A higher-order configuration of the heterodimeric DOT1L–AF10 coiled-coil domains potentiates their leukemogenenic activity

Xiaosheng Song[a,b,c,1], Liuliu Yang[b,c,1], Mingzhu Wang[d], Yue Gu[a], Buqing Ye[b], Zusen Fan[b,c], Rui-Ming Xu[b,c,2], and Na Yang[a,2]

[a]State Key Laboratory of Medicinal Chemical Biology, College of Pharmacy and Tianjin Key Laboratory of Molecular Drug Research, Nankai University, 300353 Tianjin, China; [b]National Laboratory of Biomacromolecules, Chinese Academy of Sciences Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 100101 Beijing, China; [c]School of Life Sciences, University of Chinese Academy of Sciences, 100049 Beijing, China; and [d]Institutes of Physical Science and Information Technology, Anhui University, 230601 Hefei, Anhui, China

Edited by Dinshaw J. Patel, Memorial Sloan Kettering Cancer Center, New York, NY, and approved August 27, 2019 (received for review March 18, 2019)

Chromosomal translocations of *MLL1* (Mixed Lineage Leukemia 1) yield oncogenic chimeric proteins containing the N-terminal portion of MLL1 fused with distinct partners. The MLL1–AF10 fusion causes leukemia through recruiting the H3K79 histone methyltransferase DOT1L via AF10's octapeptide and leucine zipper (OM-LZ) motifs. Yet, the precise interaction sites in DOT1L, detailed interaction modes between AF10 and DOT1L, and the functional configuration of MLL1–AF10 in leukeomogenesis remain unknown. Through a combined approach of structural and functional analyses, we found that the LZ domain of AF10 interacts with the coiled-coil domains of DOT1L through a conserved binding mode and discovered that the C-terminal end of the LZ domain and the OM domain of AF10 mediate the formation of a DOT1L–AF10 octamer via tetramerization of the binary complex. We reveal that the oligomerization ability of the DOT1L–AF10 complex is essential for MLL1–AF10's leukemogenic function. These findings provide insights into the molecular basis of pathogenesis by MLL1 rearrangements.

DOT1L | MLL1–AF10 fusion | structure | leukemogenesis | octamerization

Chromosomal rearrangement of *MLL1* (Mixed Lineage Leukemia 1), which is a trithorex group gene located on chromosome 11q23, results in a number of acute leukemias including ALL (acute lymphoid leukemia) and AML (acute myeloid leukemia) (1, 2). The gene product, MLL1, is a member of the multidomain KMT2 family of histone lysine methyltransferases (HKMTs). MLL1 contains an N-terminal DNA-targeting domain and a C-terminal SET domain, which is not catalytically active on its own but becomes competent for trimethylation of histone H3 lysine-4 (H3K4) upon formation of a complex with WDR5, RbBP5, and Ash2L (3–5). Chromosomal translocations of *MLL1* generate oncogenic chimeric proteins containing the noncatalytic N-terminal portion of MLL1 (amino acids 1 to 1,395) fused with many other partners such as AF10, AF9, AF4, ENL, and ELL (6–8). These fusion proteins retain the DNA-binding domain of MLL1 and gain the functions of the fusion partner's activities. Several of these fusion partners, including AF9, AF10, and ENL, interact with and recruit a distinct HKMT, DOT1L, to the MLL1-regulated genes. DOT1L is a non-SET domain HKMT catalyzing monomethylation, dimethylation, and trimethylation of nucleosomal H3K79 (9–11). The recruitment of DOT1L by MLL1 fusion proteins result in the hypermethylation of H3K79, leading to upregulation of a number of leukemia-relevant genes, such as the *HOXA* cluster and *MEIS1*. For example, overexpressing *HOXA9* in mice induces the transformation of primary bone marrow cells that leads to AML through specific cooperation between *HOXA9* and *MEIS1a* genes (12, 13), reduced expression of *HOXA7* and *HOXA10* induces apoptosis in a human leukemia cell line (14), and transcriptional activation of *HOXA10* and *HOXA11* causes a subset of T cell ALL (15). Although the exact mechanism by which DOT1L leads to transcription activation remains to be precisely characterized, it is clear that the catalytic activity of DOT1L is required. There are also evidences that the transcriptional regulation activity of DOT1L involves the participation in several transcription elongation complexes like SEC (Super Elongation Complex) and EAP (ENL associated proteins) (16, 17).

AF10 (ALL1-Fused gene from chromosome 10 protein) is one of the MLL1-fusion partners that function in ALL (18). The N-terminal PZP domain of AF10 recognizes unmethylated H3K27 and regulates DOT1L-mediated H3K79 methylation (19). As a consequence of chromosomal rearrangement, the C-terminal portion of human AF10 (amino acids 683 to 1,068) is fused to the N-terminal 1,395-residue segment of MLL1 (Fig. 1*A*). AF10 has been shown to regulate *HOX* gene expression through H3K79 methylation in diverse AML subtypes (20) by recruiting DOT1L through direct interaction via its OM-LZ (Octapeptide Motif and Leucine Zipper) region located at the C-terminal part of AF10 (21). The AF10-interacting region in DOT1L is less well defined. Human DOT1L is also a multidomain-containing protein, with its H3K79 HKMT activity residing in the N-terminal region. A 416-residue fragment of DOT1L possesses full HKMT activity in vitro,

Significance

The leukemogenic function of the MLL1–AF10 fusion protein is caused by aberrant recruitment of histone methyltransferase DOT1L to MLL1-regulated genes, yet a precise picture of their interaction remains elusive. Here, we find that AF10 and DOT1L forms a tight binary complex via packing of the leucine zipper helix of AF10 and the first coiled-coil helix of DOT1L. Surprisingly, 4 DOT1L–AF10 heterodimers form a tetramer of heterodimers. We show that the hetero-octamer exists in solution, and the oligomeric form of the AF10–DOT1L complex is crucial for the leukemogenic activity of MLL1–AF10. This discovery reveals the molecular basis of the pathogenic function by MLL1–AF10 and may facilitate the development of more effective strategy against MLL1–AF10-mediated leukemia.

