PNAS RESEARCH ARTICLE | BIOCHEMISTRY

LYL1 facilitates AETFC assembly and gene activation by recruiting CARM1 in t(8;21) AML

Qian Chen[a], Murat A. Cevher[b,c,1], Qi Jiang[a,1], Saisai Wang[a], Xiaojian Sun (孙晓建)[d], Robert G. Roeder[b,2], and Mo Chen[a,2]

Contributed by Robert G. Roeder; received August 9, 2022; accepted September 13, 2022; reviewed by Mark Bedford and Christopher Vakoc

Transcription factors (TFs) play critical roles in hematopoiesis, and their aberrant expression can lead to various types of leukemia. The t(8;21) leukemogenic fusion protein AML1–ETO (AE) is the most common fusion protein in acute myeloid leukemia and can enhance hematopoietic stem cell renewal while blocking differentiation. A key question in understanding AE-mediated leukemia is what determines the choice of AE to activate self-renewal genes or repress differentiation genes. Toward the resolution of this problem, we earlier showed that AE resides in the stable AETFC complex and that its components colocalize on up- or down-regulated target genes and are essential for leukemogenesis. In the current study, using biochemical and genomic approaches, we show that AE-containing complexes are heterogeneous, and that assembly of the larger AETFC (containing AE, CBFβ, HEB, E2A, LYL1, LMO2, and LDB1) requires LYL1. Furthermore, we provide strong evidence that the LYL1-containing AETFC preferentially binds to active enhancers and promotes AE-dependent gene activation. Moreover, we show that coactivator CARM1 interacts with AETFC and facilitates gene activation by AETFC. Collectively, this study describes a role of oncoprotein LYL1 in AETFC assembly and gene activation by recruiting CARM1 to chromatin for AML cell survival.

AML1–ETO | LYL1 | CARM1 | acute myeloid leukemia | E proteins

Significance

AML1–ETO represents the most common fusion transcription factor in AML. AE forms a unique and stable transcription factor/cofactor complex, designated AETFC, that is critical for leukemogenesis. Although earlier studies have shown the biological importance of AETFC components, how the assembly of AETFC affects gene expression is still unclear. Additionally, the AETFC component, LYL1, has been shown to be critical for the proliferation of other types of leukemia. Thus, understanding how LYL1 contributes to AE-dependent gene expression is important to the broader field of leukemia. Moreover, establishment of a correlation between LYL1- and AE-dependent gene activation can guide discovery of coactivators for AETFC. The potential to target activator–coactivator interactions or enzymatic activities may provide novel therapeutic directions.

Development is controlled by the timely expression and function of transcription factors (TFs) that regulate gene expression and determine cell fate. Leukemia is often driven by chromosomal translocations that result in the production of novel transcription factors, produced by the abnormal fusion of two separate gene products. The most common fusion in acute myelogenous leukemia (AML) is the t(8;21)-derived AML1–ETO (AE) (1), which drives leukemia by coopting additional normal cellular transcription factors that are critical for blood stem/progenitor cell self-renewal and proliferation, thereby trapping the cells in a proliferative immature state that is the basis of the leukemia.

Most transcription factors work in combination to regulate target gene expression in a context-dependent manner (2). Usually, the cooperation between TFs depends on formation of functional TF complexes mediated by transient protein–protein interactions or closely positioned DNA sequences. AE retains the DNA binding domain of RUNX1 (AML1) and almost all of ETO, which contains several Nervy homology regions (NHRs) that are important for oligomerization and for interactions with E proteins and other cofactors (3). Toward understanding how AE regulates target gene expression, we earlier showed that AE resides within a stable complex (termed AETFC) that also contains CBFβ (binding partner of AML1/RUNX1), E protein family members HEB or E2A, the oncogenic LYL1 protein, the LIM domain protein LMO2, and the LMO2 binding partner LDB1 (4). Similarly structured complexes have been found in several other hematopoietic compartments (5). For example, GATA2 can form a complex with E proteins, TAL1 or LYL1, LMO2, and LDB1 in hematopoietic progenitor cells; and this complex can interact functionally with RUNX1 (6, 7). In contrast to the situation in these TF complexes, acquisition of the NHR2 domain-containing ETO protein in AE allows tight AE–E proteins interactions and consequent formation of a complex (AETFC) that is much more stable than many other TF complexes (4, 8). Although previous studies from us and others demonstrated the importance of AE–E protein interactions in AE-dependent leukemia (4, 9–11), a further understanding of how AETFC is assembled and how it regulates expression is of great importance.

Like many transcription factors, AE can either activate or repress target gene expression. A central question is what determines the choice between activating self-renewal genes and repressing differentiation-related genes. The dual functionality of AE probably hinges, at least partially, on its ability to interact with both transcriptional coactivators and corepressors. Repression of differentiation genes by AE depends on the

Author contributions: R.G.R. and M.C. designed research; Q.C., M.A.C., Q.J., and S.W. performed research; X.S. contributed new reagents/analytic tools; Q.C. analyzed data; and Q.C., R.G.R., and M.C. wrote the paper.

Reviewers: M.B., University of Texas, MD Anderson Cancer Center; and C.V., Cold Spring Harbor Laboratory.

The authors declare no competing interest.

[1]M.A.C. and Q.J. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: roeder@rockefeller.edu or mochen@mail.tsinghua.edu.cn.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2213718119/-/DCSupplemental.

Published October 10, 2022.

interaction between the ETO repressor and corepressor complexes (12). Known corepressors include previously identified N-CoR, HDACs, and mSin3A. AE coactivators are more recently identified and include p300 (8, 13), which acetylates both histones and AE, the histone arginine methyltransferase PRMT1 (14), and the histone demethylase JMJD1C (15). Most of these corepressors and coactivators were identified through interactions with AE rather than the AE complex(es). However, further studies have indicated that AE and E proteins generally colocalize on chromatin and bind to both up- and down-regulated target genes (4, 16, 17). Therefore, the differential recruitment of the cofactors cannot be explained by differential binding of AE–E protein complexes to target genes. Since most of the cofactors were identified through AE interactions, we reason that differential recruitment of additional cofactors may be through other AETFC components, such as LYL1. In this regard, an understanding of the biochemical and functional roles of LYL1 in AETFC requires further study.

LYL1 is a transcription factor that contains a basic helix–loop–helix (bHLH) domain and is important for hematopoietic stem/progenitor cells (6, 18). Structurally, both LYL1 and its homolog TAL1 form DNA binding heterodimers with E proteins, such as E2A and HEB, which are also bHLH factors (19). LYL1 and TAL1 also directly interact with nuclear cofactor LIM-only (LMO) proteins to form transcription complexes that drive lineage-specific gene expression in hematopoietic cells and in malignancies. For example, TAL1 forms a complex with E2A, LDB1, LMO2, GATA3, and RUNX1 that mediates a core transcriptional regulatory circuit in T cell acute lymphoblastic leukemia (20). Functionally, LYL1 has been shown to be important for many types of leukemia (21–23). This includes cooperation with LMO2 in promoting T cell leukemia (24), further underscoring the importance of LYL1/LMO2-containing complexes in leukemogenesis. Although multiple lines of evidence indicate the biological importance of LYL1 in both T cell leukemia and acute myeloid leukemia, how LYL1 cooperates with AE to regulate gene expression is still unclear.

Protein arginine methyltransferases (PRMTs) are well-known players in a myriad of biological processes, including cancer development (25, 26). Among the nine-member family, PRMT1 has been shown to directly interact with and methylate AE9a, the more aggressive and shorter spliced isoform of AE (14). PRMT4, also known as CARM1, is essential for various types of myeloid leukemogenesis, including AE-dependent leukemia (27). CARM1 regulates a set of genes involved in the cell cycle, DNA repair, and DNA replication in AML cells. How CARM1 cooperates with leukemic transcription factors, especially AETFC, is unclear.

In this study, we set out to further elucidate mechanisms that facilitate AE's ability to activate gene expression. We first show, biochemically and functionally, the existence of at least two distinct AE-containing complexes: a LYL1-containing AETFC complex, and several LYL1-free AE complexes (AECs) that are best represented by the AE–HEB subcomplex. Importantly, we then show that the larger AETFC complex preferentially binds to active enhancers of AE-activated genes. Finally, we also show that target gene activation by AETFC involves a direct interaction and chromatin recruitment of the CARM1 transcriptional coactivator, thus extending the list of AETFC coactivators.

