unassigned, single, or six and higher charge states were excluded from fragmentation selection. The dynamic exclusion was set to 30 s.

Acquired MS raw files were analyzed by MaxQuant software v1.5.3.8[4] using the Andromeda search engine and searched against the human UniProt Reference Proteome without isoforms concatenated with common known contaminants. Enzyme specificity was set to trypsin, allowing maximum four missed cleavages. Carbamidomethylation of cysteine was set as a fixed modification, while methionine oxidation and *N*-terminal acetylation were set as variable modifications. For identification of lactylation sites, a variable modification on lysine residues ($C_{11}H_{17}O_3N_3$, + 239.12699 Da) was included. The first search peptide tolerance was set to 20 ppm, main search peptide tolerance to 4.5 ppm. The allowed fragment mass deviation was 20 ppm. Minimum length of peptides was set to seven amino acids, and maximum mass was set to 4,600 Da. The false discovery rates were set to 1% for peptide spectrum match, protein, and site decoy fraction levels. Minimum score and delta score for modified peptides were set to 40 and 6, respectively. Other parameters in MaxQuant were kept at default values.

The search results from MaxQuant on lactylation sites were analyzed with Perseus v1.6.10.0[5] and Microsoft Excel. Known contaminants, reverse hits, and hits only identified by site were removed. Lactylation sites that were identified in L-lactate control samples were assigned as false positives and removed. The MS/MS spectra of YnLac-modified peptides were generated by pFind[6].

Bioinformatic analysis

Post-translational modification (PTM), Gene ontology (GO), KEGG pathway, and REACTOME pathway enrichment analyses were performed using DAVID bioinformatics resources (https://david-d.ncifcrf.gov/, 2021 update).[7] PANTHER protein class analysis was performed in PANTHER (http://www.pantherdb.org/, version 16.0).[8] Sequence preference motif was generated by pLogo (https://plogo.uconn.edu/) using the human proteome as the background.[9] Analysis of other PTMs (acetylation, ubiquitination, and methylation) on lysine residues was performed using iPTMnet[10] (https://research.bioinformatics.udel.edu/iptmnet/) and PhosphositePlus[11] (https://www.phosphosite.org/homeAction.action).

Quantification and statistical analysis

Data were presented as mean ± standard deviation determined from biological replicates. The method for determining error bars and significance is indicated in the corresponding figure legends. Statistical analysis was performed with GraphPad Prism 9.

Supplementary Figures

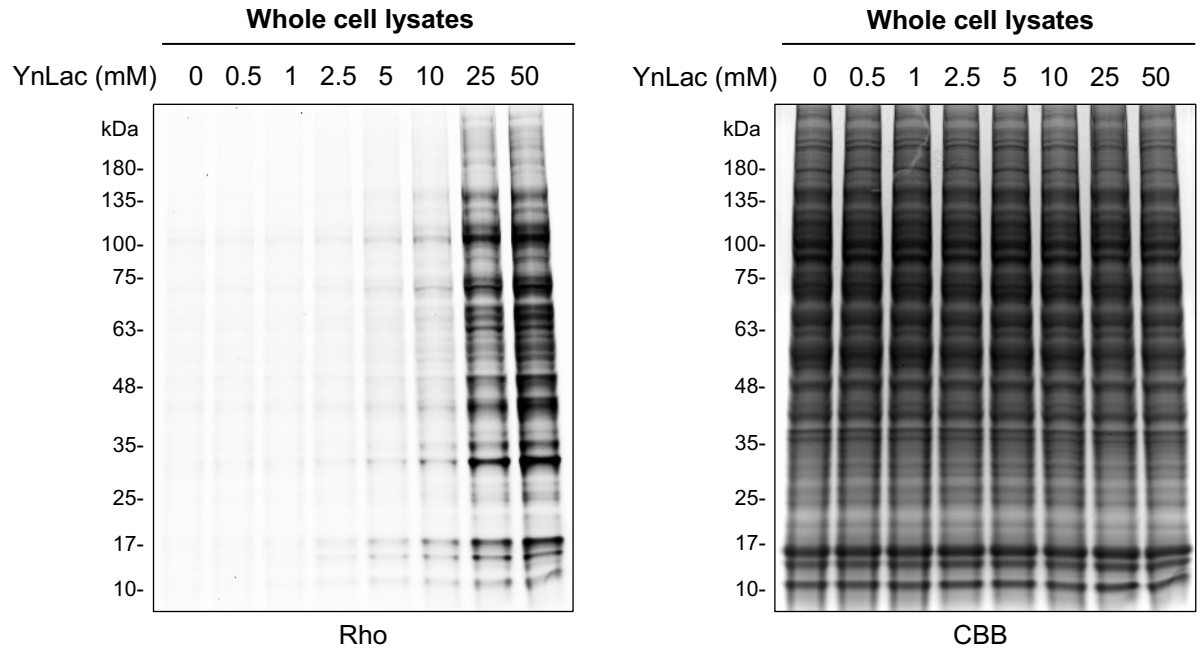


Figure S1. Concentration-dependent protein labeling by YnLac in live cells. HEK293T cells were incubated with YnLac at varying concentrations for 8 h and the whole cell lysates (lysed by SDS lysis buffer) were prepared for click reactions with az-rho and in-gel fluorescence analysis. “Rho” represents the rhodamine fluorescence channel. Coomassie brilliant blue (CBB) staining is included as protein loading control.

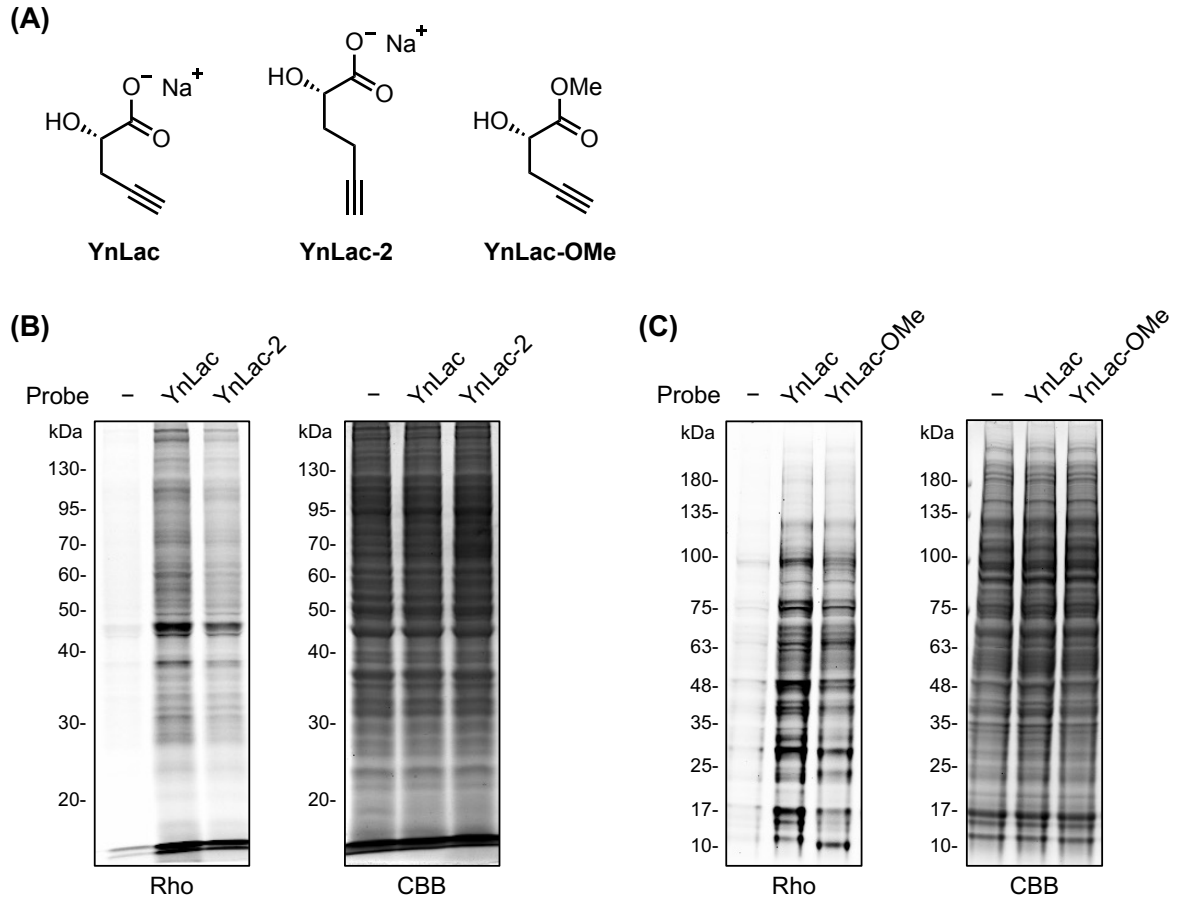


Figure S2. Investigation of L-lactate chemical reporters for metabolic labeling of proteins in live cells. (A) Structures of L-lactate chemical reporters. (B) Metabolic labeling of proteins with YnLac and YnLac-2. (C) Metabolic labeling of proteins with YnLac and YnLac-OMe. HEK293T cells were incubated with YnLac, YnLac-2, or YnLac-OMe at 10 mM concentrations for 8 h and lysed for click reactions with az-rho and in-gel fluorescence analysis. “Rho” represents the rhodamine fluorescence channel. Coomassie brilliant blue (CBB) staining is included as protein loading control. Note: A 10% acrylamide gel was used for (B), whereas a 4−20% gel was used for (C).

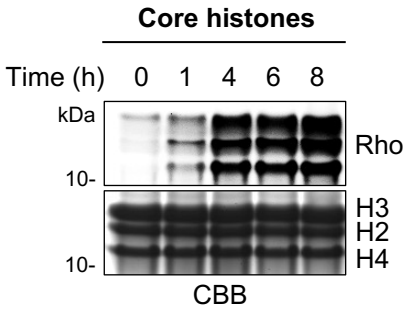


Figure S3. Time-dependent labeling of core histones by YnLac. HEK293T cells were incubated with YnLac at 20 mM for indicated time periods and the core histones were extracted for click reactions with az-rho and in-gel fluorescence analysis. “Rho” represents the rhodamine fluorescence channel. Coomassie brilliant blue (CBB) staining is included as protein loading control.

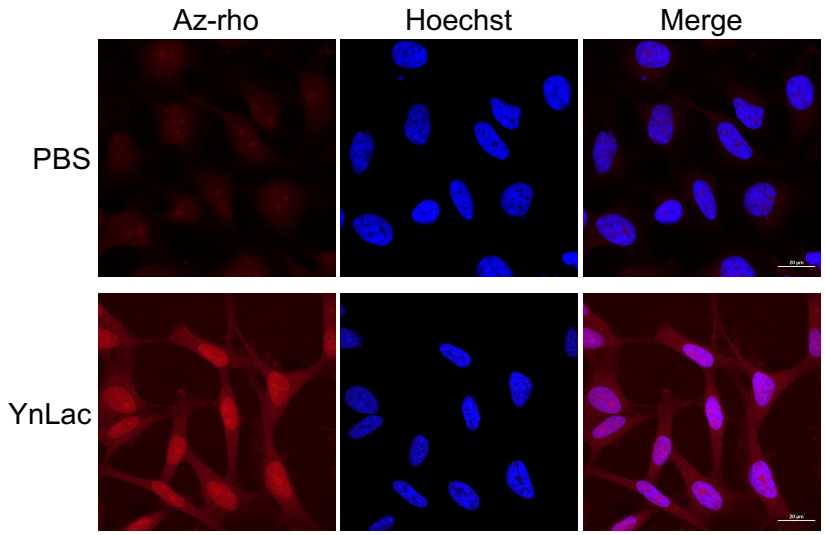


Figure S4. Fluorescence imaging of YnLac-labeled proteins in HeLa cells. Cells were incubated with YnLac (20 mM) for 12 h, fixed, and permeabilized for click reactions with az-rho and confocal fluorescence imaging. Hoechst 33342 was used for nuclear staining. Scale bars represent 20 μm.

