nature cell biology

Supplementary information

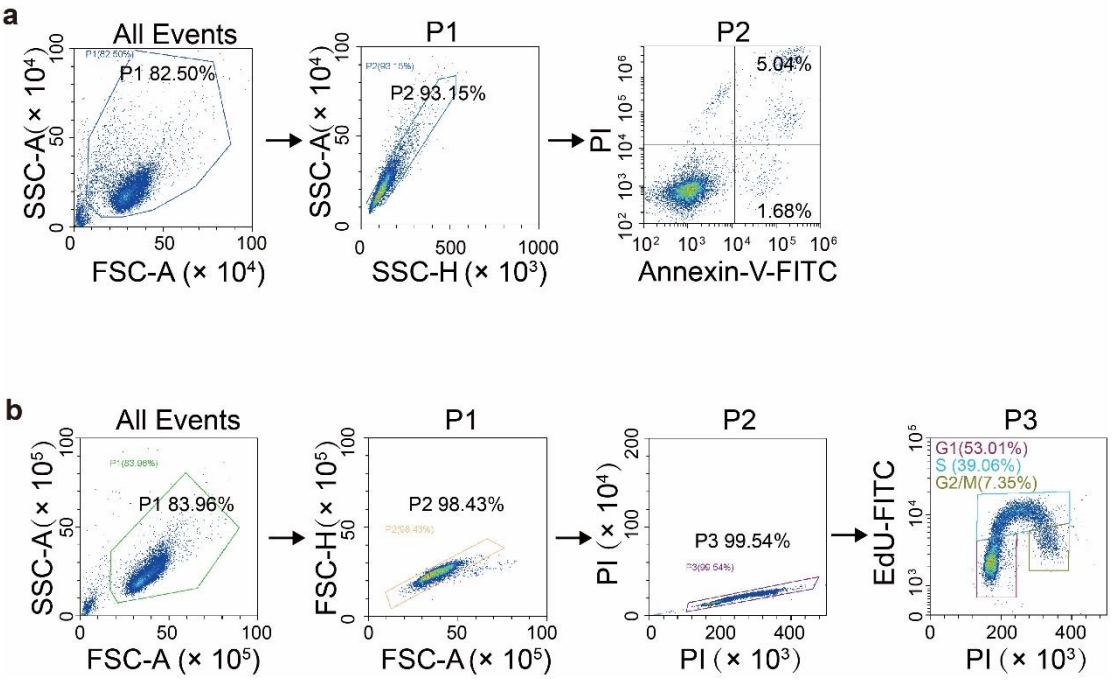
https://doi.org/10.1038/s41556-024-01548-y

NAT10-mediated mRNA N^4-acetylcytidine reprograms serine metabolism to drive leukaemogenesis and stemness in acute myeloid leukaemia

In the format provided by the authors and unedited

Supplementary Information

Supplementary Figure 1



Supplementary Figure 1 Gating strategy of flow cytometry. (a) An example showing the gating strategy for apoptosis analysis. (b) An example showing the gating strategy for EdU incorporation analysis.

Supplementary Methods

Human samples

The leukemic samples (from 3 males and 2 females) were obtained with informed consent at the time of diagnosis of AML at Sun Yat-sen University Cancer Center (SYSUCC) and Fujian Medical University Union Hospital in China. Leukemic mononuclear cells (MNCs) were purified using Ficoll (Solarbio). Normal MNCs (from 3 males) were purified from peripheral blood of healthy donors with informed consent. All human samples were collected and approved by the corresponding institutional/hospital review board. The sample size of leukemic samples was not pre-determined. The ages of these sample donors range from 30 to 70 years old.

