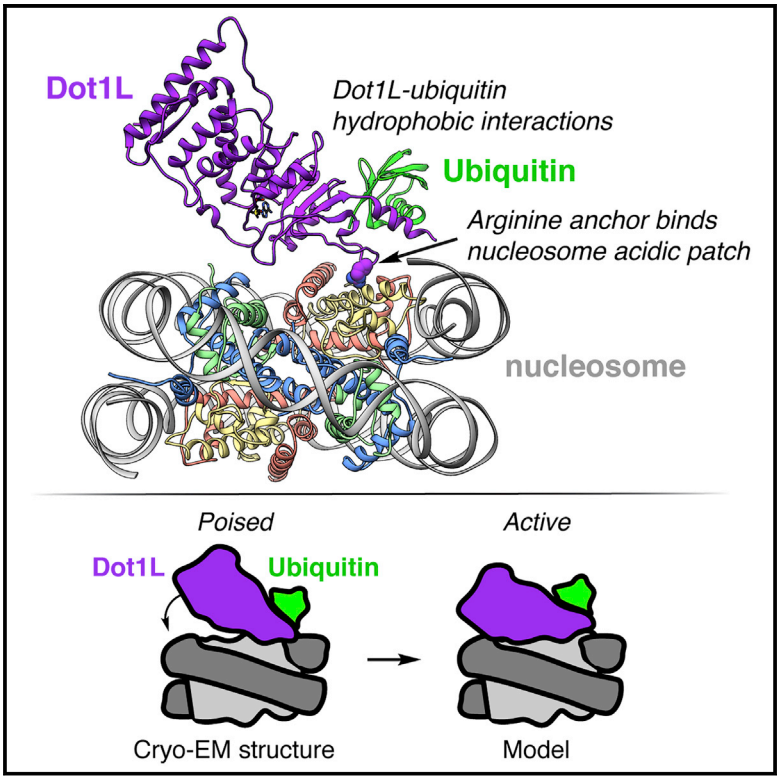
Report

Cell Reports

Structural Basis for Recognition of Ubiquitylated Nucleosome by Dot1L Methyltransferase

Graphical Abstract



Authors

Cathy J. Anderson, Matthew R. Baird, Allen Hsu, Emily H. Barbour, Yuka Koyama, Mario J. Borgnia, Robert K. McGinty

Correspondence

rmcginty@email.unc.edu

In Brief

Dot1L is a histone H3K79-specific methyltransferase that is critical to the pathogenesis of leukemia. Here, Anderson et al. report the cryo-EM structure of Dot1L in complex with a ubiquitylated nucleosome, providing molecular details of how Dot1L binds its nucleosome substrate and is activated by ubiquitin.

Highlights

- Cryo-EM structure of Dot1L bound in a poised state to a ubiquitylated nucleosome
- Dot1L engages the nucleosome acidic patch using a variant arginine anchor
- Dot1L-ubiquitin hydrophobic interactions enhance histone H3K79 methylation

Anderson et al., 2019, Cell Reports *26*, 1681–1690
February 12, 2019
https://doi.org/10.1016/j.celrep.2019.01.058

CellPress

Cell Reports
Report

Structural Basis for Recognition of Ubiquitylated Nucleosome by Dot1L Methyltransferase

Cathy J. Anderson,[1] Matthew R. Baird,[2,4] Allen Hsu,[3] Emily H. Barbour,[2] Yuka Koyama,[2] Mario J. Borgnia,[3] and Robert K. McGinty[2,5,*]

[1]Department of Biochemistry and Biophysics, School of Medicine, The University of North Carolina, Chapel Hill, Chapel Hill, NC 27599, USA
[2]Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, The University of North Carolina, Chapel Hill, Chapel Hill, NC 27599, USA
[3]Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, NIH, Department of Health and Human Services, Research Triangle Park, NC 27709, USA
[4]Present address: Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA
[5]Lead Contact
*Correspondence: rmcginty@email.unc.edu
https://doi.org/10.1016/j.celrep.2019.01.058