Author contributions: Z.F., R.-M.X., and N.Y. designed research; X.S., L.Y., M.W., and Y.G. performed research; X.S., L.Y., M.W., Y.G., B.Y., R.-M.X., and N.Y. analyzed data; and X.S., R.-M.X., and N.Y. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: Atomic coordinates and associated structure factors have been deposited in the Protein Data Bank (PDB ID code: 6JN2).

[1]X.S. and L.Y. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: rmxu@ibp.ac.cn or yangnanku@nankai.edu.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1904672116/-/DCSupplemental.

First published September 16, 2019.

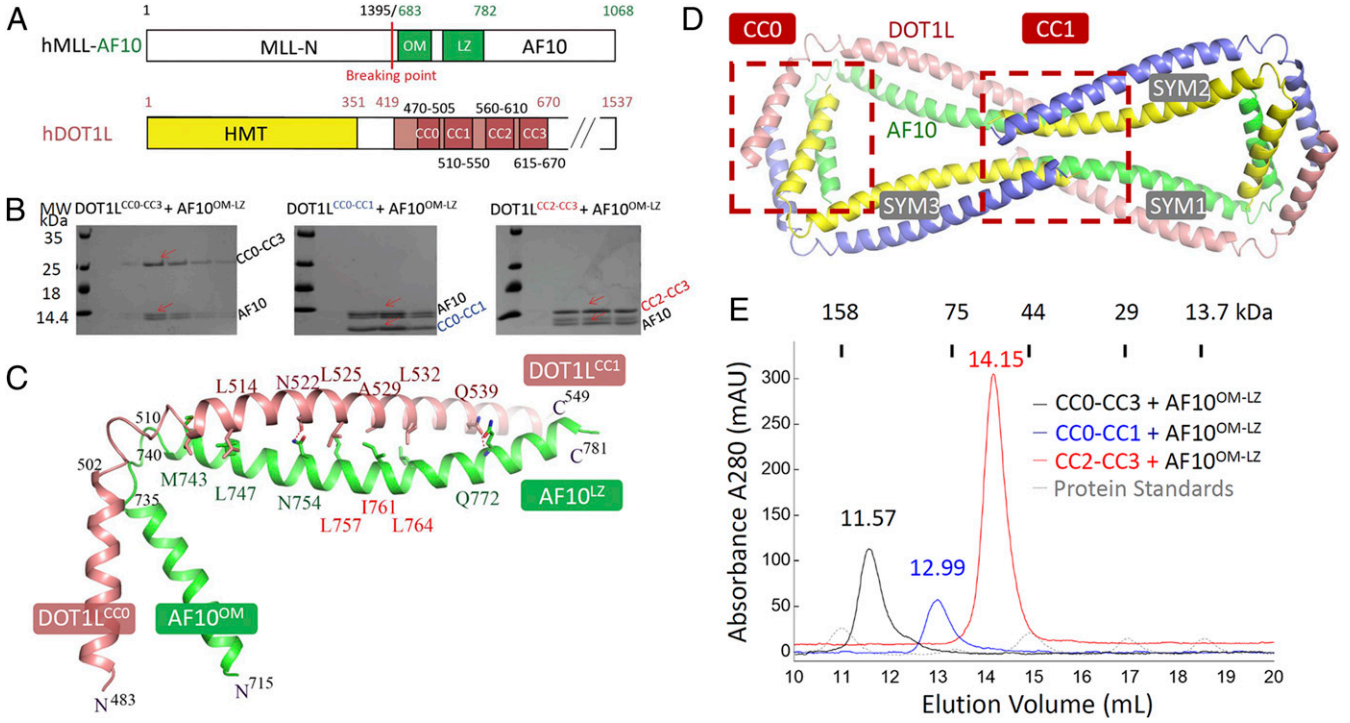

Fig. 1. Functional domains and structures of the human AF10–DOT1L complex. (*A*) Schematic diagram of domain structures of MLL1–AF10 and DOT1L. MLL1–AF10 is composed of MLL1-N (amino acids 1 to 1,395) fused by AF10-C (amino acids 683 to 1,068). DOT1L contains an N-terminal histone methyltransferase (HMT) domain (yellow) followed by a CC domain (salmon), and a large C-terminal region of unknown functions. (*B*) SDS/PAGE detection of copurification of $AF10^{OM\text{-}LZ}$ with different CC domains of DOT1L through gel-filtration analysis. (*C*) Structure of one $DOT1L^{CC0\text{-}CC1}$–$AF10^{OM\text{-}LZ}$ heterodimer in one crystallographic asymmetric unit. DOT1L is shown in salmon and AF10 in green. Interacting residues are displayed in a stick model, and the 3 key hydrophobic residues of AF10 are labeled in red. (*D*) $DOT1L^{CC0\text{-}CC1}$–$AF10^{OM\text{-}LZ}$ forms a bow tie-shaped octameric structure across the crystal lattice. The DOT1L protomer and a symmetry mate (SYM1) are shown in salmon, and the SYM2 and SYM3 symmetry mates are shown in slate, while their interacting AF10 molecules are shown in green and yellow, respectively. Two protein interaction interfaces are highlighted with red dashed boxes. (*E*) Formation of higher-order oligomers upon binding of different CC domains of DOT1L with $AF10^{OM\text{-}LZ}$. Elution profile of 5 protein standards (gray dashed curve) with indicated molecular mass is superimposed.

but only the first 351 residues form the ordered but inactive catalytic core—the remaining C-terminal segment contains patches of positively charged residues required for nucleosome binding and H3K79 HKMT activity (22, 23). The human DOT1L region spanning residues 419–670 is predicted to contain 4 tandem coiled-coil (CC) domains, including a shorter CC0 followed by 3 (CC1 to CC3) longer domains (Fig. 1*A*). A recent work demonstrated that the CC1–CC3 domains of human DOT1L are involved in AF10 binding (24). This AF10 binding region in DOT1L is distinct from that of another MLL1 fusion partner, AF9, which bears no similarity in binding mode despite also showing a weak interaction with the CC3 domain of DOT1L (25). Among CC1 to CC3 of DOT1L, it was found that CC1 has the highest affinity toward $AF10^{OM\text{-}LZ}$, ~6- and 270-fold higher than that of CC2 and CC3, respectively (24). Finally, a crystal structure of a chimeric complex of human $AF10^{OM\text{-}LZ}$ and zebrafish $DOT1L^{CC2}$ was solved, providing an understanding of the intermolecular interaction between AF10 and DOT1L, but leaving a number of questions, such as whether CC1 and CC3 interact with AF10 in a similar way and what is the role of CC0 in AF10 binding, unanswered.

