








SETD1A function in leukemia is mediated through interaction with mitotic regulators BuGZ/BUB3

Sarah Perlee^{1,2} , Sota Kikuchi³, Tomoyoshi Nakadai⁴, Takeshi Masuda^{5,6}, Sumio Ohtsuki⁵, Makoto Matsumoto³ , Bahityar Rahmutulla³ , Masaki Fukuyo³ , Paolo Cifani⁷, Alex Kentsis⁷, Robert G Roeder⁴ , Atsushi Kaneda³  & Takayuki Hoshii^{3,*} 

Abstract

The H3K4 methyltransferase SETD1A plays a crucial role in leukemia cell survival through its noncatalytic FLOS domain-mediated recruitment of cyclin K and regulation of DNA damage response genes. In this study, we identify a functional nuclear localization signal in and interaction partners of the FLOS domain. Our screen for FLOS domain-binding partners reveals that the SETD1A FLOS domain binds mitosis-associated proteins BuGZ/BUB3. Inhibition of both cyclin K and BuGZ/BUB3-binding motifs in SETD1A shows synergistic antileukemic effects. BuGZ/BUB3 localize to SETD1A-bound promoter-TSS regions and SETD1A-negative H3K4me1-positive enhancer regions adjacent to SETD1A target genes. The GLEBS motif and intrinsically disordered region of BuGZ are required for both SETD1A-binding and leukemia cell proliferation. Cell-cycle-specific SETD1A restoration assays indicate that SETD1A expression at the G1/S phase of the cell cycle promotes both the expression of DNA damage response genes and cell cycle progression in leukemia cells.

Keywords BuGZ; DNA repair; Leukemia; SETD1A; Transcription

Subject Categories Cancer; Cell Cycle; Chromatin, Transcription & Genomics

DOI 10.15252/embr.202357108 | Received 2 March 2023 | Revised 11 July

2023 | Accepted 18 July 2023 | Published online 3 August 2023

EMBO Reports (2023) 24: e57108

Introduction

Yeast SET1-like histone-lysine N-methyltransferase 2 (KMT2) family (KMT2A–D, SETD1A/B) are known histone H3K4 methyltransferases. H3K4 methylation is associated with active transcription, and many studies indicate biological roles for H3K4 methylation in development as well as cancer (Rao & Dou, 2015). Genetic deletion experiments of H3K4 methyltransferases revealed the indispensable

roles of KMT2 family proteins in various forms of cancer such as leukemia (KMT2A/B, KMT2D and SETD1A) and breast cancer (KMT2D and SETD1A/B; Santos *et al*, 2014; Chen *et al*, 2017; Toska *et al*, 2017; Wang *et al*, 2017; Hoshii *et al*, 2018). However, the functional importance of the methyltransferase activity in specific settings is still unclear.

SETD1A functions as an enzyme that facilitates H3K4me3 and is required for the proliferation of different types of stem cells, including embryonic stem cells and neural stem cells (Bledau *et al*, 2014). SETD1A is also required for hematopoietic stem cell survival (HSCs) under stress conditions, but is dispensable in steady-state conditions (Arndt *et al*, 2018). This disparity suggests that SETD1A has a critical role in the proliferative state. We previously elucidated a noncatalytic function of SETD1A in leukemia and sarcoma cell lines (Hoshii *et al*, 2018). The nonenzymatic FLOS domain of SETD1A is required for cyclin K (CCNK) binding as well as cell cycle-dependent expression of DNA damage response genes (Hoshii *et al*, 2018). Although CCNK is a regulator of RNAP2 activity, it is unclear how the noncatalytic function of SETD1A specifically regulates the transcription of DNA damage response genes in a cell cycle-dependent manner. In our previous study, we identified two functional motifs in the FLOS domain, named F1 and F2 (Hoshii *et al*, 2018). These conserved amino acid sequences are found in SETD1B as well; however, only SETD1A F1 shows a high binding affinity with CCNK. F2 is also well conserved in both proteins, but the function of F2 remains unclear. Elucidating F2 function may provide insights into the mechanism by which SETD1A/B protein complexes orchestrate gene activation and silencing.

Here, we performed detailed characterization of FLOS domain functions and identified that BuGZ/BUB3 are interacting partners of SETD1A/B that bind to the F2 motif. Mutation of both F1 and F2 motifs within the FLOS domain induces the complete disruption of the noncanonical role of SETD1A in MLL-r leukemia cells. Our data demonstrate that BuGZ regulates DNA damage response gene expression and MLL-r leukemia cell proliferation by supporting the

1 Department of Cancer Biology and Genetics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

2 Gerstner Graduate School of Biomedical Sciences, Memorial Sloan Kettering Cancer Center, New York, NY, USA

3 Department of Molecular Oncology, Graduate School of Medicine, Chiba University, Chiba, Japan

4 Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY, USA

5 Laboratory of Pharmaceutical Microbiology, Faculty of Life Sciences, Kumamoto University, Kumamoto, Japan

6 Institute for Advanced Biosciences, Keio University, Tsuruoka, Japan

7 Molecular Pharmacology Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA

*Corresponding author. Tel: +81 43 226 2039; E-mail: hoshiit@chiba-u.jp

SETD1A activation via its disordered region at G1/S phase of cell cycle. These results reveal the contribution of mitotic regulators BuGZ/BUB3 in gene expression control and show the potential efficacy of multifunctional inhibition of SETD1A in cancer therapy.

Results

SETD1A FLOS mediates nuclear import and promotes leukemia cell growth

We previously showed that the FLOS domain is required for the expression of DNA damage response genes, but it was unclear whether the FLOS domain is also functionally sufficient for leukemia cell proliferation and survival. To find the essential components of the SETD1A protein that can support leukemia cell colony formation, we constructed SETD1A-expressing vectors which encoded N-terminal (NF1–5), C-terminal (CF1–3) or FLOS-specific (FO1–3) fragments and performed a rescue experiment with SETD1A-deficient MLL-AF9 leukemia cells (Fig 1A–C). NF2, CF1, and FO3 partially rescued, and NF3–5 fully rescued the *Setd1a* knockout leukemia cells from their loss of proliferative potential (Fig 1C). These phenotypes were well correlated with the expression of *Fancd2*, one of the most highly downregulated genes in *Setd1a*-deficient MLL-AF9 leukemia cells (Fig 1D; Hoshii et al, 2018). Since NF3, but neither NF2 nor CF1, fully rescued the *Setd1a* knockout phenotypes, it demonstrates that both the RNA recognition motif (RRM) and FLOS represent functionally relevant regions on SETD1A (Figs 1C and D, and EV1A). All constructs that rescued the *Setd1a* knockout phenotypes contain the F1 functional motif (CCNK-binding site; Fig EV1B and C), further supporting the importance of SETD1A-CCNK axis in these leukemia cells (Hoshii et al, 2018). In addition to F1, we also previously identified F2 as a functional motif in the FLOS domain, although the precise function of F2 was not known (Figs 1A and EV1C; Hoshii et al, 2018). The NF2 fragment, which confers partial rescue, contains F1/2, but lacks the C-terminal end of FLOS. This suggests that the FLOS domain may encode an additional functional region. Surprisingly, the FLOS-specific fragment (FO3) also partially rescued the *Setd1a* knockout cells, revealing that the FLOS domain itself can act somewhat independently from the rest of the protein (Fig 1C and D). Alanine replacement at core

F1 and F2 motifs in the FO3 fragment was sufficient to completely disrupt the function of FO3 (Figs 1E and EV1D). The transcriptional function of FO3 suggests that the fragment localizes to the nucleus, but the nuclear localization signal (NLS) of SETD1A/B had not yet been characterized. To evaluate the nuclear localization of SETD1A, we created GFP-fused SETD1A fragments and monitored their localization in 293 T cells (Figs 1F and EV1E and F). Importantly, FO3, but not FO2, showed nuclear localization. From a prediction tool (cNLS Mapper: <http://nls-mapper.iab.keio.ac.jp>), we found a probable bipartite NLS in the C terminus of FO3 (Fig EV1G). Mutation of this NLS in FO3 is sufficient to disrupt nuclear localization in 293 T cells (Figs 1F and EV1E and F). Similar patterns of localization were observed in leukemia cells, with stably expressed FO3-GFP localizing to the nucleus while the NLS mutant showed a more diffuse distribution throughout the cytoplasm (Fig 1G). In addition, NLS-mutated FO3 failed to rescue the *Setd1a* knockout cells, suggesting a transcriptional role for the SETD1A FLOS domain (Fig 1H). Together, these data indicate that the FLOS domain is the essential component of SETD1A required for the proliferation of MLL-r leukemia cells.

Targeting of dual F1/F2 motifs in leukemia

In mouse AML cells, mutation of either the F1 or the F2 motif is insufficient to completely disrupt SETD1A function (Fig 2A–C). To evaluate the synergism of F1/2 motif mutations, we constructed a SETD1A mutant with alanine substitutions at the two functional motifs (F1/2-10A; Fig 2A). This mutant resulted in complete disruption of colony-forming ability (Fig 2B and C). These mutations did not influence the nuclear localization (Fig EV2A). Consistent with the effects on cell proliferation, the double mutation significantly disrupted the transcriptional activity of SETD1A (Figs 2D and EV2B). To confirm the effectiveness of targeting both SETD1A motifs as a potential AML therapy, we also performed a SETD1A rescue experiment *in vivo* (Fig EV2C). Although significantly reduced compared with wild-type SETD1A, single FLOS domain mutants F1-5A and F2-5A still partially rescued the engraftment of AML cells *in vivo* (Fig 2E and F). In contrast, mutation of both F1/2 motifs in combination completely depleted the leukemia cells and significantly extended the survival of recipient mice (Fig 2E and F). These data suggest that targeting multiple motifs in the SETD1A FLOS domain *in vivo* leads to significant inhibition of leukemia propagation.

Figure 1. FLOS domain is an essential component of SETD1A for leukemia cell survival.

- A The human SETD1A fragment constructs are shown as schematic illustrations. 1 and 2 on the FLOS domain represent F1 and F2 motifs, respectively.
- B Western blot analysis was performed using SETD1A fragment-expressing 293 T cells. These constructs were used for the following rescue experiments. Representative images from two independent experiments are shown.
- C *Setd1a^{fl/fl};CreER* cells were transfected with SETD1A fragment and endogenous *Setd1a* was deleted by tamoxifen and then assayed for colony-forming potential. Representative data from one out of three independent experiments with three biological replicates are shown.
- D qRT-PCR was performed at 3-day post-tamoxifen to analyze *Fancd2* expression in human SETD1A-expressing *Setd1a^{fl/fl};CreER* MLL-AF9 leukemia cells. qRT-PCR with three biological replicates was performed.
- E *Setd1a^{fl/fl};CreER* cells were transfected with SETD1A FO3 fragment and endogenous *Setd1a* was deleted by tamoxifen and then assayed for colony-forming potential. Representative data from one out of three independent experiments with three biological replicates are shown.
- F 293 T cells were transfected with SETD1A-GFP fusion constructs. Cell nucleus was visualized with DAPI. The percentage of GFP nuclear localization is indicated in each image. Asterisks in images indicate the cells with nuclear-specific SETD1A-GFP signals. Scale bars: 10 μ m.
- G, H Mouse *Setd1a^{fl/fl};CreER* MLL-AF9 leukemia cells were transfected with FO3-GFP fusion constructs. (G) Subcellular localization of GFP fusion constructs was visualized by immunofluorescence and nucleus was stained with DAPI. The percentage of GFP nuclear localization is indicated in each image. Scale bars: 10 μ m. (H) FO3-GFP fusion-expressing cells were treated with tamoxifen, and colony assay with three biological replicates was performed.

Data information: In (C–E and H), data are presented as mean \pm SD. ** $P \leq 0.01$, * $P \leq 0.05$ (Student's *t*-test).

Source data are available online for this figure.