Results

AE Forms Biochemically Distinct Complexes. Since AE binds to both up- and down-regulated target genes, we sought to determine whether these opposing transcriptional activities are mediated by a homogeneous AETFC complex or by separate AETFC subcomplexes. In this regard, we first examined whether each of the AETFC components exists outside of an AE-containing complex(es). We first generated Kasumi-1 nuclear extract (NE) and then performed immunoprecipitation (IP) experiments with antibodies against individual AETFC components. Since HEB and E2A are redundant functionally in Kasumi-1 cells (4) and have been considered biochemically similar, and for simplicity, our initial biochemical analyses were focused on HEB. Consistent with our previous results, coimmunoprecipitation of all AETFC components was observed with individual HEB, LYL1, and LDB1 antibodies at a low gel exposure (*SI Appendix*, Fig. S1*A*, lanes 1 through 7), and also with an ETO antibody at a higher gel exposure (*SI Appendix*, Fig. S1*B*). Interestingly, although similar amounts of HEB were co-IPed by HEB, LYL1, and LDB1 antibodies, less AE was co-IPed with LYL1 or LDB1 than with HEB (*SI Appendix*, Fig. S1*A*, compare lanes 2, 5, and 6). Additionally, much less non-AE AETFC components were observed with the ETO antibody compared to IPs with their own antibodies (*SI Appendix*, Fig. S1*A*, compare lane 4 with lanes 2, 5, and 6). These results suggest that: 1) LYL1 and LDB1/LMO2 only exist in a fraction of AE-containing complexes, and 2) LYL1 and LDB1/LMO2 also exist outside of AE-containing complexes. Interestingly, immunoblots of supernatants from the IPs showed that immunodepletion of LYL1, HEB, or LDB1 did not dramatically reduce the AE level in the corresponding supernatants and, importantly, that HEB antibody could deplete more AE than could LDB1 and LYL1 antibodies (*SI Appendix*, Fig. S1*C*, compare lane 3 and lanes 6 and 7). These results further suggest that HEB/LYL1/LDB1/LMO2-containing AE complexes only represent a small fraction of AE-associated protein complexes, and that HEB may form a smaller complex with AE without LYL1 and LDB1/LMO2. To further confirm this idea, we performed a serial immunodepletion assay (diagrammed in Fig. 1*A*). We first immunodepleted LYL1 from Kasumi-1 NE and confirmed its association with AE, HEB, and LMO2 (*SI Appendix*, Fig. S1*D*, lane 3). We then used IgG- or LYL1-depleted NEs (Fig. 1*B*, lanes 5 and 6) as inputs for AE IPs with ETO antibody. In this case, LYL1 and LMO2 were co-IPed with AE only in the control IgG-depleted NE (lane 9), and not from LYL1-depleted NE (lane 10). In contrast, AE and HEB were co-IPed in both conditions. The combined results from these experiments demonstrate clearly that AE can form a complex (AETFC) composed of AE, HEB, LYL1, LMO2, and LDB1, but that it also can form a smaller complex with HEB in the absence of LYL1 and other associated factors (diagram shown in Fig. 1*C*).

In order to further support this observation, we sought to examine AE complexes in a more purified system. To this end, we partially purified all AE-containing protein complexes on M2 agarose from an extract from a Kasumi-1 cell line stably expressing a flag-tagged AE at a level similar to that of endogenous AE (15). Aliquots of the preparation (M2 eluate) were then immunodepleted with LYL1, LDB1, HEB, or (as a control) IgG antibodies and the supernatants from these immunodepletion assays were examined for the presence of AETFC components. Consistent with our earlier observations, the essentially complete depletions of LYL1, LDB1, or HEB failed to deplete most of the AE from the complete group of AE-containing complexes (Fig. 1*D*). These results confirm that AE exists outside of the larger AETFC complex or even an AE–HEB complex, although the latter result may reflect an artificial enrichment of the ectopically (overexpressed) f-AE protein during affinity purification. Additionally, HEB depletion completely removed all LYL1 and LDB1 from

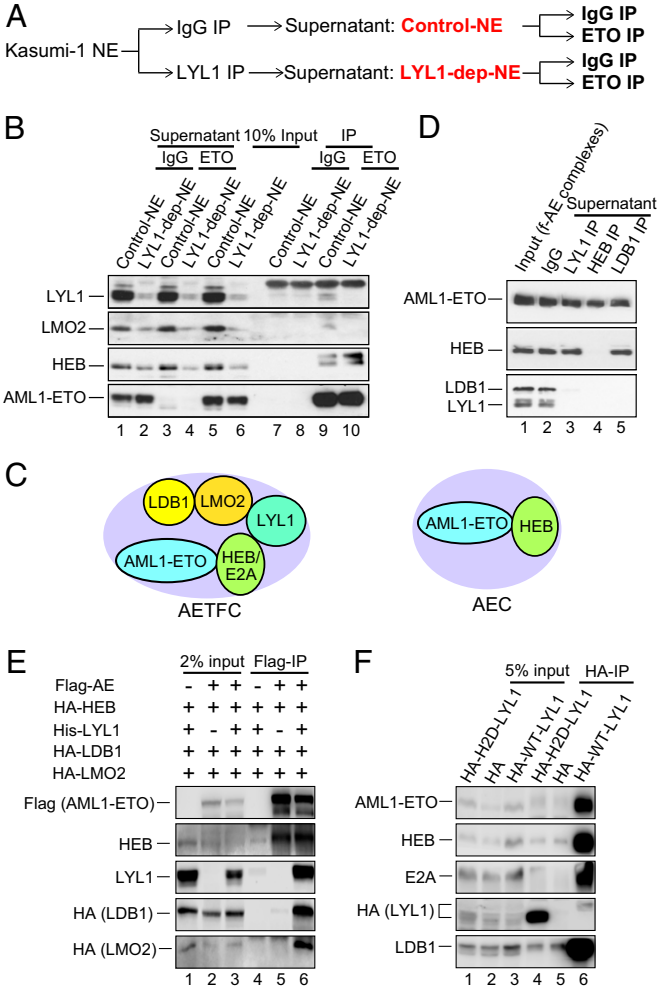


Fig. 1. AE forms biochemically distinct complexes, and LYL1 is essential for LMO2/LDB1 recruitment to form AETFC. (*A*) Diagram showing IgG or LYL1 immunoprecipitation with NE from Kasumi-1 cells. Input control NE and LYL1-dep NE are marked in red for further immunoprecipitation with IgG or ETO antibody. (*B*) Immunoprecipitation of IgG or ETO with control NE or LYL1-dep NE (described in *A*) to confirm AETFC association. Bound proteins were detected with antibodies shown on the *Left*. (*C*) Diagram showing the composition of two representative AE complexes. (*D*) Immunoblot of supernatants derived from immunoprecipitation assays using antibodies shown on *Top*. Affinity-purified crude f-AE complexes were used as input. Antibodies used for immunoblot are shown on the *Left*. (*E*) Confirmation of different AE subcomplexes in 293T cells. AETFC components were transiently expressed in different combinations in 293T cells (shown on *Top*) and AE was immunoprecipitated using M2 agarose. Associated proteins were detected with antibodies shown on the *Left*. (*F*) Immunoprecipitation with HA-agarose to detect the association of HA-LYL1 mutants and other AETFC proteins in NE from Kasumi-1 cell lines that stably express different HA-LYL1 mutants (shown on *Top*). Bound proteins were detected with antibodies shown on the *Left*.

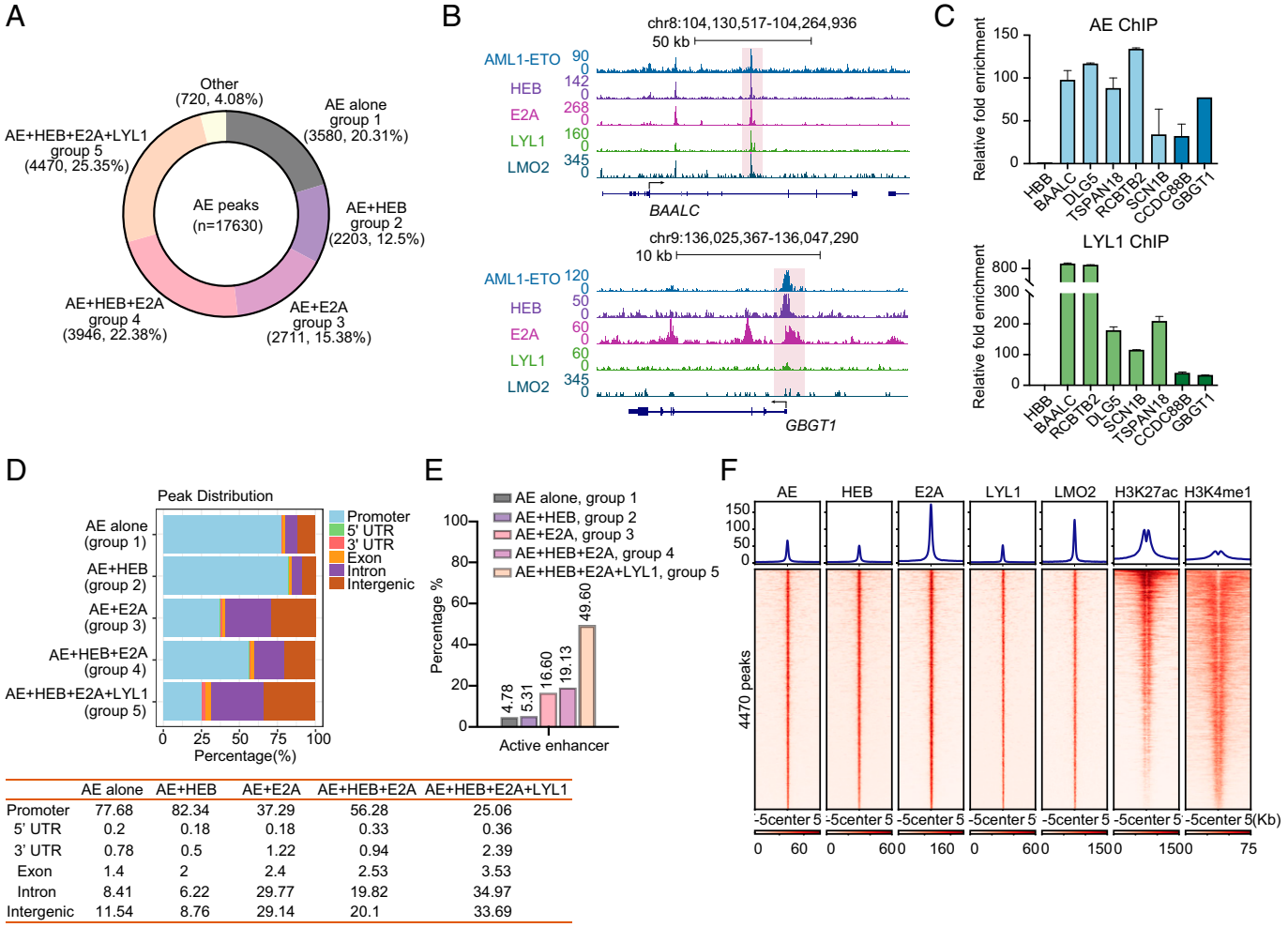
the AE-containing complex preparation—consistent with their quantitative association with HEB complexes—whereas LYL1 or LDB1 depletion failed to deplete much of the HEB (Fig. 1*D*, lanes 3 and 5). These results indicate that LYL1/LDB1/LMO2-containing complexes only represent a subset of AE- and HEB-containing complexes. Since individual LYL1 and LDB1 depletions also resulted in complete reciprocal depletions of each other, LYL1 and LDB1 must coexist in the AE-containing complexes that, again, are clearly indicated by their co-IP with AE from either nuclear extracts (*SI Appendix*, Fig. S1*A*) or purified AE complexes (Fig. 1*D*). Since E2A is thought to be redundant with HEB, we also examined its status in AETFC and AEC. Notably, HEB depletion completely depleted E2A from the AE-associated complex preparation, indicating that most of the E2A is associated with HEB in AE-containing complexes (*SI Appendix*, Fig. S1*E*, lane 3). In contrast, E2A depletion did not deplete much HEB from the AE-containing complex preparation, indicating that HEB can exist in an AE-containing complex without E2A (*SI Appendix*, Fig. S1*E*, lane 4). In addition, since LYL1 depletion resulted in a much higher fractional removal of E2A than of HEB from the AE-containing complex input (*SI Appendix*, Fig. S1*E*, lane 7), it appears that the LYL1-free AEC complex is mainly composed of AE and HEB without E2A. Therefore, we propose a model in which there are at least two AE-containing complexes: a smaller LYL1-free AE/HEB complex (AEC) and a larger LYL1-containing complex (AETFC) that contains AE, HEB, E2A, LYL1, LDB1, and LMO2 (Fig. 1*C*).

