ENL links histone acetylation to oncogenic gene expression in AML

HHS Public Access

Author manuscript

Nature. Author manuscript; available in PMC 2017 September 01.

Published in final edited form as:

Nature. 2017 March 09; 543(7644): 265–269. doi:10.1038/nature21687.

ENL links histone acetylation to oncogenic gene expression in AML

Liling Wan[1,2,*], **Hong Wen**[3,4,*], **Yuanyuan Li**[5,6,*], **Jie Lyu**[7], **Yuanxin Xi**[7], **Takayuki Hoshii**[2], **Julia Joseph**[1], **Xiaolu Wang**[3], **Yong-Hwee E. Loh**[8], **Michael A. Erb**[9], **Amanda L. Souza**[9,12], **James E. Bradner**[9,10,12], **Li Shen**[8], **Wei Li**[7], **Haitao Li**[5,6,**], **C. David Allis**[1,**,#], **Scott A. Armstrong**[2,**,#], and **Xiaobing Shi**[3,4,11,**,#]

[1]Laboratory of Chromatin Biology & Epigenetics, The Rockefeller University, New York, NY 10065, USA

[2]Cancer Biology and Genetics Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, and Department of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02215

[3]Department of Epigenetics and Molecular Carcinogenesis

[4]Center for Cancer Epigenetics, The University of Texas MD Anderson Cancer Center, Houston, TX, 77030, USA

[5]Beijing Advanced Innovation Center for Structural Biology, MOE Key Laboratory of Protein Sciences, Department of Basic Medical Sciences, School of Medicine, Tsinghua University, Beijing 100084, China

[6]Tsinghua-Peking Joint Center for Life Sciences, Tsinghua University, Beijing 100084, China

[7]Dan L. Duncan Cancer Center, Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas 77030, USA

[#]To whom correspondence should be addressed: C. David Allis: The Rockefeller University, Allis Lab, Box #78, 1230 York Avenue, New York, NY 10065, USA.; alliscd@rockefeller.edu; Scott A. Armstrong: Department of Pediatric Oncology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215-5450, USA.; Scott_Armstrong@dfci.harvard.edu; Xiaobing Shi: Department of Epigenetics and Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX, 77030, USA.; xbshi@mdanderson.org.

[*]co-first author

[**]co-senior author

[12]Present address: Novartis Institutes for BioMedical Research, Cambridge, MA 02139.

Supplementary Information is attached. The file contains the following Supplementary Tables: Differentially expressed genes, Gene ontology analysis, ChIP-seq peaks, ChIP-seq occupied genes, and lists of shRNA/sgRNA sequences, oligos, antibodies and GSEA gene sets used in this study.

Author Contributions. L.W., H.W., Y.L., H.L., C.D.A., S.A.A. and X.S. designed the study, analyzed the data and wrote the paper. L.W. and H.W. planned and performed all the molecular, cellular and genomic studies; Y.L. and H.L. performed structural and calorimetric studies; L.W., T.H., M.A.E., A.L.S, J.E.B performed mouse xenograft studies; L.W., J.L., Y.X., Y.L., L.S. and W.L. performed bioinformatics analysis; J.J. and X.W. provided technical assistance; H.L, C.D.A, S.A.A. and X.S. supervised the research.

The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. C.D.A. is a co-founder of Chroma Therapeutics and Constellation Pharmaceuticals. C.D.A. and X.S. are Scientific Advisory Board members of EpiCypher.

[8]Fishberg Department of Neuroscience and Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA

[9]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02215

[10]Department of Medicine, Harvard Medical School, Boston, MA 02115

[11]Genes and Development and Epigenetics & Molecular Carcinogenesis Graduate Programs, The University of Texas Graduate School of Biomedical Sciences, Houston, TX 77030, USA

Abstract

Cancer cells are characterized by aberrant epigenetic landscapes and often exploit chromatin machinery to activate oncogenic gene expression programs[1]. Recognition of modified histones by "reader" proteins constitutes a key mechanism underlying these processes; therefore, targeting such pathways holds clinical promise, as exemplified by the development of BET bromodomain inhibitors[2, 3]. We recently identified the YEATS domain as a novel acetyllysine-binding module[4], yet its functional importance in human cancer remains unknown. Here we show that the YEATS domain-containing protein ENL, but not its paralog AF9, is required for disease maintenance in acute myeloid leukaemia (AML). CRISPR-Cas9 mediated depletion of ENL led to anti-leukemic effects, including increased terminal myeloid differentiation and suppression of leukaemia growth *in vitro* and *in vivo*. Biochemical and crystal structural studies and ChIP-seq analyses revealed that ENL binds to acetylated histone H3, and colocalizes with H3K27ac and H3K9ac on the promoters of actively transcribed genes that are essential for leukaemias. Disrupting the interaction between the YEATS domain and histone acetylation via structure-based mutagenesis reduced RNA polymerase II recruitment to ENL target genes, leading to suppression of oncogenic gene expression programs. Importantly, disruption of ENL's functionality further sensitized leukaemia cells to BET inhibitors. Together, our study identifies ENL as a histone acetylation reader that regulates oncogenic transcriptional programs in AML and suggests that displacement of ENL from chromatin may be a promising epigenetic therapy alone or in combination with BET inhibitors for AML.

The YEATS domain-containing proteins AF9 and ENL are frequently fused with the MLL (mixed lineage leukaemia) protein as a result of *MLL* chromosomal translocations. The resultant fusion proteins are oncogenic drivers in acute myeloid leukaemia (AML) and acute lymphoid leukaemia (ALL)[5]. To assess the functional importance of the wild-type allele of *AF9* and *ENL* in *MLL*-rearranged leukaemias, we employed a CRISPR-Cas9-mediated gene editing and negative selection strategy (Fig. 1a). Cells expressing *ENL* sgRNAs were outcompeted by non-transduced cells in various *MLL*-rearranged leukaemia cells, whereas sgRNAs targeting *AF9* had little effect on cell growth (Fig. 1a and Extended Data Fig. 1a–d). Using a clonal MOLM-13 cell line in which Cas9 expression can be induced with doxycycline (Dox) administration (iCas9-MOLM-13), we observed a consistent effect of *ENL* sgRNAs on cell growth after Dox treatment (Extended Data Fig. 1e–g). CRISPR-Cas9 or shRNA-mediated depletion of ENL also impaired the clonogenic potential of MOLM-13 (Fig. 1b) and MV4;11 (Extended Data Fig. 1h, i) cells grown in cytokine-supplemented methylcellulose. ENL depletion increased surface expression of Integrin alpha M (ITGAM or CD11b), a myeloid differentiation marker (Fig. 1c) and induced a morphological change

to a macrophage-like appearance (Extended Data Fig. 1j). Importantly, *ENL* sgRNA-induced growth inhibition and differentiation were rescued by the presence of a mouse *Enl* cDNA that contains several mismatches with the human *ENL* sgRNA (Fig. 1d and Extended Data Fig. 1k). In addition to *MLL*-rearranged leukaemias, *ENL* depletion also suppressed the growth of non MLL-rearranged leukaemia cell lines such as U-937 and K562, but not the human cervical adenocarcinoma HeLa cells and human osteosarcoma U2OS cells (Extended Data Fig. 1l–o). *ENL* silencing also had minimal effect on the growth, colony-forming ability or differentiation potential of sorted lineage$^-$Sca-1$^+$c-Kit$^+$ (LSK) cells that are enriched for hematopoietic stem cells (Extended Data Fig. 1p–r).

To examine the relevance of ENL and AF9 to leukaemia progression *in vivo*, MOLM-13 cells were transduced with different sgRNAs and transplanted into immune-deficient recipient mice. The expression of *ENL* sgRNA resulted in a marked delay in leukaemia progression (Fig. 1e and Extended Data Fig. 2a) and prolonged survival of the recipient mice (Fig. 1f). Although the terminal leukaemia in the *ENL*-sgRNA group was composed of predominantly sgRNA-positive cells, these leukaemia cells had similar protein levels of ENL compared to those from the control group (Extended Data Fig. 2b, c), indicating that the mice succumbed to an outgrowth of cells that escaped ENL depletion. Together, these results suggest a critical requirement of ENL, but not AF9, for AML maintenance *in vitro* and *in vivo*.

To identify the transcriptional pathways controlled by ENL, we performed RNA-seq analyses on iCas9-MOLM-13 cells expressing different sgRNAs. While *AF9* sgRNA caused minimal changes in gene expression, *ENL* sgRNA induction led to differential expression of a subset of genes, but not to global transcriptional dysregulation (Fig. 2a and Extended Data Fig. 3a). The transcriptional programs altered by the two *ENL* sgRNAs were highly correlated (Extended Data Fig. 3b). Gene Set Enrichment Analyses (GSEA) revealed a marked upregulation of myeloid lineage differentiation signature along with a downregulation of a leukaemia stem cell (LSC) and MYC-associated gene expression signature[6] in *ENL* sgRNA-expressing MOLM-13 and MV4;11 cells (Fig. 2b and Extended Data Fig. 3c–h), suggesting that ENL is required to sustain the oncogenic gene expression programs critical for leukaemia maintenance.

ENL resides in large protein complexes termed the super elongation complex (SEC)[7], elongation assisting proteins (EAP)[8] or AF4/ENL/P-TEFb complex (AEP)[9] that contain overlapping subunits including AFF1/4, ELL, EAF1/2, ENL/AF9 and P-TEFb (for simplicity, we use the term SEC hereafter). To determine the genomic distribution of ENL, we generated stable MOLM-13 and MV4;11 cells ectopically expressing Flag-ENL at levels equivalent to the endogenous ENL proteins, and performed ChIP-seq experiments. ENL bound to a larger cohort of genes than that of MLL-fusion proteins in both cell lines, suggesting a broader role of ENL than MLL-fusion proteins in transcriptional regulation (Extended Data Fig. 4a, b). There was a greater overlap of ENL-bound genes between the leukaemia cell lines than their overlaps with ENL occupancy in the non-leukaemia HeLa cells (Extended Data Fig. 4c). Notably, the ENL-bound genes specific in leukaemia cell lines were significantly enriched in pathways implicated in cancer and hematological disease (Extended Data Fig. 4d).

ENL ChIP-seq peaks showed a strong enrichment at promoter regions (TSS±3kb) (Fig. 2c and Extended Data Fig. 4e), which overlapped with the promoter-proximal RNA polymerase II (Pol II) peak (Fig. 2d and Extended Data Fig. 4f). ENL-bound genes showed a much higher Pol II occupancy than non-ENL genes (Fig. 2e). Upon ENL depletion, Pol II ChIP-seq signals decreased more prominently on ENL-occupied genes than on non-ENL bound genes ($P < 0.0001$), including two key transcription factors in AML *MYC* and H2.0-like homeobox (*HLX*)[10] (Fig. 2e and Extended Data Fig. 4g, h). The reduction of Pol II occupancy spanned from the promoter proximal region to the 3′ end of gene body, suggesting reduced Pol II activities in both transcriptional initiation and elongation. Indeed, ChIP-seq of the elongation-specific Pol II that is phosphorylated at serine 2 (S2P)[11] showed a markedly decreased occupancy of the elongating Pol II on ENL target genes in ENL-depleted cells (Fig. 2f and Extended Data Fig. 4i). In addition, ENL depletion also resulted in decrease in the occupancy of CDK9, a SEC component[12] that phosphorylates Pol II at the serine 2 site[13], and the DOT1L-mediated H3K79me2 and H3K79me3 on ENL-target genes (Fig. 2g and Extended Data Fig. 4j, k). Together, these data suggest that ENL regulates gene expression by modulating the recruitment of Pol II likely through both the SEC and DOT1L complexes.

In light of the AF9 YEATS domain being a reader for histone acetylation[4], we hypothesized that ENL might link the Pol II transcriptional machinery to chromatin through its YEATS domain. To test this hypothesis, we performed histone peptide array screening and found that indeed the ENL YEATS domain bound specifically to a subset of acetylated histone H3 peptides (Extended Data Fig. 5a). Preferential binding to the H3K27ac, H3K9ac and H3K18ac peptides was further confirmed by peptide pull-down (Fig. 3a) and quantitative isothermal titration calorimetry (ITC) assays (Fig. 3b) with a binding K_D of 30.5, 32.2 and 50.0 μM for H3K27ac, H3K9ac and H3K18ac, respectively. To explore the underlying molecular basis of this binding specificity, we solved the co-crystal structure of human ENL YEATS bound to H3K27ac peptide at 2.7 Å (Extended Data Table 1). The ENL YEATS domain adopts an eight-stranded β-sandwich fold (Fig. 3c) with the H3K27ac peptide snugly attached to an acidic surface at the top (Fig. 3d) and the flat acetylamide group of K27ac sandwiched by aromatic residues F59 and Y78 (Fig. 3e). Recognition of the H3K27ac peptide is further stabilized by a network of hydrogen bonds in addition to hydrophobic contacts (Extended Data Fig. 5b). Acetylated histone H3 at K9, K18 and K27 share a common “R-Kac” recognition signature, in which the “−1” arginine forms a charge-stabilized hydrogen bond with D103 of ENL (Fig. 3e). Structure-based mutagenesis followed by ITC titrations and peptide pull-down assays revealed a 2.3- to > 17-fold affinity drop or binding elimination upon alanine mutation of the “R-Kac”-binding residues (Fig. 3f, g).