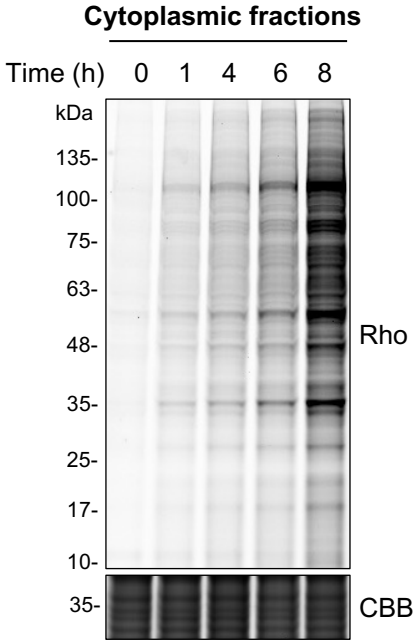


Figure S5. Time-dependent labeling of cytoplasmic proteins by YnLac. HEK293T cells were incubated with YnLac at 20 mM for indicated time periods and the cytoplasmic fractions were prepared for click reactions with az-rho and in-gel fluorescence analysis. “Rho” represents the rhodamine fluorescence channel. Coomassie brilliant blue (CBB) staining is included as protein loading control.

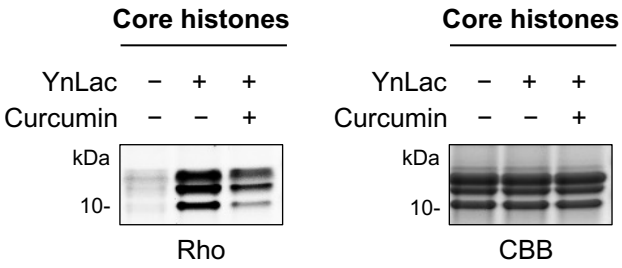


Figure S6. Metabolic labeling of core histones by YnLac in the presence of p300 inhibitor curcumin. HEK293T cells were pre-treated with 60 μM curcumin for 6 h then labeled with 10 mM YnLac for additional 6 h. Cells were lysed to separate the core histones for click reactions with az-rho and in-gel fluorescence analysis. "Rho" represents the rhodamine fluorescence channel. Coomassie brilliant blue (CBB) staining is included as protein loading control.

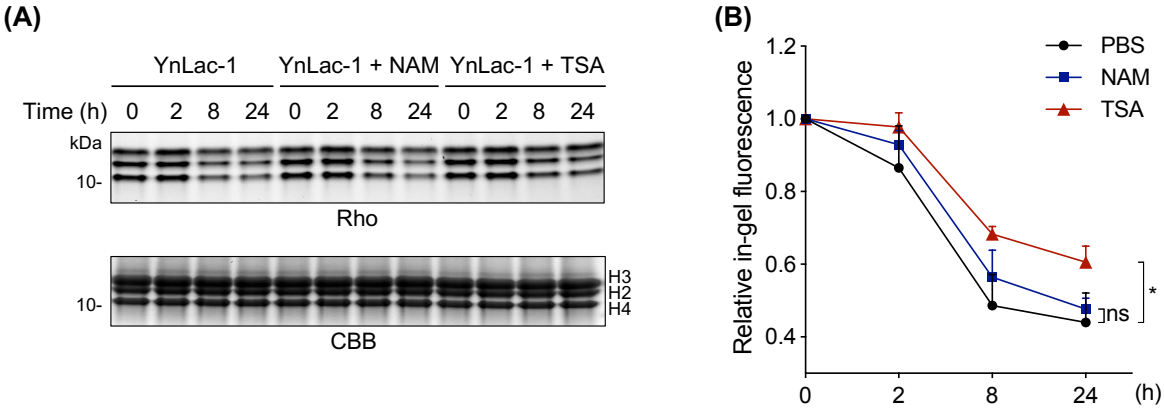


Figure S7. Analysis of the turn-over of YnLac protein labeling in the presence of deacetylase inhibitors. (A) In-gel fluorescence analysis of YnLac protein labeling in the presence of TSA or NAM. (B) Quantification of relative rhodamine fluorescence in (A). Fluorescence intensities of histone bands were quantified and normalized to the CBB intensities. Data are represented as mean ± s.d., *n* = 3, and * indicates a p-value <0.05, calculated by one-way ANOVA. HEK293T cells were labeled with 20 mM YnLac for 4 h and then treated with 0.4 μM TSA or 10 mM NAM for indicated periods. Cells were lysed to separate the core histones for click reactions with az-rho and in-gel fluorescence analysis. "Rho" represents the rhodamine fluorescence channel. Coomassie brilliant blue (CBB) staining is included as protein loading control.

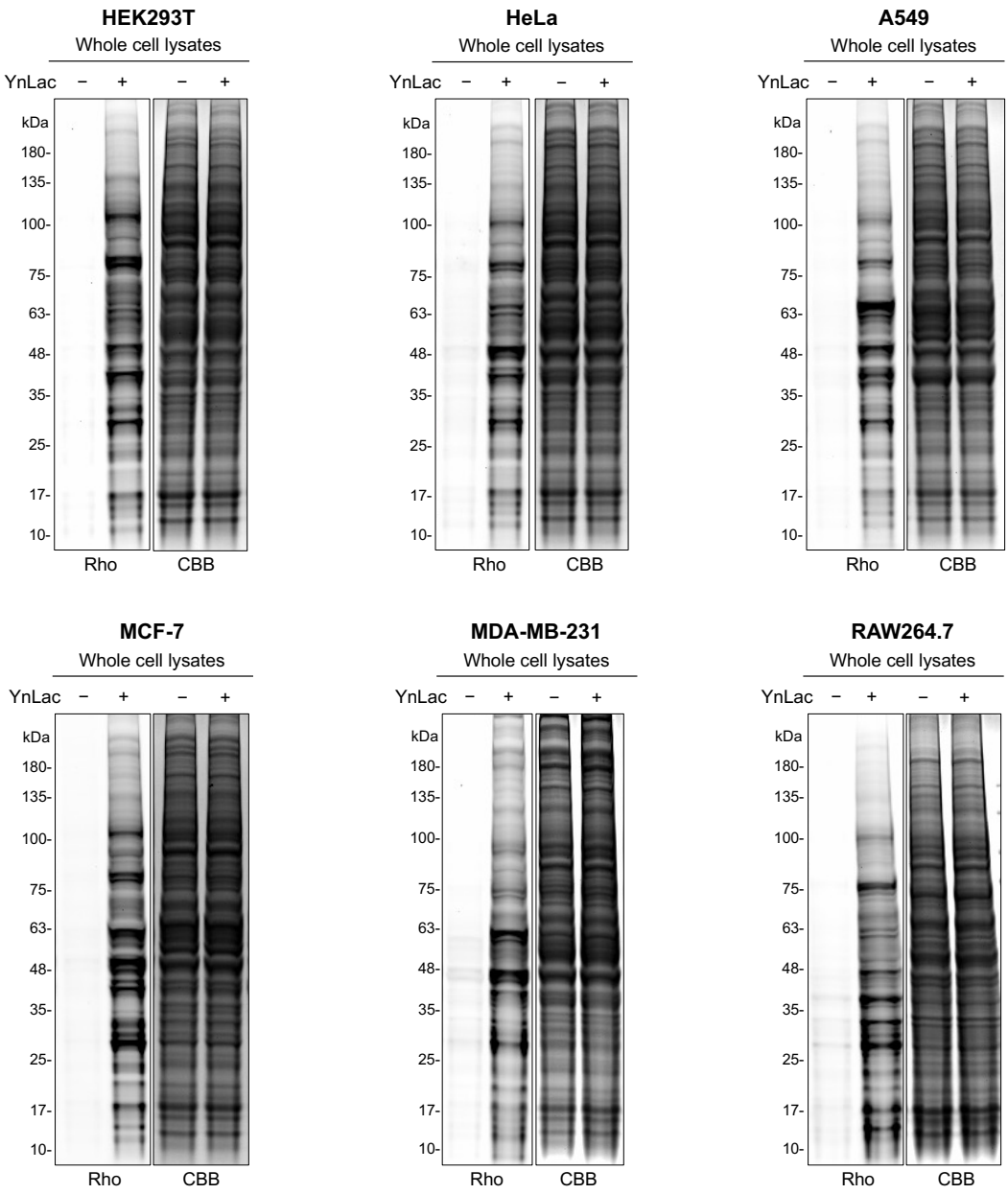
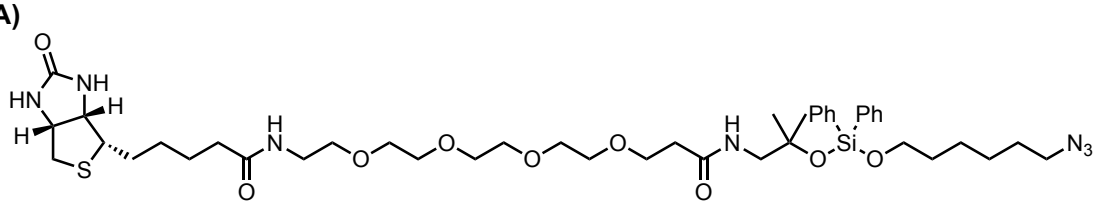


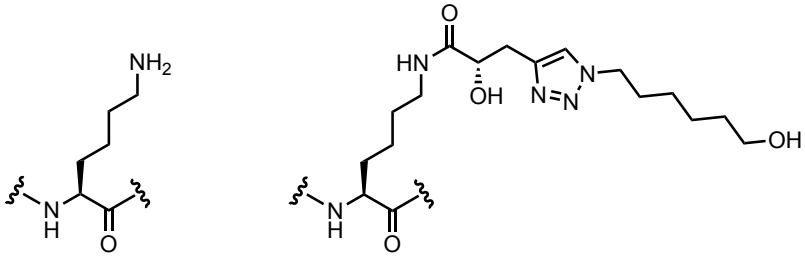
Figure S8. Metabolic labeling of proteins by YnLac in different cell types. Cells were incubated with YnLac (20 mM) for 8 h and the whole cell lysates were prepared for click reactions with az-rho and in-gel fluorescence analysis. “Rho” represents the rhodamine fluorescence channel. Coomassie brilliant blue (CBB) staining is included as protein loading control.

(A)



Azido-DADPS-biotin

(B)

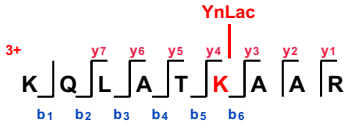


Chemical Formula: $C_6H_{13}N_2O$
Exact Mass: 129.1028

Chemical Formula: $C_{17}H_{30}N_5O_4$
Exact Mass: 368.2298

(C)

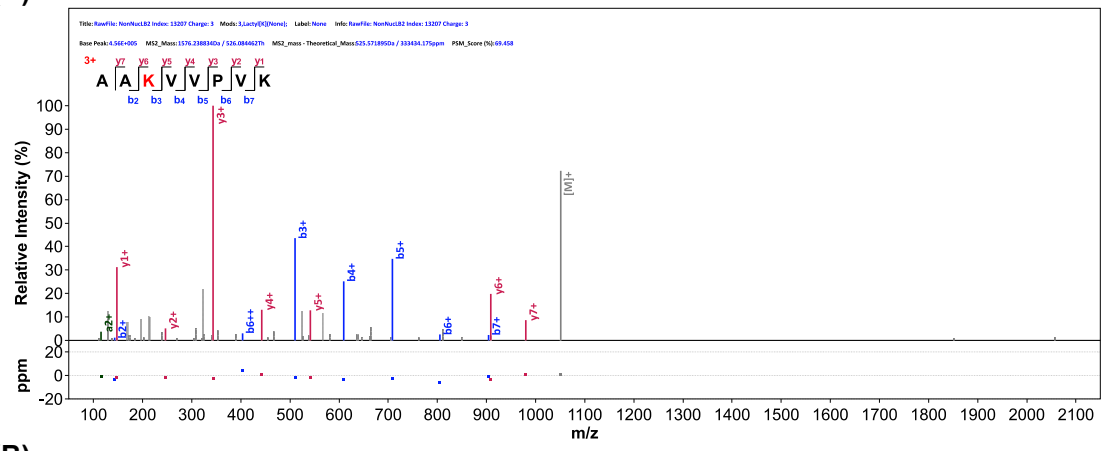
H3K23YnLac:



Ion Type	Experimental m/z	Relative intensity	Mass error (ppm)
b1+	129.1022322	100	1.06760261
b2+	257.1608036	5.48466201	-1.1026957
b3+	370.2448615	2.004579373	4.88468597
b4+	441.2819718	1.586676473	-4.4682771
b5+	542.3296454	0.98365212	-1.0056063
b6++	455.2794288	1.22330664	-8.6732891
y1+	175.1189433	35.85702202	0.09553438
y2+	246.1560536	21.70516424	0.55424142
y3+	317.1931639	43.00752627	-0.2644134
y4+	684.4150996	20.46219158	0.4535696
y4++	342.711188	13.79992635	0.23920982
y5+	785.4627732	30.14479608	-0.004036
y5++	393.2350248	28.07690417	-0.0885479
y6++	428.75358	20.67179522	-0.3031348
y7++	485.2956089	7.146290096	-0.0389868

Figure S9. Proteomic analysis of lactylation sites. (A) Structure of the acid cleavable azido-DADPS-biotin.[12] (B) Structure and molecular mass of the fragment left on the YnLac-labeled lysine residue following click reaction and acid cleavage. (C) Summary of the assigned MS/MS fragment ion peaks of the H3K23YnLac peptide shown in Figure 3B.