LYL1 Is Critical for LDB1/LMO2 Assembly to Form AETFC. Since LYL1 is codepleted with LDB1 from AE-containing complexes, we hypothesized that assembly of the larger AETFC complex relies on LYL1. To test this hypothesis, we cotransfected 293T cells with AETFC components in the presence or absence of LYL1. Our results show that Flag-AE IP exclusively co-IPs HEB in the absence of LYL1, confirming formation of an AE–HEB complex, but that it co-IPs LDB1 and LMO2, along with LYL1 and HEB, only in the presence of LYL1 (Fig. 1*E*, lanes 5 and 6), indicating that LYL1 is required for the connection of AE/HEB to LDB1/LMO2. In a similar experiment using Sf9 cells infected with baculoviruses expressing AETFC components, and consistent with the transfection results, LDB1 was only co-IPed by AE when LYL1 was present, while HEB was pulled down by AE equally well in the absence or presence of LDB1 (*SI Appendix*, Fig. S1*F*, lanes 2 and 3). These results confirmed our hypothesis that LYL1 is critical for the assembly of LDB1/LMO2 within AETFC and led us to predict that disruption of HEB–LYL1 heterodimerization would disrupt AETFC formation without affecting AEC formation. Previous work has identified a similar complex containing the LYL1 homolog TAL1, E2A, LMO2, and LDB1. TAL1 interacts with E2A through the second helix of its helix–loop–helix domain and a TAL1 Y235A point mutation completely disrupts TAL1/E2A dimerization (5). Given the similarity between LYL1 and TAL1, we aligned LYL1 and TAL1 protein sequences and found that Y198 (corresponding to Y235 in TAL1) in LYL1 is conserved (*SI Appendix*, Fig. S1*G*). Reasoning that a second helix deletion mutant (H2D) or a Y198A mutant might also disrupt the LYL1–HEB interaction, we expressed HEB and mutant LYL1 proteins in 293T cells and analyzed HEB IPs. Our results confirmed loss of HEB heterodimerization with these LYL1 mutants, whereas another LYL1 mutant (F201A/H180A, corresponding to the F238A/H217A in TAL1) that disrupts LMO2 binding did not affect LYL1–HEB dimerization (*SI Appendix*, Fig. S1*H*). In order to examine whether LYL1–HEB heterodimerization is a prerequisite for AETFC assembly, we established Kasumi-1 cell lines that stably express either wild type (WT), Y198A, or H2D mutant proteins. Due to the difficulty in expressing the Y198A mutant (*SI Appendix*, Fig. S1*I*), we focused on a comparison of H2D mutant and WT LYL1 in Kasumi-1 cells. Consistent with results from our overexpression experiment, the stably expressed H2D mutant failed to form a complex with AE (Fig. 1*F*). The results of these complementary analyses clearly indicate that LYL1 and HEB heterodimerization is critical for the assembly of AETFC.

Distinct AE-Containing Complexes Occupy Different Genomic Sites. Our own and other studies have indicated global colocalization of AE and E proteins on chromatin in Kasumi-1 cells

(4, 15, 28). To further examine how E proteins and LYL1 colocalize with AE, we first performed chromatin immunoprecipitation sequencing (ChIP-seq) analyses for AETFC DNA binding components AE, HEB, E2A, and LYL1 (4, 15). As summarized in the pie chart of Fig. 2*A*, our results indicate that only about 20% of all AE occupied regions are bound by AE alone (group 1) and that the rest are mostly cooccupied by various E proteins. The latter regions were grouped into four types of AE-containing complexes dependent upon E2A, HEB, and LYL1 occupancies: group 2, AE + HEB; group 3, AE + E2A; group 4, AE + HEB + E2A; and group 5, AE + HEB + E2A + LYL1 (the aforementioned AETFC). Interestingly, although LYL1 only occupies a subset of AE-bound regions, AE- and LYL1-cobound regions always contain both HEB and E2A peaks, which is consistent with our biochemical data (*SI Appendix*, Fig. S1*E*, lane 7). Additionally, genomic data also identified an AE/E2A-only group 3, which was not present as a major population with our biochemical purification. We reason that the presence of group 3 and group 4 peaks might reflect the higher quality of E2A ChIP-seq data in view of the observation of 35,283 peaks for E2A and only 18,452 peaks for HEB. These data are consistent with the presence of physically and functionally distinct AE-containing complexes, comparable to those identified in the biochemical assays (Fig. 1*B*), that associate with distinct genomic regions. However, they do not eliminate the possibility that some members of the group of colocalized transcription factors can bind independently to respective DNA recognition elements at the sites.

Since LYL1-bound regions only represent a quarter of all AE peaks, we wished to exclude the possibility that the lower number of LYL1- and AE-cobound sites is attributable to low antibody affinity. Thus, we further analyzed LYL1 ChIP-seq peaks to examine how AE and E proteins bind to LYL1-bound regions. We obtained a total of 10,800 LYL1 peaks, with about 50% of them showing no AE occupancy. This result confirms that LYL1-containing AETFCs indeed represent only a subpopulation of all AE-containing complexes and that the results are not an artifact of (wrongly suspected) low-quality ChIP-seq for



	AE alone	AE+HEB	AE+E2A	AE+HEB+E2A	AE+HEB+E2A+LYL1
Promoter	77.68	82.34	37.29	56.28	25.06
5' UTR	0.2	0.18	0.18	0.33	0.36
3' UTR	0.78	0.5	1.22	0.94	2.39
Exon	1.4	2	2.4	2.53	3.53
Intron	8.41	6.22	29.77	19.82	34.97
Intergenic	11.54	8.76	29.14	20.1	33.69

Fig. 2. Distinct AE subcomplexes distribute differentially on the genome. (*A*) Pie chart depicting AE peak numbers cobound by different AE-associated TFs. Each group is classified by the composition of TFs (groups 1 through 5). Peak numbers and percentages of each group are indicated below the group number. (*B*) ChIP-seq tracks for AE (light blue), HEB (purple), E2A (pink), LYL1 (green), and LMO2 (dark blue) at the *BAALC* gene locus with enhancers bound by LYL1-containing AETFC (*Upper*) and at the *GBGT1* gene locus with its promoter bound by AEC (*Lower*). Track names are indicated on the *Left*. Gene names are shown below each snapshot. (*C*) ChIP-qPCR analyses of AE (*Upper*) and LYL1 (*Lower*) occupancy on target genes. AE and LYL1 peaks from AETFC-bound genes *BAALC*, *DLG5*, *TSPAN18*, *RCBTB2*, and *SCN1B* are shown (light blue, *Upper* and light green, *Lower*). AE and LYL1 peaks from AEC-bound genes *CCDC88B* and *GBGT1* are also shown (dark blue, *Upper* and dark green, *Lower*). The HBB site served as a negative control and is set as 1. Data are represented as mean ± SD. (*D*) Genomic distribution of AE peaks bound by AE complexes classified in *A*. Promoter localization is defined as ±3 kb from the TSS. (*E*) Percentages of peak number for each group bound at active enhancers. Active enhancers are defined as H3K4me1- and H3K27ac-positive sites that are not within 1 kb upstream of any annotated TSS. (*F*) Heatmap of ChIP-seq reads for AETFC proteins (AE, HEB, E2A, LYL1, and LMO2) and histone modifications (H3K27ac and H3K4me1) rank ordered from high to low by AE occupancy centered in a ±5-kb window around the TSSs of all genes. Color density reflects read density.