ChIP-seq analyses of both MOLM-13 and MV4;11 cells revealed that ENL-bound peaks were associated with regions of H3K9ac or H3K27ac modifications (Fig. 3h and Extended Data Fig. 6a). Consistently, the average distribution of ENL correlated extremely well with that of H3K9ac and H3K27ac in the region immediately downstream of the TSS (Fig. 3i and Extended Data Fig. 6b, c). Importantly, compared to wild-type (WT) ENL, the YEATS domain mutants deficient in acetyl-binding exhibited markedly reduced occupancy at ENL-bound genes (Fig. 3j and Extended Data Fig. 6d–f), whereas their interactions with other

SEC components were not affected (Extended Data Fig. 6g, h). The ENL YEATS domain is also required for Pol II recruitment, as cells expressing the YEATS mutants exhibited reduced Pol II occupancy on ENL target genes such as *MYC* (Extended Data Fig. 6i). Together, these findings revealed an important "reader" function of the YEATS domain in mediating ENL's chromatin localization and the recruitment of the Pol II transcriptional machinery.

To test whether the YEATS domain is required for ENL-dependent maintenance of oncogenic gene expression in leukaemias, we performed rescue experiments with ectopically expressed murine WT or mutant ENL proteins in cells expressing a human *ENL* sgRNA. Transcriptional profiling analyses revealed that WT ENL, but not the F59A or Y78A mutant, restored the transcriptional changes caused by the *ENL* sgRNA (Fig. 4a). GSEA analyses further demonstrated the incompetency of these ENL mutants to revert the effect of ENL depletion on the expression of genes involved in critical cellular processes (Fig. 4b and Extended Data Fig. 7a). Consequently, the mutants were unable to rescue the growth and differentiation defects caused by ENL inhibition (Fig. 4c, d and Extended Data Fig. 7b). In line with these observations, replacement of ENL YEATS domain with that of AF9 retained the full functionality of ENL (Extended Data Fig. 7c, d). Finally, mice transplanted with *ENL* sgRNA-transduced cells expressing mutant ENL proteins exhibited a marked decrease in leukaemia burden and prolonged survival as compared to the WT counterpart (Fig. 4e and Extended Data Fig. 7e). Taken together, these findings demonstrate a critical role of the ENL YEATS-histone acetylation interaction in regulating oncogenic gene expression and leukaemia maintenance.

Strategies to disrupt the interaction between acetylated histones and their reader proteins, such as the BET family of bromodomain-containing proteins, have been shown to be effective in clinical trials for treatment of a variety of cancers, including hematologic malignancies[14]. Interestingly, similar to ENL, BRD4 also associates with P-TEFb and modulates the Pol II transcriptional machinery to regulate key oncogenes such as *MYC*[15]. This finding raised the intriguing possibility that ENL and BRD4 might cooperate to sustain optimal oncogenic gene expression programs critical for leukaemias. In line with this speculation, depletion of ENL or other SEC components sensitized MOLM-13 cells to a BET bromodomain inhibitor JQ1[2] (Fig. 4f and Extended Data Fig. 8a). Importantly, point mutations disrupting the interaction between the ENL YEATS domain and histone acetylation were sufficient to confer cells with greater sensitivity to JQ1 (Fig. 4f). Increased JQ1 sensitivity by *ENL* silencing was also observed in a panel of leukaemia cell lines, but not in HeLa or U2OS cells (Extended Data Fig. 8b–d). Furthermore, ENL depletion coupled with JQ1 treatment resulted in a significant survival advantage compared with single treatments *in vivo* (Fig. 4g, h and Extended Data Fig. 8e). Mechanistically, ENL depletion markedly enhanced JQ1-induced transcriptional changes in MOLM-13 cells, including *MYC* (Fig. 4i, j and Extended Data Fig. 8f–i), inhibition of which has been shown to critically contribute to the sensitivity to BET inhibitors[16,17]. Together, these findings suggest that inhibition of ENL's functionality is a potential therapeutic strategy to synergize with BET inhibitors.

Cancer cells often co-opt chromatin regulatory pathways and the general transcriptional machinery to sustain the oncogenic state; hence chromatin and transcriptional regulators are being explored as promising candidate drug targets. Targeting epigenetic readers represents a new class of anti-cancer therapy that holds clinical promise. Our study reveals ENL as a chromatin reader that regulates oncogenic programs through linking histone acetylation and the transcription apparatus in AML, thus establishing the ENL YEATS domain as a potential drug target, either alone or in combination with BET inhibitors. Interestingly, recurrent mutations in the ENL YEATS domain were recently reported in Wilms Tumors[18], suggesting potentially critical and broad roles of the "reader" function of ENL in other cancer types that warrant future investigations.

Methods

Plasmids

Murine wild-type or mutant *Enl* were cloned into pMSCV-IRES-GFP. Human ENL, DOT1L, AFF4, ELL2 and CDK9 cDNA in pENTR or pDONR were purchased from Open Biosystems. The coding sequences in pENTR were subsequently cloned into p3FLAG, pCDH-Flag, or pCAG-Myc destination vectors using Gateway techniques (Invitrogen). Constitutive (pLKO5d.EFS.SpCas9.P2A.BSD, Addgene #57821, a gift from Benjamin Ebert) and inducible (pCW-CAS9, Addgene #50661, a gift from Eric Lander & David Sabatini) Cas9 vectors were used to generate Cas9-expressing cells. All shRNAs targeting human and murine *ENL* were obtained from Sigma. sgRNAs were cloned into pLKO5.sgRNA.EFS.tRFP657 vector (Addgene # 57824, a gift from Benjamin Ebert). All shRNA and sgRNA sequence used are provided in Supplementary Information: Table S11. Oligo.

Cell culture, virus transduction and functional assays

Human leukaemia cell lines MOLM-13, MV4;11, ML-2, SEMK2, U-937, K562 cells were maintained in RPMI 10% FBS supplemented with 2mM L-glutamine and 100 U ml^{-1} penicillin/streptomycin. HeLa and U2OS cell lines were maintained in DMEM 10% FBS supplemented with 2mM L-glutamine and 100 U ml^{-1} penicillin/streptomycin. $Lin^{-}Sca1^{+}cKit^{+}$ (LSK) cells were freshly sorted from the bone marrow of 6–8 week-old C57BL/6J male mice and cultured in IMDM plus 15% FBS supplemented with 20 ng/ml mouse SCF (PeproTech), 10 ng/ml mouse IL-3 (PeproTech) and 10 ng/ml mouse IL-6 (PeproTech). All cell culture medium contained L-Glutamine (2mM; Gibco), penicillin (100 units/ml; Gibco), streptomycin (100 μg/ml; Gibco) and plasmocin (5 μg/ml; InvivoGen). All human cell lines were mycoplasma-negative and were tested for authentication by short tandem repeat (STR) profiling performed by ATCC or by the MDACC CCSG-funded Characterized Cell Line Core, NCI # CA016672.

Lentivirus and retrovirus packaging were performed in HEK293T and Platinum-E cells (Cell Biolabs), respectively, using Lipofectamine® 2000 reagent (Invitrogen) in accordance with the manufacturer's instructions. Medium containing virus was concentrated using PEG-it Virus Precipitation Solution (System Biosciences). Spin infection was performed at 1500 rcf

at 35°C for 90 mins and transduced cell populations were usually selected or sorted 48 h post infection.

Competitive proliferation assays using sgRNAs: for constitutive Cas9 expression, cells were analyzed for RFP^+ expression 3 days post infection. For Dox-regulated Cas9 expression cells, cells were analyzed for RFP^+ expression before and after Dox treatment (1 μg ml^{-1}). The percentage of sgRNA-expressing cells (RFP^+) was measured over time using flow cytometry and normalized to the starting time point.

Colony-forming assays were performed in Methocult (StemCell Technologies, No. H4435 for human leukaemia cells and Cat. No. M3234 supplemented with 20 ng/ml mouse SCF, 10 ng/ml mouse IL-3 and 10 ng/ml mouse IL-6 for murine LSKs). Indicated number of cells was resuspended in 100ul PBS and added to 900 μl Methocult. After 7 days different types of colonies were counted. Cell viability assays for JQ1 treatment were performed using CellTiter-Glo® Luminescent Cell Viability Assay (promega/VWR) according to manufacturer's guidelines.

For flow cytometry immunophenotyping, cultured or freshly dissociated cells were collected and stained using the indicated antibodies. Stained samples were analyzed on an LSR Fortessa (BD) flow cytometer. Data analysis was performed using FlowJo software.

Antibodies

All antibodies used are provided in Supplementary Information: Table S12. Antibody.

qRT-PCR analyses

Total RNA was isolated using the RNeasy kit (Qiagen) and reverse transcribed with the high capacity cDNA reverse transcription kit (Applied Biosystems) in accordance with the manufacturer's instructions. Quantitative PCR was performed using the SYBR Green PCR Master Mix (Applied Biosystems) with the StepOne™ System (Applied Biosystems). qPCR primers used are provided in Supplementary Information: Table S11. Oligo.

RNA sequencing

RNA was isolated from human leukaemia cells using the RNeasy Mini Kit (Qiagen, Gaithersburg, MD) and then samples were prepared as instructed using the TruSeq RNA Sample Preparation Kit v2 (Illumina, San Diego, CA) in accordance with the manufacturer's instructions. RNA-seq samples were sequenced using Illumina NextSeq 500. Raw reads were mapped to human reference genome (hg19) and transcriptome using the TopHat2 package[19]. Transcript abundances were quantified using htseq-count v0.6.0[20]. Differentially expressed genes were determined using DESeq2[21] Genes with a fold change of 1.5 were selected as differential genes. Gene ontology analysis used the Panther[22] Classification System (http://www.pantherdb.org).

Gene set enrichment analysis

Gene set enrichment analysis[23] was performed using GSEA v2.2.2 software with 1,000 gene set permutations. Gene sets used were obtained from the Molecular Signatures Database

v4.0 (MSigDB, http://www.broadinstitute.org/gsea/msigdb/index.jsp, C2 curated gene sets or C6 oncogenic signatures) or were manually curated from published dataset. A detailed description of GSEA methodology and interpretation can be found at http://www.broadinstitute.org/gsea/doc/GSEAUserGuideFrame.html. Normalized RNA-seq data were rank-ordered by expression fold change. Enrichment cores are defined as the members of the gene set that lie before or at the running sum peak (i. e. the enrichment score) of the ranked gene list. The normalized enrichment score (NES) provides "the degree to which a gene set is overrepresented at the top or bottom of a ranked list of genes". The nominal *P* value (Nom p-val) describes "the statistical significance of the enrichment score". The false discovery rate *q* value (FDR *q*-val) is "the estimated probability that a gene set with a given NES represents a false positive finding". All gene sets used in this study are provided in Supplementary information: Table S13. GSEA gene sets.

ChIP and ChIP-seq analysis

ChIP analysis was performed essentially as described previously[4]. Briefly, cells were cross-linked with 1% formaldehyde for 10 min and stopped with 125 mM glycine. The isolated nuclei were resuspended in nuclei lysis buffer and sonicated using a Bioruptor Sonicator (Diagenode). The samples were immunoprecipitated with 2–4 μg of the appropriate antibodies overnight at 4°C. Protein A/G beads (Millipore) were added and incubated for 15 min, and the immunoprecipitates were washed twice, each with low salt, high salt and LiCl buffer. Eluted DNA was reverse-crosslinked, purified using PCR purification kit (Qiagen), and analyzed by quantitative real-time PCR on the ABI 7500-FAST System using the Power SYBR Green PCR Master Mix (Applied Biosystems).

ChIP-seq samples were sequenced using the Illumina Solexa Hiseq 2000 or NextSeq 500. The raw reads were mapped to human reference genome NCBI 36 (hg19) by Solexa data processing pipeline, allowing up to 2 mismatches. The genome ChIP-seq profiles were generated using MACS 1.3.6 with only unique mapped reads[24]. Clonal reads were automatically removed by MACS. The ChIP-seq profiles were normalized to 20,000,000 total tag numbers, and peaks were called at p-values $\leq 1\times10^{-8}$. Peaks within 1kbp neighborhood were merged together across multiple samples in the peak overlap analysis.