Cell culture

U937, NB4, HL-60 and MOLM13 cells were maintained in RPMI-1640 medium (Thermo Fisher Scientific) containing 10% FBS (Thermo Fisher Scientific) and 1% penicillin-streptomycin (Biosharp). Kasumi-1 cells required 20% FBS in the same medium. For MonoMac6 cells, 1 × Non-Essential Amino Acid (Thermo Fisher Scientific), 1 mM sodium pyruvate (Thermo Fisher Scientific), 10 mM HEPES (Thermo Fisher Scientific) and 9 μg/mL insulin (Thermo Fisher Scientific) were added to the regular RPMI-1640 medium. For THP-1, 2.5g/L D-Glucose Solution (Thermo Fisher Scientific), 0.11g/L sodium pyruvate, 2.383g/L HEPES and 1 × 2-Mercaptoethanol (Thermo Fisher Scientific) were added to the regular RPMI-1640 medium. HEK293T cells were grown in DMEM medium (Thermo Fisher Scientific) supplemented with 10% FBS and 1% penicillin-streptomycin. All human cell lines were authenticated by STR profiling. All cell lines were routinely tested for mycoplasma contamination. Primary MNCs from AML patients or healthy donors were cultured in IMDM medium (Thermo Fisher Scientific) supplemented with 20% FBS and 100 ng/mL of SCF (PeproTech) and Flt-3L (PeproTech), 50 ng/mL of TPO (PeproTech), 10 ng/mL of IL-3 (PeproTech) and IL-6 (Sino Biological). MLL-ENL-ERtm-immortalized leukemic cells[1] were grown in regular RPMI-1640 supplemented with 10 ng/mL of

murine recombinant IL-3 (Sino Biological), IL-6 (PeproTech), and GM-CSF (Sino Biological) and 100 ng/mL of murine recombinant SCF (Sino Biological). The (Z)-4-Hydroxytamoxifen (4-OHT) (Sigma-Aldrich) was added at a final concentration of 100 nmol/L to induce expression of MLL-ENL. MLL-ENL-ERtm was a kind gift from Dr. Jianjun Chen (City of Hope). HEK293T (CRL-3216), U937 (CRL-1593.2), THP-1 (TIB-202), Kasumi-1 (CRL-2724), C1498 (TIB-49) and HL-60 (CCL-240) were purchased from the American Type Culture Collection (ATCC). MOLM13 (ACC-554), MonoMac6 (ACC-124), NB4 (ACC-207) were purchased from German Collection of Microorganisms and Cell Cultures (DSMZ). All of the cell lines we used have been authenticated by STR analysis.

Cell proliferation/growth and apoptosis assays

MTT (M8180, KYD Bio) or cell counting using Trypan blue or AO/PI staining was used to assess cell proliferation and growth. For apoptosis assays, Annexin V-FITC/PI apoptosis kit (100-101, GOONIE) was used following the manufacturer's manuals. Briefly, cells were harvested and washed with cold PBS, followed by staining with FITC-Annexin V and propidium iodide (PI). Apoptosis was examined by flow cytometry on a CytoFLEX LX Flow Cytometer (Beckman). CytExpert version 2.4 was used to analyze the flow cytometry data according to manufacturer's instructions.

RNA extraction and real time qPCR (RT-qPCR) analysis

Total RNA was extracted using the TRIzol reagent (Thermo Fisher Scientific) or Eastep™ Super Total RNA Extraction Kit (LS1040, Promega) according to the manufacturer's instructions. For cDNA synthesis, 200 or 500 ng of total RNA was reverse-transcribed into cDNA using HiScript III RT Super-Mix for qPCR (+gDNA wiper) (Vazyme Biotech Co.,Ltd) according to the manufacturer's instructions. Quantitative real-time PCR (qPCR) was then performed using ChamQ SYBR qPCR Master Mix (Vazyme Biotech Co.,Ltd) on the LightCycler® 480 Instrument II (Roche).

GAPDH, TBP or ACTB was used as internal control. The primers used for qPCR analysis were listed in Supplementary Table 3.

RNA stability assays

MOLM13 cells with or without NAT10 knockdown were treated with actinomycin D (A9415, Sigma-Aldrich) at a final concentration of 5 μg/mL for indicated time points and harvested. Total RNA was extracted by TRIzol reagent (Thermo Fisher Scientific) or Eastep™ Super Total RNA Extraction Kit (Promega) according to the manufacturer's instructions, then analyzed by RT-qPCR. The half-life of mRNA was calculated as previously reported[2].