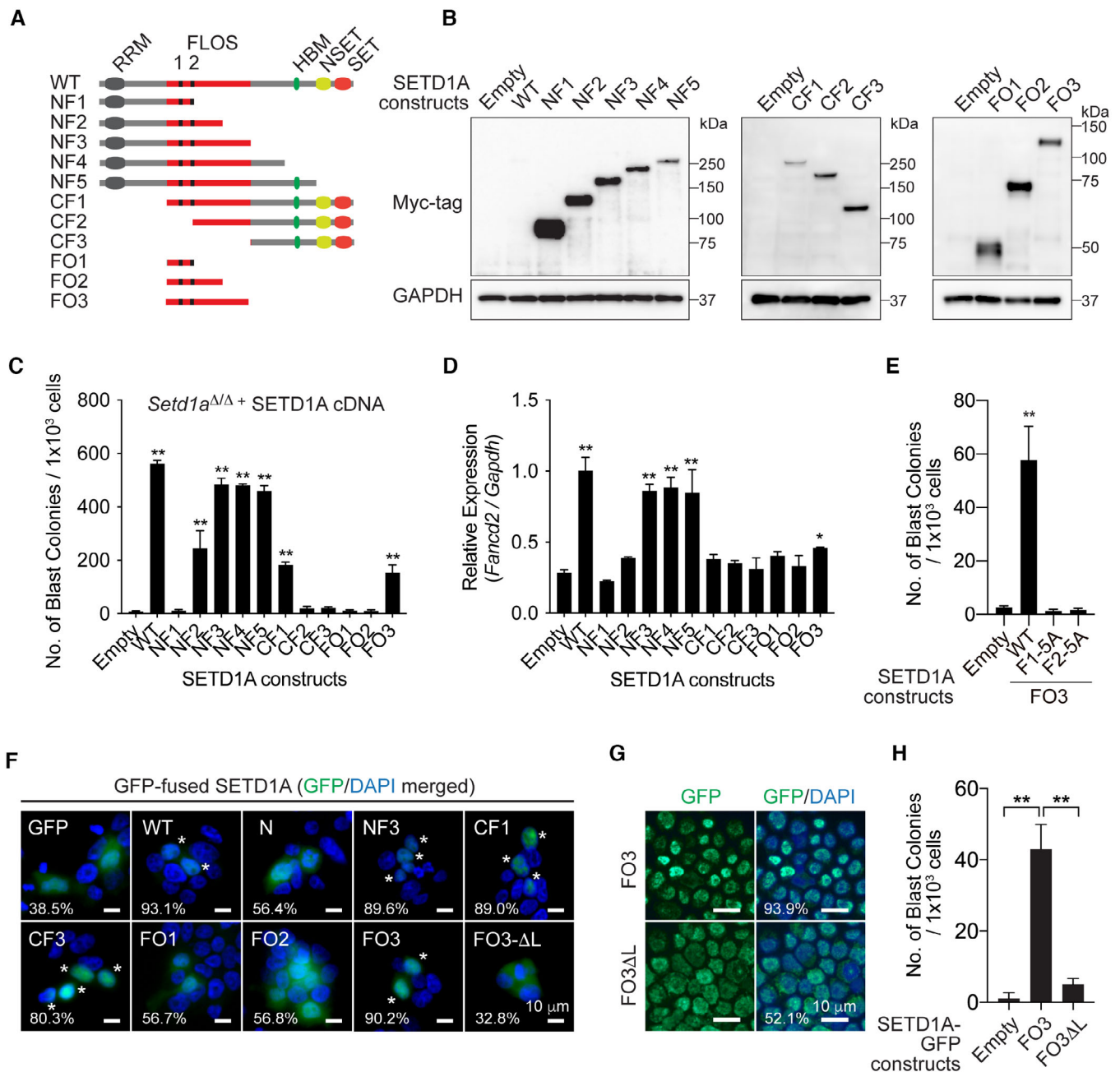


Figure 1.

SETD1A F2 motif binds BuGZ

We previously identified the function of F1 as a CCNK-binding site; however, the function of F2 had not been explored. To determine whether F2 also represents a protein-binding site, we first performed an immunoprecipitation-mass spectrometry (IP-MS) experiment with both flag-tagged full-length SETD1A and FO3 to identify FLOS domain-binding partners. We found 15 proteins, including CCNK, that were commonly identified from two IP-MS analyses (Fig 3A; Appendix Table S1). Two of the identified proteins were Bub3-interacting GLEBS-motif-containing ZNF207 (BuGZ) and BUB3,

which are especially interesting because these proteins form a complex in the mitotic phase of the cell cycle (Toledo *et al*, 2014; Jiang *et al*, 2014a). These proteins were also observed in our previous IP-MS result with full-length SETD1A (Hoshii *et al*, 2018). To identify the specific binding site of BuGZ/BUB3 on SETD1A FLOS domain, we performed an immunoprecipitation with FLOS domain fragments and identified that deletion of the F2 motif inhibits binding with BuGZ/BUB3 (Fig EV3A and B). BuGZ/BUB3 binding is also disrupted by alanine substitution of F2 in the full-length SETD1A protein (Fig 3B). However, this substitution does not affect binding with the known SETD1A partner proteins, CCNK and WDR5, which

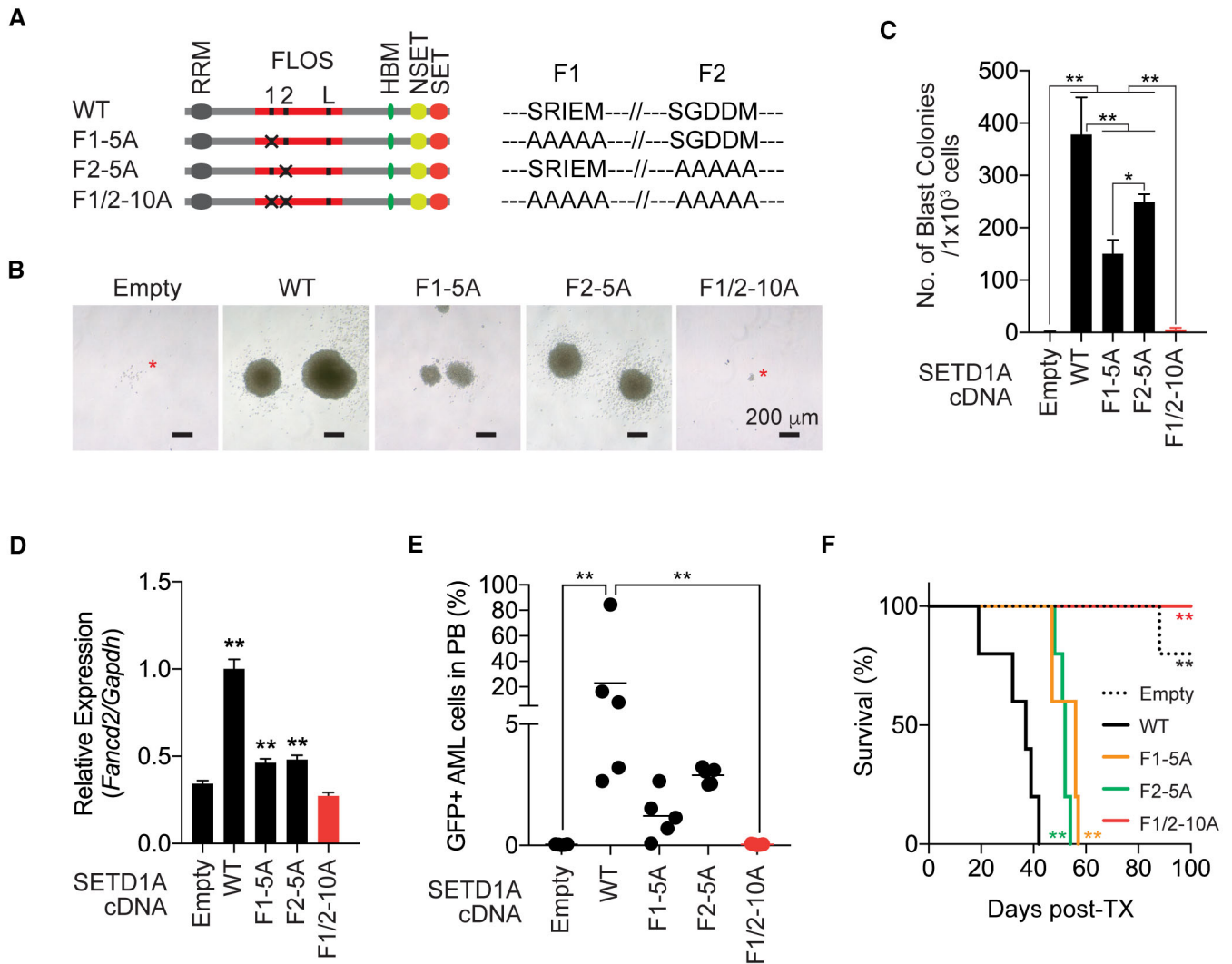


Figure 2. Dual motifs-targeting in FLOS domain synergistically inhibits the SETD1A function in leukemia cells.

- A The human SETD1A alanine-mutated constructs are shown as schematic illustrations. 1, 2 and L on the FLOS domain represent the F1 motif, the F2 motif and NLS, respectively.
- B *Setd1a^{fl/fl};CreER* MLL-AF9 leukemia cells were transfected with SETD1A mutant constructs. The cells were treated with tamoxifen and then assayed for colony-forming potential. Representative images of colonies are shown. Scale bars: 200 μ m.
- C Number of blast colonies in CFU-C assay (B) is shown. Colony formation assay with three biological replicates was performed.
- D Relative expression of *Mlh1* in SETD1A rescued MLL-AF9 leukemia cells. qRT-PCR with three biological replicates was performed.
- E Percentage of GFP+ SETD1A mutant-expressing leukemia cells in PB was analyzed at 3-week post-transplant ($n = 5$ /group).
- F Survival of recipient mice harboring SETD1A mutant-expressing leukemia cells after transplantation (TX) was plotted ($n = 5$ /group).

Data information: In (C, D), data are presented as mean \pm SD. ** $P \leq 0.01$, * $P \leq 0.05$ (Student's t-test for (D); One-way ANOVA for (C and E); Kaplan–Meier for (F)). Source data are available online for this figure.

bind with F1 and NSET domain, respectively (Fig 3B). To assess for direct protein–protein interactions between SETD1A and BuGZ/BUB3, we individually purified the SETD1A protein and BuGZ/BUB3 protein complex from insect cells. These purified proteins formed a protein complex *in vitro* and the alanine substitutions on F2, but not F1, disrupted the protein complex formation (Fig 3C). We also observed this protein interaction in MOLM-13 human leukemia cells (Fig EV3C). IP-MS analysis of SETD1A-binding proteins in MOLM-13 cells revealed that BuGZ is the most enriched protein

by SETD1A IP (Fig EV3D; Appendix Table S2). Since the F2 motif is also conserved in SETD1B, we also assessed SETD1B binding to BuGZ/BUB3. Unlike CCNK, BuGZ and BUB3 are associated with SETD1B in addition to SETD1A (Fig EV3E). To confirm the nuclear co-localization of the SETD1A/BuGZ/BUB3 complex, we visualized the subcellular localization of these three proteins in MLL-AF9 leukemia cells. SETD1A primarily localizes to interphase nuclei and is co-localized with both BuGZ and BUB3 (Fig 3D and E). All of these proteins were found to be expressed more in interphase nuclei

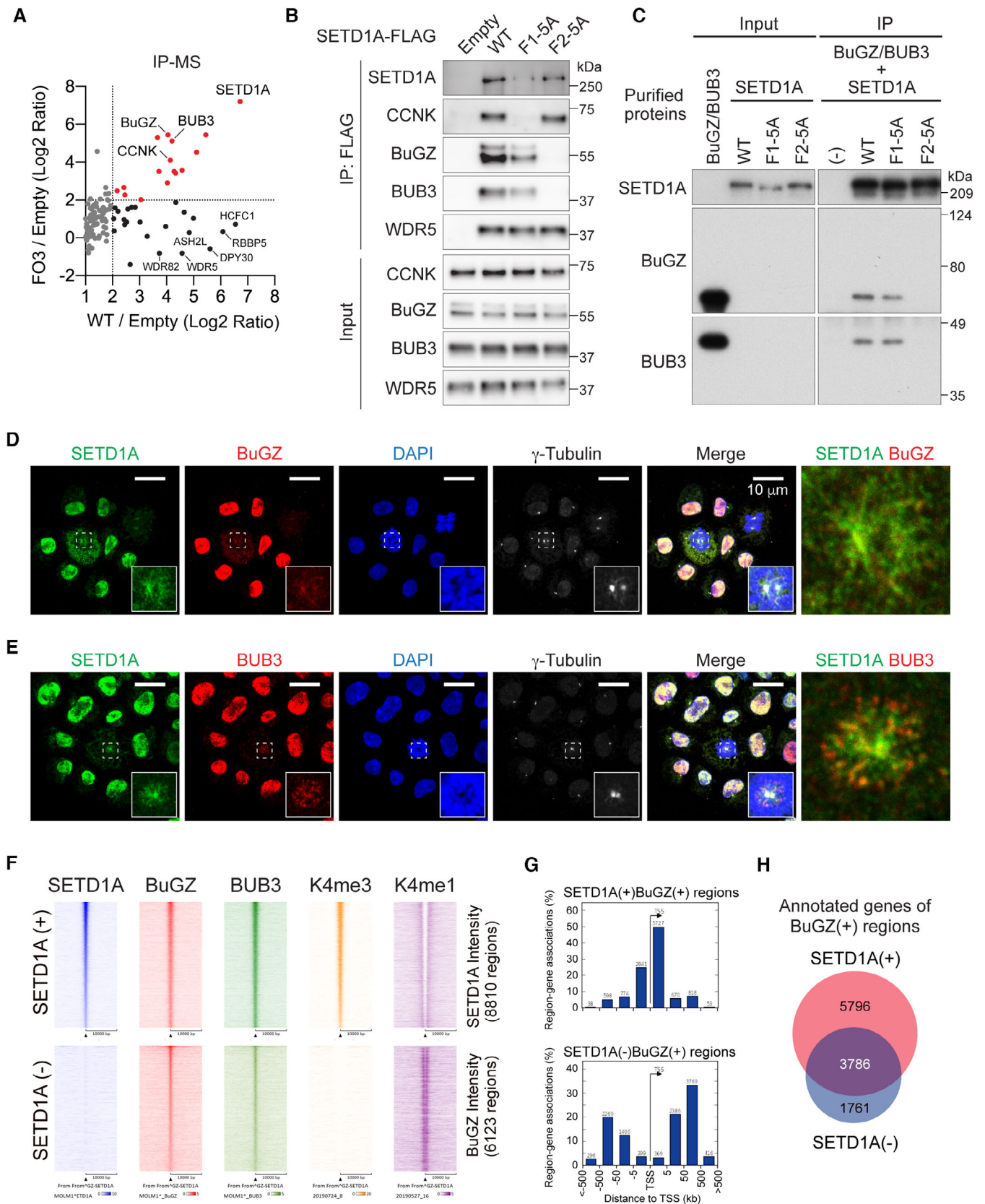


Figure 3.