(A)

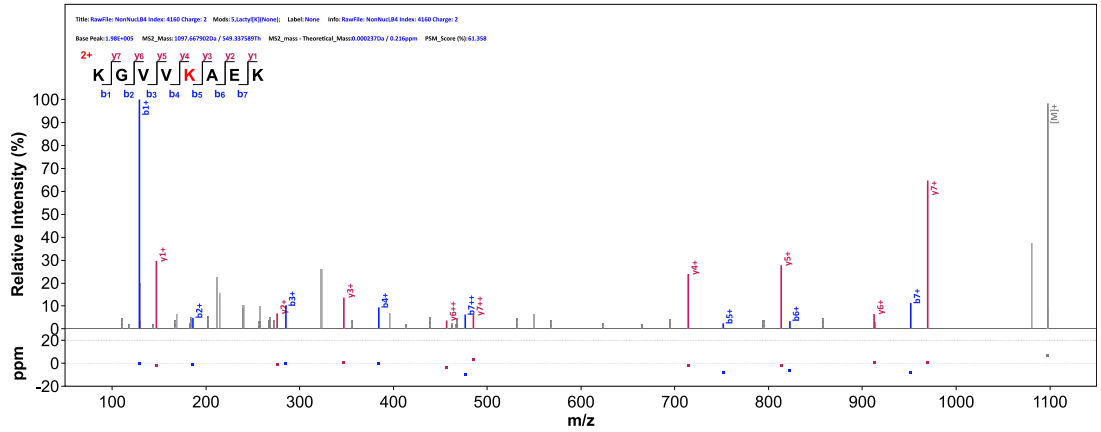


(B)

Ion Type	Experimental m/z	Ion Type	Experimental m/z
b2+	143.0815	y1+	147.1128
b3+	510.3034	y2+	246.1812
b4+	609.3718	y3+	343.2340
b5+	708.4403	y4+	442.3024
b6+	805.4930	y5+	541.3708
b7+	904.5614	y6+	908.5927
b6++	403.2501	y7+	979.6298

Figure S10. MS/MS analysis of a representative YnLac-modified peptide from NCL. (A) The MS/MS peptide spectrum showing the identification of YnLac modification on the K223 residue (in red) of NCL. (B) Summary of the assigned MS/MS fragment ion peaks of the YnLac-modified NCL peptide shown in (A).

(A)

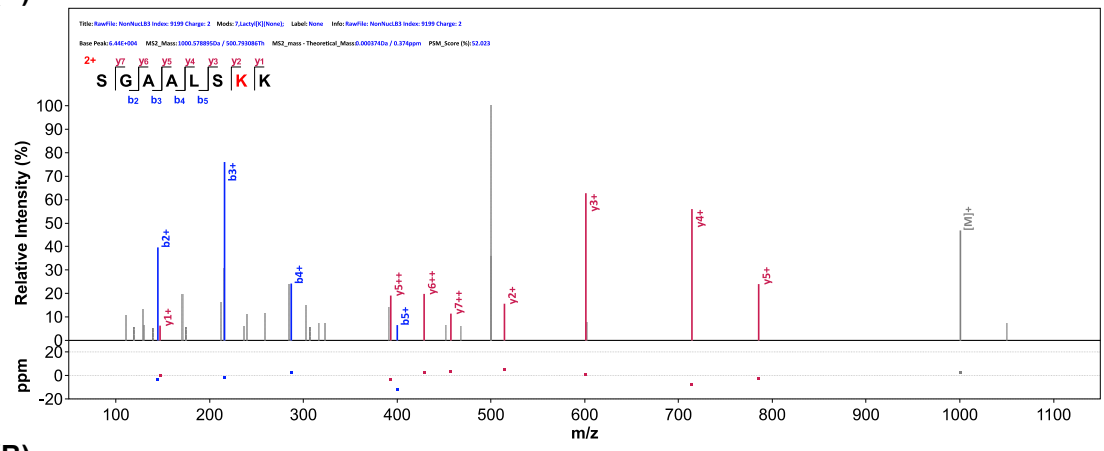


(B)

Ion Type	Experimental m/z	Ion Type	Experimental m/z
b1+	129.1022	y1+	147.1128
b2+	186.1237	y2+	276.1554
b3+	285.1921	y3+	347.1925
b4+	384.2605	y4+	714.4144
b5+	751.4824	y5+	813.4828
b6+	822.5196	y6+	912.5512
b7+	951.5621	y7+	969.5727
b7++	476.2847	y6++	456.7793
		y7++	485.2900

Figure S11. MS/MS analysis of a representative YnLac-modified peptide from HMGB1. (A) The MS/MS peptide spectrum showing the identification of YnLac modification on the K177 residue (in red) of HMGB1. (B) Summary of the assigned MS/MS fragment ion peaks of the YnLac-modified HMGB1 peptide shown in (A).

(A)

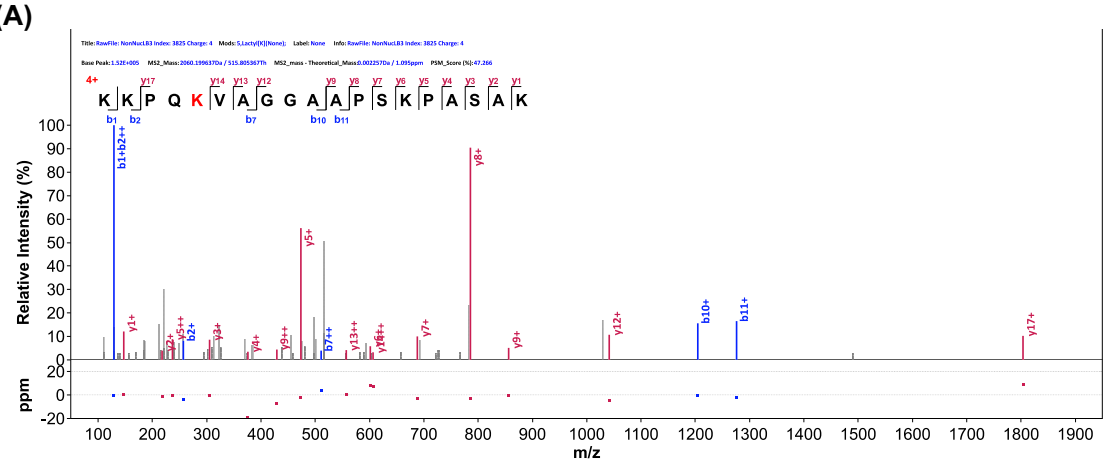


(B)

Ion Type	Experimental m/z	Ion Type	Experimental m/z
b2+	145.0608	y1+	147.1128
b3+	216.0979	y2+	514.3347
b4+	287.1350	y3+	601.3668
b5+	400.2190	y4+	714.4508
		y5+	785.4879
		y5++	393.2476
		y6++	428.7662
		y7++	457.2769

Figure S12. MS/MS analysis of a representative YnLac-modified peptide from PARP1. (A) The MS/MS peptide spectrum showing the identification of YnLac modification on the K505 residue (in red) of PARP1. (B) Summary of the assigned MS/MS fragment ion peaks of the YnLac-modified PARP1 peptide shown in (A).

(A)

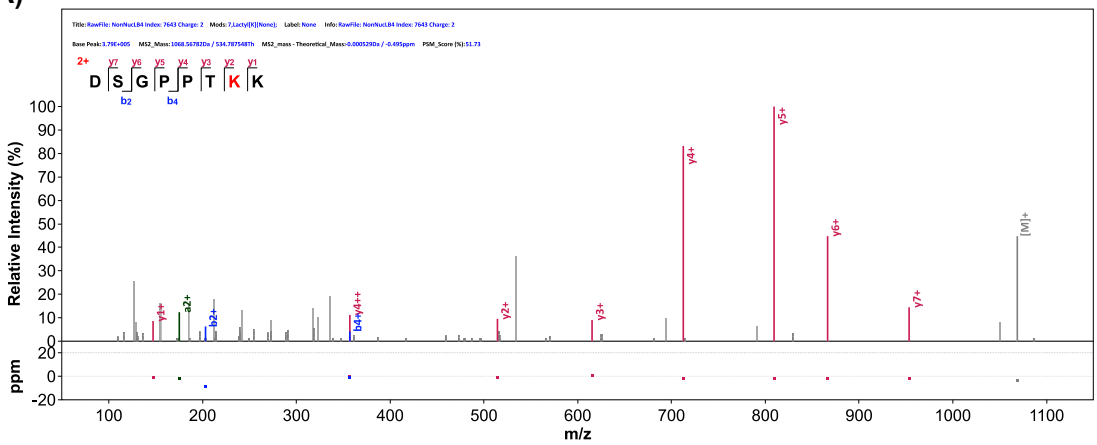


(B)

Ion Type	Experimental m/z	Ion Type	Experimental m/z
b1+	129.1022	y1+	147.1128
b2+	257.1972	y2+	218.1499
b10+	1204.7160	y3+	305.1819
b11+	1275.7531	y4+	376.2190
b2++	129.1022	y5+	473.2718
b7++	510.3216	y6+	601.3668
		y7+	688.3988
		y8+	785.4515
		y9+	856.4887
		y12+	1041.5687
		y17+	1804.0075
		y5++	237.1395
		y9++	428.7480
		y13++	556.8065
		y14++	606.3407

Figure S13. MS/MS analysis of a representative YnLac-modified peptide from NOLC1. (A) The MS/MS peptide spectrum showing the identification of YnLac modification on the K496 residue (in red) of NOLC1. (B) Summary of the assigned MS/MS fragment ion peaks of the YnLac-modified NOLC1 peptide shown in (A).

(A)

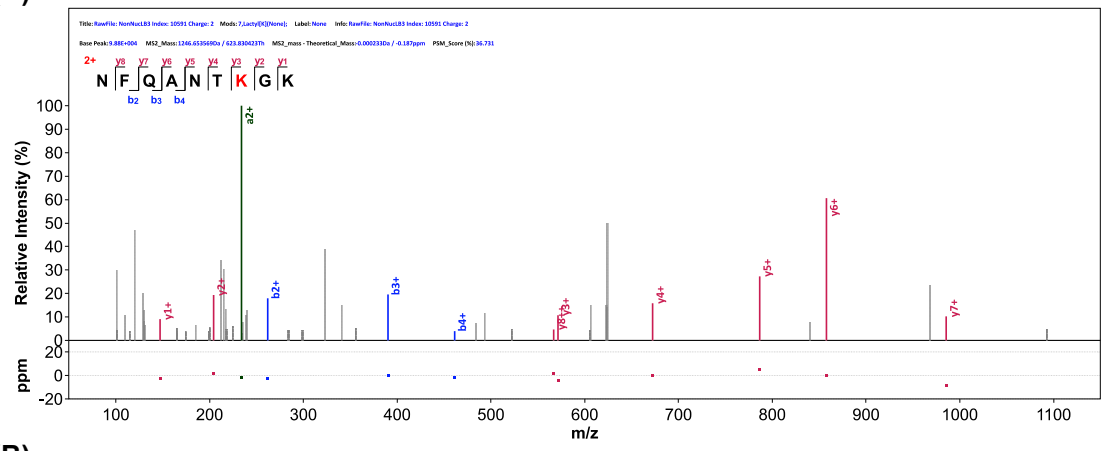


(B)

Ion Type	Experimental m/z	Ion Type	Experimental m/z
b2+	203.0662	y1+	147.1128
b4+	357.1405	y2+	514.3347
		y3+	615.3824
		y4+	712.4352
		y5+	809.4879
		y6+	866.5094
		y7+	953.5414
		y4++	356.7212

Figure S14. MS/MS analysis of a representative YnLac-modified peptide from NUCKS1. (A) The MS/MS peptide spectrum showing the identification of YnLac modification on the K35 residue (in red) of NUCKS1. (B) Summary of the assigned MS/MS fragment ion peaks of the YnLac-modified NUCKS1 peptide shown in (A).