LYL1 (*SI Appendix*, Fig. S2*A*). In addition, about 90% of all LYL1 peaks overlap with at least one E-protein peak, suggesting that LYL1 binds to DNA in the form of a heterodimer with HEB or E2A. The composition of all LYL1 ChIP peaks is consistent with the AETFC ChIP-seq analyses in that all of the joint LYL1- and AE-cobound regions are occupied by HEB or E2A (over 99%, *SI Appendix*, Fig. S2*B*). These results suggest that LYL1-containing AE complexes tend to include both HEB and E2A in Kasumi-1 cells and are consistent with our previous analysis (Fig. 2*A* and *SI Appendix*, Fig. S1*E*). Next, to confirm the conclusions from the ChIP-seq results, we monitored LYL1 and AE binding at AETFC- and AEC-protein-bound sites by RT-qPCR. Representative genes preferentially bound by AETFC (Fig. 2 *B*, *Upper*) or AE/E complexes (Fig. 2 *B*, *Lower*) in the ChIP-seq data were chosen for analysis. To further verify our findings, we included several genes, including *BAALC*, *DLG5*, *RCBTB2*, and *SCN1B*, known to be involved in leukemia or other cancer types, as well as a few genes, including *TSPAN18*, *CCDC88B*, and *GBGT1*, not yet reported to be associated with cancer for the LYL1 and AE ChIP-qPCR analyses. Our results confirmed that LYL1 is more enriched at genes (e.g., *BAALC*) with AETFC sites (light green bars in Fig. 2 *C*, *Lower*) relative to genes (e.g., *GBGT1*) without AETFC sites (dark green bars in Fig. 2 *C*, *Lower*), whereas AE binding is more similar at all selected sites (light and dark blue bars in Fig. 2 *C*, *Upper*). Notably, a small fraction of LYL1- and AE-containing regions only showed binding of E2A but not HEB (12.49%, *SI Appendix*, Fig. S2*B*), although this could again reflect the higher quality of E2A ChIP-seq data. In summary, our genomic results are consistent with our earlier biochemical data demonstrating that AE-containing complexes are heterogeneous and that the larger LYL1-containing AETFC represents a fraction of the total AE-containing complexes.

LYL1-Containing AETFC Binds to Active Enhancers. Next, we examined how the five groups of functional AE complex–bound regions distribute on the genome. Strikingly, group 1 (AE alone) and group 2 (AE + HEB) showed predominant promoter binding with 77.68% and 82.34% of peaks located to gene promoter regions (Fig. 2*D*). On the other hand, the peaks jointly occupied by AE and LYL1 as well as HEB and E2A (group 5, the biochemically defined AETFC) are enriched in intergenic and intronic regions, with only 25.06% of the peaks on promoters (Fig. 2*D*). The genomic distributions of groups 3 and 4 lie somewhere in between those of group 2 and group 5. Next, we analyzed the binding strength of LYL1 across the five groups and, consistent with our co-IP experiments, observed distinctively higher LYL1 binding in group 5 (*SI Appendix*, Fig. S2 *C*, *Left*). Notably, groups 3 and 4 showed slightly higher LYL1 than groups 1 and 2, suggesting that LYL1 might bind to these regions weakly. Consistent with our earlier results in Fig. 1*D*, an analysis of LMO2 ChIP-seq data (4) revealed the highest LMO2 binding for the LYL1-containing group 5 regions, indicative of a LYL1 requirement for genomic association of LMO2 (*SI Appendix*, Fig. S2 *C*, *Right*). Interestingly, group 5 regions also showed the strongest AE binding, which suggests that the larger complex is more stable on chromatin and is consistent with the notion that transcription factors work cooperatively (*SI Appendix*, Fig. S2*D*).

Because of the unique genomic distribution (intergenic and intragenic) of LYL1 (*SI Appendix*, Fig. S2*E*) and its presence in the larger AETFC assembly, we hypothesized that LYL1-containing AETFC preferentially binds to active enhancers. To test this hypothesis, we first defined active enhancer regions as H3K4me1- and H3K27ac-positive regions that are not within 1 kb upstream of any annotated transcription start site (TSS). Interestingly, about 50% of group 5 (AETFC) sites are positive with respect to active enhancer histone marks H3K4me1 and H3K27ac, whereas, in contrast, only about 5% of group 1 (AE alone) and group 2 (AE + HEB) peaks contain active enhancer marks (Fig. 2*E*). In agreement, individual analyses of AE, HEB, E2A, and LYL1 total ChIP-seq peaks also indicate that LYL1 is mostly enriched at active enhancers (*SI Appendix*, Fig. S2*F*). Heatmaps of AETFC components show that AE, HEB, E2A, LYL1, and LMO2 all colocalize on AE peaks on group 5 sites and correlate well with H3K27ac levels (Fig. 2*F*). In summary, our results above suggest that LYL1, LMO2, and LDB1, upon forming a complex with AE, may be directed, with AE, to active enhancers.

LYL1-Containing AETFC Preferentially Activates Gene Transcription. Since LYL1-containing AETFC tends to preferentially localize to active enhancers, we next tested whether AE and LYL1 coactivate gene expression. To this end, we knocked down AE with shRNAs, performed RNA-seq, and identified genes that are directly activated or repressed by AE. We selected genes that are either down- or up-regulated upon AE depletion (shAE vs. shNC) by twofold, with false discovery rate (FDR) <0.1, and that have nearby regulatory elements that are bound by AE. We then examined how their expression responds to LYL1 depletion. Gene set enrichment analysis (GSEA) of the RNA-seq data (shLYL1 vs. shNC) showed that genes directly activated by AE are significantly down-regulated in the LYL1 knockdown (KD) group (*SI Appendix*, Fig. S3*A*), while genes directly repressed by AE did not show significant preference for up- or down-regulation upon LYL1 depletion (*SI Appendix*, Fig. S3*B*). These results suggest that LYL1 preferentially coactivates gene expression with AE.

As RNA expression changes upon LYL1 depletion could be indirect effects, we next examined how AETFC affects gene expression of targets that have AETFC binding to nearby enhancers. To this end, we first identified genes that are close to any AETFC-bound enhancers, and then examined how these genes respond to AE depletion. GSEA analyses showed that target genes of these sites are generally down-regulated by AE depletion, indicating strong correlation of AETFC-bound enhancers and gene activation by AE (Fig. 3*A*). To determine whether there might also be a preference for gene activation for AE/E protein complexes, we next analyzed target genes of other complexes as a control. As earlier results (above) indicated that LYL1 and LMO2 can still bind, to some extent, to group 3 (primarily AE–E2A bound) and group 4 (primarily AE–E2A–HEB bound) targets, we favor the idea that group 2 targets (primarily AE–HEB bound) most likely reflect occupancy by the biochemically identified AEC that does not include any LYL1 (*SI Appendix*, Fig. S2*C*). Therefore, for simplicity and to avoid peaks with weak LYL1 binding, from this point on we focused on the AE–HEB-enriched group 2 targets (now considered AEC targets) for the following analyses. Considering that the AE–HEB (AEC) complex is predominantly enriched at gene promoters (Fig. 2*D*, group 2), we used target genes of AE–HEB-bound promoters for GSEA analysis. The GSEA results indicate that genes with AE–HEB complex (AEC) binding at promoters do not show any significant preference in gene activation or repression (Fig. 3*B*). Thus, the overall GSEA results suggest that LYL1-containing AETFC preferentially regulates a set of AE-activated genes while AEC does not have such a clear preference. In confirmation of these observations, RT-qPCR analyses showed that LYL1 depletion reduces expression of AE-activated genes with enhancer-bound