Purification of poly(A) RNA, rRNA and tRNA

Polyadenylated (poly(A)) RNA was purified from total RNA through two rounds of oligo(dT) capturing using PolyATtract® mRNA Isolation Systems (Promega) and subsequently validated by Qsep Bio-Fragment Analyzer and 18S rRNA RT-qPCR.

rRNA was purified using 1.2% formaldehyde agarose gel. Briefly, 6 μg of total RNA was mixed with 3-fold volume of RNA incubation buffer (62.5% formamide, 22.5% formaldehyde, 1 × MOPS and 1 × GelRed) and denatured at 65℃ for 5 min, then chilled on ice for 2 min. After mixing with loading buffer, RNA samples were separated using 1.2% formaldehyde agarose gel in 1 × MOPS running buffer. After cutting off the bands of 28S rRNA and 18S rRNA, RNA recovery was performed using Zymoclean™ Gel RNA Recovery Kit (R1011, Zymo Research).

tRNA was purified from total RNA by Urea-PAGE electrophoresis and size selection. Briefly, 10 μg of total RNA was mixed with the same volume of 2 × RNA incubation buffer (98% formamide, 10 mM EDTA (pH 8.0) and 3.7 mM Xylene Cyanol FF) and denatured at 65℃ for 5 min. After 15 minutes of pre-electrophoresis, the RNA sample was subjected to electrophoresis in 12% urea-PAGE gel in a running buffer of 0.5 × TBE. After excising the band of appropriate size (~80 nt), crush the gel into small debris and soak them into 0.3 M NaCl solution with shaking at 4℃ overnight. Next day, collect the supernatant and precipitate RNA by ethanol. Both rRNA and tRNA were

validated by the Qsep Bio-Fragment Analyzer.

Generation of modification-free RNA (IVT control RNA)

Generation of modification-free RNA (IVT control RNA) was conducted as described previously with some modifications[3]. Briefly, cellular mRNA was purified using the PolyATtract® mRNA Isolation Systems (Promega). 100 ng mRNA was first annealed with 1 μL poly d(T) primer (12 μM). Next, the first-strand cDNA synthesis was performed by adding 2 μL 5 × first-strand buffer for SMARTScribe Reverse Transcriptase (Takara, 639536), 0.5 μL 100 mM dithiothreitol (Sigma-Aldrich), 0.5 μL RNasin® Plus RNase Inhibitor (Promega, N2615), 1 μL TSO T7 primer (10 μM), 1 μL 50 × Advantage UltraPure PCR Deoxynucleotide Mix (Takara, 639125) and 1 μL SMARTScribe Reverse Transcriptase (Takara, 639536) and incubated at 42℃ for 1.5 h and stopped by heating at 68℃ for 10 min. Template switch was performed by adding 10 μL 10× Advantage 2 PCR buffer, 1 μL Advantage 2 PCR Polymerase Mix (Takara, 639207), 2 μL 50 × Advantage UltraPure PCR Deoxynucleotide Mix (Takara, 639125), 2 μL T7 extension primer (10 μM), 2 μL RNase H (New England Biolabs, M0297) and RNase-free water to 100 μL, the reaction was conducted at 37℃ for 15 min, 95℃ for 2 min, 60℃ for 1 min and 68℃ for 10 min. Double-strand cDNA was purified using AMPure XP reagent (Beckman Coulter, A63881). The *in vitro* transcription step was performed using the HiScribe T7 High Yield RNA Synthesis kit (New England Biolabs, E2040) and NTPs from Takara Bio. TURBO DNase (Thermo Fisher Scientific, AM2238) was added and then the IVT RNA was purified using RNA Clean & Concentrator™-5 kit (Zymo Research). TSO T7 primer, 5'-ACTCTAATACGACTCACTATAGGGAGAGGGCrGrG+G-3', r indicates ribonucleic acid base, and + indicates locked nucleic acid modification; poly d(T) primer, 5'-T (30) VN-3'; T7 extension primer, 5'-GCTCTAATACGACTCACTATAGG-3'.