Peptide microarray and peptide pull-down assay

Peptide microarray and peptide pull-down assays were performed as described previously[25]. Briefly, biotinylated histone peptides were printed in triplicate onto a streptavidin-coated slide (PolyAn) using a VersArray Compact Microarrayer (Bio-Rad). After a short blocking with biotin (Sigma), the slides were incubated with the GST-ENL YEATS domain in binding buffer (50 mM Tris-HCl 7.5, 250 mM NaCl, 0.1% NP-40, 1 mM PMSF, 20% fetal bovine serum) overnight at 4°C with gentle agitation. After being washed with the same buffer, the slides were probed with an anti-GST primary antibody and then a fluorescein-conjugated secondary antibody and visualized using a GenePix 4000 scanner (Molecular Devices). For the peptide pull-down assays, 1 μg of biotinylated histone peptides with different modifications were incubated with 1–2 μg of GST-fused proteins in binding buffer (50 mM Tris-HCl 7.5, 300 mM NaCl, 0.1% NP-40, 1 mM PMSF) overnight. Streptavidin beads (Amersham) were added to the mixture, and the mixture was incubated for 1 hr with

rotation. The beads were then washed three times and analyzed using SDS-PAGE and Western blotting.

Co-immunoprecipitation (Co-IP)

Co-IP was performed essentially as described previously[4]. Cells were lysed in cell lysis buffer containing 50 mM Tris-HCl pH 7.4, 250 mM NaCl, 0.5% Triton X-100, 10% glycerol, 1 mM DTT, and a complete protease inhibitor tablet (Roche). Antibodies conjugated with protein A/G beads (Millipore) or anti-FLAG M2-conjugated agarose beads (Sigma) were incubated with the lysates overnight at 4°C. The beads were then washed 3–6 times with cell lysis buffer, and the bound proteins were eluted in SDS buffer and analyzed by Western blotting.

Crystallographic studies

ENL YEATS domain (1–148) and the mutants were recombinantly expressed in *E.* coli and purified as N-terminally His-Sumo-tagged proteins. After Ni column chromatography followed by ulp1 digestion for tag removal, ENL proteins were further purified via ion exchange and size-exclusion chromatography. Prior to crystallization, ENL sample was mixed with $H3_{15–39}$K27ac peptide in 1:2 molar ratio for about half an hour. Crystallization was performed under 4°C *via* vapor diffusion method under the condition: 0.2 M K/Na tartrate, 0.1 M sodium citrate tribasic dihydrate, pH 5.6 and 2M ammonium sulfate. Diffraction data were collected at Shanghai Synchrotron Radiation Facility BL17U under cryo-conditions. All structures were determined by molecular replacement and refined by the program PHENIX with iterative model building by the program COOT. Detailed data collection and refinement statistics are summarized in Extended Data Table 1.

ITC titration

All ITC titrations were performed under 15°C using a MicroCal iTC200 instrument. The titration curves were fitted according to the “One Set of Sites” binding model. A high salt titration buffer (500 mM NaCl, 25 mM Tris·HCl, pH 7.5, 2 mM β-ME) was used for ENL YEATS in order to avoid protein aggregation.

Animal experiments

All the mouse experiments were approved by the Institutional Animal Care and Use Committee (IACUC) at Memorial Sloan Kettering Cancer Center and the Dana-Farber Cancer Institute (DFCI). Leukaemia cells (iCas9-MOLM-13) were transplanted by tail-vein injection of 1×10^6 cells into sub-lethally (2 Gy) irradiated female NSG recipient mice at the age of 6–8 weeks ($n = 10$ per group). Dox-containing chow and water were supplied during the course of the experiments. For JQ1 *ex vivo* treatment trials, leukemia cells were treated with 1ug / ml doxycycline for two days and then treated with 100 nM JQ1 and 1ug / ml doxycycline for next two days before equal number of viable cells were transplanted. For JQ1 *in vivo* treatment trials, JQ1 was resuspended in DMSO at 200 mg / mL (for 50 mg/kg dosing) and then diluted 1:20 with 10.5% (2-hydroxypropyl)-beta-cyclodextrin in 0.9% sterile saline. The solution was sonicated until homogenous. Vehicle was prepared and administered the same with the exception that no JQ1 was added. 100 μL doses were

administered by IP injection for 5 consecutive days with 2-day rest for 30 days starting from day 10 after transplantation. The animals were randomly assigned to the experimental groups. Kaplan-Meier survival curves of mice and the Log-rank test were performed using the Prism 6 software program (GraphPad). No animal was excluded from any of the analyses. The investigators were not blinded to allocation during experiments and outcome assessment.

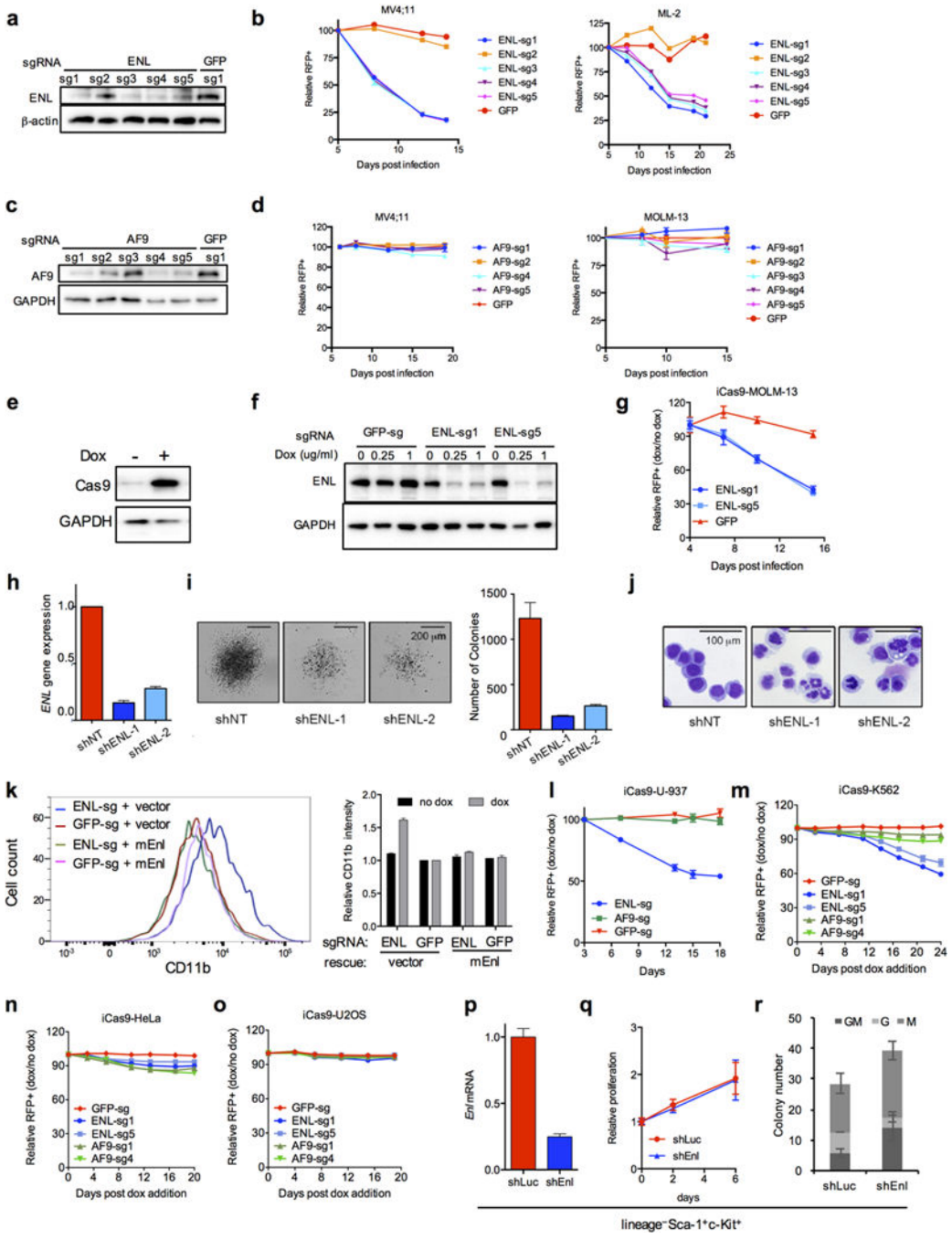
Statistical analyses

Experimental data are presented as means ± standard deviation of the mean unless stated otherwise. Statistical significance was calculated by two-tailed unpaired *t*-test on two experimental conditions with $P < 0.05$ considered statistically significant unless stated otherwise. Statistical significance levels are denoted as follows: $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; $****P < 0.0001$. No statistical methods were used to predetermine sample size.

Data availability

Structure data has been deposited in Protein Data Bank under accession number 5J9S. The ChIP-seq and RNA-seq data have been deposited in the Gene Expression Omnibus database under accession numbers GSE80779 and GSE80774, respectively. All other raw data generated or analyzed during this study are included in this published article (and its Supplementary Information files).