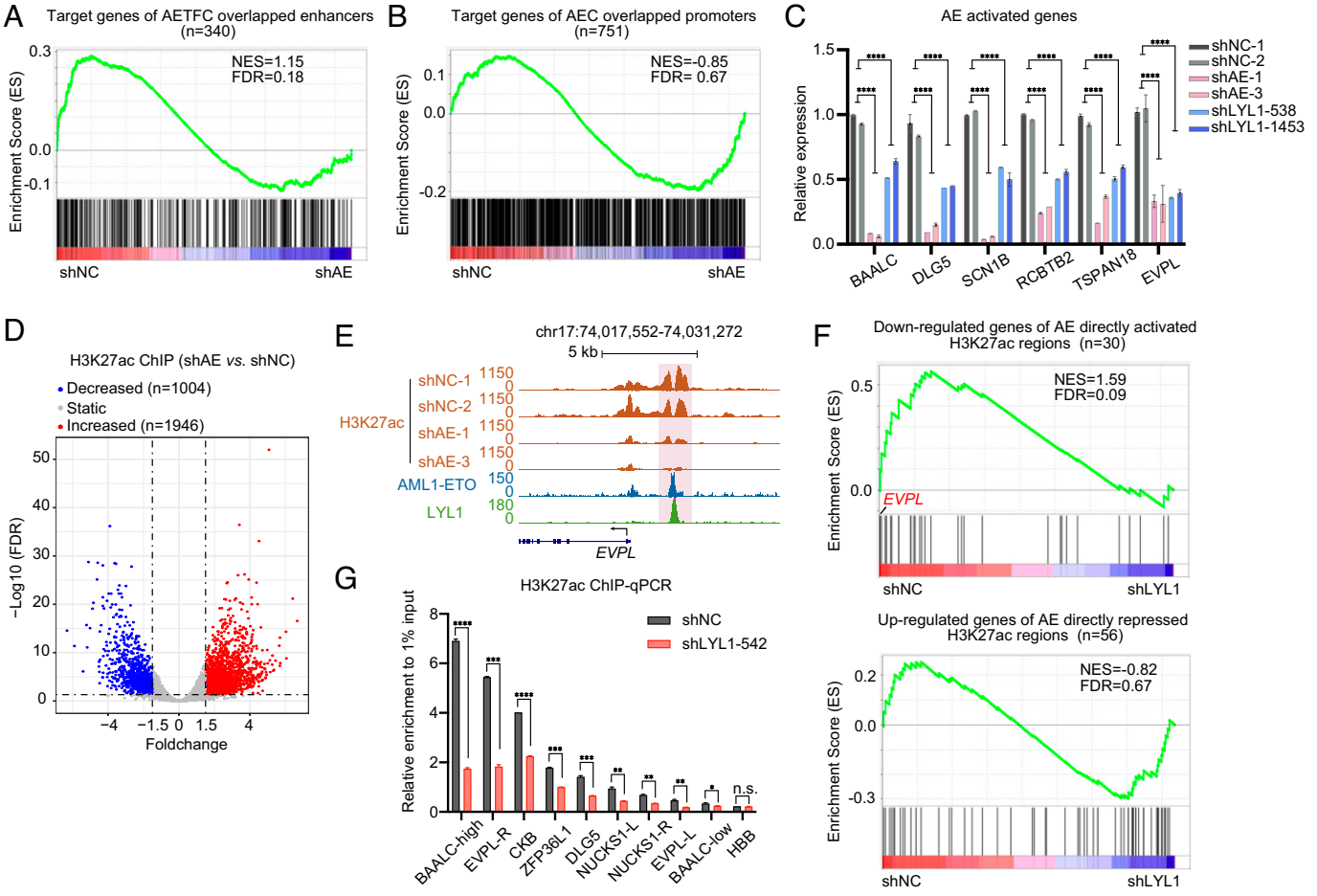


Fig. 3. LYL1-containing AETFC activates AE target genes by regulating H3K27ac levels. (*A* and *B*) GSEA analysis was used to determine the enrichment of target genes with enhancers bound by AETFC (*A*) and target genes with promoters bound by AEC (*B*) in RNA-seq data from Kasumi-1 cells treated with shRNA targeting AE (shAE) compared with cells treated with control shRNA (shNC). AETFC indicates the AE, HEB/E2A, LYL1, LMO2, and LBD1 complex. AEC indicates the AE and HEB complex. (*C*) RT-qPCR analyses of RNA levels in Kasumi-1 cells treated with either control shRNA (gray), two separate AE shRNAs (pink), or two separate LYL1 shRNAs (blue). AE-activated genes were selected for validation. Data are presented as mean ± SD. *P* values were determined using unpaired two-tailed Student's *t* test; *****P* < 0.0001. (*D*) Volcano plot showing the H3K27ac ChIP-seq peaks with decreased (blue dots), increased (red dots), or static (gray dots) signals between AE-depleted and control Kasumi-1 cells (shAE vs. shNC). (*E*) The *EVPL* gene locus indicated below was chosen as a representative region. AE-activated H3K27ac peaks (light pink shaded) on the *EVPL* gene are enriched for AE and LYL1 binding. ChIP-seq signal tracks for H3K27ac from Kasumi-1 cells treated with control and two separate AE shRNAs (brown), for AE (blue), and for LYL1 (green) are shown. Track names are indicated on the *Left*. (*F*) GSEA analysis of RNA-seq data (shAE vs. shNC) was used to determine: 1) the enrichment of AE-activated genes that are regulated by AE-activated H3K27ac regions (*Upper*, the bar of *EVPL* gene is marked) and 2) the enrichment of AE-repressed genes regulated by AE-repressed H3K27ac regions (*Lower*). (*G*) ChIP-qPCR analyses of H3K27ac enrichment on target genes *BAALC* (*BAALC*-high and *BAALC*-low), *EVPL* (*EVPL*-L and *EVPL*-R), *CKB*, *ZFP36L1*, *DLG5*, and *NUCKS1* (*NUCKS1*-L and *NUCKS1*-R) in Kasumi-1 cells treated with control (gray bars) or LYL1-542 (red bars) shRNA. High and low stand for different peaks in the BAALC gene. L and R stand for left and right peaks, respectively, near the NUCKS1 and EVPL genes. AE-activated H3K27ac peaks were selected for validation. Data are presented as mean ± SD. *P* values were determined using unpaired two-tailed Student's *t* test; *****P* < 0.0001; ****P* < 0.001; ***P* < 0.01; **P* < 0.05; n.s. represents no significance.

AETFC (Fig. 3*C* and *SI Appendix*, Fig. S3*C*) but has no effect on AE-repressed genes (*SI Appendix*, Fig. S3*D*).

Since LYL1-containing AETFC binds to active enhancers with high levels of H3K27ac (Fig. 2 *E* and *F*), we reasoned that this complex may affect enhancer activities to regulate the expression of its target genes. In this regard, we first performed H3K27ac ChIP-seq upon AE depletion and identified 1,946 increased and 1,004 decreased peaks with fold change >1.5 and FDR <0.05 (Fig. 3*D*). This result is consistent with the established view that AE is mainly a repressor, but on the other hand demonstrates that AE can also activate transcription, presumably through the action of LYL1 and associated factors at enhancer sites. In Fig. 3*E* we show an enhancer near the *EVPL* gene, as a representative region of an AE-dependent enhancer enriched with the H3K27ac mark and cooccupied by AE and LYL1. Next, in order to identify enhancers that are directly regulated by AE through H3K27ac modulation, we overlapped differential H3K27ac peaks upon AE knockdown and AE-bound regions identified by AE ChIP-seq. We then obtained the nearby genes of AE-regulated H3K27ac regions. GSEA analysis showed that AE target genes with a nearby H3K27ac peak that is directly activated by AE have a significant preference for activation by AE (*SI Appendix*, Fig. S3*E*). However, the AE target genes with a nearby H3K27ac peak that is repressed by AE do not show a significant preference for repression by AE (*SI Appendix*, Fig. S3*F*). To examine whether LYL1 is involved in regulating expression of genes with enhancers activated by AE, we performed GSEA to determine whether those genes are associated with gene activation by LYL1. The results indicate that LYL1 indeed activates AE target genes whose enhancers are activated (increased H3K27ac) by AE (Fig. 3 *F*, *Upper*). Again, LYL1 showed no preferential activation or repression of AE target genes whose enhancers are directly repressed by AE (Fig. 3 *F*, *Lower*). Considering that LYL1 is an important factor for AETFC complex formation and function, we hypothesized that it is required for H3K27ac enrichment at AETFC sites. We

therefore performed H3K27ac ChIP-qPCR upon LYL1 depletion and examined whether AE directly activated enhancers would be repressed (as evidenced by loss of H3K27ac). Indeed, H3K27ac levels on enhancers that are activated by AE were also decreased (repressed) upon LYL1 KD (Fig. 3*G* and *SI Appendix*, Fig. S3*G*). In summary, our analyses support a model in which LYL1 facilitates AE-dependent gene activation through enhancer activation by increasing H3K27ac levels.

AETFC Recruits CARM1 to Activate AE Target Genes. What is the mechanism that distinguishes AETFC from other AE-containing complexes in its transcriptional activities? Transcription factors recruit cofactors to either activate or repress gene expression. Therefore, we set out to examine whether AETFC and AEC bind differentially to cofactors. To this end, we first purified AETFC and AEC complexes following baculovirus-mediated expression in Sf9 cells and verified their composition by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) with Coomassie staining (*SI Appendix*, Fig. S4*A*). Corresponding AE complexes, containing HA tags, were then incubated with leukemic nuclear extract in search of cofactors that preferentially associate with AETFC relative to AEC. Associated proteins were purified on anti-HA agarose and then analyzed by mass spectrometry. This analysis identified the previously reported AE-interacting coactivator PRMT1 (14) and transcription factor RUNX1 (16), both of which showed comparable associations with AETFC and AEC (Fig. 4*A*). Although we did not observe previously identified cofactors p300 (8) and JMJD1C (21), this likely reflects the limiting amount of nuclear extract used for the current experiments relative to our previous studies. Interestingly, however, the arginine methyltransferase CARM1, a well-studied coactivator (26, 29), was specifically found in AETFC-associated proteins (Fig. 4*A*). To examine whether LYL1 contributed to the CARM1 association with AETFC, we performed an interaction assay with purified HEB/LYL1 heterodimer and CARM1. For this analysis, mCherry and mCherry-CARM1 proteins were purified, separately immobilized on glutathione beads bound with GST-mCherry nanobody (30), and then incubated with a purified HEB/LYL1 complex. The results confirm a direct HEB/LYL1–CARM1 interaction (Fig. 4*B*).

Next, we set out to examine the chromatin association of CARM1 with AETFC. To this end, we performed ChIP-seq for CARM1 in Kasumi-1 cells and found a significantly higher enrichment of CARM1 at AETFC-bound sites compared to AEC sites. These results are consistent with our mass spectrometry and purified CARM1-LYL1/HEB interaction data and, altogether, suggest an AETFC involvement in CARM1 recruitment to chromatin (Fig. 4*C* and *SI Appendix*, Fig. S4*B*). The data also confirm CARM1 as a cofactor selectively associated with AETFC, relative to AEC, on chromatin. As CARM1 has previously been shown to be critical for leukemogenesis but dispensable for normal hematopoiesis (27), we performed shRNA KD in Kasumi-1 cells and confirmed its requirement for survival of these cells (Fig. 4*D*). We next asked whether, as expected, CARM1 regulates expression of AETFC-regulated genes. Notably, depletion of CARM1 indeed reduced the expression of AETFC-activated genes (*SI Appendix*, Fig. S4 *C* and *D*). Since CARM1 interacts directly with AETFC in vitro and colocalizes genomically with AETFC, we proposed, mechanistically, that AETFC recruits CARM1 to chromatin. To test this hypothesis, we performed ChIP-seq for CARM1 in Kasumi-1 cells upon AE depletion and identified CARM1-bound sites that are sensitive to AE depletion. Notably, of all the CARM1-bound sites that change upon AE depletion by 1.5-fold and FDR < 0.05, over 95% (1,264 peaks) are reduced upon AE depletion (Fig. 4 *E*, blue dots), clearly indicating that AE complexes play an important role in CARM1 recruitment. Importantly, of all the AE-dependent CARM1 sites with AE binding, more than 75% are also cooccupied by LYL1 (Fig. 4*F*), suggesting that the larger AETFC plays a key role in CARM1 chromatin association. Using *DLG5* (31) as a representative AE–LYL1 target region, the data clearly show that CARM1 peaks that overlap with AE peaks are reduced upon AE depletion (Fig. 4*G*). Given the CARM1–AETFC interaction and in order to directly show that CARM1 chromatin association is sensitive to the presence of LYL1, we performed CARM1 ChIP-qPCR upon LYL1 depletion. The results show, as anticipated, that ablation of LYL1 leads to decreases in CARM1 binding on the enhancer regions of AETFC targets (Fig. 4*H* and *SI Appendix*, Fig. S4*E*).