Data analysis of RacRIP-seq

Raw reads were trimmed by cutadapt v.4.1 to remove low quality bases and adaptor

sequences. The clean reads were then mapped to the human genome (hg38) with HISAT2 v.2.1.1[4] with "--rna-strandness RF" parameters, and supplying the GENCODE Release 39 gene annotation. A post-alignment filter was preformed to remove alignments to ribosomal RNA (rRNA), secondary alignment and non-concordant mate pairs. To inspect the reproducibility between biological replicates, we broke the genome into 10 kb continuous fragments and calculated reads abundance of each segment for each sample. Pearson's correlation between reads abundance in pairs of biological replicates was estimated. ac4C peaks were called using homemade Perl scripts according to the strand, separated by "-f 83 (and 163)" or "-f 99 (and 147) with SAMtools v.1.16.1[5]", they located, to obtain highly accurate coordinate positions of peaks. Briefly, we firstly obtained the reads coverage of each single base pair (bp) nucleotide of the respective transcript of genes both in the input and IP samples with deepTools v.3.5.2[6]. Sites with at least 10 reads in both two input samples were kept for further analysis. Coverage of each site was then normalized and the differences between input and IP samples were evaluated with DESeq2 v.1.40.2[7] based on the negative binomial distribution using the total mapped reads of each sample as the normalization factor. Sites covered 1.5-fold more reads in the IP samples than the input with adjusted *P*-value < 0.05 were considered as significant sites. Adjacent significant sites within 50 bp were further collapsed into regions, and enriched ac4C peaks were identified as regions concentrated with significantly more normalized reads (fold change > 1.5 and adjusted *P*-value < 0.05) in IP than their corresponding input samples. In order to give these ac4C peaks to a stringent filtering for artifacts such as non-specific binding, we slightly relaxed the enrichment values as fold change > 1.2 and adjusted *P*-value < 0.05 for peaks in the IVT control, then intersected ac4C peaks with peaks called in IVT, and kept only regions that had no coordinate overlap. Distribution of the ac4C peaks was illustrated with circlize v.0.4.15[8]. The candidate peaks were assigned to the nearest genes base on their summits, which defined as sites piled with most normalized read count. Peaks defined previously with a difference of log2 fold change > 0.15 or < -0.15 were considered as differential ac4C peaks. Occurrence of over-represented motifs within ac4C peaks were identified with *de novo* motif discovery algorithm by

findMotifs.pl of HOMER v.4.11.1[9]. Sequence logos presented were produced with motifStack v.4.3.1[10].

Dot blot assay

The indicated amount of total RNA was mixed with 3-fold volume of RNA incubation buffer (65.7% formamide (Dingguo), 7.77% formaldehyde (Guangzhou chemical reagent factory) and 1.33 × MOPS (Sangon biotech)) and denatured at 65℃ for 5 min, then placed on ice for 2 min and mixed with 1 volume of 20 × SSC. RNA samples were then loaded onto Hybond-N+ membranes (RPN303B, G-CLONE) with a Bio-Dot Apparatus (Bio-Rad). After UV crosslinking, the membrane was stained with 0.02% Methylene Blue trihydrate (MB) (Macklin) in 0.3M sodium acetate (Sigma-Aldrich). Membranes were then washed with Tris Buffered Saline (Biosharp) with Tween (Solarbio) buffer (TBST) and blocked with 5% non-fat milk (Beyotime) in TBST for 1 h at room temperature. After three times washing with TBST, membranes were incubated overnight with anti-ac4C antibody in 2% FBS (1:1000 dilution, Abcam) at 4℃. Membranes were next washed three times with TBST and incubated with HRP-conjugated secondary anti-rabbit IgG in 5% non-fat milk (1:10000 dilution, Proteintech) at room temperature for 1 h. Following three times washing with TBST, membranes were then detected by immunoblotting with the Femto-Sensitive ECL Solution (MK-01000, MIKX). The displayed grayscale quantitative values were the relative levels after loading control correction (MB).