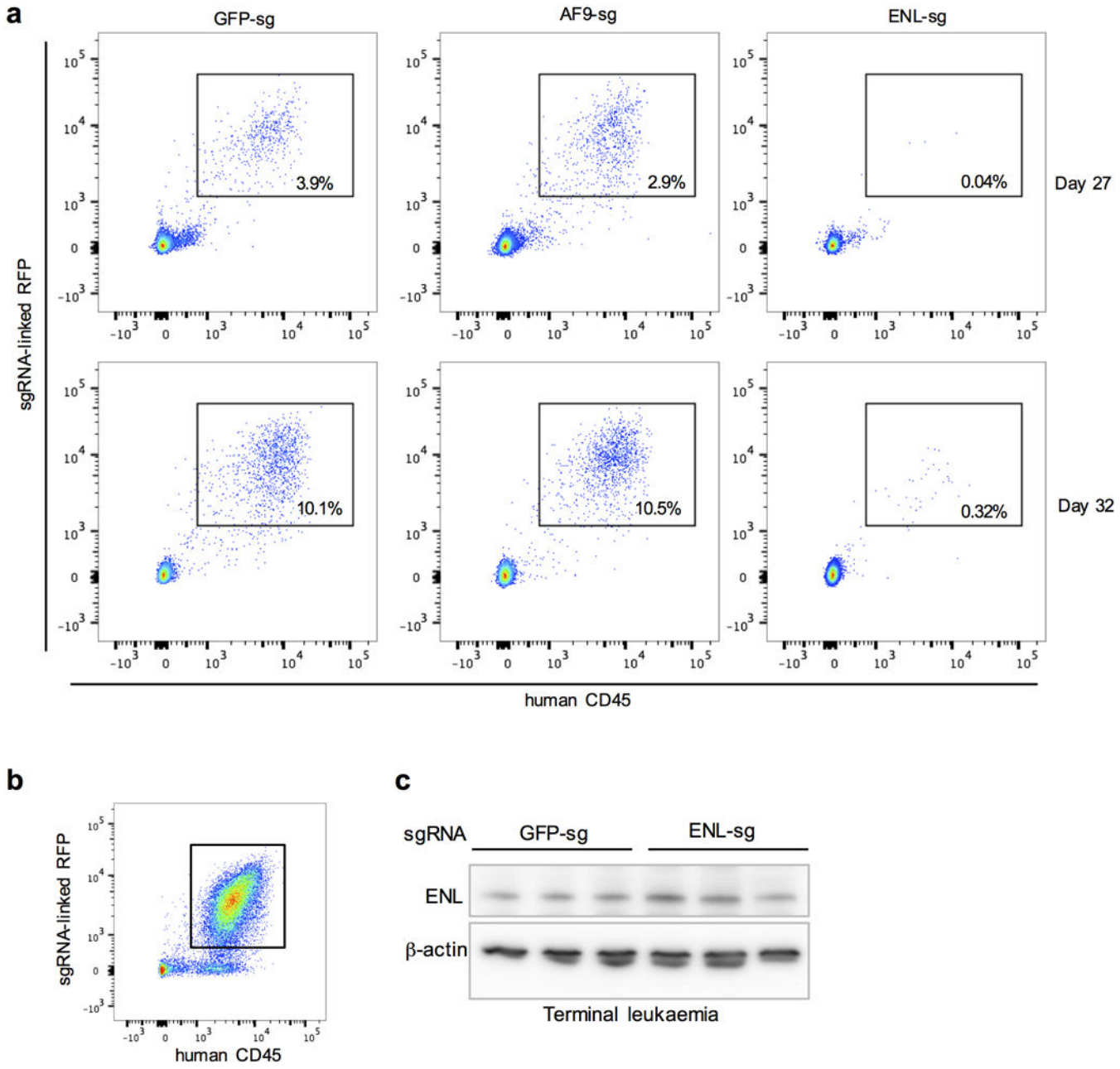
Extended Data



Extended Data Figure 1. Depletion of ENL impairs the growth of AML

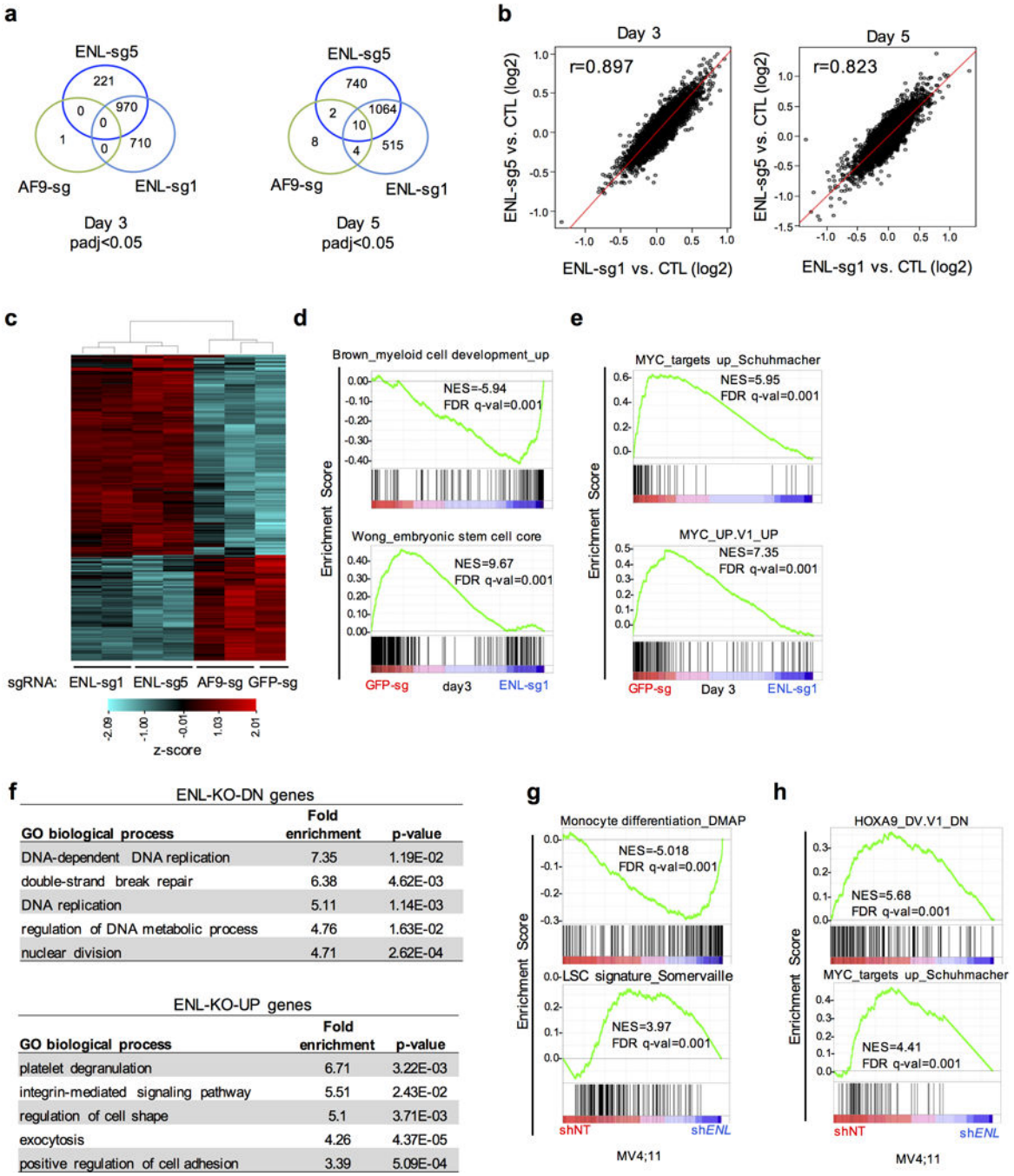
a, Western blot demonstrating the knockdown efficiency of five independent sgRNAs targeting *ENL*. **b**, Competition assay that plots the relative % of $RFP^{+}sgRNA^{+}$ cells following transduction of leukaemia cells with indicated sgRNAs. n = 3. **c**, Western blot demonstrating the knockdown efficiency of five independent sgRNAs targeting *AF9*. **d**, Competition assay that plots the relative % of $RFP^{+}sgRNA^{+}$ cells following transduction of

leukaemia cells with indicated sgRNAs. n = 3. **e**, Western blot demonstrating the induction of Cas9 expression upon Dox treatment in iCas9-MOLM-13 cells. **f**, Western blot demonstrating the decrease of ENL protein levels upon Dox treatment in iCas9-MOLM-13 cells. **g**, Competition assay that plots the relative % of RFP^+sgRNA^+ cells following Dox treatment in iCas9-MOLM-13 cells. n = 3. **h**, Relative *ENL* mRNA levels determined by qRT-PCR quantification in MV4;11 cells transduced with non-targeting (NT) control or *ENL* shRNAs. **i**, Representative images (left) and quantification (right) of colonies formed by MV4;11 cells transduced with indicated shRNAs. **j**, Light microscopy of May-Grunwald/Giemsa-stained MV4;11 leukaemia cells transduced control or *ENL* shRNAs. **k**, FACS analysis of CD11b surface expression after 4 days of Dox-induced Cas9 expression (left) and quantification of CD11b median intensity (right) in iCas9-MOLM-13 cells transduced with indicated sgRNAs and rescue constructs. n = 3. **l-o**, Negative selection competition assay that plots the relative % of RFP^+sgRNA^+ cells following Dox treatment in iCas9-U-937 (l), iCas9-K562 (m), iCas9-HeLa (n) and iCas9-U2OS (o) cells. n = 3. **p-r**, $lineage^-Sca\text{-}1^+c\text{-}Kit^+$ (LSK) cells were sorted from bone marrow of C57BL/6 mice and transduced with luciferase or *Enl* shRNA. **p**, Relative *Enl* mRNA levels determined by qRT-PCR quantification after 3 days of drug selection. **q**, Relative proliferation of control or *Enl*-KD LSKs. n = 4**. r**, Quantification of colonies formed by LSK cells cultured for 7 days. n = 4. GM, colony-forming unit containing granulocyte and macrophage. M, macrophage. G, granulocyte. All error bars represent mean ± s.d. See Supplementary Figure 1 for western blot gel source data.



Extended Data Figure 2. ENL is required for AML growth *in vivo*

a, Representative flow cytometry plots of donor-derived (human CD45+) peripheral blood cells 27 or 32 days post transplantation. The gate shown includes $RFP^{+}sgRNA^{+}$ human leukaemia cells. **b**, Representative flow cytometry plots of bone marrow cells in terminally diseased mice receiving cells transduced with *ENL* sgRNA. The majority of outgrowing leukaemia cells were $RFP^{+}sgRNA^{+}$. **c.** Western blot of sorted $RFP^{+}sgRNA^{+}$ leukaemia cells from terminally diseased mice (n = 3) receiving cells transduced with control or *ENL* sgRNA. See Supplementary Figure 1 for western blot gel source data.



f

ENL-KO-DN genes

GO biological process	Fold enrichment	p-value
DNA-dependent DNA replication	7.35	1.19E-02
double-strand break repair	6.38	4.62E-03
DNA replication	5.11	1.14E-03
regulation of DNA metabolic process	4.76	1.63E-02
nuclear division	4.71	2.62E-04

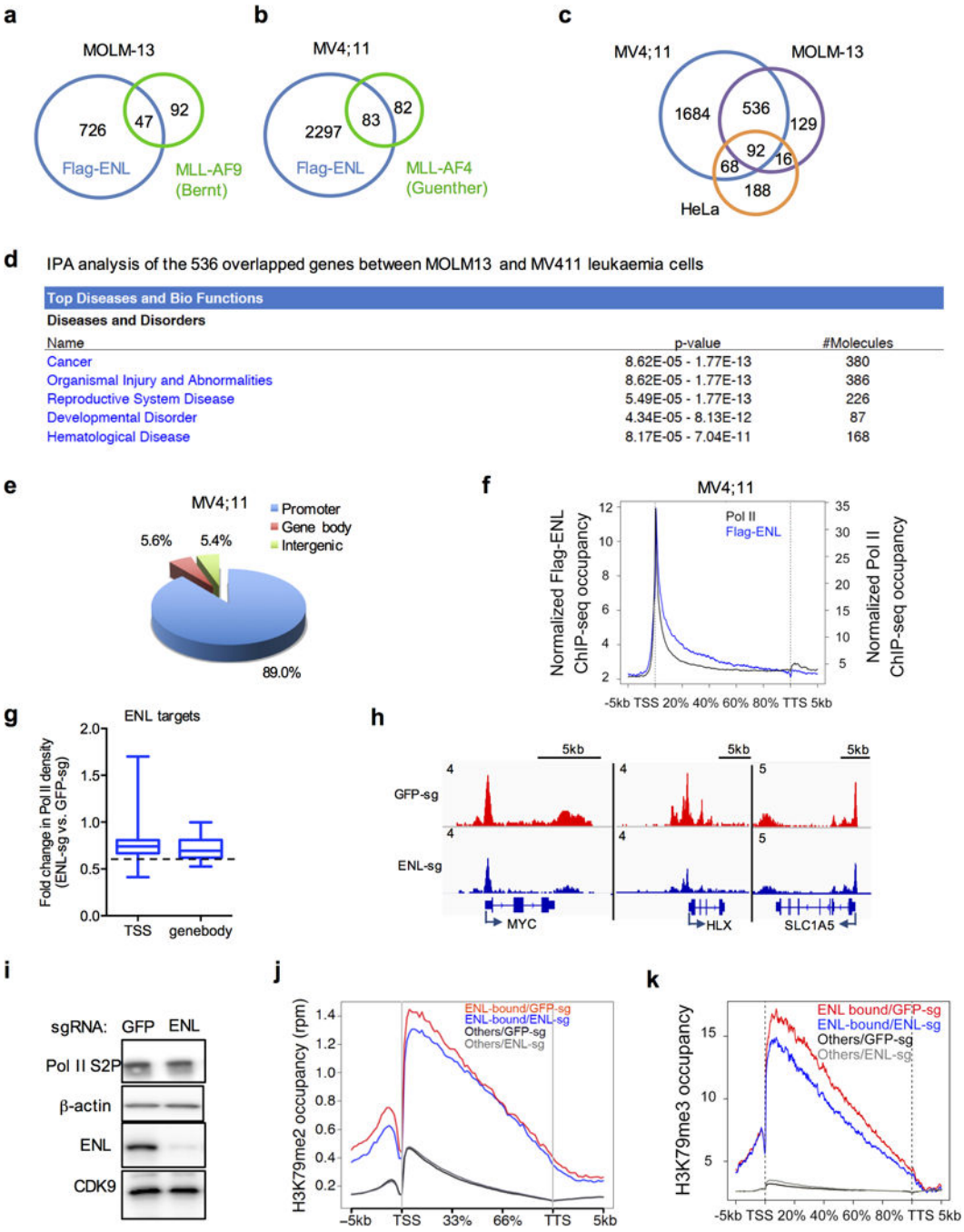
ENL-KO-UP genes

GO biological process	Fold enrichment	p-value
platelet degranulation	6.71	3.22E-03
integrin-mediated signaling pathway	5.51	2.43E-02
regulation of cell shape	5.1	3.71E-03
exocytosis	4.26	4.37E-05
positive regulation of cell adhesion	3.39	5.09E-04

Extended Data Figure 3. Depletion of ENL deregulates core cellular processes and oncogenic pathways that are required for AML maintenance

a–f, RNA for RNA-seq experiments was obtained from sorted $RFP^{+}sgRNA^{+}$ iCas9-MOLM-13 cells after 3 or 5 days of Dox treatment. **a**, Venn diagram showing the number and the overlap of genes whose expression is significantly changed (padj < 0.05) upon expression of indicated sgRNAs as compared to GFP control. **b**, Dot plots showing a strong correlation of transcriptional changes (log2 fold change over GFP control) caused by two independent sgRNAs targeting *ENL*. r, correlation coefficient. **c**. Heatmap representation of