In this study, we report the crystal structure of the OM-LZ domain of human AF10 in complex with the CC0 and CC1 domains of human DOT1L. The structure shows that although CC1 interacts with the LZ domain of AF10 in a manner similar to that of zebrafish CC2, the CC0 domain of DOT1L together with the OM domain of AF10 play an important role in the formation of an octameric structure of the human $DOT1L^{CC0\text{-}CC1}$ and $AF10^{OM\text{-}LZ}$ complex. Through a combined structural and functional analysis, we found that both the direct interaction between DOT1L and AF10 and the octameric configuration of the DOT1L–AF10 complex are important for the leukemogenic activity of the MLL1–AF10 chimera. These results provide a more complete picture of the interaction between DOT1L and AF10, as well as deeper insights for the development of therapeutic strategies against leukemia targeting the DOT1L–AF10 complex.

Results

AF10-Interacting Domains of DOT1L. A previous study showed that the leukemogenic OM-LZ domain of AF10 (amino acids 683 to 782) is necessary and sufficient for AF10–DOT1L interaction by a mammalian 2-hybrid assay, and the AF10-interacting region in DOT1L is mapped to its N-terminal 670 residues (21). We then started to map precise interaction regions of DOT1L with $AF10^{OM\text{-}LZ}$ by in vitro copurification. Gel-filtration and SDS/PAGE results show that human DOT1L (amino acids 419 to 670) forms a stable complex with the OM-LZ domain of AF10, judged by coelution from a Superdex 200 16/60 XK size exclusion column (*SI Appendix*, Fig. S1). However, the *Escherichia coli*-expressed protein fragments appear to be not very stable, as evidenced by the presence of degradation bands in the elusion fractions (*SI Appendix*, Fig. S1). For the purpose of structural and biochemical analyses, we tested several shorter truncations of the 2 proteins based on secondary structure predictions and found that 3 shorter constructs of DOT1L, $DOT1L^{CC0\text{-}CC3}$ (amino acids 470 to 670), $DOT1L^{CC0\text{-}CC1}$ (amino acids 470 to 550), and $DOT1L^{CC2\text{-}CC3}$ (amino acids 551 to 670), can interact with a shortened $AF10^{OM\text{-}LZ}$ (amino acids 699 to 782), which will be referred to as $AF10^{OM\text{-}LZ}$ hereafter when no confusion is expected, under the same condition as in the previous copurification procedure, albeit with a Superdex 200 increase 10/300 GL gel-filtration

column (Fig. 1*B*). Our findings are consistent with the newly reported results that CC1, CC2, and CC3 can individually interact with the OM-LZ domain of AF10, and they bind with the dissociate constants of 0.1 μM, 0.6 μM, and 27.1 μM, respectively (24).

Overall Structure of the Human DOT1L–AF10 Complex. We crystallized the human $DOT1L^{CC0\text{-}CC1}$–$AF10^{OM\text{-}LZ}$ complex and solved a 3.6-Å structure. Although the structure only has a medium resolution, the anomalous difference maps of mutant selenomethionyl proteins reassured correct assignment of the protein sidechains (*SI Appendix*, Fig. S2). The structure shows that both the DOT1L and AF10 fragments comprise 2 α-helices, a shorter CC0 helix (amino acids 483 to 502) and a longer CC1 helix (amino acids 510 to 549) for DOT1L, and a shorter OM helix (amino acids 715 to 735) and a longer LZ helix (amino acids 740 to 781) for AF10 (Fig. 1*C*). The $DOT1L^{CC1}$ and $AF10^{LZ}$ helices are parallelly aligned and interact intimately with each other to form a heterodimer that closely resembles the leucine zipper structural motif, which is a common dimerization motif originally found in DNA binding transcription factors. The 2 helices interact with each other mainly via hydrophobic interactions involving leucine residues and several other residues such as isoleucine and alanine (Fig. 1*C*). The centrally located L757, I761, and L764 residues of AF10 appear to be the major contact area between the 2 helices, as they directly contact L525, A529, and L532 of DOT1L, respectively. Polar interactions can also be detected. Both the sidechain amino group of $AF10^{N754}$ and the hydroxyl group of $DOT1L^{N522}$, and the mainchain amide group of $AF10^{Q772}$ and the sidechain hydroxyl group of $DOT1L^{Q539}$, form hydrogen bonds (Fig. 1*C*). In contrast, there is little contact between the OM and CC0 domains of the 2 proteins.

Although there is only one DOT1L–AF10 heterodimer in one crystallographic asymmetric unit, a tetramer of heterodimer is observed by symmetry operation (Fig. 1*D*). In the bowtie-shaped octameric structure, 4 AF10 molecules interact with each other to form the inner frame of the tetra-heterodimer, and 4 DOT1L molecules surround outside. To verify the hetero-octameric complex in solution, we examined the oligomeric state of the CC0-CC1 and CC2-CC3 domains of DOT1L when forming complex with $AF10^{OM\text{-}LZ}$. The elution volumes of $DOT1L^{CC0\text{-}CC1}$–$AF10^{OM\text{-}LZ}$ and $DOT1L^{CC2\text{-}CC3}$–$AF10^{OM\text{-}LZ}$ complexes from a Superdex 200 increase 10/300 GL column are 12.99 mL and 14.15 mL, corresponding to estimated molecular masses of ∼84 and 59 kDa, respectively (Fig. 1*E*). These data indicate that the CC0-CC1 fragment of DOT1L forms a hetero-octamer with $AF10^{OM\text{-}LZ}$, while the CC2-CC3 fragment forms a hetero-tetramer with $AF10^{OM\text{-}LZ}$ in solution, in agreement with our structural result and the recent published crystal structure of the CC2–$AF10^{OM\text{-}LZ}$ complex, which was found to form a heterotetramer in solution due to the dimerization of AF10 (24). Furthermore, we find that the complex of the CC0-CC3 fragment of DOT1L, $DOT1L^{CC0\text{-}CC3}$, with $AF10^{OM\text{-}LZ}$ eluded from the column at 11.57 mL, which corresponds to an estimated molecular mass of about 131 kDa, a value ∼4 times of the calculated molecular mass of the DOT1L–AF10 heterodimer (Fig. 1*E*), indicating the formation of a hetero-octamer in solution with longer fragments of DOT1L.