(A)

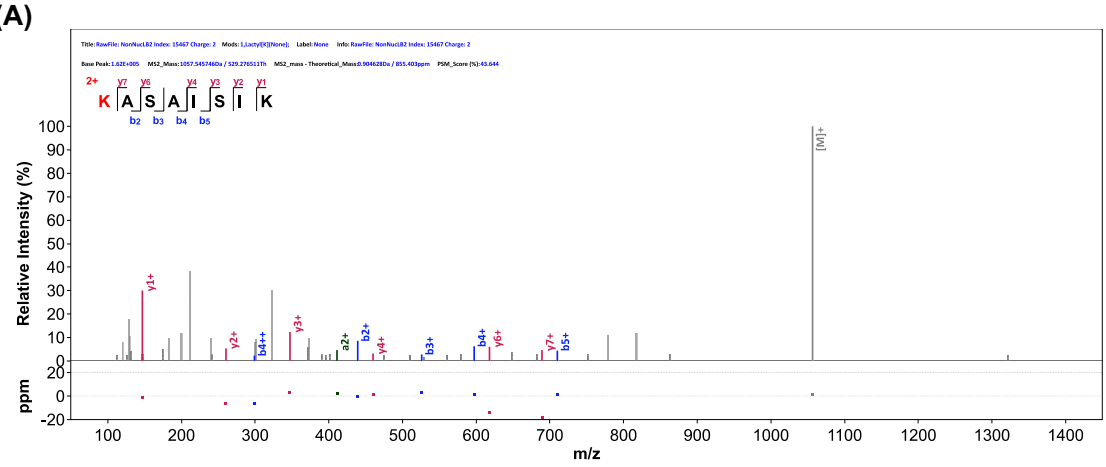


(B)

Ion Type	Experimental m/z	Ion Type	Experimental m/z
b2+	262.1186	y1+	147.1128
b3+	390.1772	y2+	204.1343
b4+	461.2143	y3+	571.3562
		y4+	672.4039
		y5+	786.4468
		y6+	857.4839
		y7+	985.5425
		y8++	566.8091

Figure S15. MS/MS analysis of a representative YnLac-modified peptide from DNAJC8. (A) The MS/MS peptide spectrum showing the identification of YnLac modification on the K231 residue (in red) of DNAJC8. (B) Summary of the assigned MS/MS fragment ion peaks of the YnLac-modified DNAJC8 peptide shown in (A).

(A)



(B)

Ion Type	Experimental m/z	Ion Type	Experimental m/z
b2+	439.2663	y1+	147.1128
b3+	526.2983	y2+	260.1969
b4+	597.3355	y3+	347.2289
b5+	710.4195	y4+	460.3129
b4++	299.1714	y6+	618.3821
		y7+	689.4192

Figure S16. MS/MS analysis of a representative YnLac-modified peptide from PCNP. (A) The MS/MS peptide spectrum showing the identification of YnLac modification on the K82 residue (in red) of PCNP. (B) Summary of the assigned MS/MS fragment ion peaks of the YnLac-modified PCNP peptide shown in (A).

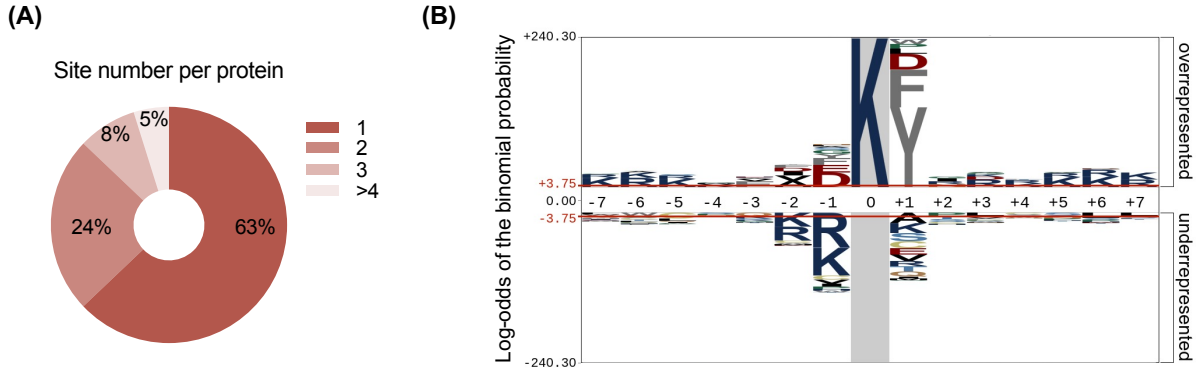


Figure S17. Analysis of lactylation sites. (A) Distribution of identified lactylation sites on proteins. (B) Sequence motif analysis of acetylated peptides. The red horizontal lines denote thresholds of $p < 0.05$. Acetylation sites were identified in a previous report.[13]

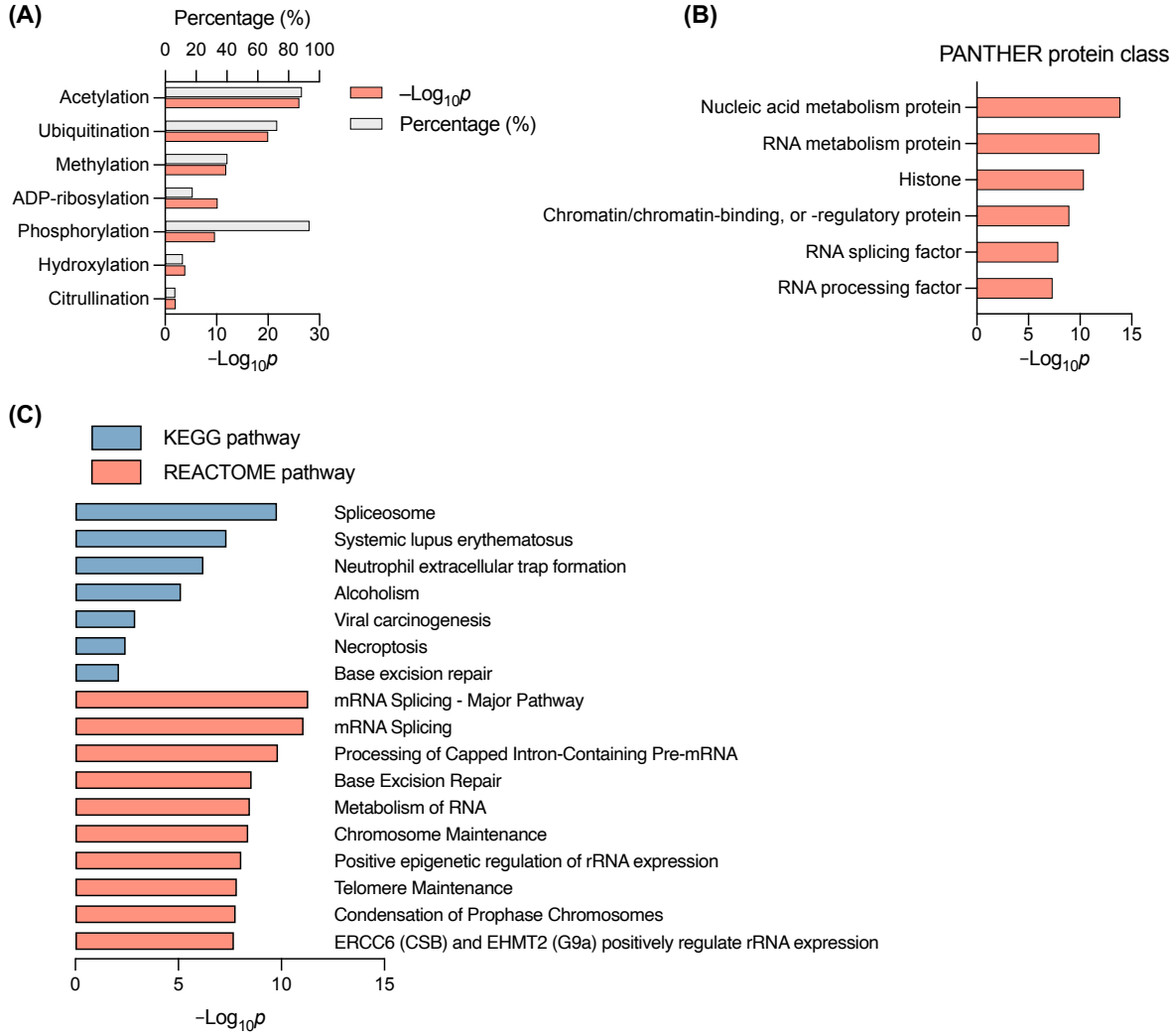


Figure S18. Bioinformatic analysis of lactylated proteins. (A) Post-translational modification (PTM) enrichment analysis of lactylated proteins. (B) PANTHER protein class enrichment analysis of lactylated proteins. (C) KEGG and REACTOME pathway enrichment analysis of lactylated proteins.

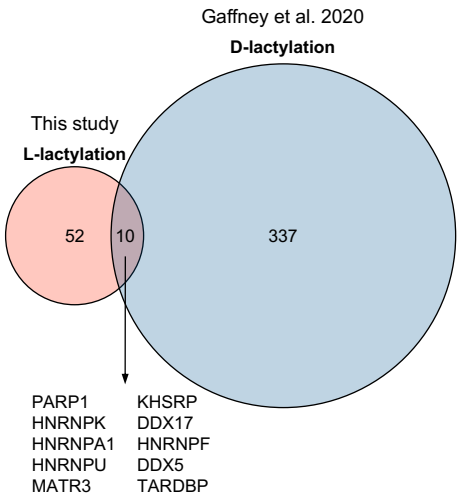


Figure S19. Cross-reference of L-lactylated and D-lactylated proteins. D-lactylated proteins were reported in a previous study.[14]

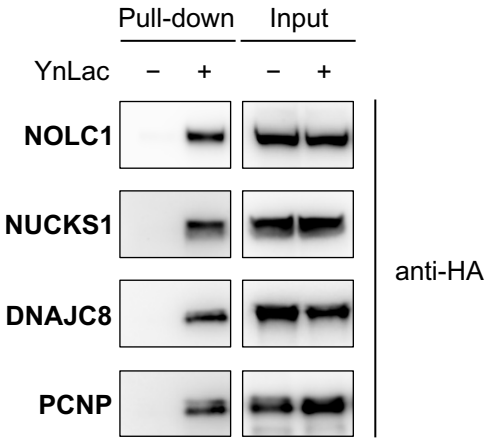


Figure S20. Validation of the enrichment of newly identified lactylated proteins. HEK293T cells expressing HA-tagged candidate lactylated proteins were labeled with YnLac (20 mM) for 8 h and lysed for click reactions with az-biotin. After pull-down with streptavidin beads, the enriched proteins were eluted for Western blotting analysis.