Figure 3. BuGZ binds with SETD1A F2 motif.

- A SETD1A FLOS domain-binding proteins were identified by Co-IP-MS. Full-length SETD1A and FO3 fragments were used as baits and overlapped proteins (red dots) were identified as SETD1A FLOS domain-binding proteins.
- B 293 T cells were transfected with full-length SETD1A alanine mutants, and protein extracts were used for the Co-IP experiments.
- C BuGZ/BUB3 complex and SETD1A protein were independently purified from insect cells, and these samples were mixed *in vitro*. His-tagged SETD1A mutants were used for Co-IP.
- D, E Immunofluorescence images of SETD1A, BuGZ, and BUB3 in SETD1A-GFP fusion-expressing mouse MLL-r leukemia cells. Upper panels (D) show the co-localization with BuGZ and lower panels (E) show the co-localization with BUB3. Right panels show the co-localization in higher magnification of mitotic cells. Scale bars: 10 μ m.
- F MOLM-13 leukemia cells were examined by ChIP-seq analyses against HA-tagged SETD1A, BuGZ, BUB3, H3K4me3, and H3K4me1. Heatmaps of SETD1A(+)/BuGZ(+) regions (upper) and SETD1A(-)/BuGZ(+) regions (lower) are shown.
- G Distance to TSS from SETD1A(+)/BuGZ(+) regions (upper) and SETD1A(-)/BuGZ(+) regions (lower) is shown.
- H The overlap of annotated genes against the SETD1A(+)/BuGZ(+) and SETD1A(-)/BuGZ(+) peaks in ChIP-seq studies is shown.
- Source data are available online for this figure.

rather than in centromeres or gamma-tubulin-positive centrosomes of dividing leukemia cells (Fig 3D and E). To examine the co-localization of SETD1A, BuGZ, and BUB3 on chromatin, we performed ChIP-seq analysis with the human MOLM-13 leukemia cell line (Fig 3F). Consistent with the immunoprecipitation studies, both BuGZ and BUB3 broadly co-localize on SETD1A- and H3K4me3-positive transcriptional start sites (TSS; Figs 3F and G, and EV3F). Unexpectedly, BuGZ and BUB3 are also distributed on SETD1A-negative H3K4me1-positive enhancers (Figs 3F and G, and EV3F). Lists of annotated genes of SETD1A-positive and SETD1A-negative BuGZ-binding regions exhibit considerable overlap (Fig 3H). These results suggest that BuGZ/BUB3 complex may contribute to transcriptional activation of these genes by forming a complex with SETD1A/B on TSS and neighboring enhancers.

BuGZ has an indispensable function in leukemia cells

To examine the effects of BuGZ/BUB3 suppression in leukemia, we performed a CRISPR targeting of *BuGZ* in human MOLM-13 cells, U937 cells, and K562 cells expressing doxycycline (dox)-inducible Cas9 (iCas9). We detected strong toxicity from the BuGZ sgRNAs,

which target the zinc finger domain (sgBuGZ-1 and sgBuGZ-2; Figs 4A and EV4A–C). The *BuGZ*-knockout MOLM-13 cells showed apoptosis and cell cycle arrest at not only in G2/M phase but also in the G1 phase of the cell cycle (Fig 4B and C). The *BuGZ* knockout did not effect on either the protein expression of SETD1A/B or the chromatin distribution of SETD1A, but induced the accumulation of p-H3 (Ser10) and p53 (Figs 4D and EV4C). Consistent with the previous work, the *BuGZ* knockout reduced the BUB3 protein level (Fig EV4C; Toledo *et al*, 2014; Jiang *et al*, 2014a). While the dual knockout of *BuGZ* and *BUB3* has additive effect, the *BuGZ* knockout showed the stronger effect than the *BUB3* knockout on cell proliferation in MOLM-13 cells (Fig EV4C and D). To confirm this phenotype *in vivo*, we established doxycycline-inducible shRNAs to knock down either *Bugz* or *Bub3* in a mouse MLL-AF9 leukemia model (Figs 4E and EV4E). We observed a significant knockdown of SETD1A target genes as well as growth suppression after induction of shRNAs (Fig EV4F and G). *Bugz* knockdown showed the stronger effect on gene expression compared with *Bub3* knockdown (Fig EV4F). To evaluate the effects of *Bugz* knockdown *in vivo*, we transplanted the mouse *Bugz* shRNA-expressing leukemia cells to sublethally irradiated recipient mice. Doxycycline treatment of these

Figure 4. BuGZ regulates the cell proliferation without affecting the chromosomal stability in mouse MLL-r leukemia cells.

- A BuGZ sgRNAs with a GFP reporter were expressed in doxycycline-inducible Cas9 (iCas9)-expressing MOLM-13 (A), U937 (B) and K562 (C) leukemia cells. sgRNA for the essential replication gene RPA3 was used as a positive control for the competitive cell growth assay. Data from six biological replicates are shown.
- B Annexin V + DAPI- population in sgRNA-expressing MOLM-13 cells was analyzed at 5-day post-dox. Data from five biological replicates are shown.
- C Cell cycle distribution was analyzed by the 2 h pulse EdU-labeling method and 7AAD. Data were acquired by using flow cytometry. Data from five biological replicates are shown.
- D Distribution of SETD1A in sgRNA-expressing MOLM-13 cells was analyzed using Cut&Tag assay at 5-day post-dox.
- E Relative expression levels of *Bugz* were analyzed in *Bugz* shRNA-expressing mouse MLL-AF9 leukemia cells at day 3 post-treatment, respectively. Data from three biological replicates are shown.
- F Percentage of GFP+ shRNA-expressing leukemia cells in BM was analyzed at 3-week post-transplant ($n = 5$ /group).
- G Annexin V + DAPI- population in shRNA-expressing leukemia cells in BM was analyzed at 3-week post-transplant ($n = 5$ /group).
- H Survival of recipient mice harboring dox-inducible shRNA-expressing leukemia cells was plotted ($n = 8$ –10/group).
- I Images are shown of representative karyotypes in control, *Bugz*, and *Bub3* shRNA-expressing leukemia cells. Asterisks in images indicate the trisomy 15 in all cells analyzed. Red arrows indicate chromosomal aneuploidies. Representative images of sister-chromatid separation or aneuploidy in *Bub3* shRNAs are shown in lower panels.
- J Percentage of aneuploidy in control, *Bugz*, and *Bub3* shRNA-expressing leukemia cells are shown. Data from three independent shRNAs were plotted ($n = 3$ /group for *Ren* and *Bub3*, $n = 6$ /group for *BuGZ*).
- K RNA-seq analysis was performed with iCas9-MOLM-13 leukemia cells harboring *BuGZ* sgRNAs or *SETD1A* sgRNA at 6-day post-dox. The overlap of downregulated genes and annotated genes against the SETD1A(+)/BuGZ(+) peaks (Fig 3H) is shown. Two independent *Bugz* shRNAs were used.
- L The pathway analysis was performed with Enrichr tool using downregulated 719 genes between *BuGZ* knockdown cells and *SETD1A* knockdown cells in Fig 4J.

Data information: In (A–C and E–G), data are presented as mean \pm SD. $**P \leq 0.01$ (Student's *t*-test for (A, B, E–G and J); two-way ANOVA for (C); Kaplan–Meier for (H)). Source data are available online for this figure.

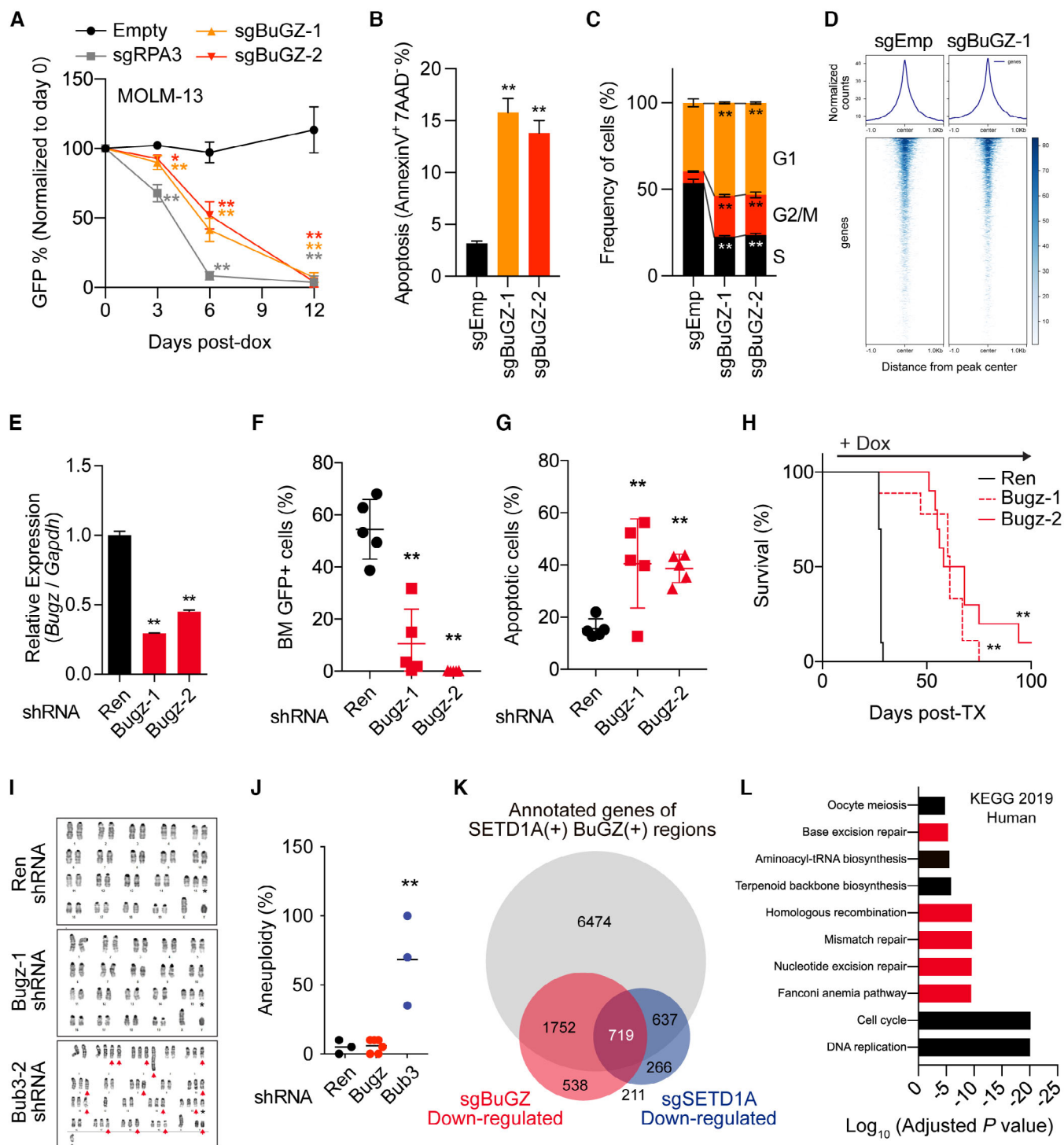


Figure 4.

mice resulted in a significant reduction of leukemia cells as well as the induction of apoptosis in their bone marrow (Fig 4F and G). Bugz shRNAs significantly extend the survival of these recipient mice (Fig 4H). Thus, BuGZ is indispensable in both mouse and human leukemia models. Although both BuGZ and BUB3 are mitotic regulators, we only detected the aberrant mitotic phenotypes, such

as aneuploidy as well as sister-chromatid separation, in *Bub3* knock-down cells (Figs 4I and J, and EV4H and I). This suggests that BuGZ might have a chromosome segregation-independent function in leukemia cells. To assess changes in gene expression we performed RNA-seq on *BuGZ* or *SETD1A* knockdown iCas9-MOLM-13 leukemia cells and detected suppression of 719 genes which are classified as

the downstream genes of both SETD1A and BuGZ from the overlap with SETD1A and BuGZ ChIP-seq results (Figs 4K and EV4J and K). Both SETD1A and BuGZ are required for the transcriptional upregulation of genes associated with the cell cycle and DNA repair, which have previously been described as downstream targets of SETD1A (Fig 4L). In contrast, SETD1A-independent BuGZ targets are associated with RNA transport or the spliceosome (Fig EV4L). Taken together, our data suggest that BuGZ acts as a transcriptional regulator in leukemia cell survival.