Protein Interaction Interfaces. Interactions responsible for tetramerization of the heterodimers occur at 2 interfaces, as indicated by the boxed regions in Fig. 1*D*. One interface is located at the very C-terminal region of the LZ-helix of AF10 (CC1 interface), and another involves the OM helix of AF10 and a symmetry-related CC0 helix of DOT1L (CC0 interface). In the CC1 interface, the very C-terminal end of $AF10^{LZ}$ (amino acids L773 to V780) plays a crucial role for homotetramerization of AF10 (Fig. 2*A*). Specifically, L774 and L778 from 1 AF10 protomer engage L778 and L774 from a symmetry-related AF10, AF10–SYM1, via hydrophobic interactions, while L773 and V780 from the same AF10 protomer interact with V780 and L773 from a second symmetry-related AF10, AF10–SYM2. A third symmetry-related AF10, which is related to the AF10 protomer by a 2-fold symmetry around an axis perpendicular to the plane, does not contact the AF10 protomer directly but it interacts with AF10–SYM1 and AF10–SYM2 in a manner identical to the one above (Fig. 2*A*). In addition, L778 of all 4 AF10 molecules at the interface interact with L546 of the adjacent DOT1L molecule lined the outside rim of the bowtie-shaped DOT1L–AF10 hetero-octamer. At the CC0 interface, it is interesting to find that although the OM helix of AF10 does not interact with the CC0 helix of DOT1L in the same heterodimer, it interacts with the CC0 helix of a symmetry-related DOT1L (DOT1L–SYM3) mainly through hydrophobic interactions. Specifically, L719 and F730 of AF10 interact with F498 and L487 of DOT1L–SYM3, respectively. The OM helix of AF10 also interacts with the OM helix of AF10–SYM3 by hydrophobic interactions through 2 pairs of L720 and L731 residues (Fig. 2*B*). These hydrophobic interactions constitute the major octamerization force of the DOT1L–AF10 complex besides the interactions within the heterodimer.

To confirm the role of the 2 hydrophobic interfaces in octamer formation, we designed several AF10 mutants to test their effects in octamer formation with $DOT1L^{CC0\text{-}CC3}$. The L773R/L774R/L778R/V780R quadruple mutant of AF10 is located at the CC1 interface, and the L719R/F730R and L720R/L731R double mutants are located at the CC0 interface. Velocity sedimentation analytical ultracentrifugation experiments show that the mutant complexes all appeared significantly smaller than the WT complex, indicating that the mutations impacted the formation of the DOT1L–AF10 hetero-octamer (Fig. 2*C*). These results support the structural findings that key hydrophobic interactions at the CC1 and CC0 interfaces are important for the formation of the higher-order AF10–DOT1L complex.

Determinants for Human DOT1L–AF10 Interaction. To definitively test the roles of amino acid residues involved in intermolecular interactions, it would be ideal to carry out in vitro pull-down experiments with purified proteins. However, DOT1L fragments containing the CC0 domain cannot be expressed individually in soluble forms, although expression of DOT1L can be detected when it forms a complex with AF10 or its derivatives. Alternatively, we tested their binding by an intracellular fluorescence colocalization approach, which offers the advantage of being close to a physiological environment. Consistent with the structural findings, the CC0-CC1 domain of DOT1L colocalizes with $AF10^{OM\text{-}LZ}$ and $AF10^{\Delta OM}$, but does not interact with $AF10^{OM}$ alone, indicating that the LZ domain of AF10 is responsible for interaction with DOT1L (Fig. 3*A*, first 3 rows). A previous work showed that L757, I761, and L764 of the AF10 LZ domain interact with the CC2 domain of DOT1L (24), and the same set of amino acids are found to be important for interaction with the CC1 domain of DOT1L in our study here (Fig. 1*C*).

Our earlier results showed that mutating AF10 residues at the tetramerization interfaces of the DOT1L–AF10 heterodimers disrupt the octamer formation (Fig. 2*C*). However, they should not affect the binary interaction in the heterodimer. Indeed, when the L773R/L774R/L778R/V780R mutant of AF10 at the CC1 interface and the L719R/F730R mutant at the CC0 interface were used in the colocalization experiments, they retain the ability to bind DOT1L (Fig. 3*A*, fourth and fifth rows). Similar results were obtained using a longer DOT1L fragment containing all of the 4 CC domains (CC0 to CC3), and these interactions can also be detected in in vitro copurification experiments (Fig. 3*B* and *SI Appendix*, Fig. S3). Statistics of cells showing colocalization of AF10 or its mutants with DOT1L are shown in *SI Appendix*, Fig. S4. Altogether, the fluorescence colocalization data fully corroborate the structural and biochemical results, namely, formation of the DOT1L–AF10 heterodimer is principally mediated by hydrophobic