Having demonstrated a key function for LYL1 in the assembly of the large AETFC, the selective binding of AETFC to active enhancers (high H3K27ac), and the regulation of cognate genes, the CARM1 results are consistent with these observations and pinpoint a key part of the underlying gene activation mechanism. Specifically, they show a preferred CARM1 association with AETFC relative to AEC, a direct CARM1 interaction with AETFC (including LYL1/HEB), and a LYL1- and AE-dependent recruitment of CARM1 to AETFC target genes. These results lead to a model (*SI Appendix*, Fig. S4*F*), whose physiological relevance is supported by the disruption of AETFC-activated gene expression and the inhibition of cell survival and growth by CARM1 depletion. The CARM1 coactivator function likely complements the functions of previously documented AE coactivators that include p300, also shown to bind preferentially to AETFC relative to AE (4), as well as PRMT1 (14) and JMJD1C (15).

Discussion

AML1-ETO (AE), the product of the t(8;21) AML driver oncogene, was earlier shown to exist in an unusually stable transcription factor complex AETFC (4). In this study, we first show, biochemically, that AE-containing complexes from Kasumi-1 cells are heterogeneous and composed of complete AETFC complexes as well as LYL1-free subcomplexes containing AE in association with HEB and/or E2A. These results are also consistent with the differential genomic distribution of the distinct complexes. We further show that LYL1 is the key factor that mediates assembly of the large AETFC. Notably, the LYL1-containing AETFC preferentially binds to active enhancers and mediates gene activation. With a focus on gene activation by AE and toward a further understanding of the underlying molecular mechanism of action of AETFC, we identified an interacting coactivator, CARM1, whose recruitment mechanism and essential functions are established. Our work provides insights into mechanisms of gene regulation, most notably gene activation, by AETFC and valuable information for potential therapeutic targets.

Heterogeneity of AE-Containing Complexes. Transcription factors usually work in combination to regulate gene expression. Unlike most functional transcription factor complexes, AETFC is a stable complex, reflecting in part an interaction between the NHR2 domain of AE and E proteins (HEB or E2A) (4, 8). However, our earlier identification of AETFC involved a one-step affinity purification scheme; and no previous studies have carefully examined the potentially complex nature and heterogeneity of AE-containing complexes. E proteins belong to the basic helix–loop–helix (bHLH) group of transcription factors. The class I

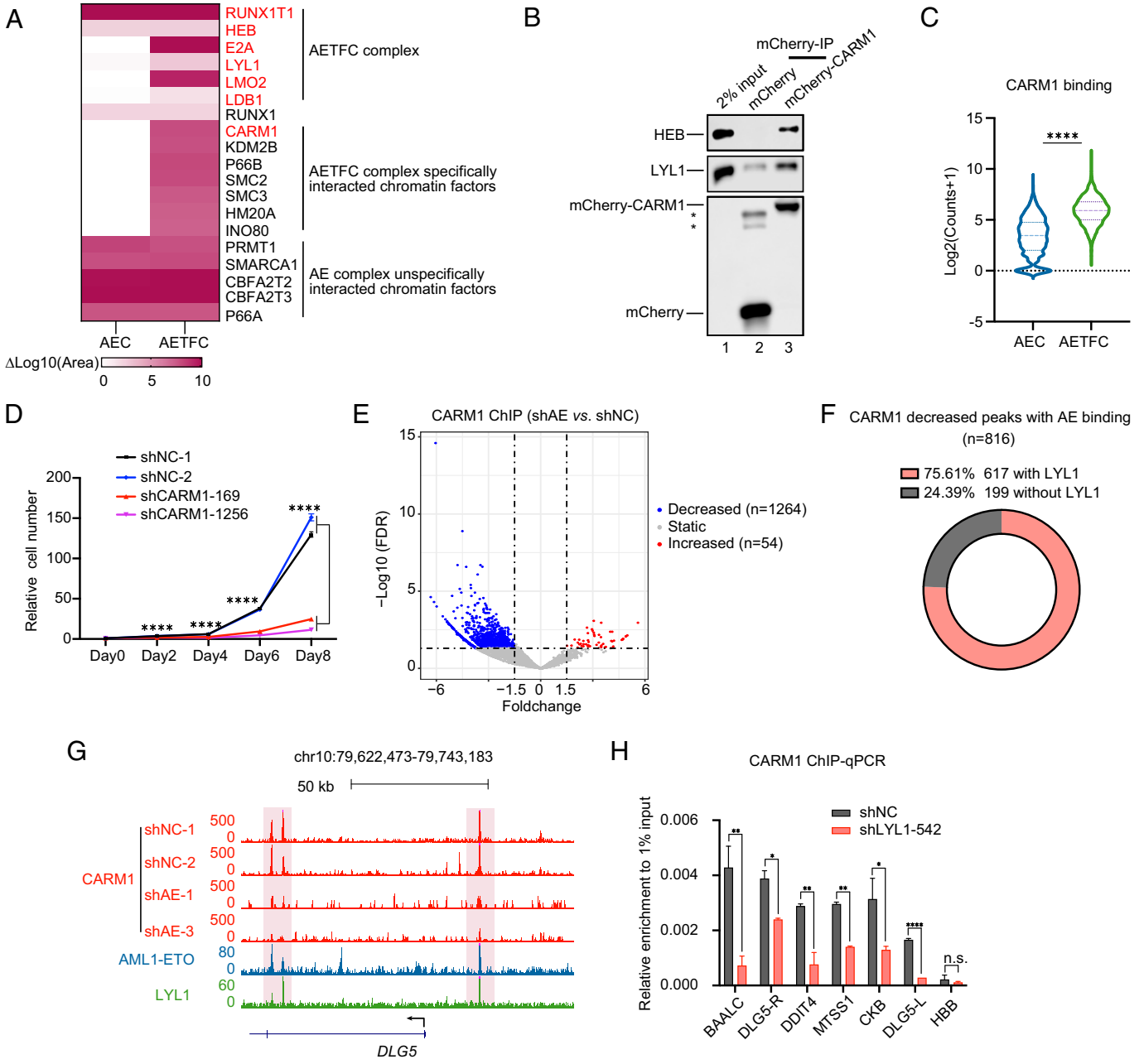


Fig. 4. AETFC recruits coactivator CARM1 to activate target gene expression. (*A*) Heatmap indicating the abundance of AETFC components and associated chromatin factors, monitored by MS, following IP of purified AE/HEB and AETFC complexes (with HA-tagged AE) incubated with Kasumi-1 nuclear extract. Δlog10(area) is the log10(area) value of each protein identified in each IP sample subtracted by the log10(area) value of the same protein in the control sample. (*B*) Direct interaction between f-LYL1/HEB heterodimer and mCherry-CARM1 protein. Purified LYL1/HEB heterodimer was incubated with mCherry or mCherry-CARM1 immobilized on GST beads through a GST–anti-mCherry nanobody fusion protein. Immunoprecipitates were analyzed by immunoblot with antibodies indicated on the *Left*. (*C*) Violin plots indicating the binding strength of CARM1 on AEC and AETFC peaks. The three thin dashed lines indicate quartile positions. *P* values were calculated by unpaired two-tailed Student's *t* test; ****$P < 0.0001$. (*D*) Assessment of proliferation of Kasumi-1 cells treated with either control shRNAs or two separate CARM1 shRNAs. Data are presented as mean ± SD. *P* values were determined using unpaired two-tailed Student's *t* test; ****$P < 0.0001$. (*E*) Volcano plot showing the CARM1 ChIP-seq peaks with decreased (blue dots), increased (red dots), or static (gray dots) signals between AE-depleted and control Kasumi-1 cells (shAE vs. shNC). (*F*) Pie chart showing the percentages of LYL1 occupancy on CARM1 peaks that are sensitive to AE depletion. Total peaks here represent CARM1 peaks that are sensitive to AE depletion and are also occupied by AE. (*G*) Representative region of the *DLG5* gene locus showing AE-dependent CARM1 peaks that are enriched for AE and LYL1 binding. ChIP-seq signal tracks for AE (blue), LYL1 (green), and CARM1 from Kasumi-1 cells treated with control and AE shRNAs (red) are shown. Track names are indicated on the *Left* and the gene name is indicated below. (*H*) ChIP-qPCR analyses of CARM1 occupancy on target genes *BAALC*, *DLG5* (*DLG5-L* and *DLG5-R*), *MTSS1*, *CKB*, or *DDIT4* in Kasumi-1 cells treated with control (gray bars) or LYL1-542 (red bars) shRNA. AE-dependent peaks were selected for validation. Data are presented as mean ± SD. *P* values were determined using unpaired two-tailed Student's *t* test; **$P < 0.01$; *$P < 0.05$; n.s. represents no significance.

bHLH factors HEB and E2A can form homodimers or heterodimers, as well as heterodimers with class II bHLH factors such as LYL1 and TAL1 (32). The myriad of dimerization possibilities, along with other strong protein–protein interactions, raised the possibility that the originally characterized natural AETFC could reflect a group of heterogeneous complexes, an issue not previously addressed. Here, our biochemical experiments have unequivocally demonstrated the presence of at least two complexes: the AE–class I bHLH E protein complex (AEC) that lacks LYL1/LMO2/LDB1, and the larger AETFC in which LYL1 heterodimerizes with AE-bound E2A and/or HEB and in turn recruits LMO2 and LDB1 to the complex. Our demonstration of a stark

difference between the preferential genomic (chromatin) localizations of LYL1-containing AETFC to active enhancers and LYL1-lacking AE/HEB complexes to promoters (discussed below) is further indicative of functional differences between these two types of complexes. Exactly how LYL1 prompts preferential enhancer association by AETFC and what promotes preferential promoter localization of the AE/HEB complex are both important questions for future studies.