Gene ontology and functional category analyses

Gene ontology (GO) analysis of designated gene set, such as ac4C peak associated genes or differentially expressed genes, was carried out with Metascape v.3.5.20230501[11]. Gene Set Enrichment Analysis (GSEA) with differentially expressed genes was performed to identify enriched pathway using c2.cgp (c2.cgp.v2023.1.Hs.symbols.gmt) as the reference gene set[12, 13]. Protein-Protein Interaction (PPI) Networks Functional Enrichment Analysis was performed by

STRING v.11.5[14] with default minimum required interaction score and illustrated with Cytoscape v.3.10.1[15]. For visualization, only representative terms with most significant enrichment were displayed, and redundant categories were discarded.

TMT labeling and quantitative proteomics

Sample preparation was performed as described with slight modifications[16]. Briefly, PBS washed cells were lysed in urea extraction buffer (8 M urea, 100 mM HEPES, pH 8.0) containing Protease Inhibitor Cocktail (Roche) and sonicated in the Bioruptor Pico (Diagenode SA), followed by centrifugation (20,000 g, 4 °C, 20 min) to collect the supernatant. Equal amounts of protein samples were reduced with 10 mM dithiothreitol for 30 min at room temperature (RT) and alkylated with 20 mM iodoacetamide (30 min at RT in the dark). Subsequently, the protein mixtures were consecutively incubated with Lys-C (enzyme/protein, 1:100, w/w) for 4 h and trypsin (enzyme/protein, 1:50, w/w) for 12 h. Equal amounts of peptides were labeled with TMTpro^{TM} 16plex reagents (Thermo Scientific) according to the manufacturer's protocols. The labeled peptides were pooled, dried, resolved with 0.1% formic acid (FA), and desalted with an Oasis HLB column (Waters). The resulting peptides were dried and dissolved in 2% Acetonitrile (ACN)/98% H2O (pH 10), followed by fractionation on an Ultimate 3000 equipped with an XBridge BEH C18 column (Waters). Ten fractions were collected and analyzed by LC-MS by combining an Easy-nLC 1200 connected online to an Orbitrap Eclipse mass spectrometer (Thermo Fisher Scientific). A 250 mm Acclaim PepMap100 C18 column (Thermo Fisher Scientific) with internal diameter of 75 μm was used to separate the peptides with mobile phase A (0.1% FA in water) and mobile phase B (0.1% FA in 80% ACN) at a 90 min gradient: 5–7% B in 1 min, 7–25% B in 72 min, 25–35% B in 10 min, 35–90% B in 4 min, and then kept B at 90% for 3 min. The flow rate was set as 300 nL/min. The Orbitrap Eclipse mass spectrometer was operated in a data-dependent acquisition mode. MS1 data were collected using the Orbitrap (120,000 resolution; AGC target “standard”, maximum injection time “auto”). The MS2 stage consisted of fragmentation by HCD (normalized collision energy 32%)

and analysis using the Orbitrap (AGC 200%, maximum injection time 120 ms, isolation window 0.7 m/z, resolution 30,000 with TurboTMT on). FAIMS voltage was set as −45 v and −65 v and the cycle time was set at 1 s per CV. Three biological replicates were performed.

The raw data were processed using Proteome Discoverer software v2.5 with Sequest HT search engine using a human database downloaded from UniProt (including 20397 entries, downloaded on 16 October 2022). Max missed cleavage sites was set to 2 and minimum peptide length to 6. Precursor Mass Tolerance was set to 10 ppm and Fragment Mass Tolerance to 0.02 Da. The percolator algorithm was used for FDR analysis. Peptides with FDR <1% were set as high confidence peptides. Normalization was set to total peptide amount and confidence to high. Differentially expressed proteins (absolute fold change > 1.2, *P*-value < 0.05) were calculated by DEP package v.1.27.0[17] using the protein-wise linear models combined with empirical Bayes statistics.

Mass spectrometry analysis for histone acetylation

For histone extraction, cells were lysed with high salt NETN buffer (20 mM Tris, pH8.0, 500 mM NaCl, 0.5% NP40, 1 mM EDTA) containing protease inhibitors, and then the pellets were lysed with 0.2 M HCl and neutralized with Tris-HCl pH8.0. The lysates were separated using SDS-PAGE and then stained with Coomassie blue. Bands were excised and subjected to identification of acetylation sites on histone protein via mass spectrometer on a nanoscale EASY-nLC 1200UHPLC system (Thermo Fisher Scientific) connected to an Orbitrap Fusion Lumos equipped with a nanoelectrospray source (Thermo Fisher Scientific). Proteome Discoverer (PD, version 2.2) was used to analyze the LC-MS/MS data.