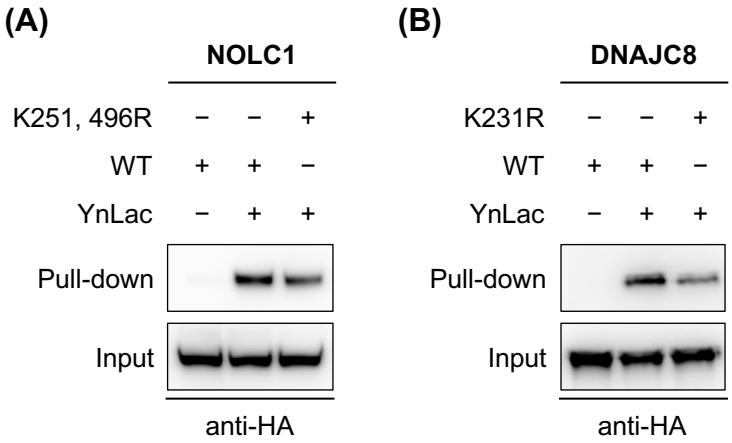


Figure S21. Validation of the identification of YnLac-modified sites on NOLC1 and DNAJC8. HEK293T cells expressing HA-tagged candidate lactylated proteins or KR mutants were labeled with YnLac (20 mM) for 8 h and lysed for click reactions with az-biotin. After pull-down with streptavidin beads, the enriched proteins were eluted for Western blotting analysis. (A) NOLC1 wild-type protein and its K251/496R mutant. (B) DNAJC8 wild-type protein and its K231R mutant.

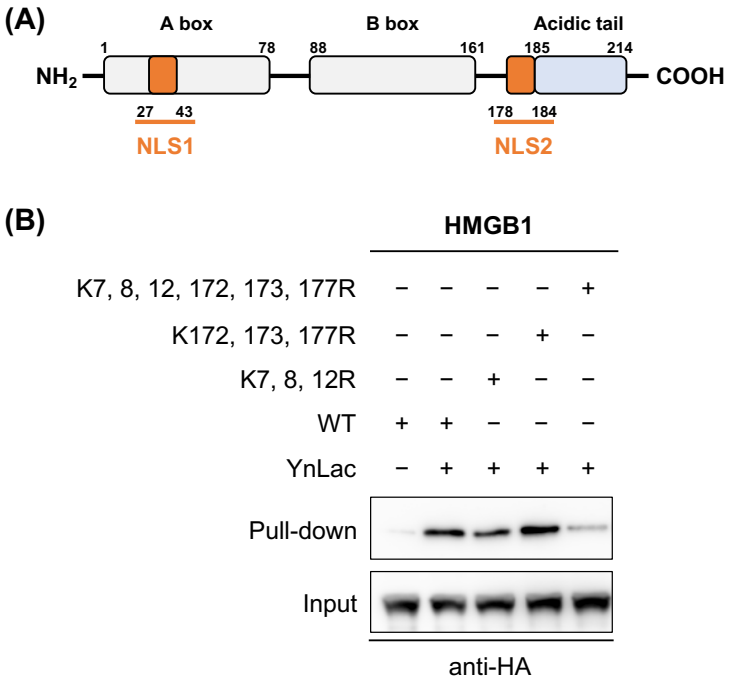


Figure S22. Investigation on YnLac-modified sites on HMGB1. (A) HMGB1 protein domains. NLS1: the first nuclear localization signal (residues 27–43); NLS2: the second nuclear localization signal (residues 178–184). (B) Pull-down of YnLac-labelled HMGB1 and mutants. HEK293T cells expressing HA-tagged HMGB1 or KR mutants (K7/8/12R, K172/173/177R, and K7/8/12/172/173/177R) were labelled with YnLac (20 mM) for 8 h and lysed for click reactions with az-biotin. After pull-down with streptavidin beads, the enriched proteins were eluted for Western blotting analysis.

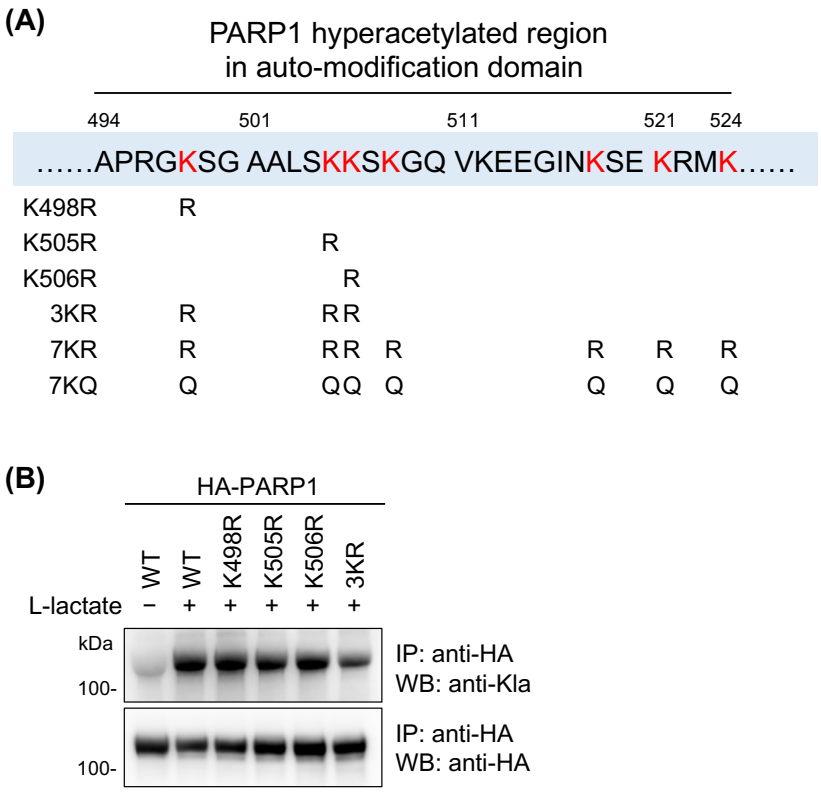


Figure S23. Analysis of PARP1 lactylation. (A) Schematic of the PARP1 disordered segment in auto-modification domain and its mutants. 7KQ is a hyperacetylated mimic.[15] (B) Western blotting analysis of lactylation of PARP1 and mutants. HEK293T cells expressing HA-tagged PARP1 or mutants were treated with or without sodium L-lactate (25 mM) and lysed for immunoprecipitation. The immunoprecipitates were analyzed by Western blotting using a pan anti-lactyllysine antibody (anti-Kla). Anti-HA blotting shows the protein loading for each lane.

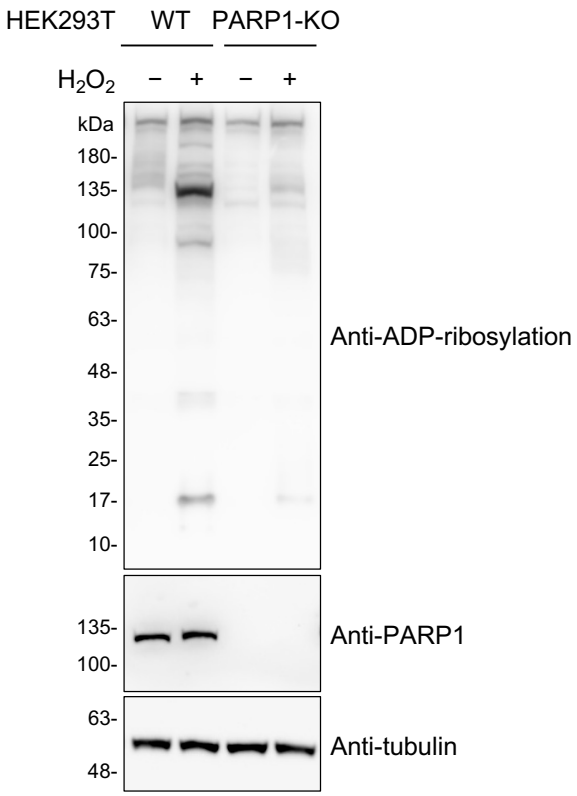


Figure S24. Characterization of PARP1 knockout HEK293T cells. Wild-type (WT) and PARP1 knockout cells were treated with or without 2 mM H_2O_2 for 10 min and lysed for Western blot analysis.

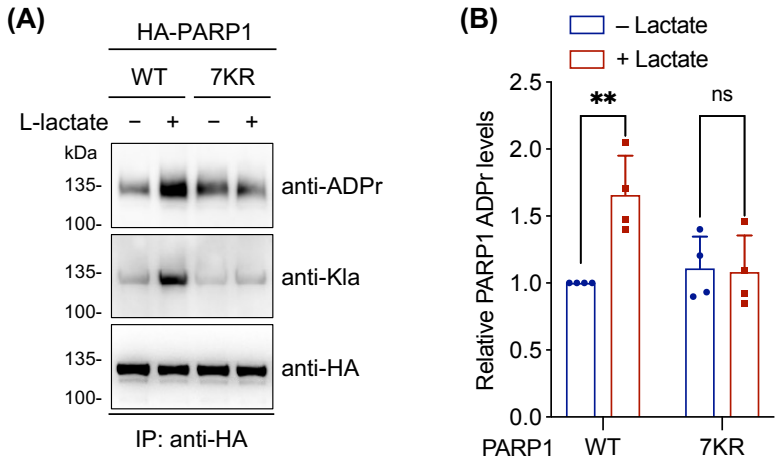


Figure S25. Investigation on PARP1 auto-ADP-ribosylation. (A) Analysis of auto-ADP-ribosylation of PARP1 and mutants. PARP1 knockout HEK293T cells were transfected with HA-tagged PARP1 or 7KR mutant (K498/505/506/508/518/521/524R), treated with or without sodium L-lactate (10 mM), and stimulated with 2 mM H_2O_2 for 10 min. Cells were lysed for immunoprecipitation and Western blot analysis using an anti-ADP-ribosylation antibody (anti-ADPr). (B) Quantification of auto-ADP-ribosylation levels of PARP1 relative to anti-HA signals shown in (A). Data are represented as mean ± s.d., n = 4. ** indicates a p-value <0.01, and ns indicates a p-value >0.05, calculated by Two-way ANOVA test.

Supplementary Tables

Table S1. Summary of the lactylation sites identified in the nuclear fraction. The complete list is shown in Table S2 in the associated excel file.
[a] For the "Position within protein", the *N*-terminal Met residue of each protein is counted as Met1.

Protein	Protein name	Gene name	Position within protein[a]	Modified Sequence and Lactyl (K) Probabilities	PEP	Score	Delta score	Score for localization	Reported? [16]
Q16778	Histone H2B type 2-E	HIST2H2BE	21	KAVTK(1)AQK	0	168.26	71.191	168.26	Y
Q16778	Histone H2B type 2-E	HIST2H2BE	6	PEPAK(1)SAPAPK	0	136.7	117.04	136.7	Y
Q16778	Histone H2B type 2-E	HIST2H2BE	12	SAPAPK(0.998)K(0.002)GSK	6.78E-141	84.244	51.773	82.417	Y
Q99878	Histone H2A type 1-J	HIST1H2AJ	10	GKQGGK(1)AR	3.50E-19	71.877	19.194	71.877	N
P05114	Non-histone chromosomal protein HMG-14	HMGN1	5	K(1)VSSAEGAAKEEPK	1.16E-22	69.92	50.605	69.92	N
P06748	Nucleophosmin	NPM1	202	DTPAK(1)NAQK	8.18E-06	68.224	52.582	68.224	N
Q71UI9	Histone H2A.V	H2AFV	5	AGGK(1)AGK	2.12E-37	88.029	55.136	88.029	N
Q93079	Histone H2B type 1-H	HIST1H2BH	6	PDPAK(1)SAPAPK	0	125.54	108.31	125.54	Y
Q93079	Histone H2B type 1-H	HIST1H2BH	12	SAPAPK(0.998)K(0.002)GSK	6.78E-141	84.244	51.773	82.417	Y
P58876	Histone H2B type 1-D	HIST1H2BD	6	PEPTK(1)SAPAPK	0	95.094	81.285	91.961	Y
P58876	Histone H2B type 1-D	HIST1H2BD	12	SAPAPK(0.998)K(0.002)GSK	6.78E-141	84.244	51.773	82.417	Y
P62805	Histone H4	HIST1H4A	9	GGK(1)GLGK	0	94.262	27.269	80.462	Y
P62805	Histone H4	HIST1H4A	13	GLGK(1)GGAK	0	98.458	45.521	98.458	Y
P84243	Histone H3.3	H3F3A	24	KQLATK(1)AAR	3.16E-08	81.525	33.565	81.525	Y

P84243	Histone H3.3	H3F3A	28	K(1)SAPSTGGVK	0.00191255	56.548	30.105	56.548	Y
Q53EL6	Programmed cell death protein 4	PDCD4	115	K(1)GGAGGK	8.61E-206	86.086	18.346	86.086	N
Q92576	PHD finger protein 3	PHF3	957	AAK(1)VATK	9.37E-263	88.029	18.715	88.029	N
Q99880	Histone H2B type 1-L	HIST1H2BL	6	PELAK(1)SAPAPK	0.00136829	60.255	49.884	60.255	Y
Q99880	Histone H2B type 1-L	HIST1H2BL	12	SAPAPK(0.998)K(0.002)GSK	6.78E-141	84.244	51.773	82.417	Y

Table S2. Complete list of the lactylation sites identified in the nuclear fraction. All database search information is listed. See the associated excel file.