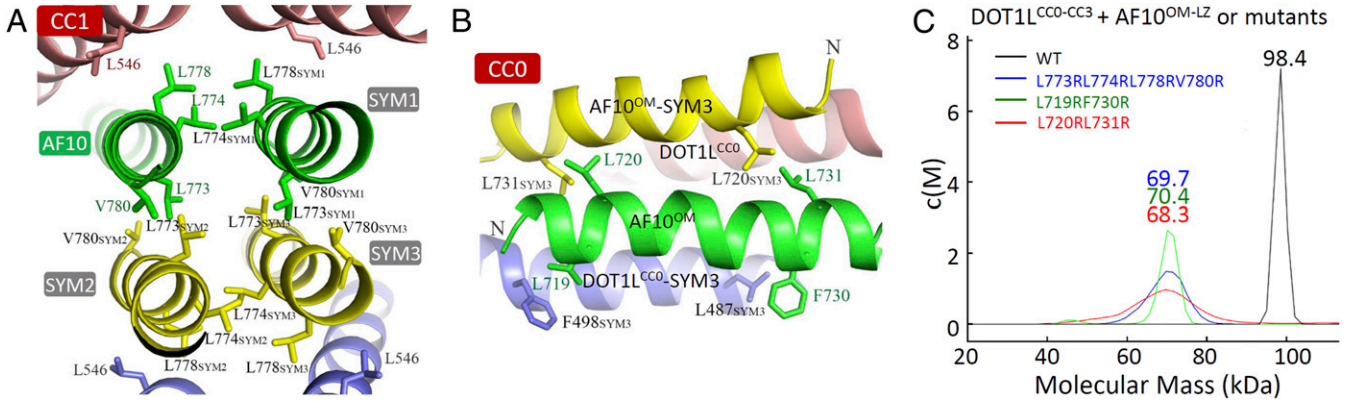

Fig. 2. Detailed interactions at the 2 oligomeric interfaces of human DOT1L–AF10 complex. (*A*) Hydrophobic interactions between 4 AF10 molecules at the CC1 interface located at the C-terminal end of the LZ helix of AF10. (*B*) Detailed interactions at the CC0 interface. The OM helix of AF10 interacts with the CC0 helix of DOT1L–SYM3 and the OM helix of AF10–SYM3. Key residues are highlighted in a stick model. (*C*) AUC results showing distinct sedimentation properties, expressed in estimated molecular mass, of complexes between $DOT1L^{CC0\text{-}CC3}$ and $AF10^{OM\text{-}LZ}$ or its mutants.

interaction between the DOT1L CC1 and the AF10 LZ motifs, and the heterodimer formation is independent of the formation of higher oligomers. These findings reveal an assembly pathway of the observed DOT1L–AF10 octamer through tetramerization of their heterodimers.

Role of the Higher-Order DOT1L–AF10 Complex in Leukemogenesis. Previous studies have linked the interaction between DOT1L and MLL1–AF10 to leukemogenesis (21, 24), but given our finding that DOT1L and AF10 form a hetero-octamer, it is natural to ask whether oligomerzation of the DOT1L–AF10 heterodimers has any pathogeneic implications. To test this, serial colony formation assays were used to evaluate the importance of interactions responsible for heterodimer and octamer formations in leukemogenesis. Hematopoietic progenitor cells isolated from murine bone marrow were transduced with retrovirus carrying MLL1-N, MLL1-AF10, or MLL1–AF10 mutants. Cells were serially plated over 3 rounds, and colony numbers were counted. Consistent with the previous reports, cells expressing the MLL1–AF10 fusion protein showed increasing numbers of colonies through 3 rounds of plating, indicating the association with leukemic transformation (Fig. 4*A*). Cells expressing $MLL1\text{–}AF10^{L757R/I761R/L764R}$, which is a mutant disrupting the binary interaction between DOT1L and AF10, exhibited remarkably reduced transformation activities, a result consistent with that of the previously reported $AF10^{L757D/I761D/L764D}$ mutant (24). The functional importance of L757, I761, and L764 again proves that the LZ domain of AF10 is crucial for the recruitment of DOT1L in the leukemogenesis function of the MLL1–AF10 fusion protein.

Next, we examined the role of tetramerization of the AF10–DOT1L heterodimer in leukemogenesis. As depicted in Fig. 1*D*, 2 interfaces are involved in the formation of AF10–DOT1L hetero-octamer. One is the CC1 interface involving L773, L774, L778, and V780 residues located at the very C-terminal region of the LZ helix of AF10 (Fig. 2*A*). Cells expressing the $MLL1\text{–}AF10^{L773R/L774R/L778R/V780R}$ mutant significantly lost its transformation ability, as exhibited by the low number of colonies after 3 rounds of plating (Fig. 4*A*). The other is the CC0 interface involving L719 and F730 of the OM helix of AF10 that interact with the CC0 helix of DOT1L–SYM3, and L720 and L731 with the symmetry-related OM helix of AF10–SYM3 (Fig. 2*B*). Cells

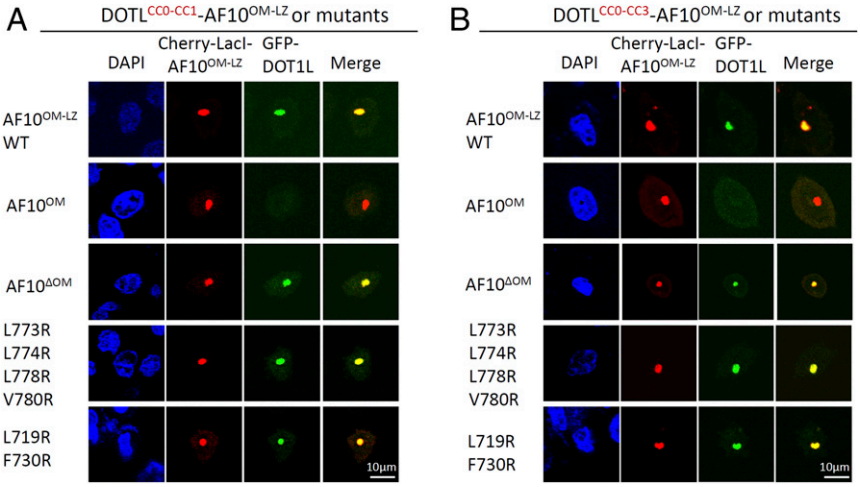

Fig. 3. Fluorescence colocalization of $AF10^{OM\text{-}LZ}$ or its mutants with different CC domains of DOT1L. Cell nuclei are stained blue with DAPI, AF10 and its mutants are fused with Cherry and colored red, DOT1L fragments with GFP tag are colored green. Overlap of red and green (yellow) spots indicates colocalization. (*A*) The CC0-CC1 domain of DOT1L colocalize with the OM-LZ domain of AF10 or its mutants disrupted of octamer formation, but not with the OM domain alone or OM-LZ mutants that affect formation of the heterodimer. (*B*) Similar results obtained when all 4 CC domains of DOT1L are present. Statistics of cells showing colocalization are shown in *SI Appendix*, Fig. S4.