Cleavage Under Targets & Release Using Nuclease (CUT&RUN)

CUT&RUN was performed using Hyperactive pG-MNase CUT&RUN Assay Kit (Vazyme Biotech Co.,Ltd) according to the manufacturer's instructions. Briefly, collect cells and incubate with ConA Beads at room temperature for 10 minutes. Place the tube

on the magnetic stand and discard the supernatant. Resuspend cells with antibody buffer containing antibody and incubate cells at 4℃ overnight. After discarding the supernatant, wash cells with Dig-wash Buffer twice and incubate cells with MNase Dilution Buffer containing pG-MNase Enzyme at 4℃ for 1 h. After that, perform the fragmentation and DNA extraction step and perform qPCR using specific primers. The primers used for qPCR analysis were listed in Supplementary Table 3. The antibodies used for CUT&RUN were listed in Supplementary Table 4.

Western blot and protein dot blot assay

Cells were collected and washed twice with PBS, followed by direct lysed in 1 × SDS buffer at 95℃ for 10 min. Equal amounts of cell lysates were separated by SDS-PAGE and transferred onto PVDF membranes (Merck Millipore). Protein dot blot was performed by loading lysates onto PVDF membranes with a Bio-Dot Apparatus (Bio-Rad). Membranes were blocked with 5% non-fat milk (Beyotime) in TBST for 1h at room temperature followed by three times washing with TBST. Membranes were then incubated overnight with primary antibodies at 4℃ and washed three times with TBST. After incubated with HRP-conjugated secondary antibodies (1:10000 dilution, Proteintech) at room temperature for 1 h, membranes were washed three times with TBST and detected by immunoblotting with the Femto-Sensitive ECL Solution (MK-01000, MIKX). Signal from western blots and dot blots were quantified using Tanon Image software. The displayed grayscale quantitative values were the relative levels after internal reference correction, while for DARTS, CETSA, and IP assays, the relative grayscale levels of each band were displayed. The antibodies used for western blot and protein dot blot assays were listed in Supplementary Table 4.

Cellular thermal shift assay (CETSA)

CETSA was performed as previously described with some modifications[18]. Briefly, 1 × 10^7 cells were collected from MonoMac6 pretreated with 500 μM Remodelin (S7641, Selleck), 500 μM Fludarabine (Selleck) or DMSO (Sigma-Aldrich) for 2 h. Equal amount cells were collected and washed twice with PBS, and then resuspended in 0.7

mL PBS supplemented with protease inhibitor cocktail (Roche). Samples were aliquoted to 12 different PCR tubes with 50 μL in each tube. Samples were then heated for 3 min using a pre-heated thermal cycler at indicated temperatures, and immediately allowed to equilibrate to room temperature. Finally, the samples were subjected to freeze-thaw cycles with dry ice and room temperature for 3-5 times to lyse cells, and centrifuged at 20,000 ×g for 20 min at 4℃. The supernatant was boiled with 5 × SDS loading buffer for western blot analysis. The band intensities of the target protein were quantified using Tanon Image software and plotted.

Drug affinity responsive target stability (DARTS) assays

DARTS assays were performed to evaluate the direct binding of NAT10 protein with compounds according to a published protocol[19] with modifications. Briefly, 1×10^7 cells were collected and washed with cold PBS, then lysed in 200 μL M-PER buffer (78501, Thermo Fisher Scientific) supplemented with protease inhibitor cocktail (Roche) and phosphatase inhibitor cocktail (Beyotime) for 10 min on ice. After centrifuging at 18,000 ×g for 10 min at 4℃, the supernatant was transferred into a new tube and mixed with 10 × TNC buffer (500 mM Tris-HCl, pH 8.0, 500 mM NaCl, 100 mM $CaCl_2$). BCA assays (Biosharp) were used to determine protein concentration. 300 μg of the lysate was incubated with Remodelin (S7641, Selleck), Fludarabine Phosphate (P-F-ara-A, MedChem Express) or Fludarabine triphosphate (F-ara-ATP, MedChem Express) from 0 to 200 μM for 1-2 h at room temperature. Pronase (Roche, 1:400-1:500) digestion was then conducted for 30 min at room temperature. Samples were then immediately placed on ice. SDS loading buffer was added and samples were subjected to western blot analysis. ACTB (Beta Actin, Proteintech) and METTL3 (Proteintech) were used as a negative control.