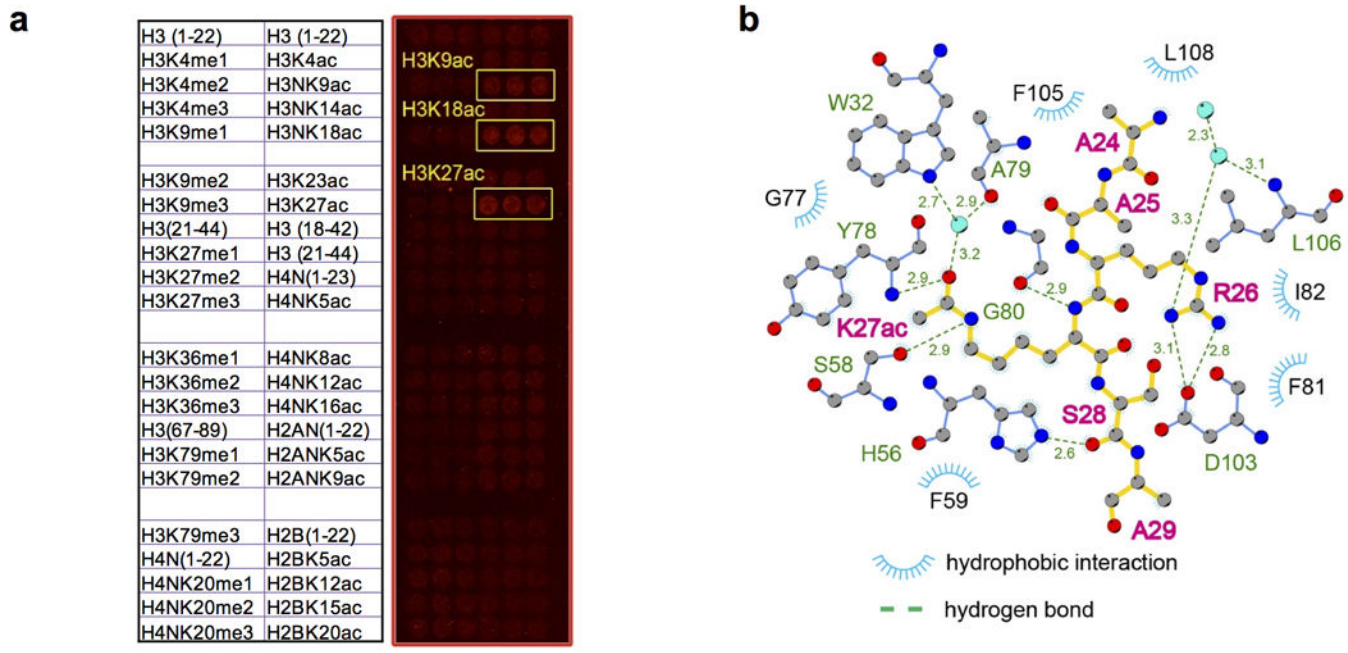
genes differentially expressed between iCas9-MOLM-13 cells expressing sgRNAs targeting *GFP* control, *ENL* or *AF9* (fold change > 1.5 and padj < 0.05) after 3 days of Dox induction. **d and e**, GSEA plots evaluating the changes in monocyte differentiation and LSC gene signatures (d) and the *MYC* pathways (e) upon ENL depletion. **f**, GO term analyses of downregulated (ENL-KO-DN, top) or upregulated (ENL-KO-UP, bottom) genes in *ENL* sgRNA-expressing cells. The top 5 biological processes that each group of genes were enriched in were shown (details in Supplementary Table S2). Fold enrichment and *P* values are shown at the right. **g and h**, RNA for RNA-seq experiments was obtained from MV4;11 transduced with non-targeting (NT) or *ENL* shRNAs. GSEA plots evaluating the changes in monocyte differentiation and LSC gene signatures (g) and the oncogenic pathways (h) upon *ENL* knockdown.



Extended Data Figure 4. ENL depletion decreases the occupancies of total Pol II and Pol II S2P on ENL-bound genes

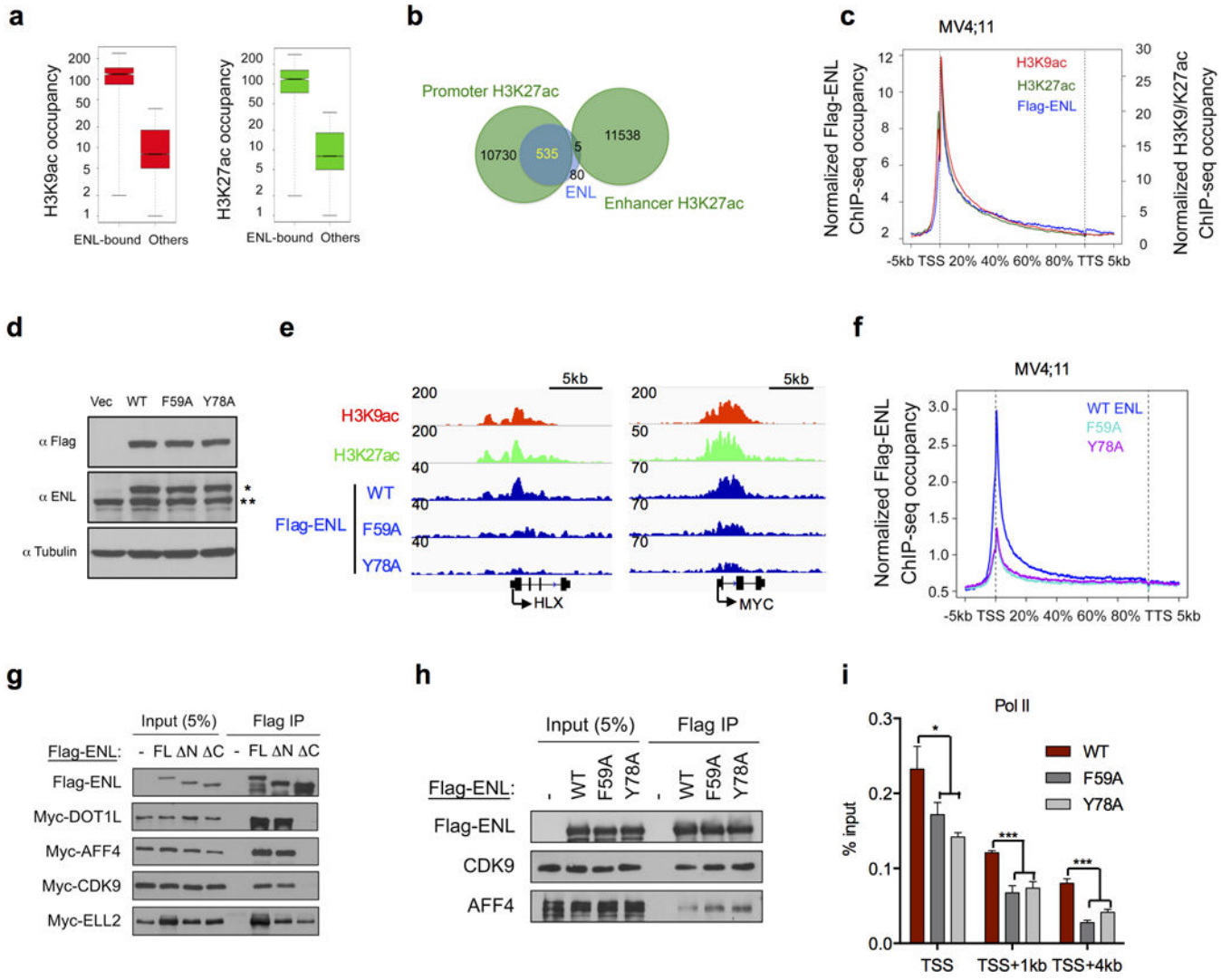
a and b, Venn diagram showing overlaps of Flag-ENL-occupied genes with those of MLL-AF9 in MOLM-13 (Bernt et al, Cancer Cell 2011) (a) or MLL-AF4 in MV4;11 cells (Guenther et al. Genes&Dev. 2008) (b), respectively**. c**, Venn diagram showing overlaps of Flag-ENL occupied genes in MOLM13, MV4;11 and HeLa cells. More details shown in Supplementary Table S7. **d**, IPA analysis of ENL-bound genes overlapped among leukaemia cells but not HeLa cells. **e**, Genomic distribution of Flag-ENL ChIP-seq peaks in MV4;11

cells. The peaks are enriched in the promoter regions ([TSS] ±3kb). $P < 1\times10^{-300}$ (binomial test). More details shown in Supplementary Table S6. **f**, Average occupancies of Flag-ENL (blue) and Pol II (black) on Flag-ENL-bound genes in MV4;11 cells along the transcription unit. **g**, Boxplots showing the fold changes (normalized to GFP control) of Pol II occupancy at TSS (TSS – 30bp to TSS + 300bp) or the rest of the gene body on ENL-bound and activated genes upon the expression of *ENL* sgRNA. The log2 fold changes at both TSS and genebody were significantly lower than 0 ($P < 0.0001$ by one sample, two-tailed Student's *t*-test). **h**, The genome browser view of Pol II signals in a few of ENL-bound genes (*MYC*, *HLX*, *SLC1A5*) in cells expressing sgRNAs targeting *GFP* (red) or *ENL* (blue). TSS is indicated with arrow. **i**, Western blot showing comparable cellular levels of Pol II S2P in MOLM-13 cells expressing sgRNAs targeting *GFP* or *ENL*. See Supplementary Figure 1 for gel source data. **j and k**, Average H3K79me2 (j) and H3K79me3 (k) occupancy on Flag-ENL-bound or non-ENL bound genes (others) in cells expressing sgRNAs targeting *GFP* control or *ENL*.



Extended Data Figure 5. Binding specificity and detail of H3K27ac-bound ENL YEATS complex

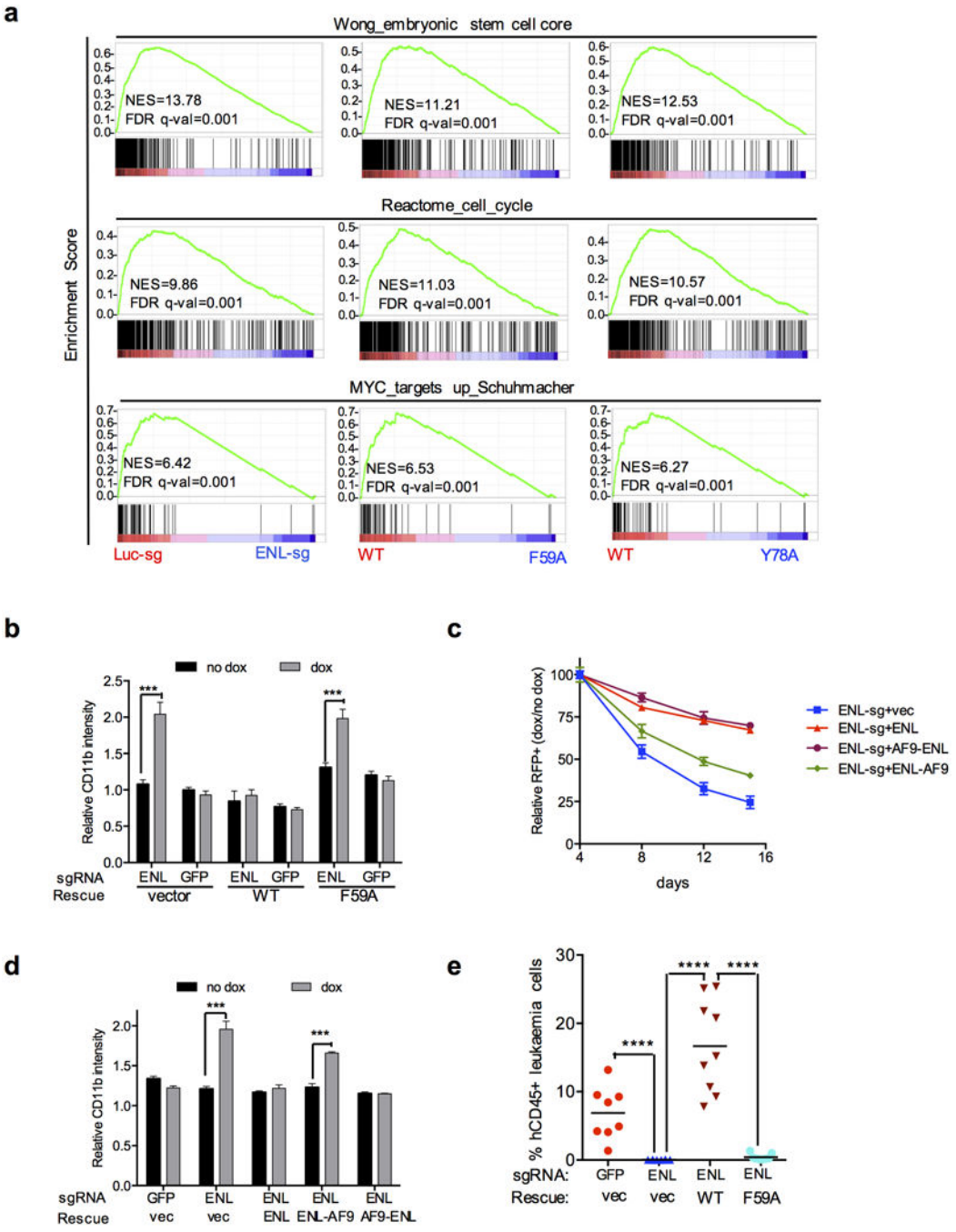
a, Histone peptide microarray (detailed annotations on the left) probed with anti-GST antibody against GST-ENL YEATS domain. H3K9ac, H3K18ac and H3K27ac are highlighted in yellow boxes. **b**, LIGPLOT diagrams of H3K27ac-ENL YEATS complex, listing interactions between H3 peptide and ENL YEATS. H3 segments (orange) and key residues of ENL YEATS (blue) are depicted in ball-and-stick mode. Green ball, carbon; Blue ball, nitrogen; Red ball, oxygen; Big cyan sphere, water molecule. Hydrogen bonds are indicated as green dashed lines with bond length shown as values in the unit of angstrom. Hydrophobic contacts are represented by an arc with spokes radiating towards the ligand atoms they contact, and the contacted atoms are shown with spokes radiating back.