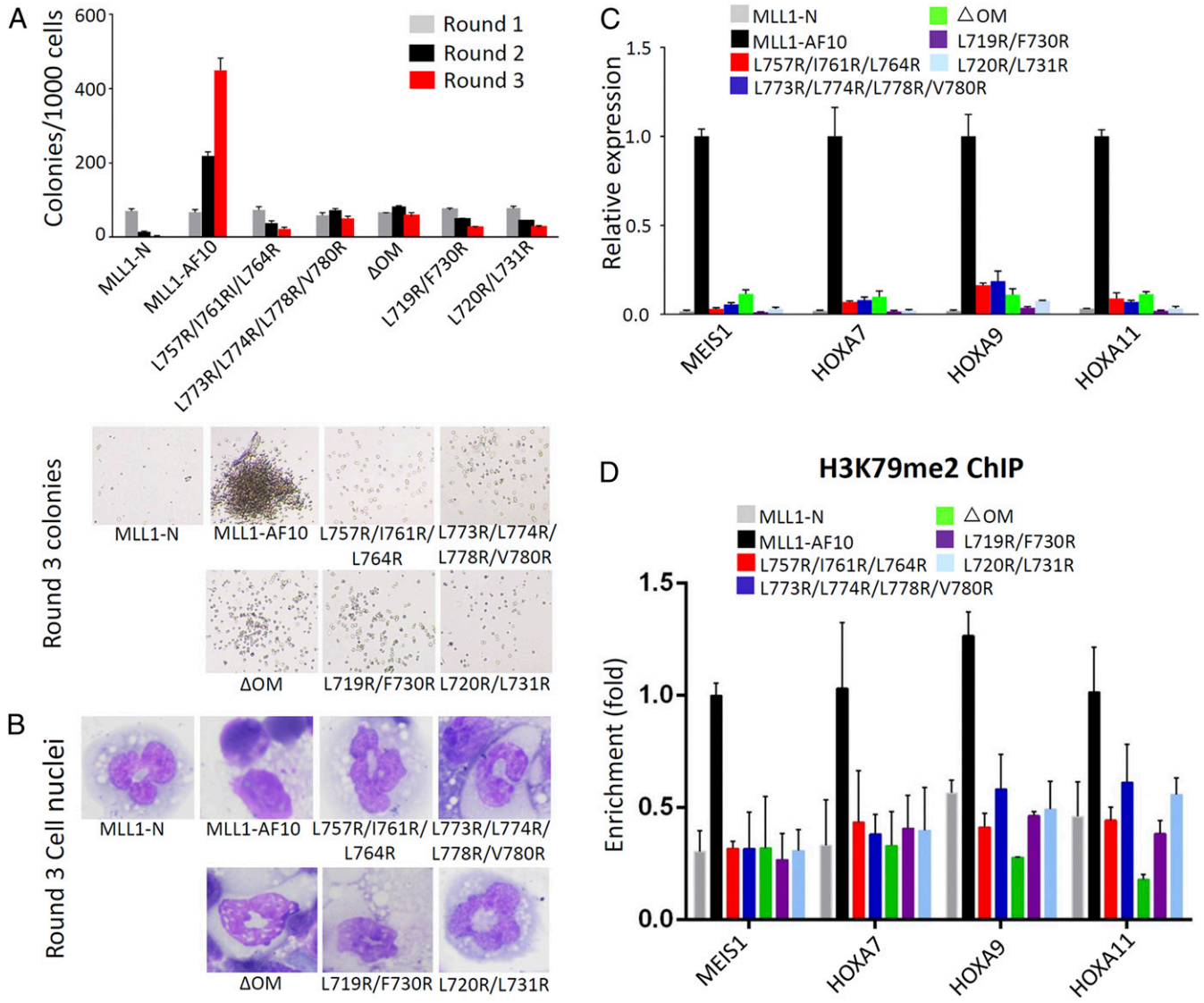

Fig. 4. Function of the higher-order DOT1L–AF10 complex in leukemogenesis. (*A*) Colony formation assays through 3 rounds of plating bone marrow cells expressing MLL1–AF10 or its mutants. MLL1-N was used as a negative control. Colony numbers from each round were counted as shown (*Upper*). Pictures of colonies from the third round are shown by 10-fold magnification (*Lower*). (*B*) Wright staining of the nuclei of third-round plated cells expressing MLL1–AF10 or mutants (100-fold). (*C*) qRT-PCR results of the expression levels of *MEIS1*, *HOXA7*, *HOXA9*, and *HOXA11* genes in the second-round plated cells expressing MLL1-N, MLL1–AF10, or its mutants. (*D*) ChIP-qPCR results show the different level of enrichment of DOT1L at *MEIS1*, *HOXA7*, *HOXA9*, and *HOXA11* gene regions in the second-round plated cells expressing MLL1-N, MLL1–AF10, or its mutants.

expressing MLL1–AF10$^{L719R/F730R}$ or MLL1–AF10$^{L720R/L731R}$ mutants both lost the colony formation activities, to a similar degree as that of deletion of the OM helix in the MLL1–AF10$^{\Delta OM}$ mutant (Fig. 4*A*). Wright staining of the third-round cells showed that the cells carrying the MLL1–AF10 mutants had diffuse colonies and exhibited the characteristics of differentiated diffused nuclei, while cells expressing WT MLL1–AF10 showed compact colonies and blast-like undifferentiated cell nuclei as previously reported (Fig. 4 *A* and *B*) (21).

Leukemogenic activities of MLL1–AF10 are associated with aberrant up-regulation of a number of genes. We next tested whether the MLL1–AF10 fusion protein carrying the aforementioned mutations in AF10 differ in their abilities from the WT fusion protein in activating the relevant genes. qRT-PCR was employed to detect the expression levels of *MEIS1*, *HOXA7*, *HOXA9*, and *HOXA11* genes using total RNA extracted from the second-round plating cells. Consistent with the results of the colony formation assay, cells expressing the MLL1–AF10 mutants showed a dramatically down-regulated expression level of all 4 selected genes (Fig. 4*C*). Furthermore, chromatin immunoprecipitation (ChIP) was carried out to explore the role of the higher-order DOT1L–AF10 complex in directing DOT1L methylase activity to MLL1–AF10-targeted genomic loci. Antibody against H3K79me2 was used to precipitate DNA fragments associated with DOT1L activity, which is followed by qPCR amplification to examine enrichment of DOT1L on the aforementioned leukemia-associated genes, including *MEIS1*, *HOXA7*, *HOXA9*, and *HOXA11*. As expected, much lower levels of DOT1L activity could be detected at the 4 examined genes in cells expressing the MLL1–AF10$^{L757R/I761R/L764R}$ mutant, which is defective for interaction with DOT1L. Interestingly, we also find remarkable reduction of H3K79 methylation at the 4 selected genomic loci in cells expressing MLL1–AF10 mutants, including L773R/L774R/L778R/V780R, ΔOM, L719R/F730R,

and L720R/L731R, that still maintain interaction with DOT1L but incapable of forming a DOT1L–AF10 octamer (Fig. 4*D*).