Table S3. Summary of the lactylation sites identified in the cytoplasmic fraction. The complete list is shown in Table S4 in the associated excel file.

[a] For the "Position within protein", the *N*-terminal Met residue of each protein is counted as Met1.

Protein	Protein name	Gene name	Position within protein[a]	Modified Sequence and Lactyl (K) Probabilities	PEP	Score	Delta score	Score for localization
A0A024R1R8	Translation machinery-associated protein 7	hCG_2014768	59	GPLATGGIK(0.5)K(0.5)	0.00042973	53.756	28.197	53.756
A0A024R1R8	Translation machinery-associated protein 7	hCG_2014768	60	GPLATGGIK(0.5)K(0.5)	0.00042973	53.756	28.197	53.756
O00193	Small acidic protein	SMAP	174	SNYK(1)MMFVK	0.00583565	56.916	46.911	56.916
O15042	U2 snRNP-associated SURP motif-containing protein	U2SURP	88	AFSIGK(0.952)MSTAK(0.048)	5.58E-12	55.314	36.22	55.314
O43719	HIV Tat-specific factor 1	HTATSF1	239	K(1)LSMQQK	0.0156274	41.892	18.124	41.892
O43768	Alpha-endosulfine	ENSA	107	K(1)SSLVTSK	0.0214605	44.753	11.861	44.753
O43768	Alpha-endosulfine	ENSA	80	NK(1)QLPSAGPDK	0.0035582	49.934	36.897	49.934
O60869	Endothelial differentiation-related factor 1	EDF1	23	KKGPTAAQAK(0.731)SK(0.269)	1.58E-06	48.532	31.619	48.532
O60869	Endothelial differentiation-related factor 1	EDF1	25	SK(1)QAILAAQR	3.41E-14	74.841	38.446	74.841
O75937	DnaJ homolog subfamily C member 8	DNAJC8	231	NFQANTK(0.898)GK(0.102)	6.83E-28	80.688	55.794	74.173
P05455	Lupus La protein	SSB	354	GK(1)VQFQGK	2.78E-155	98.458	74.503	98.458
P06748	Nucleophosmin	NPM1	202	DTPAK(1)NAQK	1.53E-12	74.301	36.743	67.494
P06748	Nucleophosmin	NPM1	233	KQEK(1)TPK	0	116.25	67.19	109.71
P06748	Nucleophosmin	NPM1	141	LLSISGK(1)R	0.00798037	60.91	11.552	41.448
P06748	Nucleophosmin	NPM1	150	SAPGGGSK(0.999)VPQK(0.001)	3.23E-05	51.286	36.275	51.286

P06748	Nucleophosmin	NPM1	223	SK(1)GQESFK	5.66E-88	91.867	35.745	91.867
P06748	Nucleophosmin	NPM1	239	TPK(1)GPSSVEDIK	1.66E-05	61.235	37.28	61.235
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	HNRNPC	157	AVVPSK(1)R	7.11E-79	86.866	29.764	86.866
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	HNRNPC	94	GK(1)AGVKR	0.0196069	64.757	26.624	64.757
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	HNRNPC	176	SGFNSK(1)SGQR	1.23E-133	93.839	83.12	93.839
P09429	High mobility group protein B1	HMGB1	177	KGVVK(1)AEK	0	120.31	68.519	120.31
P09651	Heterogeneous nuclear ribonucleoprotein A1	HNRNPA1	3	SK(1)SESPKEPEQLR	7.82E-06	63.185	50.177	63.185
P09661	U2 small nuclear ribonucleoprotein A'	SNRPA1	172	GAQLAK(1)DIAR	0.0107642	57.785	42.525	57.785
P09874	Poly [ADP-ribose] polymerase 1	PARP1	498	GK(1)SGAALSK	2.66E-48	82.008	50.073	82.008
P09874	Poly [ADP-ribose] polymerase 1	PARP1	505	SGAALSK(0.692)K(0.308)	0.0046151	64.42	37.283	64.42
P09874	Poly [ADP-ribose] polymerase 1	PARP1	506	SGAALSK(0.429)K(0.571)	0.00504922	57.59	32.752	50.04
P17844	Probable ATP-dependent RNA helicase DDX5	DDX5	523	GYSSLLK(1)R	0.00761281	61.962	32.813	52.716
P17844	Probable ATP-dependent RNA helicase DDX5	DDX5	33	K(1)FGNPGEK	0.00088372	62.546	30.076	62.546
P19338	Nucleolin	NCL	95	AAATPAK(0.5)K(0.5)	0.00565823	63.565	33.822	63.565
P19338	Nucleolin	NCL	96	K(1)TVTPAK	0.00248219	67.981	34.43	67.981
P19338	Nucleolin	NCL	223	AAK(1)VVPVK	1.10E-204	103.83	52.867	103.83
P19338	Nucleolin	NCL	87	AAVTPGK(0.819)K(0.181)	0.0207502	46.813	25.131	46.813
P19338	Nucleolin	NCL	88	K(1)AAATPAKK	9.53E-27	78.113	48.775	78.113
P19338	Nucleolin	NCL	9	AGK(1)NQGDPK	3.28E-05	68.371	42.325	68.371

P19338	Nucleolin	NCL	124	ALVATPGK(0.5)K(0.5)	1.02E-75	89.698	66.679	89.698
P19338	Nucleolin	NCL	125	K(1)GAAIPAK	2.77E-228	109.11	79.961	64.265
P19338	Nucleolin	NCL	110	K(1)GATPGK	1.69E-12	73.26	25.411	73.26
P19338	Nucleolin	NCL	116	GATPGK(1)ALVATPGK	9.89E-12	56.872	43.809	52.527
P19338	Nucleolin	NCL	55	K(1)AAATSAK	0	122.28	72.691	122.28
P19338	Nucleolin	NCL	63	K(1)VVVSPTK	2.73E-77	86.136	30.766	86.136
P19338	Nucleolin	NCL	80	K(1)AAVTPGKK	1.64E-228	109.71	54.184	70.056
P19338	Nucleolin	NCL	102	TVTPAK(1)AVTTPGK	1.55E-87	90.412	62.129	90.412
P19338	Nucleolin	NCL	71	K(1)VAVATPAK	4.33E-06	68.557	52.018	68.557
P19338	Nucleolin	NCL	79	VAVATPAK(0.909)K(0.091)	0	123.72	55.74	123.72
P27694	Replication protein A 70 kDa DNA-binding subunit	RPA1	163	AYGASK(0.995)TFGK(0.005)	1.72E-18	73.499	55.275	73.499
P27694	Replication protein A 70 kDa DNA-binding subunit	RPA1	88	FIVNTLK(1)DGR	0.00233473	56.404	46.4	47.302
P37108	Signal recognition particle 14 kDa protein	SRP14	43	K(1)GTVEGFEPADNK	1.90E-05	44.511	34.711	44.511
P39748	Flap endonuclease 1	FEN1	361	RKEPEPK(0.998)GSTK(0.002)	0.00028173	50.354	26.677	50.354
P43243	Matrin-3	MATR3	702	K(1)DGSASAAAK	1.16E-15	57.532	28.383	57.532
P51965	Ubiquitin-conjugating enzyme E2 E1	UBE2E1	43	NSK(1)LLSTSAK	0.00044788	64.82	45.034	64.82
P52597	Heterogeneous nuclear ribonucleoprotein F	HNRNPF	224	YIGIVK(1)QAGLER	6.92E-22	73.435	48.582	71.176
P53999	Activated RNA polymerase II transcriptional coactivator p15	SUB1	53	ALSSSK(1)QSSSSR	2.47E-38	80.69	53.209	80.69
P53999	Activated RNA polymerase II transcriptional coactivator p15	SUB1	29	K(0.998)QVAPEK(0.001)PVK(0.001)K	0.00245052	40.154	21.197	40.154

P61978	Heterogeneous nuclear ribonucleoprotein K	HNRNPK	405	DLAGSIIGK(1)GGQR	5.74E-06	55.261	41.861	49.448
P62304	Small nuclear ribonucleoprotein E	SNRPE	12	VQK(1)VMVQPINLIFR	1.04E-05	59.227	51.66	59.227
Q5VTE0	Putative elongation factor 1-alpha-like 3	EEF1A1P5	457	SAQK(1)AQK	1.77E-169	104.37	45.016	104.21
Q5VTE0	Putative elongation factor 1-alpha-like 3	EEF1A1P5	453	VTK(1)SAQK	3.92E-15	77.282	39.637	72.789
P78347	General transcription factor II-I	GTF2I	353	EFNFEK(1)WNAR	0.0118696	51.762	42.44	51.762
Q00839	Heterogeneous nuclear ribonucleoprotein U	HNRNPU	215	K(1)AEGGGGGGRPGAPAAGDGK	9.49E-20	54.964	44.194	54.964
Q01105	Protein SET	SET	189	SSQTQNK(1)ASR	5.13E-131	97.813	83.134	87.568
Q07666	KH domain-containing, RNA-binding, signal transduction-associated protein 1	KHDRBS1	432	APPARPVK(1)GAYR	9.54E-05	47.302	25.752	47.302
Q08945	FACT complex subunit SSRP1	SSRP1	540	K(0.5)GK(0.5)DPNAPK	0.0130152	45.022	29.878	45.022
Q08945	FACT complex subunit SSRP1	SSRP1	542	K(0.5)GK(0.5)DPNAPK	0.0130152	45.022	29.878	45.022
Q13148	TAR DNA-binding protein 43	TARDBP	84	K(1)MDETDASSAVK	0.00049464	60.196	48.613	60.196
Q13185	Chromobox protein homolog 3	CBX3	92	K(1)SLSDSESDDSK	7.68E-23	73.927	63.283	73.927
Q13185	Chromobox protein homolog 3	CBX3	20	SK(0.955)K(0.045)VEEAEPEEFVVEK	6.83E-11	58.246	46.663	58.246
Q13185	Chromobox protein homolog 3	CBX3	10	TTLQK(0.996)MGK(0.004)	3.99E-33	82.671	53.522	82.671
Q13435	Splicing factor 3B subunit 2	SF3B2	877	K(1)AQPQDSR	0.0164184	48.794	31.282	48.794
Q14978	Nucleolar and coiled-body phosphoprotein 1	NOLC1	496	K(0.199)K(0.199)PQK(0.602)VAGGAA PSKPASAK	2.08E-15	47.088	29.972	43.696
Q14978	Nucleolar and coiled-body phosphoprotein 1	NOLC1	251	K(0.003)QVVAK(0.997)APVK	0.00182515	49.06	16.59	49.06
Q15233	Non-POU domain-containing octamer-binding protein	NONO	467	AAPGAEFAPNK(1)R	0	119.39	105.21	119.39
Q15691	Microtubule-associated protein RP/EB family member 1	MAPRE1	174	TAAAPK(1)AGPGVVR	9.78E-06	47.062	35.015	41.242