Extended Data Figure 6. The YEATS domain is required for ENL's chromatin localization

a, Boxplots showing H3K9ac (red) and H3K27ac (green) occupancy in ENL-bound or unbound genes (others) in MOLM-13 cells. $P < 8.1\times10^{-152}$ (H3K9ac) and $P < 2.2\times10^{-136}$ (H3K27ac) by two-tailed unpaired Student's *t*-test. **b**, Venn diagram showing the overlap of Flag-ENL (blue) and H3K27ac ChIP-seq peaks (green) at promoter or enhancer regions. Promoter H3K27ac is defined as H3K27ac peaks at [TSS] ±3kb regions co-occupied with H3K4me3; enhancer H3K27ac is defined as non-promoter H3K27ac peaks co-occupied with H3K4me1. There is a significant overlap between Flag-ENL and H3K27ac ChIP-seq peaks at TSS ($p = 5.7\times10^{-105}$, two-way Fisher Exact Test) but not at enhancer ($p = 1$, two-way Fisher Exact Test). **c**, Average genome-wide occupancies of Flag-ENL (blue), H3K9ac (red), H3K27ac (green) at Flag-ENL-bound genes along the transcription unit in MV4;11 cells. More details shown in Supplementary Table S8, S9. **d**, Western blot showing the protein levels of ectopically expressed WT or mutant Flag-ENL (marked as *) and endogenous ENL (marked as **). **e**, The genome browser view of H3K27ac, H3K9ac, Flag-ENL signals in a few of ENL-bound genes (*MYC*, *HLX*). TSS is indicated with arrow. **f**, Average occupancies of WT, F59A or Y78A mutant Flag-ENL on ENL-bound genes along the

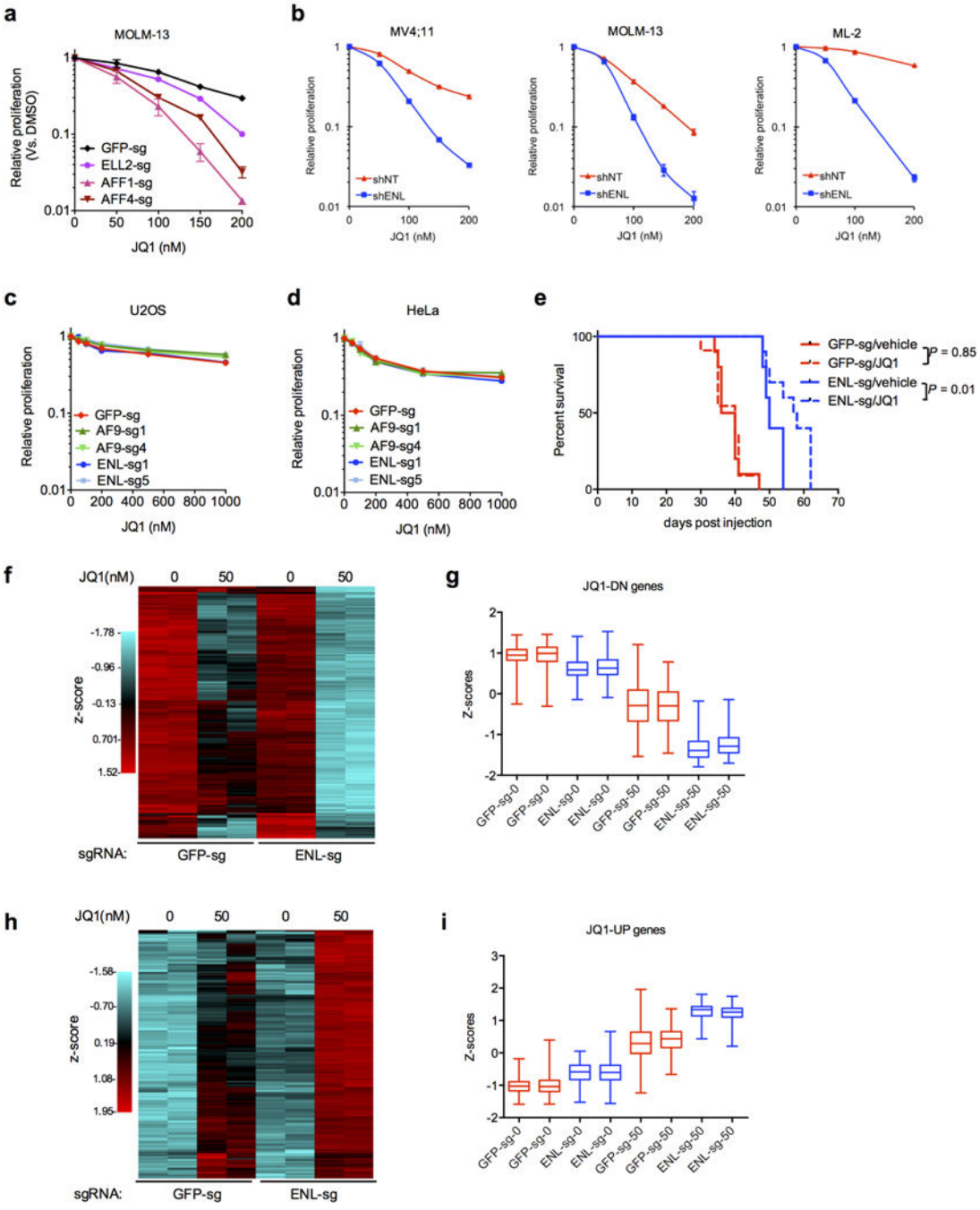
transcription unit in MV4;11 cells. **g**, Western blot analysis of coIP using the M2 anti-Flag antibody in cells expressing Flag-ENL and Myc-tagged DOT1L, AFF4, CDK9 or ELL2 proteins. FL: full- length; ΔN: deletion of aa1-113; ΔC: deletion of aa430–559 of ENL. **h**, Western blot analysis of IP using the M2 anti-Flag antibody in cells expressing WT or mutant Flag-ENL. Endogenous CDK9 and AFF4 were assessed. **i**, qPCR analysis of the Pol II ChIP signal in *MYC* gene in *ENL* sgRNA-expressing cells rescued by WT or mutant (F59A or Y78A) murine ENL. Error bars represent mean ± s.e.m. * $P < 0.5$, *** $P < 0.001$ (two-tailed unpaired Student's *t*-test). See Supplementary Figure 1 for gel source data.



Extended Data Figure 7. The YEATS domain-histone acetylation interaction is required for ENL's roles in leukaemias

a, GSEA plots evaluating the enrichment of signatures related to stem cells, cell cycle or the *MYC* pathway in the indicated comparisons. **b**, Quantification of CD11b median intensity 4 days after Dox induction in iCas9-MOLM-13 cells transduced with indicated sgRNAs and rescue constructs. n = 3. ***$P < 0.001$ by two-tailed unpaired Student's *t*-test. **c**, Negative-selection competition assay that plots the relative % of $RFP^{+}sgRNA^{+}$ cells following transduction of leukaemia cells with indicated constructs. n = 3. **d**, Quantification of CD11b

median intensity 6 days after Dox induction in iCas9-MOLM-13 cells transduced with indicated sgRNAs and rescue constructs. n = 3. *** $P < 0.001$ by two-tailed unpaired Student's *t*-test. **e**, Percentage of human $CD45^+$ cells in peripheral blood of mice transplanted with MOLM-13 cells expressing indicated sgRNAs and rescue constructs 30 days post injection (n≥8). **** $P < 0.0001$ by two-tailed unpaired Student's *t*-test. All error bars represent mean ± s.d.



Extended Data Figure 8. Depletion of ENL increases sensitivity to JQ1 by potentiating JQ1-induced transcriptional changes

a, Effect of JQ1 on the proliferation (normalized to DMSO control) of MOLM-13 cells transduced with indicated sgRNAs targeting SEC components. n = 5. **b**, Effect of JQ1 on the proliferation of indicated *MLL*-rearranged leukaemia cells transduced with shNT (red) or shENL (blue) shRNAs. (n = 3). **c and d**, Effect of JQ1 on the proliferation of indicated non-leukaemia cells (U2OS and HeLa) transduced with *GFP*, *AF9* or *ENL* sgRNAs. n = 5. **e**, Kaplan-Meier survival curves of mice (n = 10 per group) transplanted with iCas9-MOLM-13 cells expressing indicated sgRNAs and pretreated with doxycycline for 4 days and JQ1 (or DMSO control) for 2 days *ex vivo*. *P* values were calculated using a Log-rank test. **f-i**, RNA for RNA-seq experiments was obtained from sorted $RFP^{+}sgRNA^{+}$ iCas9-MOLM-13 cells treated with DMSO (marked as "0") or 50 nM JQ1 for 24 hours. Row-normalized heatmap (f and h) and boxplots of relative expression levels (z-scores, g and i) of genes found to be 2-fold downregulated (f and g) or upregulated (h and i) following JQ1 treatment in *ENL* sgRNA-expressing cells. All error bars represent mean ± s.e.m.

Extended Data Table 1

Data collection and refinement statistics

	ENL YEATS-H3K27ac[b] (PDB code: 5J9S)
Data collection	
Space group	P321
Cell dimensions	
a, *b*, *c* (Å)	105.0, 105.0, 45.1
α, *β*, *γ* (°)	90, 90, 120
Resolution (Å)	50.00–2.70(2.75–2.70)[a]
R_{sym}	19.5(53.3)
$I/\sigma(I)$	8.5(2.3)
Completeness (%)	98.5(97.7)
Redundancy	4.2(4.0)
Refinement	
Resolution (Å)	40.4–2.7
No. reflections	8,050
R_{work} / R_{free}	18.5/21.6
No. atoms	
Protein	1,249
Peptide/ion(sulfate)	44/20
Water	51
B factors	
Protein	22.4
Peptide/ion(sulfate)	34.9/42.4
Water	22.9
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.611

[a]Values in parentheses are for highest-resolution shell.

[b] One crystal was used for the data collection.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank H. Li, M. Armstrong, C-W. Chen, M. Yu, H. Chu, K. Tanaka, D. Peng and T. Scott for discussions and technical supports. We are grateful to The Rockefeller University Genomic Resource Center, Flow Cytometry Resource Center, The MD Anderson Science Park Next-Generation Sequencing Facility (CPRIT RP120348), The Sequencing and Microarray Facility and The Flow Cytometry and Cellular Imaging Core Facility (NIH/NCI P30CA016672), The Shanghai Synchrotron Radiation Facility BL17U beamline, and The China National Center for Protein Sciences Beijing for facility supports. The research was supported by funds from NIH, the Leukaemia and Lymphoma Society (LLS-SCOR 7006-13), and The Rockefeller University to C.D.A.; NIH grants (CA66996 and CA140575) and the Leukaemia and Lymphoma Society to S.A.A.; grants from NIH/NCI (1R01CA204020-01), Cancer Prevention and Research Institute of Texas (RP160237 and RP170285) and Welch Foundation (G1719) to X.S.; grants from the Major State Basic Research Development Program in China (2016YFA0500700 and 2015CB910503), and the Tsinghua University Initiative Scientific Research Program to H.L.; and grants from NIH (R01HG007538 and R01CA193466) to W.L. X.S. is a recipient of Leukaemia & Lymphoma Society Career Development Award and a R. Lee Clark Fellow and Faculty Scholar of MD. Anderson Cancer Center. L.W. is a fellow of the Jane Coffin Childs Memorial Fund. Y.L. is a Tsinghua Advanced Innovation fellow.