BuGZ GLEBS motif binds the MELT-like SETD1A F2 motif

BuGZ encodes a protein containing a N-terminal C2H2 zinc finger motif and the BUB3-binding Gle2-binding-sequence (GLEBS) motif at the C terminus. BUB3 also binds with other GLEBS-motif-containing proteins BUBR1 and BUB1 and is recruited to the KNL1 MELT repeats to form ternary complexes for chromosome segregation in the mitotic phase (Bolanos-Garcia & Blundell, 2011). Interestingly, the F2 motif has a similar structure to MELT repeats (Fig 5A). Phosphorylation of the threonine residue on MELT repeats contributes to BUB3 binding on KNL1 (Vleugel et al, 2015). The serine residue (S586) on SETD1A F2 motif seems to be homologous to this threonine residue on MELT repeats. While a peptide containing S586 was not detected from our IP-MS data that used MOLM-13 cell line, perhaps due to the high density of acidic residues at the F2 motif, the phosphorylation of S586 has been reported in the Jurkat T-cell leukemia cell line from the phosphorylated peptides enriched library (Fig EV1G and PhosphoSitePlus, ID: 23143264). To examine the function of the SETD1A S586, we replaced the serine residue (S586) with either alanine (S586A) or aspartic acid (S586D), to mimic the unphosphorylated or phosphorylated state, respectively (Fig 5B). We also constructed a SETD1A mutant that has alanine substitutions on both F2 and the MELT-like sequence (F2-8A; Fig 5A and B). These alanine substitution mutants disrupted the binding with both BuGZ and BUB3 and failed to rescue the colony-forming ability of *Setd1a*-deficient AML cells (Figs 5B and EV5A). Consistent with the *in vitro* data, the cDNA rescue experiment indicated that S586A, but not S586D mutant inhibits leukemia propagation *in vivo* (Fig 5C). Although the upstream effectors of this pathway have yet to be identified, these results suggest that a post-translational modification of SETD1A F2 may regulate leukemia growth by facilitating complex formation with BuGZ/BUB3.

CRISPR-tiling screening is a useful method to evaluate the functional region on proteins (Shi et al, 2015). To further characterize BuGZ domain structure, we performed a functional CRISPR-tiling screening and detected a strong dependency on the zinc finger domain as well as the GLEBS domain (Fig 5D). We then performed immunoprecipitation studies with various BuGZ and BUB3 mutants (Figs 5E and F, and EV5B and C). Alanine substitution or deletion of the BuGZ GLEBS motif was found to severely diminish the binding to SETD1A. The BUB3 sites S18/S19 (phosphorylated by BUB1), R202 (corresponding to ScBub3 R217, MELT repeat interaction site), Y207 (PKM2 interaction site), and F221/K222 (nuclear and kinetochore localization; Appendix Fig S5B; Niikura et al, 2010; Jiang et al, 2014b; Zhu et al, 2015) were examined, but only alanine substitutions of F221/K222 disrupted the binding of BUB3 with both BuGZ and SETD1A (Fig 5F).

To confirm the function of the SETD1A-binding site on BuGZ in MLL-r leukemia cells, we performed a cDNA rescue experiment with two mouse *Bugz*-specific sgRNAs and the GLEBS-motif-mutated human BuGZ-expression vector (Figs 5G and EV5D). The introduction of human wild-type BuGZ, but not the GLEBS mutant, reinstated cellular proliferation in mouse MLL-r leukemia cells expressing *Bugz* sgRNA (Fig 5G). Additionally, we observed a downregulation of *Fancd2* in the leukemia cells expressing GLEBS mutant (Fig 5H). Furthermore, co-immunoprecipitation studies demonstrated a reduced binding affinity between GLEBS mutant and SETD1A, BUB3, and CCNK (Fig 5I). While the BuGZ is indispensable for BUB3 stabilization, the introduction of exogenous BUB3 failed to rescue the knockout effect in *Bugz*-knockout cells (Appendix Fig S5E). Furthermore, we established a stable mouse leukemia cell line expressing the BUB3 F221/K222 mutant. This mutant exhibited decreased protein level and was unable to rescue the effects of BUB3 knockout (Fig EV5F and G). These findings reveal that SETD1A interacts with BuGZ/BUB3 complex and the GLEBS motif of BuGZ is essential for the complex formation and the leukemia cell growth.

A disordered region of BuGZ promotes SETD1A interaction and leukemia cell growth

The C terminus of the BuGZ protein is comprised of GLEBS motif as well as an intrinsically disordered region (IDR), which is known to be important for phase transition and microtubule formation

Figure 5. F2 motif is the BuGZ/BUB3-binding site on SETD1A.

- Amino acid sequences of human KNL MELT repeats and F2 motif were compared. F2-5 and F2-8 are regions that are substituted by alanine in (B).
- 293 T cells were transfected with SETD1A F2 alanine mutant constructs. Myc-DDK-tagged SETD1A was used for Co-IP.
- Percentage of GFP+ SETD1A mutant-expressing leukemia cells in PB was analyzed at 3-week post-transplant ($n = 4-5/\text{group}$).
- CRISPR-tiling screening for the mouse *Bugz* gene in MLL-AF9 leukemia cells. One hundred and fifty-two sgRNAs were designed and transfected to Cas9-expressing MLL-AF9 leukemia cells.
- 293 T cells were transfected with BuGZ mutants. Myc-DDK-tagged BuGZ mutants illustrated on Appendix Fig S5B were used for Co-IP.
- 293 T cells were transfected with BUB3 mutants and protein extracts were used for the Co-IP experiments. Myc-DDK-tagged BUB3 mutants illustrated on Appendix Fig S5C were used.
- Human BuGZ-expressing mouse MLL-AF9 leukemia cells were transfected with mouse *Bugz*-targeting sgRNAs, and the competitive cell growth assay was performed. Data from three biological replicates are shown.
- Relative expression level of *Fancd2* was analyzed in BuGZ-rescued MLL-AF9 leukemia cells at 3-day post-dox. Data from three biological replicates are shown.
- 293 T cells were transfected with Myc-DDK-tagged human BuGZ constructs, and the protein extracts were used for a Co-IP experiment.

Data information: In (G, H), data are presented as mean \pm SD. ** $P \leq 0.01$, * $P \leq 0.05$. (One-way ANOVA for (C); two-way ANOVA for (G); Student's *t*-test for (H)). Source data are available online for this figure.

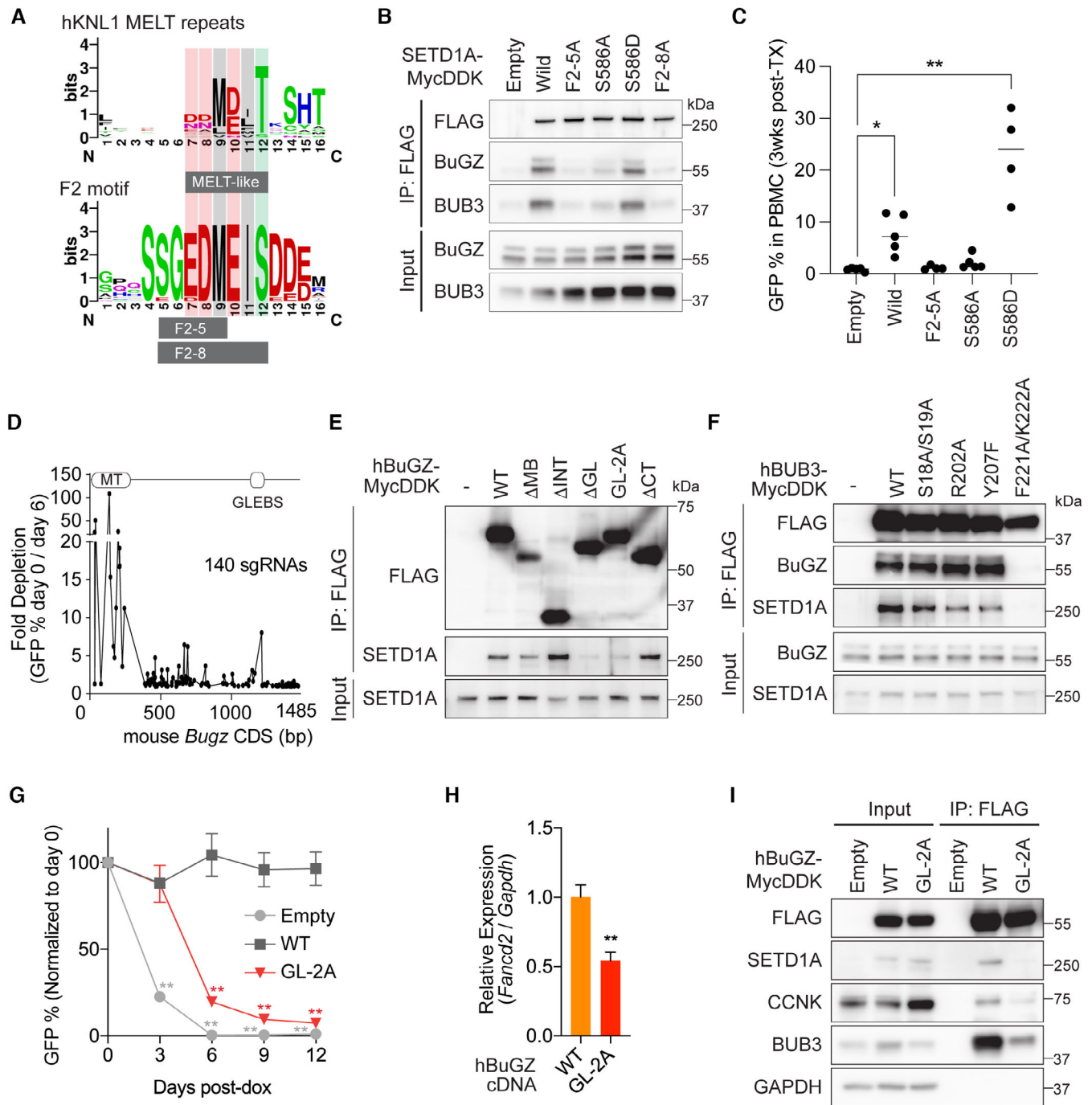


Figure 5.

(Fig 6A; Jiang *et al*, 2015). Both human and mouse *BuGZ* genes have common transcriptional variants at IDR (isoform A – C). A previous study suggested that isoform C functions as a transcriptional regulator while isoform B is involved in regulation of the spindle assembly checkpoint during mitosis in embryonic stem cells (Fang *et al*, 2018). One tempting explanation for the strikingly diverse roles of *BuGZ* in regulating both mitotic chromosome alignment and gene expression in leukemia cells is that different *BuGZ* isoforms are responsible for these disparate functions. In contrast,

our RNA-seq data showed the dominant expression of isoform B in both mouse and human leukemia cells (Fig 6B; Appendix Fig S1A–C). To address the functional difference of these isoforms in leukemia cells, we performed immunoprecipitation and rescue assays with various *BuGZ* isoforms. Surprisingly, all isoforms of *BuGZ* were able to bind to SETD1A and rescue the phenotypes of *Bugz* sgRNA knockdown (Fig 6C and D). This suggests that *BuGZ* dependency in the context of leukemia cells is isoform independent.

We detected limited effects induced by sgRNAs which target the C terminus of BuGZ in Fig 5D, but a CRISPR-tiling screening is likely not suitable for the analysis of a nonstructured amino acid sequence (He *et al*, 2019). Conserved phenylalanine (F) and tyrosine (Y) residues in the BuGZ IDR are required for phase transition (Jiang *et al*, 2015). To determine the function of the BuGZ IDR in leukemia cells, we designed a BuGZ mutant (FY>S) to disrupt the IDR at the C terminus (Appendix Fig S1D). The FY>S mutant was stably

expressed in mouse MLL-r leukemia cells, but the mutant failed to rescue the growth defect in Bugz sgRNA-expressing cells (Fig 6E). The FY>S mutant also failed to bind SETD1A and BUB3, despite the intact GLEBS motif (Fig 6F). To further evaluate the importance of the BuGZ IDR in SETD1A-BuGZ complex formation, 1,6-hexanediol was used to disrupt the weak hydrophobic interactions through the IDR. Interestingly, administration of 1,6-hexanediol resulted in the disruption of SETD1A-BuGZ binding to a greater extent than the

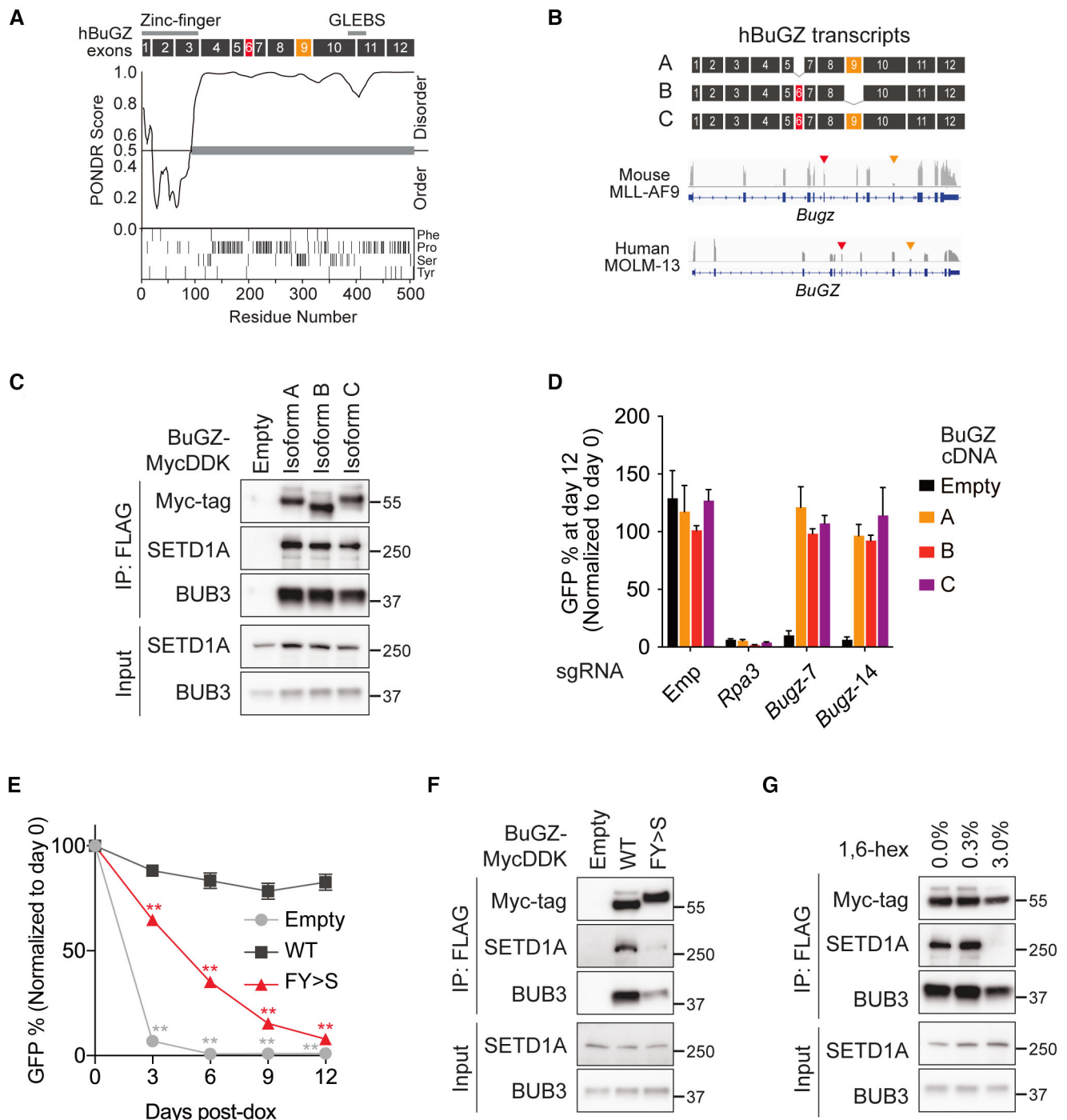


Figure 6.