Q4G0X9	Coiled-coil domain-containing protein 40	CCDC40	881	ETIK(0.89)MQDK(0.11)	0.018897	54.157	25.008	54.157
Q8WW12	PEST proteolytic signal-containing nuclear protein	PCNP	82	K(1)ASAISIK	3.96E-19	74.173	33.291	72.607
Q8WXF1	Paraspeckle component 1	PSPC1	14	IEK(1)NPAR	1.24E-41	68.312	40.138	40.268
Q92841	Probable ATP-dependent RNA helicase DDX17	DDX17	109	K(1)FGNPGER	0.00761281	61.962	48.954	61.962
Q92945	Far upstream element-binding protein 2	KHSRP	71	K(1)DAFADAVQR	0.00393654	44.543	30.385	44.543
Q96AT1	Uncharacterized protein KIAA1143	KIAA1143	99	K(0.814)PVK(0.186)HPSDEK	0.00146283	51.276	27.923	51.276
Q96I25	Splicing factor 45	RBM17	41	AALTQAK(1)SQR	0.0208151	42.21	31.98	40.496
Q96I25	Splicing factor 45	RBM17	245	IMQK(1)YGFR	6.95E-06	55.676	35.156	45.081
Q96I25	Splicing factor 45	RBM17	24	NFK(1)LLQSQLQVK	0.00156423	42.843	25.631	42.843
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	NUCKS1	186	ATVTPSPVK(0.441)GK(0.559)	0.00056024	54.259	34.473	54.259
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	NUCKS1	35	DSGPPTK(0.507)K(0.493)	0.00058113	68.016	51.476	68.016
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	NUCKS1	36	DSGPPTK(0.424)K(0.576)	0.0107158	57.59	47.511	57.59
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	NUCKS1	188	GK(1)VGRPTASK	0.00016745	55.353	31.323	55.353
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	NUCKS1	196	VGRPTASK(1)ASK	2.95E-37	78.705	51	78.705
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	NUCKS1	9	K(1)VVDYSQFQESDDADEDYGR	1.28E-75	83.423	82.459	83.423
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	NUCKS1	175	LK(1)ATVTPSPVK	3.81E-12	71.176	52.081	71.176
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	NUCKS1	98	QAASK(1)AASK	2.23E-05	68.847	38.235	68.847
Q9H444	Charged multivesicular body protein 4b	CHMP4B	6	SVFGK(1)LFGAGGGK	1.09E-07	54.343	47.656	54.343
Q9HB71	Calcyclin-binding protein	CACYBP	212	TINK(1)AWVESR	0.00974291	47.302	33.493	47.302

Q9NR30	Nucleolar RNA helicase 2	DDX21	97	K(0.492)K(0.492)K(0.016)EPIEK	0.0133739	45.368	22.727	45.368
Q9NR30	Nucleolar RNA helicase 2	DDX21	98	K(0.492)K(0.492)K(0.016)EPIEK	0.0133739	45.368	22.727	45.368
Q9UHX1	Poly(U)-binding-splicing factor PUF60	PUF60	454	HMVMQK(1)LLR	0.0134816	41.448	34.994	41.448
Q9Y2W1	Thyroid hormone receptor-associated protein 3	THRAP3	527	VTAYK(1)AVQEK	0.00044788	64.82	41.144	64.82
Q9Y5B9	FACT complex subunit SPT16	SUPT16H	1021	K(1)ASVHSSGR	0.00197861	47.187	39.62	47.187

Table S4. Complete list of the lactylation sites identified in the cytoplasmic fraction. All database search information is listed. See the associated excel file.

Table S5. PTM enrichment analysis of lactylated proteins. See the associated excel file.

Table S6. PANTHER protein class enrichment analysis of lactylated proteins. See the associated excel file.

Table S7. GO enrichment analysis of lactylated proteins. See the associated excel file.

Table S8. KEGG and REACTOME pathway enrichment analysis of lactylated proteins. See the associated excel file.

References

[1] G. Charron, M. M. Zhang, J. S. Yount, J. Wilson, A. S. Raghavan, E. Shamir, H. C. Hang, *J. Am. Chem. Soc.* **2009**, *131*, 4967-4975.

[2] K. Suzuki, P. Bose, R. Y. Y. Leong-Quong, D. J. Fujita, K. Riabowol, *BMC Res. Notes* **2010**, *3*, 294.

[3] D. Shechter, H. L. Dormann, C. D. Allis, S. B. Hake, *Nat. Protoc.* **2007**, *2*, 1445-1457.

[4] J. Cox, M. Mann, *Nat. Biotechnol.* **2008**, *26*, 1367-1372.

[5] S. Tyanova, T. Temu, P. Sinitcyn, A. Carlson, M. Y. Hein, T. Geiger, M. Mann, J. Cox, *Nat. Methods* **2016**, *13*, 731-740.

[6] H. Chi, C. Liu, H. Yang, W.-F. Zeng, L. Wu, W.-J. Zhou, R.-M. Wang, X.-N. Niu, Y.-H. Ding, Y. Zhang, Z.-W. Wang, Z.-L. Chen, R.-X. Sun, T. Liu, G.-M. Tan, M.-Q. Dong, P. Xu, P.-H. Zhang, S.-M. He, *Nat. Biotechnol.* **2018**, *36*, 1059-1061.

[7] D. W. Huang, B. T. Sherman, R. A. Lempicki, *Nat. Protoc.* **2009**, *4*, 44-57.

[8] H. Mi, D. Ebert, A. Muruganujan, C. Mills, L.-P. Albou, T. Mushayamaha, P. D. Thomas, *Nucleic Acids Res.* **2020**, *49*, D394-D403.

[9] J. P. O'Shea, M. F. Chou, S. A. Quader, J. K. Ryan, G. M. Church, D. Schwartz, *Nat. Methods* **2013**, *10*, 1211-1212.

[10] H. Huang, C. N. Arighi, K. E. Ross, J. Ren, G. Li, S.-C. Chen, Q. Wang, J. Cowart, K. Vijay-Shanker, C. H. Wu, *Nucleic Acids Res.* **2018**, *46*, D542-D550.

[11] P. V. Hornbeck, B. Zhang, B. Murray, J. M. Kornhauser, V. Latham, E. Skrzypek, *Nucleic Acids Res.* **2014**, *43*, D512-D520.

[12] J. Szychowski, A. Mahdavi, J. J. L. Hodas, J. D. Bagert, J. T. Ngo, P. Landgraf, D. C. Dieterich, E. M. Schuman, D. A. Tirrell, *J. Am. Chem. Soc.* **2010**, *132*, 18351-18360.

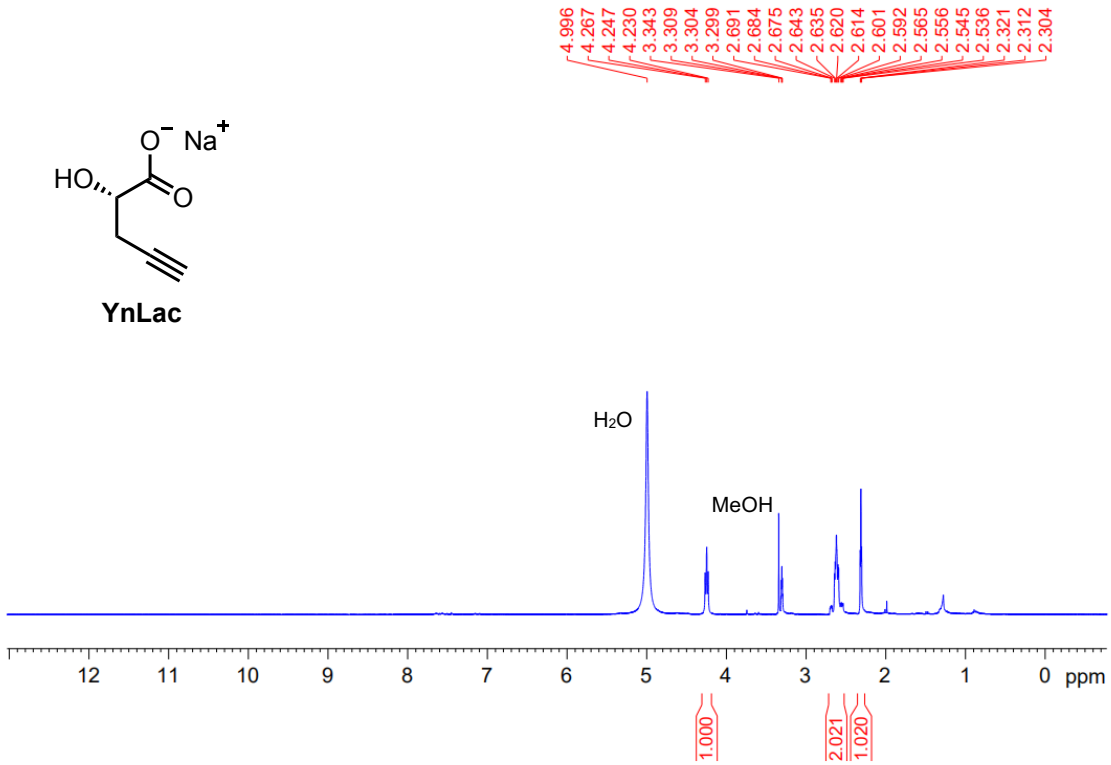
[13] H. Huang, S. Tang, M. Ji, Z. Tang, M. Shimada, X. Liu, S. Qi, J. W. Locasale, R. G. Roeder, Y. Zhao, X. Li, *Mol. Cell* **2018**, *70*, 663-678.e666.

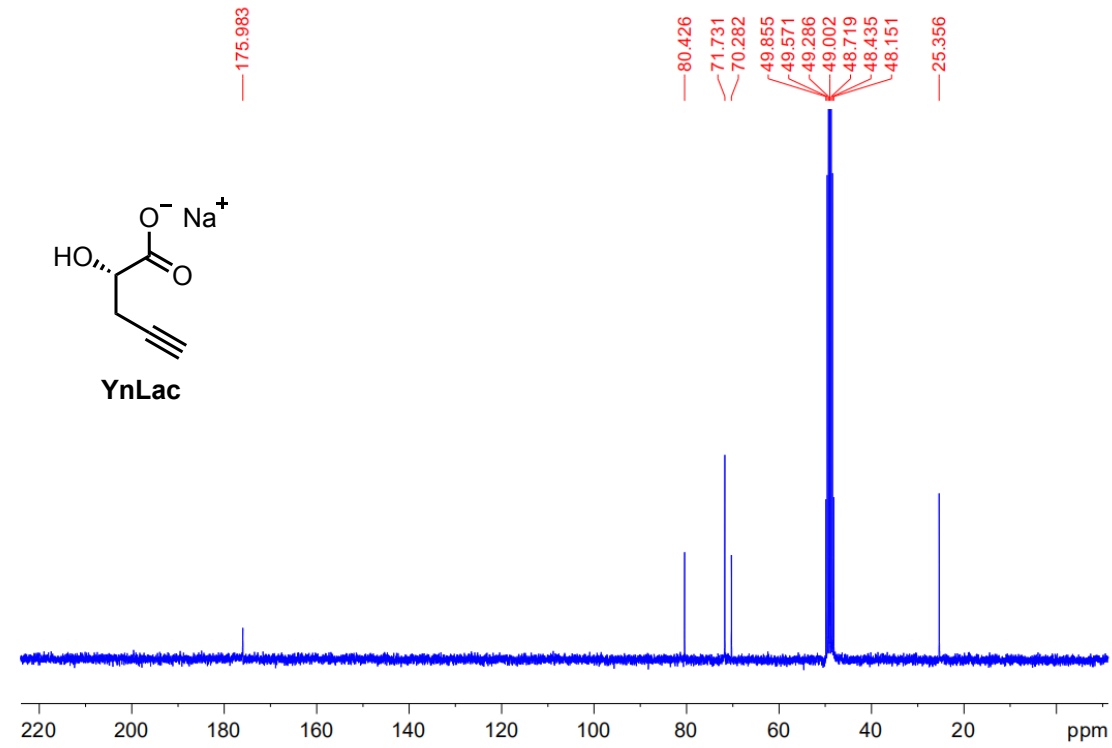
[14] D. O. Gaffney, E. Q. Jennings, C. C. Anderson, J. O. Marentette, T. Shi, A.-M. Schou Oxvig, M. D. Streeter, M. Johannsen, D. A. Spiegel, E. Chapman, J. R. Roede, J. J. Galligan, *Cell Chem. Biol.* **2020**, *27*, 206-213.e206.