References

1. Garraway LA, Lander ES. Lessons from the cancer genome. Cell. 2013; 153:17–37. [PubMed: 23540688]
2. Filippakopoulos P, et al. Selective inhibition of BET bromodomains. Nature. 2010; 468:1067–1073. [PubMed: 20871596]
3. Dawson MA, et al. Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. Nature. 2011; 478:529–533. [PubMed: 21964340]
4. Li Y, et al. AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation. Cell. 2014; 159:558–571. [PubMed: 25417107]
5. Deshpande AJ, Bradner J, Armstrong SA. Chromatin modifications as therapeutic targets in MLL-rearranged leukemia. Trends Immunol. 2012; 33:563–570. [PubMed: 22867873]
6. Somervaille TCP, et al. Hierarchical maintenance of MLL myeloid leukemia stem cells employs a transcriptional program shared with embryonic rather than adult stem cells. Cell Stem Cell. 2009; 4:129–140. [PubMed: 19200802]
7. Lin C, et al. AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. Mol Cell. 2010; 37:429–437. [PubMed: 20159561]
8. He N, et al. Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. Proc Natl Acad Sci USA. 2011; 108:E636–45. [PubMed: 21873227]
9. Yokoyama A, Lin M, Naresh A, Kitabayashi I, Cleary ML. A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription. Cancer Cell. 2010; 17:198–212. [PubMed: 20153263]
10. Kawahara M, et al. H2.0-like homeobox regulates early hematopoiesis and promotes acute myeloid leukemia. Cancer Cell. 2012; 22:194–208. [PubMed: 22897850]
11. Sims RJ, Belotserkovskaya R, Reinberg D. Elongation by RNA polymerase II: the short and long of it. Genes Dev. 2004; 18:2437–2468. [PubMed: 15489290]
12. Luo Z, et al. The super elongation complex family of RNA polymerase II elongation factors: gene target specificity and transcriptional output. Mol Cell Biol. 2012; 32:2608–2617. [PubMed: 22547686]
13. Peterlin BM, Price DH. Controlling the elongation phase of transcription with P-TEFb. Mol Cell. 2006; 23:297–305. [PubMed: 16885020]

14. Filippakopoulos P, Knapp S. Targeting bromodomains: epigenetic readers of lysine acetylation. Nat Rev Drug Discov. 2014; 13:337–356. [PubMed: 24751816]
15. Jang MK, et al. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. Mol Cell. 2005; 19:523–534. [PubMed: 16109376]
16. Rathert P, et al. Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. Nature. 2015; 525:543–547. [PubMed: 26367798]
17. Fong CY, et al. BET inhibitor resistance emerges from leukaemia stem cells. Nature. 2015; 525:538–542. [PubMed: 26367796]
18. Perlman EJ, et al. MLLT1 YEATS domain mutations in clinically distinctive Favourable Histology Wilms tumours. Nat Commun. 2015; 6:10013. [PubMed: 26635203]
19. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009; 25:1105–1111. [PubMed: 19289445]
20. Anders S, Pyl PT, Huber W. HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015; 31:166–169. [PubMed: 25260700]
21. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15:550. [PubMed: 25516281]
22. Mi H, Poudel S, Muruganujan A, Casagrande JT, Thomas PD. PANTHER version 10: expanded protein families and functions, and analysis tools. Nucleic Acids Res. 2016; 44:D336–42. [PubMed: 26578592]
23. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005; 102:15545–15550. [PubMed: 16199517]
24. Zhang Y, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008; 9:R137. [PubMed: 18798982]
25. Wen H, et al. ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. Nature. 2014; 508:263–268. [PubMed: 24590075]

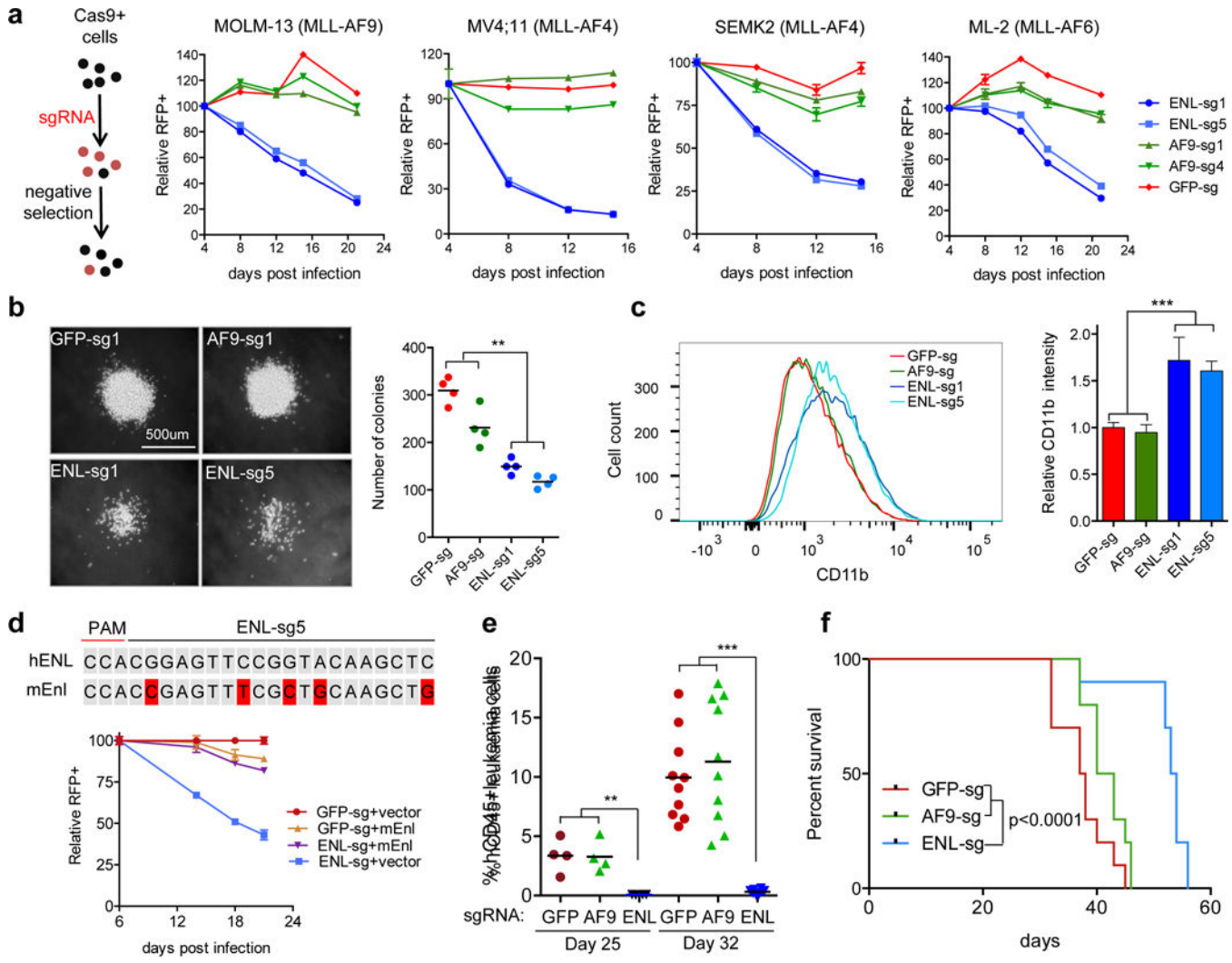


Figure 1. AML growth is sensitive to ENL depletion *in vitro* and *in vivo*

a, Negative-selection competition assay that plots the relative percentage of RFP^+sgRNA^+ cells over time following transduction of different *MLL*-rearranged leukaemia cell lines with indicated sgRNAs. n = 3. **b**, Representative images (left) and quantification (right) of colonies formed by MOLM-13 cells transduced with indicated sgRNAs. n = 4. **c**, FACS analysis of CD11b surface expression after 4 days of Dox treatment (left) and quantification of CD11b median intensity (right). n = 4. **d**, (top) Comparison of mouse and human *ENL* sequences at the indicated sgRNA recognition sites. Red nucleotides indicate mismatches. PAM, protospacer-adjacent motif. (bottom) Negative-selection competition assay that plots the relative % of RFP^+sgRNA^+ cells following transduction of leukaemia cells with indicated constructs. n = 3. **e**, Percentage of human $CD45^+$ cells in the peripheral blood of mice receiving MOLM-13 cells transduced with indicated sgRNAs at 27 (n = 4) or 32 (n = 10) days post injection. **f**, Kaplan–Meier survival curves of recipient mice (n = 10 per group) transplanted with MOLM-13 cells expressing indicated sgRNAs. $P < 0.0001$ using a Log-rank test. All error bars represent mean ± s.d. and statistical significance was calculated using two-tailed unpaired Student's *t*-test unless noted otherwise. $^{**}P < 0.01$, $^{***}P < 0.001$.

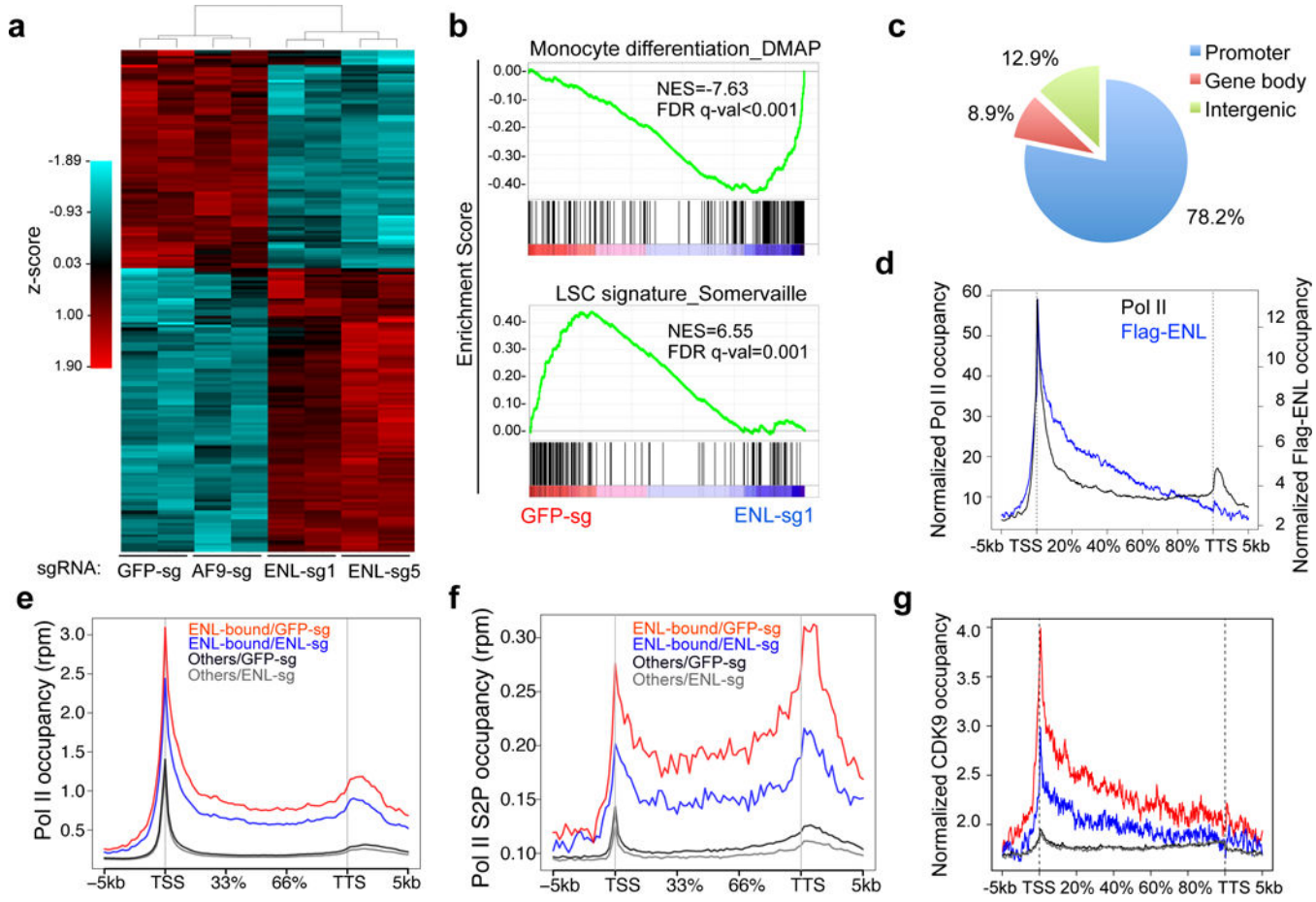


Figure 2. ENL modulates the recruitment of Pol II to activate oncogenic gene expression

a, Heatmap representation of genes differentially expressed in iCas9-MOLM-13 cells expressing sgRNAs targeting *GFP* control, *ENL* or *AF9* (fold change > 1.5 and padj < 0.05) 5 days post Dox treatment. Red and green indicate relative high and low expression, respectively. More details shown in Supplementary Table S1. **b**, GSEA plots evaluating the changes in monocyte differentiation and LSC gene signatures upon ENL depletion. NES, normalized enrichment score; FDR q-val, false discovery rate q value. **c**, Genomic distribution of Flag-ENL ChIP-seq peaks in MOLM-13 cells. The peaks are enriched in the promoter regions ([TSS] ± 3kb). $P < 1\times10^{-300}$ (binomial test). More details shown in Supplementary Table S3. **d**, Average genome-wide occupancies of Flag-ENL (blue) and Pol II (black) on Flag-ENL-bound genes along the transcription unit. The gene body length is aligned by percentage from the TSS to TTS. 5 kb upstream of TSS and 5 kb downstream of TTS are also included. **e-g**, Average occupancy of Pol II (e), Pol II Ser2P (f), or CDK9 (g) on Flag-ENL-bound or non-ENL bound genes (others) in iCas9-MOLM-13 cells expressing sgRNAs targeting *GFP* control or *ENL*. The gene body length is aligned as in d. rpm, reads per million.