Figure 6. BuGZ IDR regulates the formation of protein complex with SETD1A.

- A Sequence features of human BuGZ. VSL2 predictor was used to predict a natural disordered region. The line at 0.5 distinguishes disordered and ordered predictions.
- B The BuGZ splicing variants are shown as schematic illustrations. Histograms depict BuGZ mRNA peaks in mouse MLL-AF9 leukemia cells and human MOLM-13 leukemia cells. Arrowheads indicate exon 6 (red) or exon 9 (orange).
- C 293 T cells were transfected with Myc-DDK-tagged BuGZ variant constructs, and the protein extracts were used for the Co-IP experiment.
- D BuGZ variant-expressing mouse iCas9 MLL-AF9 leukemia cells were transfected with mouse *Bugz*-targeting sgRNA. Percentage of GFP+ sgRNA-expressing cells were analyzed at 0- and 12-day post-dox. Data from three biological replicates are shown.
- E Human BuGZ-expressing mouse MLL-AF9 leukemia cells were transfected with mouse *Bugz*-targeting sgRNAs, and the competitive cell growth assay was performed. Data from three biological replicates are shown.
- F 293 T cells were transfected with Myc-DDK-tagged BuGZ constructs, and the protein extracts were used for a Co-IP experiment.
- G 293 T cells were transfected with Myc-DDK-tagged wild-type BuGZ, and the cells were treated with 1,6-hexanediol for 3 h. The protein extracts were used for the Co-IP experiments.

Data information: In (D–E), data are presented as mean \pm SD. $^{**}P \leq 0.01$. (Two-way ANOVA for (E)).
Source data are available online for this figure.

BuGZ-BUB3 binding (Fig 6G). Our results suggest that the BuGZ IDR promotes leukemia growth and that hydrophobic interaction may be an important step in the SETD1A-BuGZ complex formation.

The noncanonical SETD1A activation in G1/S phase supports intact cell cycling

BuGZ interaction with the SETD1A FLOS domain suggested a potential role for the noncanonical SETD1A-BuGZ complex in the mitotic phase, but our previous work indicated SETD1A is an important mediator of gene expression control in the G1/S phase (Hoshii *et al*, 2018). To examine the role of SETD1A in different cell cycle phases, we established a cell-cycle-dependent SETD1A expression system by applying a Cdt (30–120) or Geminin (1–110)-tag, which are both derived from the FUCCI cell cycle indicator, to the SETD1A-GFP fusion construct (Fig 7A). The SETD1A-GFP fusion protein (SG) expresses in both G1/S and G2/M phases (Fig 7B). The Cdt-tagged SETD1A-GFP fusion protein (SGC) is mainly expressed in G1/S phase (Fig 7B). In contrast, the Geminin-tagged SETD1A-GFP fusion protein (GSG) is observed in G2/M phase (Fig 7B). Interestingly, while the original FUCCI reporter construct did not show any effect, the cell-cycle-dependent expression of SETD1A in G1/S (Cdt-tag), but not G2/M phase (Geminin-tag), completely rescued the defective colony-forming ability of *Setd1a* deficient leukemia cells (Fig 7C). Additionally, SETD1A-Cdt expression restored the transcriptional levels of SETD1A downstream genes, which are associated with DNA repair and aminoacyl-tRNA biosynthesis and are downregulated in *BuGZ*-knockout leukemia cells (Fig 7D and E). These results demonstrate that spatiotemporal activation of the noncanonical SETD1A complex in the G1/S phase is crucial for the expression of DNA damage response genes and cell cycle progression in leukemia cells.

Discussion

FLOS is an essential functional domain of SETD1A

Our results demonstrated that the SETD1A FLOS domain can rescue SETD1A-deficient MLL-AF9 leukemia cells. In our previous study, we identified the noncatalytic function of SETD1A, but it was still unclear how the SETD1A FLOS domain can bind to chromatin without a specific DNA-binding domain or histone reader domain (Hoshii *et al*, 2018). Here, this study indicates that the FLOS

domain is sufficient for retention in the nucleus and to maintain the activity for the noncanonical SETD1A function in leukemia cells. The interaction of SETD1A with BuGZ/BUB3 and CCNK through the FLOS domain is required for the maintenance of SETD1A's noncatalytic function. Although the FLOS domain is an essential component of SETD1A required for MLL-r leukemia cell proliferation, our rescue experiments indicated that the N-terminus region is also indispensable for normal SETD1A function in leukemia cells. The SETD1A N terminus recruits WDR82, a specific subunit in the SETD1A/SETD1B complex (Lee & Skalnik, 2008). WDR82 is a seven WD40 repeat-containing protein and has structural similarity with BUB3 as well as WDR5. WD40-repeat proteins are essential components for multiprotein complex assemblies to regulate pre-mRNA processing, transcription, cell cycle and various other processes (Schapira *et al*, 2017). These components might partially compensate each other by directly replacing the functions of other proteins or by forming similar protein complexes. Further multiple mutagenesis studies in the N-terminus region and FLOS domains on SETD1A could provide a mechanistic insight into the chromatin-binding module of the SETD1A protein. The functional relevance of SETD1A in embryonic stem (ES) cells and embryogenesis is well characterized by using genetically engineered murine ES cells (Bledau *et al*, 2014; Sze *et al*, 2017). The dispensable role of SETD1A SET domain and the indispensable role of SETD1A itself in undifferentiated ES cells suggest the essential role of the FLOS domain function in the maintenance of the ES cell growth. Importantly, BuGZ is also reported as an essential factor for the CDK activity and DNA damage response in ES cells (Fang *et al*, 2018). MLL-r leukemia stem cells have the similar properties with ES cells, suggesting a common role of BuGZ-SETD1A FLOS-CCNK axis in the maintenance of cell growth in leukemia and ES cells.

Functional suppression of multiple SETD1A FLOS domains has a synergistic effect on leukemia cells

We used the retrovirus-based expression system for *SETD1A* or *BuGZ* cDNA in leukemia cells, but it did not reproduce the expression levels of the endogenous genes; therefore, the loss of function of mutant proteins might be underestimated by the overexpression of these proteins. However, a rescue experiment featuring truncated and mutated forms of SETD1A demonstrated that the orchestrating of this protein complex is crucial for maintaining intact cell proliferation of MLL-r leukemia cells. Both F1 and F2 motifs are located

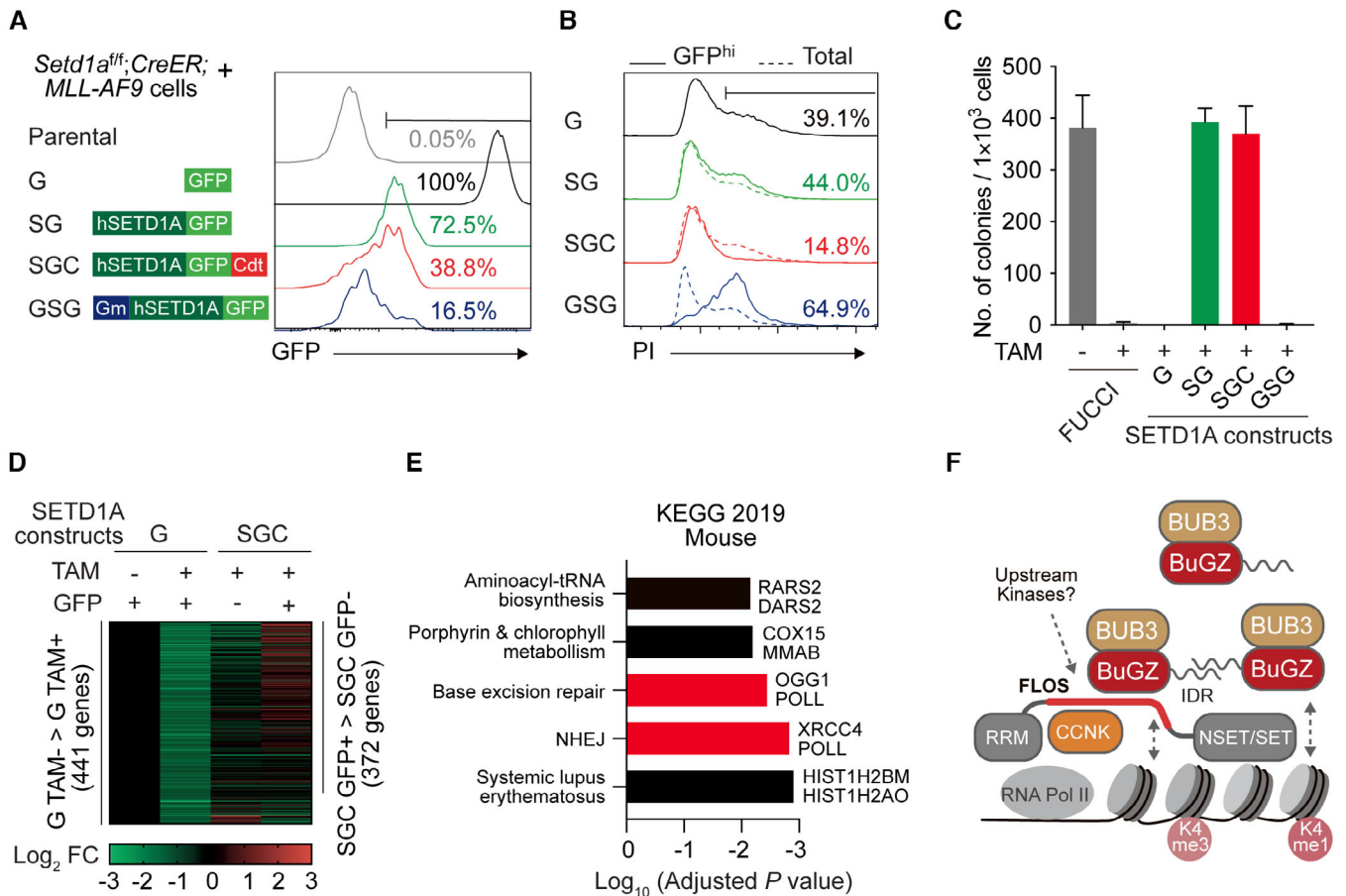


Figure 7. SETD1A-CCNK-BuGZ complex formation in G1 phase of the cell cycling.

A, B The single-cell clones of SETD1A-GFP-Cdt (SGC) or Geminin-SETD1A-GFP (GSG) fusion-expressing *Setd1a^{fl/fl}; CreER* MLL-AF9 leukemia were established and their GFP fluorescence (A) and DNA contents (PI) were analyzed by flow cytometry (B). Solid lines and dotted lines in (B) indicate the DNA contents in GFP-gated cells and total cells, respectively.

C SETD1A-GFP fusion-expressing *Setd1a^{fl/fl}; CreER* MLL-AF9 leukemia cells were treated with tamoxifen and then assayed for colony-forming potential. The FUCCI reporter (mCherry-Cdt1(30–120)/Cittrine-Geminin(1–110))-expressing cells were used as control. Data from three biological replicates are shown. Data are presented as mean ± SD.

D SGC-expressing *Setd1a^{fl/fl}; CreER* MLL-AF9 leukemia cells were treated with tamoxifen, and RNA-seq analysis was performed with sorted cells. Data are shown as log₂ fold change over tamoxifen nontreated GFP-expressing control.

E Genes that show the downregulation in *Setd1a* knockout cells and rescued in SGC-expressing GFP+ *Setd1a* knockout cells were analyzed with Enrichr tool.

F A schematic model for the SETD1A-BuGZ complex. The FLOS domain of SETD1A binds to CCNK, BuGZ, and chromatin via independent regions. BuGZ may serve as an upstream signal from kinases and enhancers in leukemia cells.