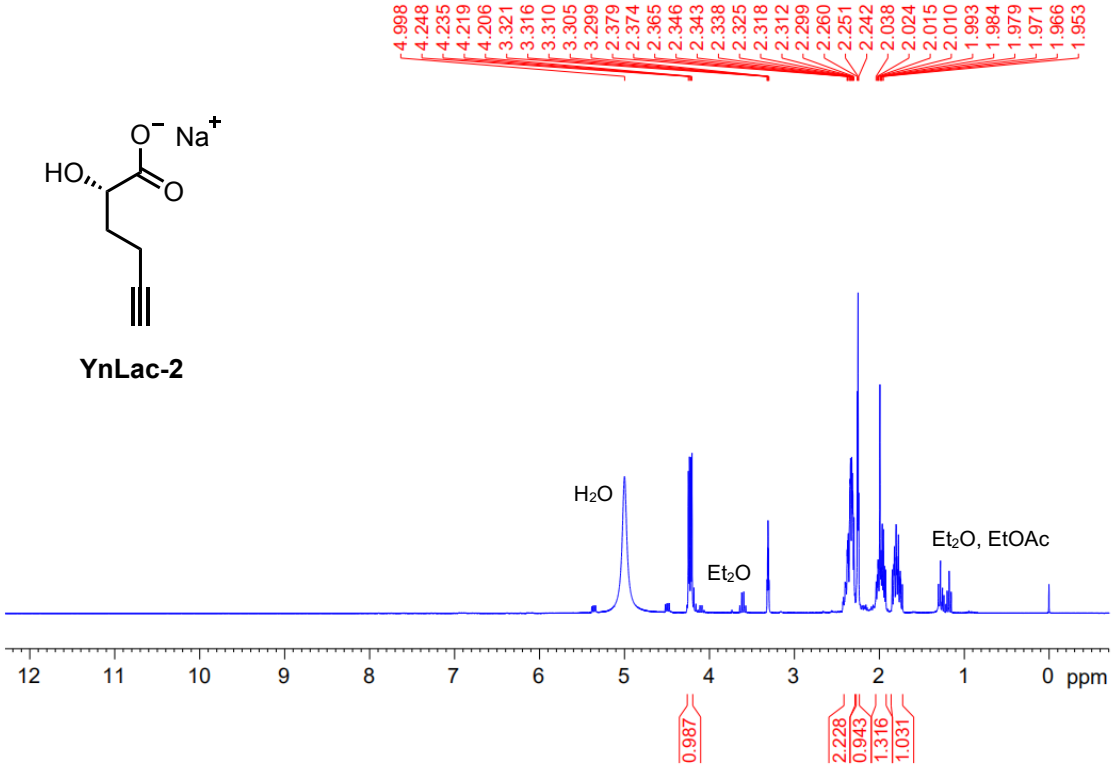
[15] G. Liszczak, K. L. Diehl, G. P. Dann, T. W. Muir, *Nat. Chem. Biol.* **2018**, *14*, 837-840.

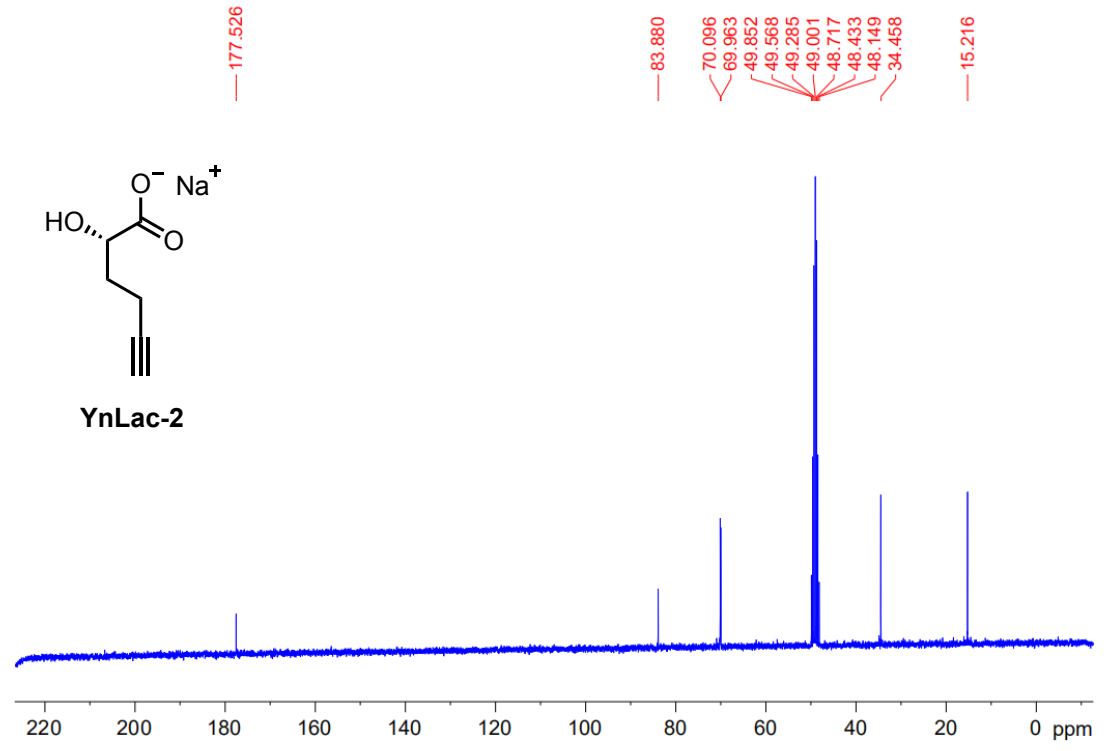
[16] D. Zhang, Z. Tang, H. Huang, G. Zhou, C. Cui, Y. Weng, W. Liu, S. Kim, S. Lee, M. Perez-Neut, J. Ding, D. Czyz, R. Hu, Z. Ye, M. He, Y. G. Zheng, H. A. Shuman, L. Dai, B. Ren, R. G. Roeder, L. Becker, Y. Zhao, *Nature* **2019**, *574*, 575-580.

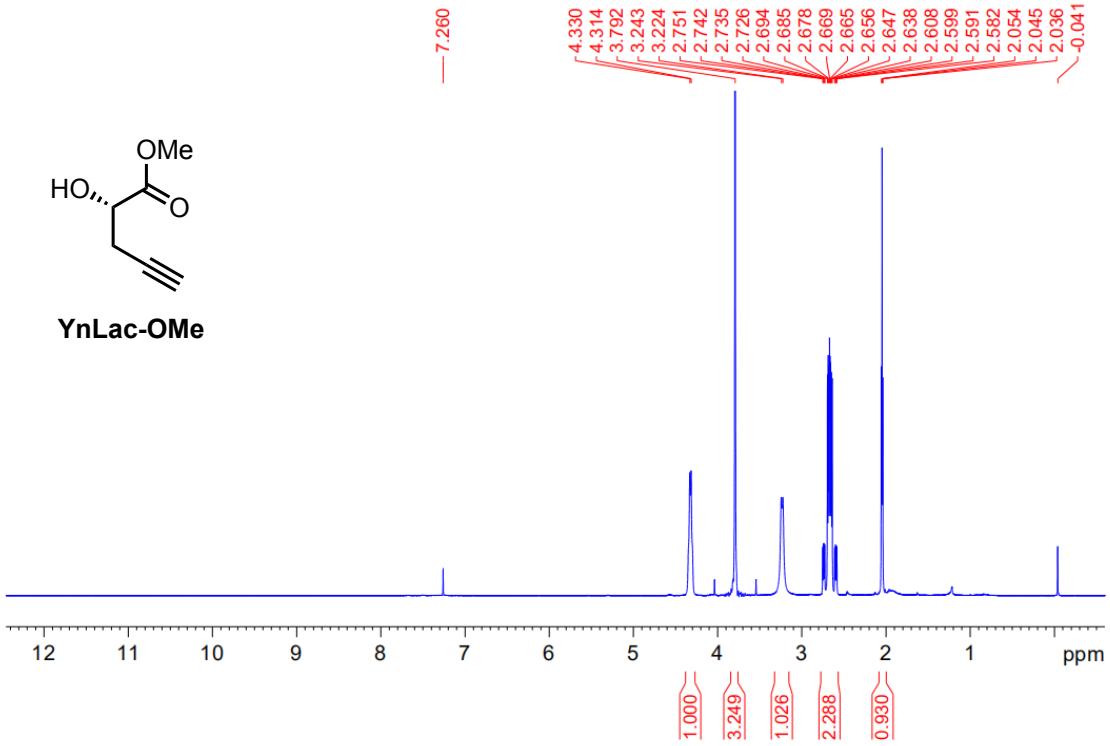
NMR spectra


4.996
4.267
4.247
4.230
3.343
3.309
3.304
3.299
2.691
2.684
2.675
2.643
2.635
2.620
2.614
2.601
2.592
2.565
2.556
2.545
2.536
2.321
2.312
2.304
O⁻ Na⁺
HO
O
YnLac
H₂O
MeOH
12 11 10 9 8 7 6 5 4 3 2 1 0 ppm
1.000
2.021
1.020

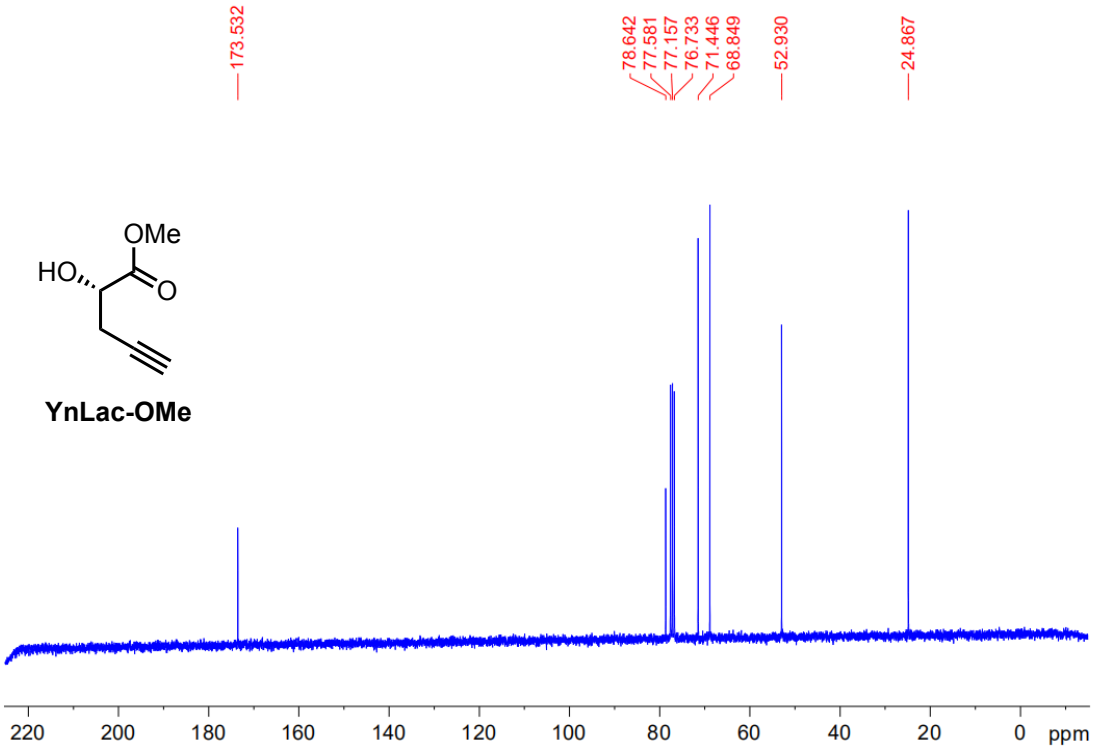

175.983
80.426
71.731
70.282
49.855
49.571
49.286
49.002
48.719
48.435
48.151
25.356
O⁻ Na⁺
HO
O
YnLac
220 200 180 160 140 120 100 80 60 40 20 ppm


4.998
4.248
4.235
4.219
4.206
3.321
3.316
3.310
3.305
3.299
2.379
2.374
2.365
2.346
2.343
2.338
2.325
2.318
2.312
2.299
2.260
2.251
2.242
2.038
2.024
2.015
2.010
1.993
1.984
1.979
1.971
1.966
1.953
O⁻ Na⁺
HO
O
YnLac-2
H₂O
Et₂O
Et₂O, EtOAc
12 11 10 9 8 7 6 5 4 3 2 1 0 ppm
0.987
2.228
0.943
1.316
1.031


177.526
83.880
70.096
69.963
49.852
49.568
49.285
49.001
48.717
48.433
48.149
34.458
15.216
O⁻ Na⁺
HO
O
YnLac-2
220 200 180 160 140 120 100 80 60 40 20 0 ppm



YnLac-OMe



YnLac-OMe