AETFC May Promote Gene Activation by Mediating Enhancer–Promoter Contacts. As mentioned above, our data show that LYL1-containing AETFC is significantly enriched at active enhancers. Enhancers are typically tens to hundreds of kilobases away from target gene promoters (33), and their function is thought to be mediated by three-dimensional DNA looping that brings an enhancer physically close to the promoter that it regulates (34). Recent studies have shown that TFs and cofactors can orchestrate chromatin loops that play critical roles in regulating target gene expression in various cell types (35–39). In this regard, LDB1 previously was implicated in chromatin loop formation (39–42). Therefore, because it contains LDB1, AETFC can probably activate gene transcription via long-range interactions of enhancers and promoters. Anchors of TF-mediated loops were reported to be frequently enriched for the H3K27ac mark (35, 38, 43–45). In this regard, our data show that H3K27ac marks are highly enriched at AETFC peaks, which suggests that enhancer-bound AETFC may mediate long-range interactions.

Notably, in contrast to the correlation of LYL1-containing (AETFC) complexes with active enhancers and gene activation, LYL1-free (AEC) complexes could not be selectively correlated with either active or inactive genes. However, the greater abundance of AEC complexes relative to AETFC complexes and the apparently more prominent role of AE in gene repression than in activation (46) raise the possibility that the AE repression functions may be mediated primarily by AEC complexes.

Potential for AETFC to Form Functional Complexes in Association with Additional TFs. The dysregulation of key hematopoietic TFs has long been known to play a role in leukemogenesis and, in recent years, increasing numbers of transcription factors and cofactors have proved essential both for hematopoiesis and for leukemogenesis (47). Here, in an extension of our earlier studies (4, 15), we have demonstrated that oncogenic transcription factor LYL1 is required for AETFC assembly and function. In addition, in exploring the possibility that other TFs are involved in AETFC function, motif enrichment analyses of AETFC- and AEC-bound genomic sites have shown that the GATA2 motif sequence is specifically enriched at AETFC peaks relative to AEC peaks. GATA2, which is a member of the GATA family of zinc finger–containing TFs (48), has been identified as a critical regulator of hematopoietic stem cells (HSCs) during oncogenesis in the hematopoietic system. In this regard, deletion of GATA2 in Meis1a/Hoxa9-driven AML impedes maintenance and self-renewal of leukemic stem cells (LSCs) (49). Therefore, and in light of our observation of an association of AETFC with the key hematopoietic transcription factor RUNX1, AETFC could function in part by competing with GATA2 for binding of RUNX1, thus leading to the dysregulation of RUNX1/GATA2 target genes and consequent enhanced proliferation and blocked differentiation of leukemia cells.

AETFC Recruits CARM1 to Activate Gene Expression. Whether a transcription factor activates or represses a gene largely depends on what protein factors it interacts with at its site of action. In previous studies, AE-associated (co)factors were identified primarily through coimmunoprecipitation–mass spectrometric methods that may have overlooked heterogeneity of AE complexes. Here, using a modified approach that compared proteins interacting with independent AE–HEB (AEC) and AETFC preparations, we identified a preferential association of coactivator CARM1 with AETFC relative to AEC. Complementary biochemical and genomic analyses confirmed that CARM1, through a direct interaction with LDB1/HEB, is specifically recruited to AETFC target genes by AETFC, which provides an explanation for at least one major aspect of gene activation by AETFC (*SI Appendix*, Fig. S4*D*). Importantly, we further show that CARM1 is required for expression of AETFC target genes and leukemic cell growth, consistent with a previous report that loss of CARM1 strongly impairs leukemogenesis by regulating cell-cycle progression, myeloid differentiation, and apoptosis (27). Similar phenotypes have been observed with the deletion of AE or LYL1 (6, 18, 50). Our biochemical, genetic, and genomic analyses thus establish both a mutual AETFC and CARM1 dependency and underlying molecular mechanisms.

The Methyltransferase Activity of CARM1 and CARM1 as a Potential Therapeutic Target. Protein arginine methylation is an indispensable posttranslational modification (PTM) implicated in epigenetic regulatory mechanisms (25, 26). CARM1 is a type I arginine methyltransferase enzyme that specifically effects asymmetric dimethylation of H3R17 and H3R26 residues in histones (51). CARM1 also has many nonhistone protein substrates that include transcription factor RUNX1, coactivators p300/CBP and Mediator, and components of SWI/SNF complexes (52). Interestingly, CARM1 and CBP/p300 have been shown to cooperate in effecting transcriptional activation by estrogen receptor (53–55) and by p53 (56), with a direct crosstalk between arginine methylation and lysine acetylation being most clearly demonstrated in the biochemical p53 studies. Whether the methytransferase activity of CARM1 is required for its function as coactivator of AETFC is currently unclear. It also is unclear whether CARM1 activity facilitates p300/CBP activity (including H3K27 acetylation), and potentially p300/CBP binding, or whether p300/CBP activity facilitates CARM1 activity, as in the case of p53-dependent transcription (56). Such mechanistic studies will guide further progress in therapeutic development. Small-molecule inhibitors that target the methyltransferase activity of PRMTs and show antitumor activity have been developed (57). More importantly, in contrast to its key function in transformed cells, CARM1 appears to play a modest role in normal HSPC differentiation and proliferation (27), which indicates a favorable therapeutic value for CARM1 inhibitors. Taken together, our findings demonstrate heterogeneity within AE-containing complexes and indicate essential LYL1 functions both for stable incorporation of LDB1 and LMO2 into AETFC and for CARM1 recruitment, through direct AETFC interaction, and function on AETFC target genes. These results provide a mechanism for AETFC-mediated gene activation and further establish CARM1 as a potential therapeutic target in AML.

Materials and Methods

Purification of Complexes with Different AETFC Components. The Bac-to-Bac baculovirus expression system was used to generate recombinant AETFC components. To generate expression plasmids, AETFC component sequences were cloned into the pFastbac recombinant plasmid, and resultant plasmids were transferred into DH10Bac competent cells to generate bacmids. After this, FUGENE HD (Promega) was used to transfect Sf9 insect cells with bacmids to

generate baculoviruses for individual AETFC components. Next, reconstituted protein complexes were purified from Sf9 cells coinfected with different combinations of baculoviruses, and eluted proteins were analyzed and quantified by Coomassie staining or immunoblots.

The *SI Appendix* includes additional details and descriptions of cell culture, stable cell line generation, cell preparation, nuclear extraction and coimmunoprecipitation, IP-MS, in vitro binding assays, shRNA knockdown, immunoblotting, RT-qPCR, RNA-seq, ChIP-seq, and ChIP-qPCR, RNA-seq data processing and differential gene expression analysis, ChIP-seq data processing, and quantification and statistical analysis.

Data, Materials, and Software Availability. The raw and processed data of ChIP-seq and RNA-seq data generated in this study have been deposited in the Gene Expression Omnibus (GEO) database under accession no. GSE207234. GEO numbers of previously published datasets used in this study are as follows: AML1-ETO ChIP-seq of Kasumi-1 cells, GSM1082306; HEB ChIP-seq of Kasumi-1 cells, GSM1082308; E2A ChIP-seq of Kasumi-1 cells, GSM1082309; LYL1 ChIP-seq of Kasumi-1 cells, GSM1901542; and LMO2 ChIP-seq of Kasumi-1 cells, GSM1082311. All other study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. We thank Dr. Xiangdong Lu for CARM1 baculovirus and Dr. Xiaochuan Liu for help with initial bioinformatic analysis. We thank The Rockefeller University genomic core facility, the proteomic core facility of the School of Medicine at Tsinghua University, and colleagues in the R.G.R. and M. Chen laboratories for valuable discussion. R.G.R. was supported by NIH CA163086 and by Leukemia and Lymphoma Society Specialized Center of Research Grant (1740.03-19). M.C. was supported by National Key R&D Program of China 2021YFA1300100, Natural Science Foundation of China 81870122 and 92068114, and start-up support from Tsinghua University and the Peking University–Tsinghua Center for Life Sciences. Q.C. was supported by postdoc funding from Tsinghua University and the Peking University–Tsinghua Center for Life Sciences and Postdoctoral Science Foundation of China 2021M691810.