Source data are available online for this figure.

close together at the N terminus of the FLOS domain, but these motifs recruit the independent binding partners. The synergistic effect of the multiple mutations on the SETD1A FLOS domain indicates the cooperative role of these motifs and their binding partners on SETD1A-mediated transcriptional activation. Broad distribution of BuGZ and BUB3 on chromatin over the H3K4me1-positive enhancer-like regions and SETD1A/H3K4me3-positive TSS suggests that the BuGZ/BUB3 complex may act as an upstream regulator of noncanonical SETD1A function. Since the F2-mutated SETD1A showed higher colony-forming ability than the F1 mutant in the rescue study, this suggests that F1 may have more critical role in SETD1A-dependent transcriptional activation than F2. Because of the hierarchical relationship between F1-CCNK and F2-BuGZ, the

effect of F2-single mutation may have been partly compensated by a downstream F1-CCNK activity. Although the precise targeting of F1 would be more specific for cancer therapy, the multidomain targeting strategy will provide a new tool to overcome resistance and could potentially be utilized to target noncancerous diseases. Recently, loss-of-function variants of SETD1A were found in patients with schizophrenia and other developmental disorders, and the synaptic function was evaluated in mouse models (Takata et al, 2014; Singh et al, 2016; Nagahama et al, 2020; Kummeling et al, 2021; Wang et al, 2022). Variants in SETD1B are also associated with neurodevelopmental disorder (Hiraide et al, 2018; Weerts et al, 2021). Interestingly, one third of these mutations in SETD1A were found at the exons encoding the FLOS domain, yet it remains unclear

whether truncated fragments result in disease. Since our rescue experiment showed SETD1A N-terminal truncated proteins retain partial function, the expression of truncated proteins may disturb the SETD1A function after a long latency in brain function. Our study revealed that the development of SETD1A drugs that target multiple functions on the FLOS domain would have potential efficacy in treating many diseases associated with SETD1A mutations in addition to diseases such as AML that are dependent on the SETD1A pathway.

BuGZ is a functional component of the SETD1A complex

In this study, we identified the role of BuGZ as a potential regulatory subunit for a nonenzymatic SETD1A function in the nucleus at the G1/S phase of cell cycle (Fig 7F). BuGZ and BUB3 are abundantly expressed in many cell lines *in vitro* as well as cancer cells *in vivo* (Toledo et al, 2014; Wan et al, 2015). On account of the zinc finger domain of BuGZ, several studies have implicated that BuGZ may act as a transcription factor, but the actual function of BuGZ in transcription is unclear (Wan et al, 2015; Fang et al, 2018). BuGZ binds to the kinetochore protein BUB3 through the GLEBS motif, implicating a role for BuGZ in chromosome segregation (Toledo et al, 2014; Jiang et al, 2014a, 2015). While we did not directly evaluate the DNA damage itself, we observed the reduced expression of multiple DNA damage response genes in both SETD1A- and BuGZ-knockout cells. Similar to BuGZ, recent studies revealed the essential role of BOD1L as a binding partner of SETD1A, in the DNA damage response (Higgs et al, 2018; Bayley et al, 2022). BOD1 and BOD1L were originally identified as kinetochore proteins; however, they were also found to form complexes with SETD1B and SETD1A, respectively (Porter et al, 2007; Wang et al, 2017; Higgs et al, 2018). Bifunctional roles of chromosome segregation factors in transcriptional regulations are also reported in cohesin complex subunits, and these factors are frequently mutated in leukemia patients (Kon

et al, 2013; Yan et al, 2013). Thus, BuGZ might serve other purposes independent of chromosomal segregation and could participate in transcriptional regulation through its interaction with SETD1A/B.

Although the upstream kinase has yet to be identified, the site-specific inhibition of F2 with small molecules could represent a more specific mechanism for disruption of the SETD1A-BuGZ pathway and would be a SETD1A/BuGZ-specific targeting strategy that could be utilized for cancer therapy. SETD1B also encodes the MELT-like sequence and binds BuGZ/BUB3; however, the function remains unclear. In contrast to the conservation of F1/2 motifs, phosphorylation sites are enriched only in the FLOS domain of SETD1A (Fig EV1G). Thus, post-translational modifications may contribute the functional difference between SETD1A and SETD1B. Further analysis using SETD1B-dependent cells would be required to evaluate the role of the SETD1B-BuGZ axis. BuGZ was also reported as an inducer of phase transition for microtubule assembly in the mitotic phase (Jiang et al, 2015). Interestingly, the BuGZ IDR mutant that failed to induce phase transition also failed to bind with SETD1A and did not support MLL-r leukemia cell proliferation. As previously reported, transcriptional factors and co-activators, including OCT4, BRD4, MED1, or ENL, also encode IDRs that are functionally important for phase transition, which is thought to promote efficient transcriptional regulation (Boija et al, 2018; Sabari et al, 2018; Wan et al, 2020). In this study, our data suggest that the BuGZ IDR may have an essential role for the transcriptional activation. Regarding the potential function of BuGZ in transcriptional regulation, we observed the broad chromosomal distribution of BuGZ. These results suggest that BuGZ could be important for facilitating the proper chromosomal dynamics between TSS and enhancer regions, possibly through binding with SETD1A/B. It is still unclear if and how BuGZ directly modulates the structure or functions of SETD1A complex on chromatin. Further structural analysis of the SETD1A FLOS domain will be required to elucidate the exact mechanism of how this complex regulates chromatin dynamics and gene expression.

Materials and Methods

Reagents and Tools table

Reagent/Resource	Reference or Source	Identifier or Catalog Number
Experimental models		
<i>List cell lines types etc. Indicate the species when appropriate.</i>		
C57BL/6j (M. musculus)	Jackson Lab	IMSR_JAX:000664
C57BL/6Jcl (M. musculus)	CLEA Japan	N/A
293T (Human, cell line)	ATCC	CRL-3216
Plat-E (Human, cell line)	Morita et al (2000)	N/A
MOLM-13 (Human, cell line)	DSMZ	ACC-554
K562 (Human, cell line)	ATCC	CCL-243
U937 (Human, cell line)	ATCC	CRL-1593.2
MLL-AF9 transduced AML (Mouse, cell line)	Hoshii et al (2018)	N/A
Setd1af1/fl;CreER MLL-AF9 transduced AML (Mouse, cell line)	Hoshii et al (2018)	N/A
Recombinant DNA		
pMSCV-hSETD1A-MycDDK-ires-GFP	Hoshii et al (2018)	N/A
pCMV-hBuGZ-MycDDK	Origene	RC223933

Reagents and Tools table (continued)

Reagent/Resource	Reference or Source	Identifier or Catalog Number
pCMV-hBUB3-MycDDK	Origene	RC200376
TRIN miR-E	Laboratory of Scott Lowe	N/A
RT3GEPiR miRE	MSKCC RNAi Core	N/A
pMSCV-ires-tdTomato	Laboratory of Scott Armstrong	N/A
pMSCV-ires-hyg	Laboratory of Scott Armstrong	N/A
lenti-sgRNA-hygro	Addgene	#104991
pLKO5.sgRNA.EFS.GFP	Addgene	#57822
pLKO5.sgRNA.EFS.tRFP657	Addgene	#57824
pMD2.G	Addgene	#12259
psPAX2	Addgene	#12260
pCW-Cas9	Addgene	#50661
ES-FUCCI	Addgene	#62541
Antibodies		
<i>Include the name of the antibody, the company (or lab) who supplied the antibody, the catalogue or clone number, the host species in which the antibody was raised and mention whether the antibody is monoclonal or polyclonal. Please indicate the concentrations used for different experimental procedures.</i>		
Rabbit anti-Myc-tag	Cell Signaling	#2276
Rabbit anti-GAPDH	Cell Signaling	#2118
Rabbit anti-BuGZ	Sigma	HPA017013
Rabbit anti-BuGZ	Novus Biologicals	NBP2-39010
Rabbit anti-BUB3	Abcam	ab133699
Rabbit anti-SETD1A	Cell Signaling	#61702
Rabbit anti-SETD1A	Abcam	ab70378
Rabbit anti-CCNK	Bethyl	A301-939A
Rabbit anti-WDR5	Abcam	ab178410
Rabbit anti-gamma-Tubulin	Abcam	ab11316
Rabbit anti-FLAG-tag	Cell Signaling	#14793
Rabbit anti-HA-tag	Cell Signaling	#3724
Rabbit anti-WDR82	Abgent	AP4812a-ev
Rabbit anti-phospho-H3	Cell Signaling	#53348
Rabbit anti-H3	Abcam	ab1791
Mouse anti-TP53	Santa Cruz	sc-126
Chicken anti-GFP	Aves Labs	GFP-1010
Other		
Illumina NexSeq 500	Illumina	
Illumina HiSeq 1500	Illumina	
Illumina NovaSeq 6000	Illumina	
SONY SH800	SONY	
BD FACSAria	BD Bioscience	
CytoFLEX	Beckman Coulter	

Methods and Protocols

Cell lines and cell culture

293 T, Plat-E cell lines were maintained in DMEM medium containing 10% fetal bovine serum (FBS) and 1% penicillin–streptomycin. K562 cell line was maintained in IMDM medium containing

10% FBS and 1% penicillin–streptomycin. U937 and MOLM-13 cell lines were maintained in RPMI1640 medium containing 10% FBS and 1% penicillin–streptomycin. Cells were maintained in a humidified incubator at 37°C, 5% CO₂. Cell line authentication for human cell lines was performed by STR profiling (Appendix Fig S2). Murine MLL-r leukemia cells were established previously (Hoshii *et al*,

2018). Murine MLL-r leukemia cells were maintained in RPMI1640 medium containing 10 ng/ml rMLL3, 10% FBS and 1% penicillin–streptomycin. For the 1,6-hex treatment, 293 T cells were treated with 0.3% or 3% 1,6-Hexanediol (Sigma) for 3 h.

shRNA

TRIN miR-E retroviral vector was obtained from Scott Lowe's laboratory as described previously. Mouse Bugz and Bub3 shRNA were designed previously (Fellmann et al, 2013; Appendix Table S3). Plat-E cells were transiently transfected with retroviral vector plasmids using X-tremegene HP (Roche), and culture supernatants containing retroviruses were collected 48 h after transfection. For the doxycycline-inducible TRIN miR-E shRNA system, we used the single-cell clone of MLL-AF9 leukemia cells expressing rtTA. The cells were selected with 1 mg/ml G418. The cells were treated with 1 µg/ml doxycycline and analyzed at 3-day post-doxycycline. Karyotyping was performed at the Molecular Cytogenetics Core Facility in MSKCC.

CDNA mutagenesis and rescue experiment

Myc- and DDK-tagged human SETD1A (NM_014712) cDNA in pMSCV-ires-GFP vector was constructed previously (Hoshii et al, 2018). The deletion mutants of SETD1A were constructed by inverse PCR. Myc- and DDK-tagged human *BuGZ* cDNA (NM_003457) and *BUB3* cDNA (NM_004725) were obtained from Origene. Site-directed mutagenesis and deletion mutagenesis were performed by inverse PCR, with 5'-phosphorylated primers and CloneAmp HiFi PCR premix (Clontech). Protein expression from each mutated cDNA was confirmed by western blotting. For the cDNA rescue experiment, *BuGZ* cDNA was cloned into pMSCV-ires-tdTomato vector, and retroviruses were generated in Plat-E cells. For the multiple amino acid substitutions for the BuGZ disordered region, mutated cDNA fragments were generated via gBlocks Gene Fragments (IDT) and inserted into the expression vectors. cDNA expression vectors were transduced into mouse *Setd1a*^{fllox/fllox}; *CreER*^{T2}; *MLL-AF9* leukemia cells or mouse Cas9-expressing *MLL-AF9* leukemia cells. GFP or tdTomato-positive cells were sorted by FACSAia (BD) or SH800S (SONY). Single-cell clone was established by limiting dilution. For the cell-cycle-dependent SETD1A expression, we first removed the ires sequence from pMSCV-SETD1A-ires-GFP vector and added the -Cdt-tag or Geminin-tag, which are both derived from ES-FUCCI vector (addgene #62451). We also constructed a pMSCV-FUCCI-ires-HP vector and used as a control vector for a colony-forming assay.

Immunoprecipitation/mass spectrometry

293 T cells were cultured in 6-well plates and transfected with X-tremegene HP transfection reagent. Cells were harvested at 72-h post-transfection and lysed with 1 × cell lysis buffer (Cell signaling) containing proteinase inhibitors (Sigma). Lysis was completed via ultrasonication, and proteins were quantified with BCA Protein Assay Kit (Thermo Fisher Scientific). Four hundred µg of protein lysate was loaded on prewashed FLAG M2 affinity gel (Sigma) and incubated overnight. After three TBS washes, we eluted the FLAG-tagged protein with 40 µl of TBS containing 3 × FLAG peptides (ApexBio). Binding proteins were analyzed by western blotting or mass spectrometry. For immunoprecipitation of endogenous proteins, protein lysate was mixed with Rabbit IgG or antibodies against SETD1A, BuGZ or BUB3, and incubated overnight.