Taken together, we conclude that the leukemogenic activity of the MLL1–AF10 fusion protein not only requires the binary interaction between AF10 and DOT1L, but also formation of a higher-order complex through tetramerization of the heterodimer.

Discussion

Previous studies have shown that MLL1–AF10's leukemogenic function involves the recruitment of H3K79 histone methyltransferase DOT1L through the OM-LZ domain of AF10, and the CC domains of DOT1L have been implicated in the interaction. Nevertheless, their detailed interaction modes remained unclear, as DOT1L contains 4 tandem CC domains. We found that AF10 can interact with both the N- and C-terminal halves of the CC domains of DOT1L, encompassing CC0-CC1 and CC2-CC3 regions, respectively. This result is consistent with the recent finding that AF10 could bind to CC1, CC2, and CC3 motifs of AF10 independently, albeit with descending binding affinities. The binding of AF10 to CC0 was not checked due to the difficulty of making a soluble and well-behaved protein. The structure of human $AF10^{OM\text{-}LZ}$ in complex with the CC2 domain of zebrafish DOT1L has been recently solved, showing the LZ helix of AF10 interact with the CC2 helix of DOT1L via a classical knobs-into-holes packing. However, interesting questions remain, including whether AF10 binds different CC motifs in a similar manner, and whether CC0 of DOT1L plays any role in interaction with AF10.

Our structure of $AF10^{OM\text{-}LZ}$ in complex with the CC0-CC1 fragment of DOT1L provides an opportunity to directly compare the differences between binding to CC1 and CC2. Aligning the LZ helix of AF10 from the 2 structures shows that it interacts with CC1 (salmon) and CC2 (cyan) of DOT1L in a similar manner (*SI Appendix*, Fig. S5*A*). The 3 key residues of AF10, L757, I761, and L764, respectively, interact with L525, A529, and L532 from CC1, while L757 and L764, respectively, interact with L578 and L585 of CC2 via hydrophobic interactions (*SI Appendix*, Fig. S5*A*). Please note that we number AF10 residues using the convention by Okada et al. (21), which differs from the one used in the structure of 6CKO by 16 residues due to references to different AF10 isoforms. In the CC1 complex, A529 uses its small hydrophobic sidechain to interact with I761, while in the CC2 complex, T582 of zebrafish DOT1L (S582 in human DOT1L) takes the place of A529. This residue is not strictly conserved among CC1 to CC3, and the difference may account for their different binding affinities when tested in isolation (*SI Appendix*, Fig. S5*B*).

Direct binding of an isolated CC0 of DOT1L to $AF10^{OM\text{-}LZ}$ is not tested due to difficulty of obtaining a soluble sample. In the absence of CC0, the structure of the CC2 complex shows that the OM helix of AF10 folds back and packs against the LZ helix (*SI Appendix*, Fig. S5*A*). This self-interaction of AF10 is not observed in our DOT1L–AF10 complex in the presence of CC0, as the OM helix interacts with the CC0 helix from an adjacent heterodimer to form an octameric complex. Given that the LZ domain of AF10 can interact with CC1 to CC3 independently, it is interesting to wonder which CC motif of DOT1L does the LZ domain bind under a physiological setting? We think CC1 is the most probable binding site of the LZ domain of AF10 for the following reasons: First, CC1 is reported to have the highest binding affinity compared to CC2 and CC3 (24); second, the presence of CC0 and its engagement with the OM domain of AF10 from an adjacent heterodimer in our structure with a longer fragment of DOT1L places a spatial constraint that better suited the CC1 domain to interact with the LZ helix of AF10. In fact, our in vitro experiments with the CC1 or CC2 mutant of the CC0-CC3 fragment of DOT1L individually defective in AF10 binding show that they respectively form tetrameric and octameric complexes with AF10, assayed by velocity sedimentation analytical ultracentrifugation (AUC) experiments (*SI Appendix*, Fig. S6). This result is consistent with the binding of AF10 to the CC1 domain in an in vitro setting. Obviously, we cannot rule out the possibility that the LZ helix of AF10 may directly bind CC2 or CC3 in vivo, such as conformational rearrangement in the presence of other proteins.

The most striking finding from our study is that 4 AF10–DOT1L heterodimers tetramerize to form a hetero-octamer, which was first observed in the crystal lattice, and subsequently we confirmed its existence in solution through biochemical and functional analyses. This finding provides a clue of the pathogenesis function of the MLL1–AF10 fusion protein through the formation of a higher order-oligomeric state. Interestingly, a similar finding concerning the oligomeric AML1-ETO fusion protein in leukemogenesis has been reported (26). Since we used shorter fragments of AF10 and DOT1L in our biochemical analysis, one may ask whether there is evidence of octamer formation for full-length or large fragments of AF10 and DOT1L. Several lines of evidence suggest so. First, size-exclusion column chromatography and low-resolution electron microscopy (EM) studies revealed that full-length AF10 forms a homo-tetramer in solution, which is consistent with our structural result that the OM-LZ domains of AF10 forms a tetramer and each AF10 is bound by one DOT1L in the hetero-octamer (27). Second, although we were not able to express full-length DOT1L for biochemical characterization, we did succeed in using a large fragment of DOT1L (amino acids 419 to 670) encompassing the entire 4 CC motifs, from CC0 to CC3, for complex formation with AF10. The measured molecular mass of this complex is approximately 4 times of the calculated mass of one heterodimer (*SI Appendix*, Fig. S1). Furthermore, our in vivo leukemogenic studies of the full-length MLL1–AF10 and its mutants fully corroborate with the structural results. Obviously, it would be ideal to obtain a complex structure of full-length MLL1–AF10 and DOT1L, such as using the powerful cryo-EM reconstruction, to definitively visualize the functional architecture of the protein complex. Nevertheless, our results presented here do generate deeper questions concerning the mechanisms underlying the pathogenic function of MLL1–AF10, such as how the multimeric complex serves to recruit DOT1L more effectively, to increase the dosage of DOT1L at the relevant genomic loci, or to provide a specific configuration needed for a yet-unknown process during leukemogenesis? An understanding to these questions will undoubtedly increase our knowledge about the disease process and help developing more effective intervention strategies.