Molecular docking

Since the protein structure of human NAT10 is not available, a predicted structure of NAT10, obtained by superimposing the full-length structure of NAT10 predicted by AlphaFold2[20] to the structure of NAT10 active sites predicted through homology

modeling by Dalhat et al.[21] was generated and used for molecular docking. The Protein Preparation Wizard (PrepWizard) module of Schrödinger Maestro 2022[22] was applied to prepare the protein structure, which was then energy-minimized using the OPLS3 force field. The Glide module of Schrödinger was employed to generate grid file and undertake docking procedure. After protein preparation and minimization, Fludarabine phosphate and 3P-Fludarabine (prepared by the LigPrep module of Schrödinger) were docked into the binding pocket of NAT10. Extra precision was used for the docking methods. Default settings were employed for other parameters for the grid generation and docking. Each top 1 pose of docking conformations was used for subsequent interaction analysis.

EdU incorporation assay

EdU (5-ethynyl-2′-deoxyuridine) incorporation assays were performed using the Click-iT™ Plus EdU Alexa Fluor™ 488 Kit (C10632, ThermoFisher Scientific). Briefly, cells were treated with 10 μM EdU for 2 hours before harvesting. Cells were then fixed and permeablized, and the Click-iT reaction was performed with 30 min incubation at room temperature. PI was then added to stain DNA with the addition of RNase A (20 μg/ml). Cells were analyzed on a CytoFLEX LX Flow Cytometer (Beckman).

References

1. Zeisig, B.B. *et al.* Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. *Mol Cell Biol* **24**, 617-628 (2004).
2. Weng, H. *et al.* METTL14 Inhibits Hematopoietic Stem/Progenitor Differentiation and Promotes Leukemogenesis via mRNA m(6)A Modification. *Cell Stem Cell* **22**, 191-205 e199 (2018).
3. Zhang, Z. *et al.* Systematic calibration of epitranscriptomic maps using a synthetic modification-free RNA library. *Nat Methods* **18**, 1213-1222 (2021).
4. Kim, D., Langmead, B. & Salzberg, S.L. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**, 357-360 (2015).
5. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
6. Ramirez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165 (2016).
7. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
8. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize Implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811-2812 (2014).
9. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589 (2010).
10. Ou, J., Wolfe, S.A., Brodsky, M.H. & Zhu, L.J. motifStack for the analysis of transcription factor binding site evolution. *Nat Methods* **15**, 8-9 (2018).
11. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523 (2019).
12. Mootha, V.K. *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* **34**, 267-273 (2003).
13. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).
14. Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **47**, D607-D613 (2019).
15. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504 (2003).
16. Li, J. *et al.* TMTpro reagents: a set of isobaric labeling mass tags enables simultaneous proteome-wide measurements across 16 samples. *Nat Methods* **17**, 399-404 (2020).
17. Zhang, X. *et al.* Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nat Protoc* **13**, 530-550 (2018).
18. Jafari, R. *et al.* The cellular thermal shift assay for evaluating drug target interactions in cells. *Nat Protoc* **9**, 2100-2122 (2014).
19. Lomenick, B. *et al.* Target identification using drug affinity responsive target stability (DARTS). *Proc Natl Acad Sci U S A* **106**, 21984-21989 (2009).
20. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**,

583-589 (2021).

21. Dalhat, M.H., Altayb, H.N., Khan, M.I. & Choudhry, H. Structural insights of human N-acetyltransferase 10 and identification of its potential novel inhibitors. *Sci Rep* **11**, 6051 (2021).

22. Sastry, G.M., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des* **27**, 221-234 (2013).