The lysate was loaded on prewashed Protein A/G Dynabeads, incubated for 2 h at 4°C, and then eluted by 2 × SDS sample buffer. Baculoviruses expressing FLAG-SETD1A, FLAG-ZNF207, and His-Bub3 were prepared as described (Nakadai et al, 2023). HF cells were co-infected with baculovirus expressing His-Bub3 and FLAG-ZNF207, and His-Bub3/FLAG-ZNF207 heterodimer was purified in BC300-NP0.1D0.25 buffer containing 0.4 M Imidazole by TALON and then by M2 agarose. One microliter of SETD1 antibody (Abcam, #ab70378, 1 µg/µl) was bound to 2.5 µl of Dynabeads Protein A in PBS containing 0.05% Tween20 and then mixed with 1.2 pmol of FLAG-SETD1A and 4.6 pmol of His-Bub3/FLAG-ZNF207 heterodimer in BC300-NP0.1 buffer (Nakadai et al, 2023) containing 23 µg BSA (Roche, #10711454001) at 30°C for 90 min. After five times washes with BC300-NP0.1 buffer, the bound proteins were eluted by boiling with 30 µl of SDS sample buffer. Proteins were analyzed by western blotting. For mass spectrometry, SETD1A complexes pulled down using anti-FLAG or anti-SETD1A antibody were separated by SDS-PAGE. Proteins in gel were reduced using 10 mM dithiothreitol for 30 min followed by alkylation with 50 mM iodoacetamide for 30 min. Tryptic-digested peptides were purified with SDB-XC StageTip (Rappsilber et al, 2003, 2007). NanoLC-MS/MS was conducted using a Triple TOF 5600 (SCIEX) with an Ultimate 3000 RSLCnano (Thermo Fisher Scientific). Peptides were separated with an Acclaim PepMap RSLC (2 µm, 25 cm × 75 µm) column. MS data were subjected to a search against the Uniprot Human database with Protein Pilot V.5.0 (AB Sciex) for making a spectral library. Peak area of peptides was calculated by PeakView (AB Sciex).

Western blotting

Lysate was mixed with 4 × SDS sample buffer and beta-mercaptoethanol, and proteins were denatured by boiling for 5 min. Denatured proteins were separated on a 4–12% Bis-Tris gel (Thermo Fisher Scientific) in NuPAGE MOPS SDS running buffer (Thermo Fisher Scientific) and transferred to a PVDF membrane (Merck Millipore) in NuPAGE Transfer Buffer (Thermo Fisher Scientific) with XCell II Blot Module (Thermo Fisher Scientific). Blots were blocked with 5% skim milk in TBST for 30 min and incubated with primary antibodies (which are listed in the Reagents Tools Table) at 4°C overnight. Immunocomplexes were labeled by HRP-conjugated anti-mouse IgG or anti-rabbit IgG and visualized using Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific) or Amersham ECL Prime (Cytiva). The signals were detected with ImageQuant LAS 4000 (Cytiva) or ChemiDoc Touch MP (Bio-rad).

Immunofluorescent microscopy

293 T cells were plated onto the glass slide coated with poly-L lysine and cultured for overnight. Suspension cells were centrifuged onto glass slides. Cells were fixed with 4% paraformaldehyde in PBS for 10 min at 4°C and permeabilized with 50 µg/ml of Digitonin in PBS for 5 min at RT. For the detection of GFP fluorescence in 293 T cells, cells were washed, counterstained with DAPI, and then analyzed by fluorescence microscope Axio Observer (Zeiss) or BZ-X710 (Keyence). To measure percentage of nuclear localization of SETD1A-GFP proteins, GFP and DAPI-positive areas were detected, and the overlap region was calculated by ImageJ software. For the immunofluorescence staining, after three PBS washes, we blocked the slide with 3% BSA in PBS for 30 min at RT. Primary antibodies against GFP (Aves Labs, GFP-1010), BuGZ (Novus Biologicals,

NBP2-39010), and BUB3 (Abcam, ab133699) were diluted in 3% BSA in PBS and incubated at 4°C overnight. After three PBS washes, secondary antibodies conjugated fluorescent dye are used for the detection. Slides were washed with PBS containing DAPI and mounted using the Prolong Gold Antifade Reagent (Thermo Fisher Scientific). Cells were analyzed by confocal microscope FV10i-LIV (Olympus).

RNA analyses

Total RNA was purified with RNeasy Mini kit (Qiagen) from mouse MLL-r leukemia cells or MOLM-13 leukemia cells. For RT-qPCR, cDNA was synthesized with ReverTra Ace qPCR RT Master Mix (TOYOBO). cDNA fragments were quantified by a TaqMan Gene expression assay (Applied Biosystems; Appendix Table S5) or SYBR green qPCR with gene-specific primers and Taq DNA polymerase (NEB) with a ViiA 7 Real time PCR system (Thermo Fisher Scientific) or CFX96 Touch real-time PCR detection system (Bio-Rad). Libraries were prepared by using the TruSeq Stranded mRNA Sample Prep Kit (Illumina). The DNA library was validated using TapeStation (Agilent Technologies) and was quantified using a QuantiFluor dsDNA system (Promega) and a Quantus Fluorometer (Promega). Libraries were pooled and sequenced on Illumina HiSeq 1500, NextSeq 500, or NovaSeq 6000. Data were analyzed by using HISAT2 and Cufflinks. List of interest genes was analyzed by Enrichr tool (<https://amp.pharm.mssm.edu/Enrichr/>).

ChIP and CUT&Tag analyses

The previously established MOLM-13 leukemia cells expressing HA-tagged SETD1A (Hoshii *et al*, 2022) were fixed with 2 mM DSG for 30 min followed by 1% formaldehyde for 10 min and treated with 0.125 M Glycine for 5 min. Fixed cells were washed twice with cold PBS. Washed cells were resuspended in ChIP Lysis Buffer (50 mM HEPES pH 8, 140 mM NaCl, 1 mM EDTA, 10% Glycerol, 0.5% NP-40, 0.25% Triton-X100, 1 × protease inhibitor cocktail) and centrifuged at 2,000 *g* for 5 min at 4°C to collect the nuclei pellet. The pellet was washed with ChIP wash buffer (10 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA) and resuspended in ChIP shearing buffer (0.1% SDS, 1 mM EDTA, 10 mM Tris-HCl, 1 × protease inhibitor cocktail) and then frozen at -80°C. Nuclei samples were shredded using Picoruptor (Diagenode) and precleared with Dynabeads protein G (Thermo Fisher Scientific) for 60 min at 4°C with gentle rotation. Appropriate amounts of antibodies were added into the chromatin and incubated overnight at 4°C. The immune complex was collected with Dynabeads protein G and washed sequentially in the low salt wash buffer (20 mM Tris pH8, 150 mM NaCl, 0.1% SDS, 1% Triton-X100, 2 mM EDTA), the high salt wash buffer (20 mM Tris pH8, 500 mM NaCl, 0.1% SDS, 1% Triton-X100, 2 mM EDTA), the LiCl wash buffer (10 mM Tris pH8, 250 mM LiCl, 1% NP-40, 1% Sodium Deoxycholate, 1 mM EDTA), and TE. Chromatin was eluted with elution buffer (1% SDS, 20 mM Tris-HCl, 10 mM EDTA) containing RNaseA and then reverse cross-linked with proteinase K at 65°C for overnight. DNA was purified with a PCR purification kit (Qiagen) and was quantified using a QuantiFluor dsDNA system (Promega) and a Quantus Fluorometer (Promega). Libraries were prepared by using the KAPA HyperPrep Kit (for Illumina; Kapabiosystems). For the CUT&Tag against SETD1A, 1 × 10⁵ sgBuGZ-expressing MOLM-13 leukemia cells were mixed with 1 × 10⁴ mouse NIH3T3 cells for the normalization and fixed with 0.1% formaldehyde/PBS at room temperature for 2 min. Cells

were treated with 0.125 M Glycine and washed with PBS. Cells were bound to conA beads and then incubated with SETD1A antibody (Cell Signaling, #50805) at 4°C for overnight. After the incubation with anti-rabbit secondary antibody, conA beads were washed with 100 μl of Dig-300 buffer (20 mM HEPES pH7.5, 300 mM NaCl, 0.5 mM Spermidine, 0.01% Digitonin, 1 × protease inhibitor cocktail) three times, mixed with 100 μl of pA-Tn5 adapter complex, then the tagmentation was performed at 37°C for 1 h at 600 rpm. De-crosslinking was performed by adding 3.3 μl of 0.5 M EDTA and 1 μl of 10% SDS and then incubated at 67°C at 600 rpm for overnight. Samples were treated in 2 μl of Proteinase K and purified by the phenol-chloroform extraction and ethanol precipitation. DNA was amplified using uniquely barcoded i7 and i5 primers with high-fidelity PCR polymerase and then purified with SPRIselect beads. The DNA library was validated using TapeStation (Agilent Technologies). Libraries were sequenced as described above. Data were mapped to the University of California Santa Cruz human genome assembly (hg19) using Bowtie2, and then, duplicated reads were removed by Picard tools. Peak calling was performed using HOMER software. Heatmaps or histograms of ChIP-seq data were generated by using EaSeq software (<http://easeq.net>) or deepTools. Gene annotation to ChIP-seq peaks was performed with the Genomic Regions Enrichment of Annotations Tool (GREAT) using the basal plus extension method.

CRISPR

Doxycycline-inducible Cas9-expressing mouse MLL-r leukemia cells and human leukemia cell lines are established previously (Hoshii *et al*, 2018). All sgRNA sequences used in this study are shown in Appendix Table S4. For the dual CRISPR assay, both BuGZ sgRNA with a GFP reporter and BUB3 sgRNA with a hygromycin-resistant cassette were used. The two sgRNA-expressing cells were enriched by 1 mg/ml of hygromycin followed by cell sorting for GFP-positive cells. The empty vector backbone or the *Rosa26* locus-targeting sgRNA was used for the negative control, and the mouse *Rpa3* or human *RPA3*-targeting sgRNA was used for the positive control in competitive cell growth assay. Two BuGZ sgRNAs or the SETD1A Exon 8-targeting sgRNA were used for the deletion of endogenous *BuGZ* or *SETD1A* genes, respectively, in human MOLM-13 cells (Appendix Table S4). For CRISPR-tiling screen against BuGZ, 152 sgRNA vectors against mouse *Bugz* gene with a GFP reporter were constructed with 96-well Miniprep kit (Merck Millipore; Appendix Table S4). The competitive cell growth assay in sgRNA transduced cells was performed by monitoring the ratio of reporter protein-expressing cells by using CytoFLEX flow cytometer with plate loader option (Beckman Coulter).

Protein motif and secondary structure analyses

cNLS Mapper (<http://nls-mapper.iab.keio.ac.jp>) was used for the prediction of nuclear localization signal. Conserved motif analysis for human KNL1 MELT repeats and F2 motif was performed with Weblogo (<https://weblogo.berkeley.edu/logo.cgi>). Disordered region was analyzed with the predictor of natural disordered regions (PONDR: <http://www.pondr.com>).

In vivo leukemia model

C57BL/6J female mice were used as transplantation recipients. RT3GEPiR Bugz miR-E shRNA vectors or MSCV *SETD1A* cDNA

vectors were transduced into wild-type mouse MLL-r leukemia cells or *Setd1a*^{fl/fl};CreER mouse MLL-r leukemia cells, respectively, *in vitro*. For Bugz shRNA, vector-transduced cells were selected with 2.5 µg/ml puromycin selection and single-cell clones harboring shRNA vector were established. Mice used in this study have been maintained in the DFCI Animal Resource Center, following a vertebrate animal protocol approved by DFCI's Institutional Animal Care and Use Committee (IACUC). Recipient mice were randomly divided into groups, but no blinding was performed after transplant. One million cells were transplanted to sublethally irradiated (6.0 Gy) syngenic recipient mice. For Bugz shRNA, we administered doxycycline to the recipient mice by food intake (625 mg/kg diet, Harlan). For *Setd1a* rescue experiment, GFP-positive cells were sorted at 2-day postinfection and then treated with 0.5 µg/ml 4-hydroxytamoxifen for 2 days. Mice used in this study have been maintained in the Laboratory Animal Center for Chiba University, following an animal protocol approved by Chiba University (A5-054).

Cell cycle and apoptosis assay

Cell cycle was measured with Click-iT EdU Flow Cytometry Assay Kit (ThermoFisher). Cells were treated with 10 µM EdU containing medium for 2 h and fixed with fixative buffer in the kit. Incorporated EdU was labeled with Alexa647-conjugated azide. Stained cells were counterstained with 7AAD. For apoptosis assay, cells were washed with 1× annexin V binding buffer (Abcam) and then incubated with Annexin V-APC containing buffer at room temperature for 10 min. Cells were washed with 1× annexin V binding buffer and stained with 7AAD. Stained cells were analyzed with CytoFLEX flow cytometer (Beckman Coulter).

Quantification and statistical analysis

Error bars in all of the data represent a standard deviation. Number of replicates and number of repeated experiments are reported in the figure legends. For statistical comparison, we performed a Student's *t*-test, one-way ANOVA followed by Tukey's test, two-way ANOVA followed by multiple comparisons, or Kaplan–Meier. Data with statistical significance (**P* < 0.05, ***P* < 0.01) are shown in figures. Statistical analyses were performed using Prism 9 software (GraphPad).

Data availability

The accession number for the RNA-seq and ChIP-seq data reported in this paper is NCBI GEO: GSE159146 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE159146>). Previously deposited data (GSM5737312 and GSM5737322) are used in this study and also available at NCBI GEO: GSE189894 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE189894>). The MS raw data and result files have been deposited in the ProteomeXchange Consortium (<http://www.proteomexchange.org/>, PXD022158 [<http://www.ebi.ac.uk/pride/archive/projects/PXD022158>] and PXD042848 [<http://www.ebi.ac.uk/pride/archive/projects/PXD042848>]) via the jPOST partner repository (<https://jpostdb.org>, JPST000982 and JPST002187; Okuda et al, 2017).