Author affiliations: [a]School of Medicine, Tsinghua University, Beijing 100084, China; [b]Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10065; [c]Department of Molecular Biology and Genetics, Bilkent University, 06800 Ankara, Turkey; and [d]Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics, National Research Center for Translational Medicine (Shanghai), Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

1. E. Papaemmanuil *et al.*, Genomic classification and prognosis in acute myeloid leukemia. *N. Engl. J. Med.* **374**, 2209–2221 (2016).
2. R. G. Roeder, 50+ years of eukaryotic transcription: An expanding universe of factors and mechanisms. *Nat. Struct. Mol. Biol.* **26**, 783–791 (2019).
3. K. Rejeski, J. Duque-Afonso, M. Lübbert, AML1/ETO and its function as a regulator of gene transcription via epigenetic mechanisms. *Oncogene* **40**, 5665–5676 (2021).
4. X. J. Sun *et al.*, A stable transcription factor complex nucleated by oligomeric AML1-ETO controls leukaemogenesis. *Nature* **500**, 93–97 (2013).
5. K. El Omari *et al.*, Structural basis for LMO2-driven recruitment of the SCL:E47bHLH heterodimer to hematopoietic-specific transcriptional targets. *Cell Rep.* **4**, 135–147 (2013).
6. N. K. Wilson *et al.*, Combinatorial transcriptional control in blood stem/progenitor cells: Genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* **7**, 532–544 (2010).
7. H. Osada, G. G. Grutz, H. Axelson, A. Forster, T. H. Rabbitts, LIM-only protein Lmo2 forms a protein complex with erythroid transcription factor GATA-1. *Leukemia* **11** (suppl. 3), 307–312 (1997).
8. J. Zhang, M. Kalkum, S. Yamamura, B. T. Chait, R. G. Roeder, E protein silencing by the leukemogenic AML1-ETO fusion protein. *Science* **305**, 1286–1289 (2004).
9. M. Yan, E. Y. Ahn, S. W. Hiebert, D. E. Zhang, RUNX1/AML1 DNA-binding domain and ETO/MTG8 NHR2-dimerization domain are critical to AML1-ETO9a leukemogenesis. *Blood* **113**, 883–886 (2009).
10. N. Liu *et al.*, Different roles of E proteins in t(8;21) leukemia: E2-2 compromises the function of AETFC and negatively regulates leukemogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 890–899 (2019).
11. M. M. Zhang *et al.*, Destabilization of AETFC through C/EBPα-mediated repression of LYL1 contributes to t(8;21) leukemic cell differentiation. *Leukemia* **33**, 1822–1827 (2019).
12. B. A. Hug, M. A. Lazar, ETO interacting proteins. *Oncogene* **23**, 4270–4274 (2004).
13. L. Wang *et al.*, The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation. *Science* **333**, 765–769 (2011).
14. W. J. Shia *et al.*, PRMT1 interacts with AML1-ETO to promote its transcriptional activation and progenitor cell proliferative potential. *Blood* **119**, 4953–4962 (2012).
15. M. Chen *et al.*, JMJD1C is required for the survival of acute myeloid leukemia by functioning as a coactivator for key transcription factors. *Genes Dev.* **29**, 2123–2139 (2015).
16. Y. Li *et al.*, Genome-wide studies identify a novel interplay between AML1 and AML1/ETO in t(8;21) acute myeloid leukemia. *Blood* **127**, 233–242 (2016).
17. A. Ptasinska *et al.*, Identification of a dynamic core transcriptional network in t(8;21) AML that regulates differentiation block and self-renewal. *Cell Rep.* **8**, 1974–1988 (2014).
18. G. P. Souroullas, J. M. Salmon, F. Sablitzky, D. J. Curtis, M. A. Goodell, Adult hematopoietic stem and progenitor cells require either Lyl1 or Scl for survival. *Cell Stem Cell* **4**, 180–186 (2009).
19. A. Miyamoto, X. Cui, L. Naumovski, M. L. Cleary, Helix-loop-helix proteins LYL1 and E2a form heterodimeric complexes with distinctive DNA-binding properties in hematolymphoid cells. *Mol. Cell. Biol.* **16**, 2394–2401 (1996).
20. T. Sanda *et al.*, Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia. *Cancer Cell* **22**, 209–221 (2012).
21. J. D. Mellentin, S. D. Smith, M. L. Cleary, lyl-1, a novel gene altered by chromosomal translocation in T cell leukemia, codes for a protein with a helix-loop-helix DNA binding motif. *Cell* **58**, 77–83 (1989).
22. B. J. Shields *et al.*, The NUP98-HOXD13 fusion oncogene induces thymocyte self-renewal via Lmo2/Lyl1. *Leukemia* **33**, 1868–1880 (2019).
23. Y. S. Meng, H. Khoury, J. E. Dick, M. D. Minden, Oncogenic potential of the transcription factor LYL1 in acute myeloblastic leukemia. *Leukemia* **19**, 1941–1947 (2005).
24. M. P. McCormack *et al.*, Requirement for Lyl1 in a model of Lmo2-driven early T-cell precursor ALL. *Blood* **122**, 2093–2103 (2013).
25. E. Guccione, S. Richard, The regulation, functions and clinical relevance of arginine methylation. *Nat. Rev. Mol. Cell Biol.* **20**, 642–657 (2019).
26. Y. Yang, M. T. Bedford, Protein arginine methyltransferases and cancer. *Nat. Rev. Cancer* **13**, 37–50 (2013).
27. S. M. Greenblatt *et al.*, CARM1 is essential for myeloid leukemogenesis but dispensable for normal hematopoiesis. *Cancer Cell* **34**, 868 (2018).
28. A. Gardini *et al.*, AML1/ETO oncoprotein is directed to AML1 binding regions and co-localizes with AML1 and HEB on its targets. *PLoS Genet.* **4**, e1000275 (2008).
29. M. T. Bedford, S. G. Clarke, Protein arginine methylation in mammals: Who, what, and why. *Mol. Cell* **33**, 1–13 (2009).
30. P. C. Fridy *et al.*, A robust pipeline for rapid production of versatile nanobody repertoires. *Nat. Methods* **11**, 1253–1260 (2014).
31. D. Mahadevan *et al.*, Transcriptosome and serum cytokine profiling of an atypical case of myelodysplastic syndrome with progression to acute myelogenous leukemia. *Am. J. Hematol.* **81**, 779–786 (2006).
32. C. Murre, Helix-loop-helix proteins and lymphocyte development. *Nat. Immunol.* **6**, 1079–1086 (2005).
33. M. Gasperini *et al.*, A genome-wide framework for mapping gene regulation via cellular genetic screens. *Cell* **176**, 377–390.e19 (2019).
34. E. E. M. Furlong, M. Levine, Developmental enhancers and chromosome topology. *Science* **361**, 1341–1345 (2018).
35. D. C. Di Giammartino *et al.*, KLF4 is involved in the organization and regulation of pluripotency-associated three-dimensional enhancer networks. *Nat. Cell Biol.* **21**, 1179–1190 (2019).
36. R. Stadhouders *et al.*, Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat. Genet.* **50**, 238–249 (2018).
37. T. M. Johanson *et al.*, Transcription-factor-mediated supervision of global genome architecture maintains B cell identity. *Nat. Immunol.* **19**, 1257–1264 (2018).
38. J. A. Bertolini *et al.*, Mapping the global chromatin connectivity network for Sox2 function in neural stem cell maintenance. *Cell Stem Cell* **24**, 462–476.e6 (2019).
39. W. Deng *et al.*, Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **149**, 1233–1244 (2012).
40. I. Krivega, R. K. Dale, A. Dean, Role of LDB1 in the transition from chromatin looping to transcription activation. *Genes Dev.* **28**, 1278–1290 (2014).
41. S. H. Song, C. Hou, A. Dean, A positive role for NLI/Ldb1 in long-range beta-globin locus control region function. *Mol. Cell* **28**, 810–822 (2007).
42. J. Lee, I. Krivega, R. K. Dale, A. Dean, The LDB1 complex co-opts CTCF for erythroid lineage-specific long-range enhancer interactions. *Cell Rep.* **19**, 2490–2502 (2017).
43. K. L. Bunting *et al.*, Multi-tiered reorganization of the genome during B cell affinity maturation anchored by a germinal center-specific locus control region. *Immunity* **45**, 497–512 (2016).
44. D. H. Phanstiel *et al.*, Static and dynamic DNA loops form AP-1-bound activation hubs during macrophage development. *Mol. Cell* **67**, 1037–1048.e6 (2017).
45. B. Bonev *et al.*, Multiscale 3D genome rewiring during mouse neural development. *Cell* **171**, 557–572.e24 (2017).
46. K. R. Stengel, J. D. Ellis, C. L. Spielman, M. L. Bomber, S. W. Hiebert, Definition of a small core transcriptional circuit regulated by AML1-ETO. *Mol. Cell* **81**, 530–545.e5 (2021).
47. Z. Wang *et al.*, Interplay between cofactors and transcription factors in hematopoiesis and hematological malignancies. *Signal Transduct. Target. Ther.* **6**, 24 (2021).
48. T. Nagai *et al.*, Transcription factor GATA-2 is expressed in erythroid, early myeloid, and CD34+ human leukemia-derived cell lines. *Blood* **84**, 1074–1084 (1994).
49. J. B. Menendez-Gonzalez *et al.*, Gata2 as a crucial regulator of stem cells in adult hematopoiesis and acute myeloid leukemia. *Stem Cell Reports* **13**, 291–306 (2019).
50. L. F. Peterson, D. E. Zhang, The 8;21 translocation in leukemogenesis. *Oncogene* **23**, 4255–4262 (2004).
51. B. T. Schurter *et al.*, Methylation of histone H3 by coactivator-associated arginine methyltransferase 1. *Biochemistry* **40**, 5747–5756 (2001).
52. S. Suresh, S. Huard, T. Dubois, CARM1/PRMT4: Making its mark beyond its function as a transcriptional coactivator. *Trends Cell Biol.* **31**, 402–417 (2021).
53. D. Chen, S. M. Huang, M. R. Stallcup, Synergistic, p160 coactivator-dependent enhancement of estrogen receptor function by CARM1 and p300. *J. Biol. Chem.* **275**, 40810–40816 (2000).
54. K. J. Veazey *et al.*, CARM1 inhibition reduces histone acetyltransferase activity causing synthetic lethality in CREBBP/EP300-mutated lymphomas. *Leukemia* **34**, 3269–3285 (2020).
55. S. Daujat *et al.*, Crosstalk between CARM1 methylation and CBP acetylation on histone H3. *Curr. Biol.* **12**, 2090–2097 (2002).
56. W. An, J. Kim, R. G. Roeder, Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53. *Cell* **117**, 735–748 (2004).
57. E. Chan-Penebre *et al.*, A selective inhibitor of PRMT5 with in vivo and in vitro potency in MCL models. *Nat. Chem. Biol.* **11**, 432–437 (2015).