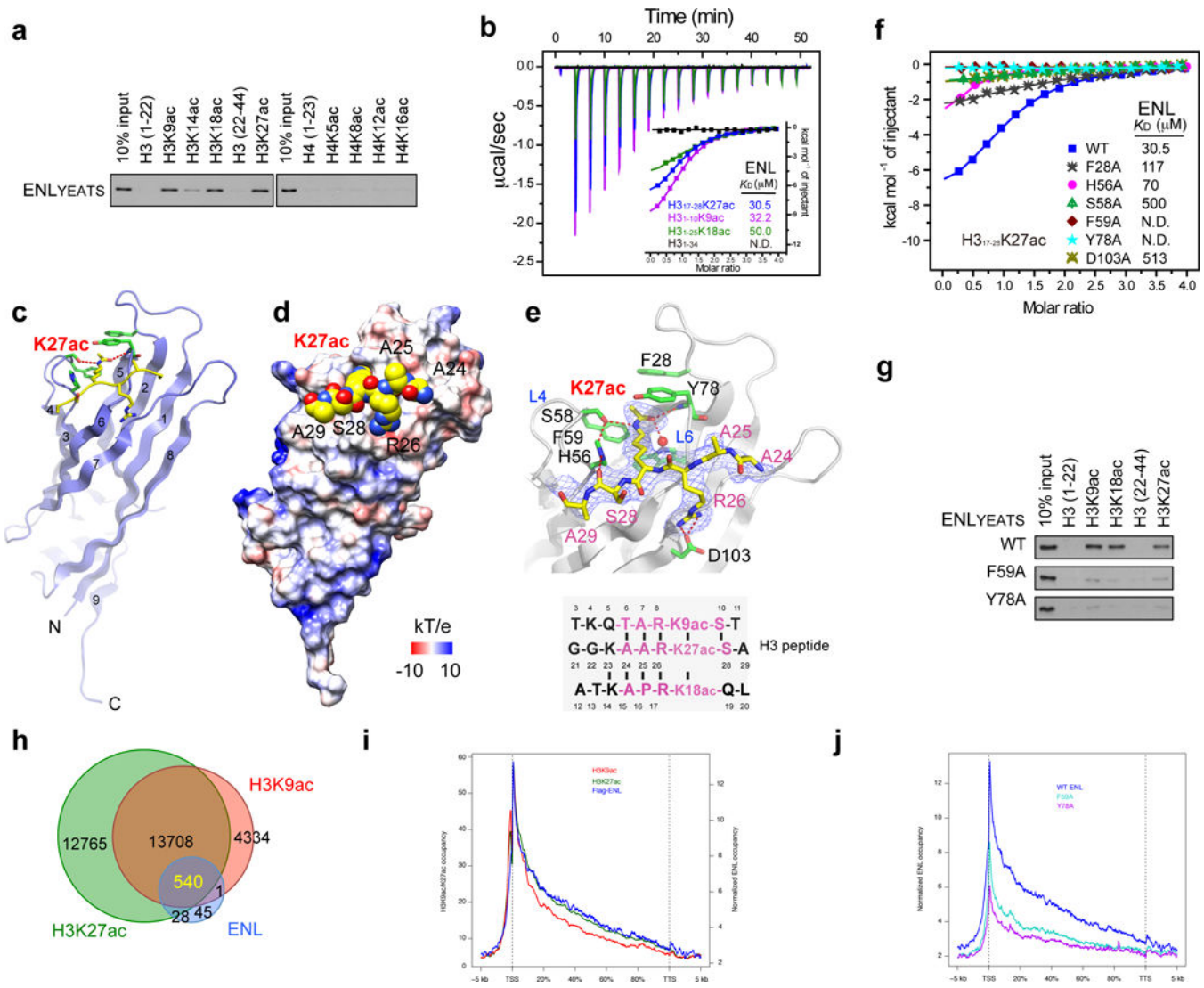


Figure 3. ENL binds and colocalizes with acetylated histone H3 genome-wide via its YEATS domain

a, Peptide pull-down assay of indicated histone peptides and ENL YEATS domain. **b**, ITC titration and fitting curves of human ENL YEATS domain titrated with $H3_{17–28}$K27ac, $H3_{1–10}$K9ac, $H3K_{1–25}$18ac or unmodified $H3_{1–34}$ peptides. **c**, Overall structure of the ENL YEATS domain bound to H3K27ac peptide. ENL YEATS is depicted as blue ribbons with key residues highlighted as green sticks. Histone H3K27ac peptide is shown as a yellow ribbon with side chains highlighted in sticks. Red dashes, hydrogen bonds; red sphere, water molecule. **d**, Electrostatic potential surface of the ENL YEATS domain ranging from −10 to 10 kT/e. Histone H3 peptide is depicted as space-filling spheres. **e**, Hydrogen bonding network between H3K27ac peptide and ENL (top). Hydrogen bonds are shown as red dashes. Key residues of ENL are depicted as green sticks and labeled black; the H3 peptide is shown as yellow sticks and labeled red. Grey meshes, Fo-Fc omit map countered at 2.0 σ level. (Bottom) Sequence alignment of histone H3 sequences flanking residues K9, K18 and K27. Conserved residues are highlighted in pink. **f**, ITC titration fitting curves of ENL

YEATS mutants with $H3_{17–28}$K27ac peptide. **g**, Peptide pull-downs of ENL YEATS mutants and indicated histone H3 peptides. See Supplementary Figure 1 for gel source data. **h**, Venn diagram showing the overlap of Flag-ENL-, H3K9ac- and H3K27ac ChIP-seq peaks in MOLM-13 cells. $P < 1\times10^{-300}$ (3-way Fisher's exact test). More details shown in Supplementary Table S4, S5. **i**, Average occupancies of Flag-ENL (blue), H3K9ac (red), H3K27ac (green) on ENL-bound genes along the transcription unit. **j**, Average genome-wide occupancies of WT (blue) and mutant ENL (F59A in cyan or Y78A in purple) on ENL-bound genes along the transcription unit.

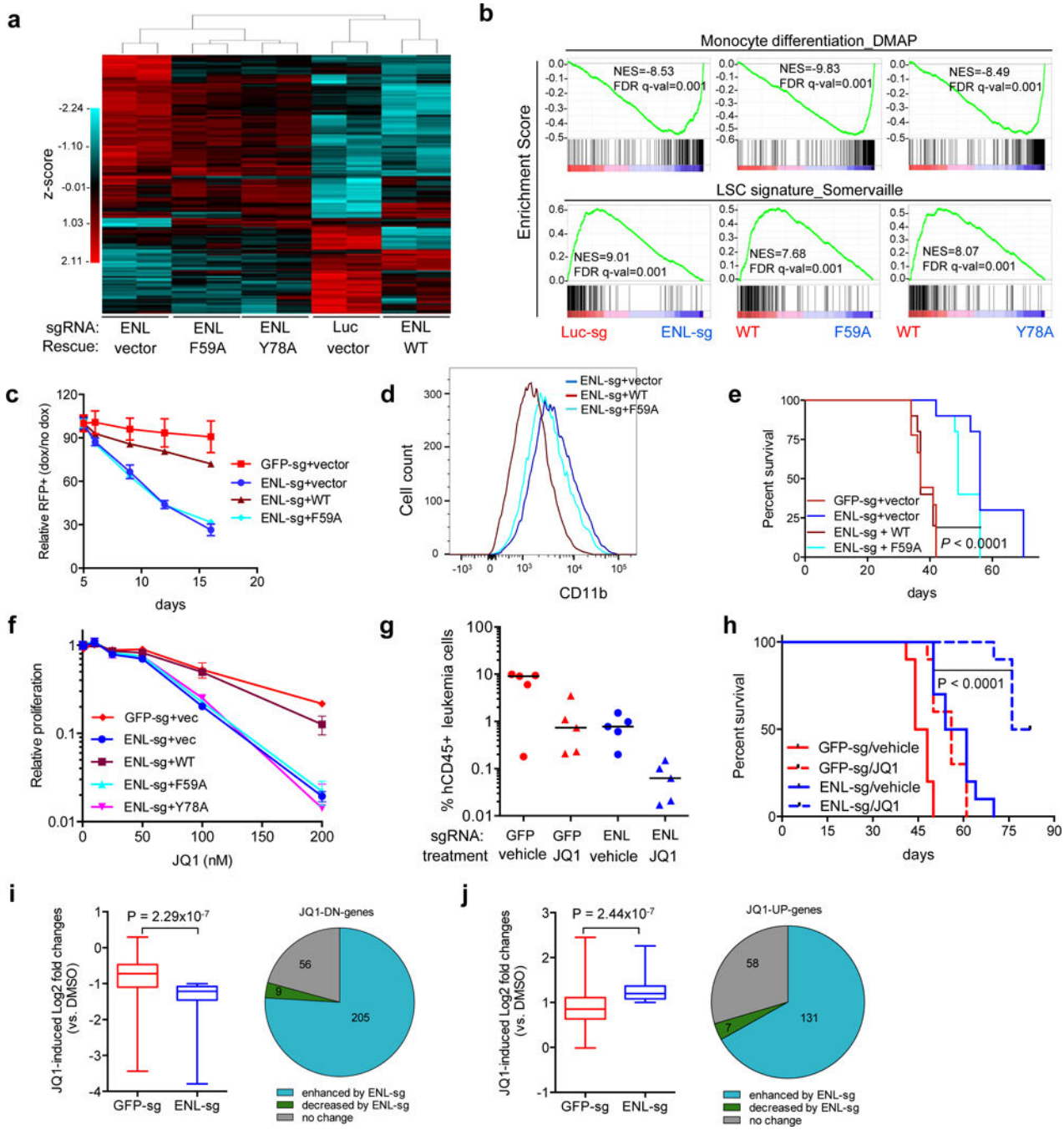


Figure 4. Disrupting the YEATS-histone acetylation interaction inhibits ENL's functionality and sensitizes leukaemia cells to BET inhibitors

a, Heatmap representation of genes differentially expressed in iCas9-MOLM-13 cells expressing sgRNAs targeting luciferase (*Luc*) control or *ENL* (fold change > 1.5 and padj < 0.05) in the indicated rescue conditions. **b**, GSEA plots evaluating the changes in monocyte differentiation and LSC gene signatures in the indicated comparisons. **c**, Negative-selection competition assay that plots the relative % of $RFP^{+}sgRNA^{+}$ cells following transduction of leukaemia cells with indicated constructs. n = 3. **d**, FACS analysis of CD11b surface

expression in iCas9-MOLM-13 cells expressing *ENL* sgRNA and indicated murine *Enl* rescue constructs after 4 days of Dox treatment. See Extended Data Fig. 7b for quantification. n = 3. **e**, Kaplan-Meier survival curves of mice (n = 10 per group) transplanted with MOLM-13 cells transduced with indicated sgRNAs and rescue constructs. $P < 0.0001$ using a Log-rank test. **f**, Effect of JQ1 on the proliferation (normalized to DMSO control) of MOLM-13 cells transduced with indicated sgRNAs and rescue constructs. n = 5. **g** and **h**, Leukemia burden (g) and survival curves (h) of mice transplanted with iCas9-MOLM-13 cells expressing indicated sgRNAs. Treatment with JQ1 (or vehicle control) and doxycycline was initiated at day 10 after transplantation. **g**, Percentage of human $CD45^+$ cells in the peripheral blood of mice (n = 5) at 40 days after transplantation. $P < 0.01$ (ENL-sg/JQ1 vs. all other groups) using Mann-Whitney test. **h**, Kaplan–Meier survival curves of control and JQ1-treated mice (n = 10 per group). $P < 0.0001$ (ENL-sg/JQ1 vs. all other groups) using a log rank test. **i** and **j**, RNA for RNA-seq experiments was obtained from sorted RFP^+sgRNA^+ iCas9-MOLM-13 cells treated with DMSO or 50 nM JQ1 for 24 hours. Genes found to be > 2-fold downregulated (i) or upregulated (j) upon JQ1 treatment in *ENL* sgRNA-expressing cells were examined. (Left) Boxplots comparing the JQ1-induced fold changes of these genes in either control (red) or *ENL* sgRNA-expressing cells (blue). Error bars indicate min to max. $P = 2.29\times10^{-7}$ (i) and $P = 2.44\times10^{-7}$ (j) by two-tailed paired Student's *t*-test. (right) Pie charts showing the categorization of these genes based on the relationship to ENL depletion. Genes whose absolute JQ1-induced fold change was > 1.2 fold higher or lower in *ENL* sgRNA-expressing cells compared to control were classified as "enhanced by ENL-sg" (blue) and "decreased by ENL-sg" (green), respectively. There are significantly more genes falling into "enhanced by ENL-sg" group than "decreased by ENL-sg" group ($P < 0.0001$ by Fisher's exact test). More details shown in Supplementary Table S10. All error bars represent mean ± s.d. (n = 3) unless noted otherwise.