Expanded View for this article is available [online](#).

Acknowledgements

We wish to thank Dr. Scott A. Armstrong for collaboration on the early stages of this work. We would like to thank Kayoko Ejima, Haruki Miyagi, Reina Sato, Kyoko Hirano, Saihanqiqiqe and Haruka Maruyama for technical support. T.H. was supported by JSPS KAKENHI (JP19H03690, JP22H03099, JP22H04684, JP22K19538), Mochida Memorial Foundation for Medical and Pharmaceutical Research, Kobayashi Foundation for Cancer Research, the Leading Initiative for Excellent Young Researchers (LEADER), and the Japan Agency for Medical Research and Development (AMED) (23ama221118h0002). A.K. was supported by NIH R01 CA204396 and P30 CA008748 and was a scholar of the Leukemia and Lymphoma Society. R.G.R. was supported by the Leukemia and Lymphoma Society Grant 7021-20. At.K. was supported by IAAR Research Support Program from Chiba University.

Author contributions

Sarah Perlee: Investigation; writing – review and editing. **Sota Kikuchi:** Investigation. **Tomoyoshi Nakadai:** Investigation. **Takeshi Masuda:** Investigation. **Sumio Ohtsuki:** Investigation. **Makoto Matsumoto:** Investigation. **Bahityar Rahmutulla:** Investigation. **Masaki Fukuyo:** Data curation. **Paolo Cifani:** Investigation. **Alex Kentsis:** Investigation. **Robert G Roeder:** Resources. **Atsushi Kaneda:** Resources; supervision; writing – review and editing. **Takayuki Hoshii:** Conceptualization; data curation; formal analysis; supervision; funding acquisition; investigation; visualization; writing – original draft.

Disclosure and competing interests statement

The authors declare that they have no conflict of interest.

References

- Arndt K, Kranz A, Fohgrub J, Jolly A, Bledau AS, Di Virgilio M, Lesche M, Dahl A, Hofer T, Stewart AF et al (2018) SETD1A protects HSCs from activation-induced functional decline *in vivo*. *Blood* 131: 1311–1324
- Bayley R, Borel V, Moss RJ, Sweatman E, Ruis P, Ormrod A, Goula A, Mottram RMA, Stanage T, Hewitt G et al (2022) H3K4 methylation by SETD1A/BOD1L facilitates RIF1-dependent NHEJ. *Mol Cell* 82: 1924–1939
- Bledau AS, Schmidt K, Neumann K, Hill U, Ciotta G, Gupta A, Torres DC, Fu J, Kranz A, Stewart AF et al (2014) The H3K4 methyltransferase Setd1a is first required at the epiblast stage, whereas Setd1b becomes essential after gastrulation. *Development* 141: 1022–1035
- Boija A, Klein IA, Sabari BR, Dall'Agnesse A, Coffey EL, Zamudio AV, Li CH, Shrinivas K, Manteiga JC, Hannett NM et al (2018) Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* 175: 1842–1855
- Bolanos-Garcia VM, Blundell TL (2011) BUB1 and BUBR1: multifaceted kinases of the cell cycle. *Trends Biochem Sci* 36: 141–150
- Chen Y, Anastassiadis K, Kranz A, Stewart AF, Arndt K, Waskow C, Yokoyama A, Jones K, Neff T, Lee Y et al (2017) MLL2, Not MLL1, plays a major role in sustaining MLL-rearranged acute myeloid leukemia. *Cancer Cell* 31: 755–770
- Fang F, Xia N, Angulo B, Carey J, Cady Z, Durruthy-Durruthy J, Bennett T, Sebastiano V, Reijo Pera RA (2018) A distinct isoform of ZNF207 controls self-renewal and pluripotency of human embryonic stem cells. *Nat Commun* 9: 4384
- Fellmann C, Hoffmann T, Sridhar V, Hopfgartner B, Muhar M, Roth M, Lai DY, Barbosa IA, Kwon JS, Guan Y et al (2013) An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep* 5: 1704–1713

- He W, Zhang L, Villarreal OD, Fu R, Bedford E, Dou J, Patel AY, Bedford MT, Shi X, Chen T *et al* (2019) De novo identification of essential protein domains from CRISPR-Cas9 tiling-sgRNA knockout screens. *Nat Commun* 10: 4541
- Higgs MR, Sato K, Reynolds JJ, Begum S, Bayley R, Goula A, Vernet A, Paquin KL, Skalnik DG, Kobayashi W *et al* (2018) Histone methylation by SETD1A protects nascent DNA through the nucleosome chaperone activity of FANCD2. *Mol Cell* 71: 25–41
- Hiraide T, Nakashima M, Yamoto K, Fukuda T, Kato M, Ikeda H, Sugie Y, Aoto K, Kaname T, Nakabayashi K *et al* (2018) De novo variants in SETD1B are associated with intellectual disability, epilepsy and autism. *Hum Genet* 137: 95–104
- Hoshii T, Cifani P, Feng Z, Huang CH, Koche R, Chen CW, Delaney CD, Lowe SW, Kentsis A, Armstrong SA (2018) A non-catalytic function of SETD1A regulates cyclin K and the DNA damage response. *Cell* 172: 1007–1021
- Hoshii T, Perlee S, Kikuchi S, Rahmutulla B, Fukuyo M, Masuda T, Ohtsuki S, Soga T, Nabet B, Kaneda A (2022) SETD1A regulates transcriptional pause release of heme biosynthesis genes in leukemia. *Cell Rep* 41: 111727
- Jiang H, He X, Wang S, Jia J, Wan Y, Wang Y, Zeng R, Yates J 3rd, Zhu X, Zheng Y (2014a) A microtubule-associated zinc finger protein, BuGZ, regulates mitotic chromosome alignment by ensuring Bub3 stability and kinetochore targeting. *Dev Cell* 28: 268–281
- Jiang Y, Li X, Yang W, Hawke DH, Zheng Y, Xia Y, Aldape K, Wei C, Guo F, Chen Y *et al* (2014b) PKM2 regulates chromosome segregation and mitosis progression of tumor cells. *Mol Cell* 53: 75–87
- Jiang H, Wang S, Huang Y, He X, Cui H, Zhu X, Zheng Y (2015) Phase transition of spindle-associated protein regulate spindle apparatus assembly. *Cell* 163: 108–122
- Kon A, Shih LY, Minamino M, Sanada M, Shiraishi Y, Nagata Y, Yoshida K, Okuno Y, Bando M, Nakato R *et al* (2013) Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nat Genet* 45: 1232–1237
- Kummeling J, Stremmelaar DE, Raun N, Reijnders MRF, Willemsen MH, Ruiterkamp-Versteeg M, Schepens M, Man CCO, Gilissen C, Cho MT *et al* (2021) Characterization of SETD1A haploinsufficiency in humans and *Drosophila* defines a novel neurodevelopmental syndrome. *Mol Psychiatry* 26: 2013–2024
- Lee JH, Skalnik DG (2008) Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. *Mol Cell Biol* 28: 609–618
- Morita S, Kojima T, Kitamura T (2000) Plat-E: an efficient and stable system for transient packaging of retroviruses. *Gene Therapy* 7: 1063–1066
- Nagahama K, Sakoori K, Watanabe T, Kishi Y, Kawaji K, Koebis M, Nakao K, Gotoh Y, Aiba A, Uesaka N *et al* (2020) Setd1a insufficiency in mice attenuates excitatory synaptic function and recapitulates schizophrenia-related behavioral abnormalities. *Cell Rep* 32: 108126
- Nakadai T, Shimada M, Ito K, Cevher MA, Chu CS, Kumegawa K, Maruyama R, Malik S, Roeder RG (2023) Two target gene activation pathways for orphan ERR nuclear receptors. *Cell Res* 33: 165–183
- Niikura Y, Ogi H, Kikuchi K, Kitagawa K (2010) BUB3 that dissociates from BUB1 activates caspase-independent mitotic death (CIMD). *Cell Death Differ* 17: 1011–1024
- Okuda S, Watanabe Y, Moriya Y, Kawano S, Yamamoto T, Matsumoto M, Takami T, Kobayashi D, Araki N, Yoshizawa AC *et al* (2017) jPOSTrepo: an international standard data repository for proteomes. *Nucleic Acids Res* 45: D1107–D1111
- Porter IM, McClelland SE, Khoudoli GA, Hunter CJ, Andersen JS, McAinsh AD, Blow JJ, Swedlow JR (2007) Bod1, a novel kinetochore protein required for chromosome biorientation. *J Cell Biol* 179: 187–197
- Rao RC, Dou Y (2015) Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat Rev Cancer* 15: 334–346
- Rappsilber J, Ishihama Y, Mann M (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* 75: 663–670
- Rappsilber J, Mann M, Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2: 1896–1906
- Sabari BR, Dall'Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, Abraham BJ, Hannett NM, Zamudio AV, Manteiga JC *et al* (2018) Coactivator condensation at super-enhancers links phase separation and gene control. *Science* 361: eaar3958
- Santos MA, Faryabi RB, Ergen AV, Day AM, Malhowski A, Canela A, Onozawa M, Lee JE, Callen E, Gutierrez-Martinez P *et al* (2014) DNA-damage-induced differentiation of leukaemic cells as an anti-cancer barrier. *Nature* 514: 107–111
- Schapiro M, Tyers M, Torrent M, Arrowsmith CH (2017) WD40 repeat domain proteins: a novel target class? *Nat Rev Drug Discov* 16: 773–786
- Shi J, Wang E, Milazzo JP, Wang Z, Kinney JB, Vakoc CR (2015) Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat Biotechnol* 33: 661–667
- Singh T, Kurki MI, Curtis D, Purcell SM, Crooks L, McRae J, Suvisaari J, Chheda H, Blackwood D, Breen G *et al* (2016) Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. *Nat Neurosci* 19: 571–577
- Sze CC, Cao K, Collings CK, Marshall SA, Rendleman EJ, Ozark PA, Chen FX, Morgan MA, Wang L, Shilatifard A (2017) Histone H3K4 methylation-dependent and -independent functions of Set1A/COMPASS in embryonic stem cell self-renewal and differentiation. *Genes Dev* 31: 1732–1737
- Takata A, Xu B, Ionita-Laza I, Roos JL, Gogos JA, Karayiorgou M (2014) Loss-of-function variants in schizophrenia risk and SETD1A as a candidate susceptibility gene. *Neuron* 82: 773–780
- Toledo CM, Herman JA, Olsen JB, Ding Y, Corrin P, Girard EJ, Olson JM, Emili A, DeLuca JG, Paddison PJ (2014) BuGZ is required for Bub3 stability, Bub1 kinetochore function, and chromosome alignment. *Dev Cell* 28: 282–294
- Toska E, Osmanbeyoglu HU, Castel P, Chan C, Hendrickson RC, Elkabets M, Dickler MN, Scaltriti M, Leslie CS, Armstrong SA *et al* (2017) PI3K pathway regulates ER-dependent transcription in breast cancer through the epigenetic regulator KMT2D. *Science* 355: 1324–1330
- Vleugel M, Omerzu M, Groenewold V, Hadders MA, Lens SM, Kops GJ (2015) Sequential multisite phospho-regulation of KNL1-BUB3 interfaces at mitotic kinetochores. *Mol Cell* 57: 824–835
- Wan Y, Zheng X, Chen H, Guo Y, Jiang H, He X, Zhu X, Zheng Y (2015) Splicing function of mitotic regulators links R-loop-mediated DNA damage to tumor cell killing. *J Cell Biol* 209: 235–246
- Wan L, Chong S, Xuan F, Liang A, Cui X, Gates L, Carroll TS, Li Y, Feng L, Chen G *et al* (2020) Impaired cell fate through gain-of-function mutations in a chromatin reader. *Nature* 577: 121–126
- Wang L, Collings CK, Zhao Z, Cozzolino KA, Ma Q, Liang K, Marshall SA, Sze CC, Hashizume R, Savas JN *et al* (2017) A cytoplasmic COMPASS is necessary for cell survival and triple-negative breast cancer pathogenesis by regulating metabolism. *Genes Dev* 31: 2056–2066

- Wang S, Rhijn JV, Akkouch I, Kogo N, Maas N, Bleeck A, Ortiz IS, Lewerissa E, Wu KM, Schoenmaker C et al (2022) Loss-of-function variants in the schizophrenia risk gene SETD1A alter neuronal network activity in human neurons through the cAMP/PKA pathway. *Cell Rep* 39: 110790
- Weerts MJA, Lanko K, Guzman-Vega FJ, Jackson A, Ramakrishnan R, Cardona-Londono KJ, Pena-Guerra KA, van Bever Y, van Paassen BW, Kievit A et al (2021) Delineating the molecular and phenotypic spectrum of the SETD1B-related syndrome. *Genet Med* 23: 2122–2137
- Yan J, Enge M, Whittington T, Dave K, Liu J, Sur I, Schmierer B, Jolma A, Kivioja T, Taipale M et al (2013) Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites. *Cell* 154: 801–813
- Zhu S, Jing R, Yang Y, Huang Y, Wang X, Leng Y, Xi J, Wang G, Jia W, Kang J (2015) A motif from Lys216 to Lys222 in human BUB3 protein is a nuclear localization signal and critical for BUB3 function in mitotic checkpoint. *J Biol Chem* 290: 11282–11292