Experimental Procedures

cDNA fragments encoding various truncations of human AF10 and DOT1L or their mutants were expressed in *E. coli* and purified by affinity, ion-exchange, and gel-filtration column chromatography. The human $AF10^{699\text{-}782}$–$DOT1L^{470\text{-}550}$ complex was crystallized by hanging drop vapor diffusion and the crystal structure was solved by the single-wavelength anomalous dispersion method. Detailed statistics for data collection and refinement are shown in *SI Appendix*, Table S1. Florescence colocalization experiments were carried out using a LacI-LacO targeting system. Colony formation, qPCR, and ChIP assays were performed as described previously (20, 21, 24). Detailed methods and associated references are available in *SI Appendix*.

ACKNOWLEDGMENTS. We thank Dr. Yi Zhang at Harvard University for the generous gift of the plasmids for colony formation assay; Dr. Chaoyang Xiong, Wensi Yang, and Xiaoxia Yu for help on the fluorescence colocalization, structure refinement, and analytical ultracentrifugation experiments; Qian Zhang for participation at an early stage of the project; Jingjing Chen for help with logistics throughout the project; and the Shanghai Synchrotron Radiation Facility beamline scientists for help with data collection. We also thank for the support from Natural Science Foundation of China and Chinese Ministry of Science and Technology Grants 31622020, 31870737, and 2018YFA0107004 (to N.Y.) and 31521002, 31430018, and 2017YFA0103304 (to R.-M.X.); Tianjin Funds for Distinguished Young Scientists Grant 17JCJQJC45900 (to N.Y.); and Beijing Municipal Science and Technology Project Grant Z171100000417001 (to R.-M.X.).

1. Y. Gu *et al*., The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to Drosophila trithorax, to the AF-4 gene. *Cell* **71**, 701–708 (1992).
2. M. J. Thirman *et al*., Rearrangement of the MLL gene in acute lymphoblastic and acute myeloid leukemias with 11q23 chromosomal translocations. *N. Engl. J. Med.* **329**, 909–914 (1993).
3. T. A. Milne *et al*., MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol. Cell* **10**, 1107–1117 (2002).
4. Y. Dou *et al*., Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.* **13**, 713–719 (2006).
5. Y. Dou *et al*., Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* **121**, 873–885 (2005).
6. P. M. Ayton, M. L. Cleary, Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. *Oncogene* **20**, 5695–5707 (2001).
7. J. F. DiMartino *et al*., A carboxy-terminal domain of ELL is required and sufficient for immortalization of myeloid progenitors by MLL-ELL. *Blood* **96**, 3887–3893 (2000).
8. D. Biswas *et al*., Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15751–15756 (2011).
9. F. van Leeuwen, P. R. Gafken, D. E. Gottschling, Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell* **109**, 745–756 (2002).
10. Q. Feng *et al*., Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr. Biol.* **12**, 1052–1058 (2002).
11. N. Lacoste, R. T. Utley, J. M. Hunter, G. G. Poirier, J. Côte, Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. *J. Biol. Chem.* **277**, 30421–30424 (2002).
12. T. R. Golub *et al*., Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–537 (1999).
13. E. Kroon *et al*., Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *EMBO J.* **17**, 3714–3725 (1998).
14. H. Kawagoe, R. Kawagoe, K. Sano, Targeted down-regulation of MLL-AF9 with antisense oligodeoxyribonucleotide reduces the expression of the HOXA7 and -A10 genes and induces apoptosis in a human leukemia cell line, THP-1. *Leukemia* **15**, 1743–1749 (2001).
15. F. Speleman *et al*., A new recurrent inversion, inv(7)(p15q34), leads to transcriptional activation of HOXA10 and HOXA11 in a subset of T-cell acute lymphoblastic leukemias. *Leukemia* **19**, 358–366 (2005).
16. E. Bitoun, P. L. Oliver, K. E. Davies, The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. *Hum. Mol. Genet.* **16**, 92–106 (2007).
17. D. Mueller *et al*., A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood* **110**, 4445–4454 (2007).
18. T. Chaplin *et al*., The t(10;11) translocation in acute myeloid leukemia (M5) consistently fuses the leucine zipper motif of AF10 onto the HRX gene. *Blood* **86**, 2073–2076 (1995).
19. S. Chen *et al*., The PZP domain of AF10 senses unmodified H3K27 to regulate DOT1L-mediated methylation of H3K79. *Mol. Cell* **60**, 319–327 (2015).
20. A. J. Deshpande *et al*., AF10 regulates progressive H3K79 methylation and HOX gene expression in diverse AML subtypes. *Cancer Cell* **26**, 896–908 (2014).
21. Y. Okada *et al*., hDOT1L links histone methylation to leukemogenesis. *Cell* **121**, 167–178 (2005).
22. J. Min, Q. Feng, Z. Li, Y. Zhang, R. M. Xu, Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. *Cell* **112**, 711–723 (2003).
23. K. Sawada *et al*., Structure of the conserved core of the yeast Dot1p, a nucleosomal histone H3 lysine 79 methyltransferase. *J. Biol. Chem.* **279**, 43296–43306 (2004).
24. H. Zhang *et al*., Structural and functional analysis of the DOT1L-AF10 complex reveals mechanistic insights into MLL-AF10-associated leukemogenesis. *Genes Dev.* **32**, 341–346 (2018).
25. A. Kuntimaddi *et al*., Degree of recruitment of DOT1L to MLL-AF9 defines level of H3K79 Di- and tri-methylation on target genes and transformation potential. *Cell Rep.* **11**, 808–820 (2015).
26. X. J. Sun *et al*., A stable transcription factor complex nucleated by oligomeric AML1-ETO controls leukaemogenesis. *Nature* **500**, 93–97 (2013).
27. B. Linder *et al*., Biochemical analyses of the AF10 protein: The extended LAP/PHD-finger mediates oligomerisation. *J. Mol. Biol.* **299**, 369–378 (2000).