SUMMARY

Histone H3 lysine 79 (H3K79) methylation is enriched on actively transcribed genes, and its misregulation is a hallmark of leukemia. Methylation of H3K79, which resides on the structured disk face of the nucleosome, is mediated by the Dot1L methyltransferase. Dot1L activity is part of a trans-histone crosstalk pathway, requiring prior histone H2B ubiquitylation of lysine 120 (H2BK120ub) for optimal activity. However, the molecular details describing both how Dot1L binds to the nucleosome and why Dot1L is activated by H2BK120 ubiquitylation are unknown. Here, we present the cryoelectron microscopy (cryo-EM) structure of Dot1L bound to a nucleosome reconstituted with site-specifically ubiquitylated H2BK120. The structure reveals that Dot1L engages the nucleosome acidic patch using a variant arginine anchor and occupies a conformation poised for methylation. In this conformation, Dot1L and ubiquitin interact directly through complementary hydrophobic surfaces. This study establishes a path to better understand Dot1L function in normal and leukemia cells.

INTRODUCTION

Histone lysine methylation contributes to the regulation of transcription by tuning the recruitment of effector proteins to specific genomic sites (Hyun et al., 2017). It exists in mono-, di-, and trimethylated (me1, me2, and me3) forms, and functional outcomes depend on both the methylated histone residue and degree of methylation (Greer and Shi, 2012). Most well-characterized sites of histone lysine methylation are found in the flexible N-terminal tails of histones (Zhao and Garcia, 2015). One counterexample is histone H3 Lys79 (H3K79), which is solvent exposed on the structured disk face of the nucleosome (Luger et al., 1997a). H3K79 methylation is observed within transcriptionally active genes, and methylation levels are highly correlated with gene expression (Schübeler et al., 2004; Wang et al., 2008; Wood et al., 2018). In human cells, H3K79me2 and H3K79me3 are enriched immediately after transcription start sites and decrease gradually across gene bodies, and H3K79me1 is distributed more broadly across the bodies of active genes (Wang et al., 2008).

Dot1L/KMT4 (disruptor of telomeric silencing-1 like/lysine methyltransferase 4) is the primary H3K79 methyltransferase in human cells and is conserved across eukaryotes (Feng et al., 2002; Lacoste et al., 2002; Ng et al., 2002a; van Leeuwen et al., 2002). Rather than having the characteristic SET (Su(var)3-9, enhancer-of-zeste, trithorax) domain found in other histone lysine methyltransferases (Dillon et al., 2005), Dot1 proteins have a catalytic domain resembling class I methyltransferase domains found in DNA and protein arginine methyltransferases (Min et al., 2003; Sawada et al., 2004). Although known to participate in several transcriptional elongation complexes (Wood et al., 2018), Dot1L can bind to and methylate H3K79 in nucleosomes in isolation (Feng et al., 2002; Min et al., 2003). Histone H3 alone is a poor substrate for Dot1L, suggesting that Dot1L requires non-H3 surfaces of the nucleosome for substrate binding and/or activity (Feng et al., 2002; Lacoste et al., 2002; Ng et al., 2002a).

Efficient methylation of H3K79 in cells requires prior ubiquitylation of H2BK120 (Briggs et al., 2002; Kim et al., 2005; Ng et al., 2002b). H3K79me2 and H3K79me3 are significantly decreased without change to H3K79me1 following knockdown of the H2BK120-targeting ubiquitin E3 ligase, Bre1, in human cells or upon mutation of H2BK120 in *S. cerevisiae* (Kim et al., 2005; Shahbazian et al., 2005). Using designer nucleosomes assembled with monoubiquitylated H2BK120 (H2BK120ub), this trans-histone crosstalk between H2BK120ub and H3K79 methylation has been shown to be direct and require only the catalytic domain of Dot1L (McGinty et al., 2008). Previous studies implicate the C-terminal tail of ubiquitin and the N-terminal tail of histone H2A in mediating ubiquitin-dependent Dot1L activity (Holt et al., 2015; Zhou et al., 2016). The N-terminal tail of H4 has also been shown to be important for Dot1L activity independent of H2B ubiquitylation (Fingerman et al., 2007; McGinty et al., 2009). In recent years, Dot1L has emerged as a potential therapeutic target for MLL-rearranged leukemias because the catalytic activity of Dot1L is required

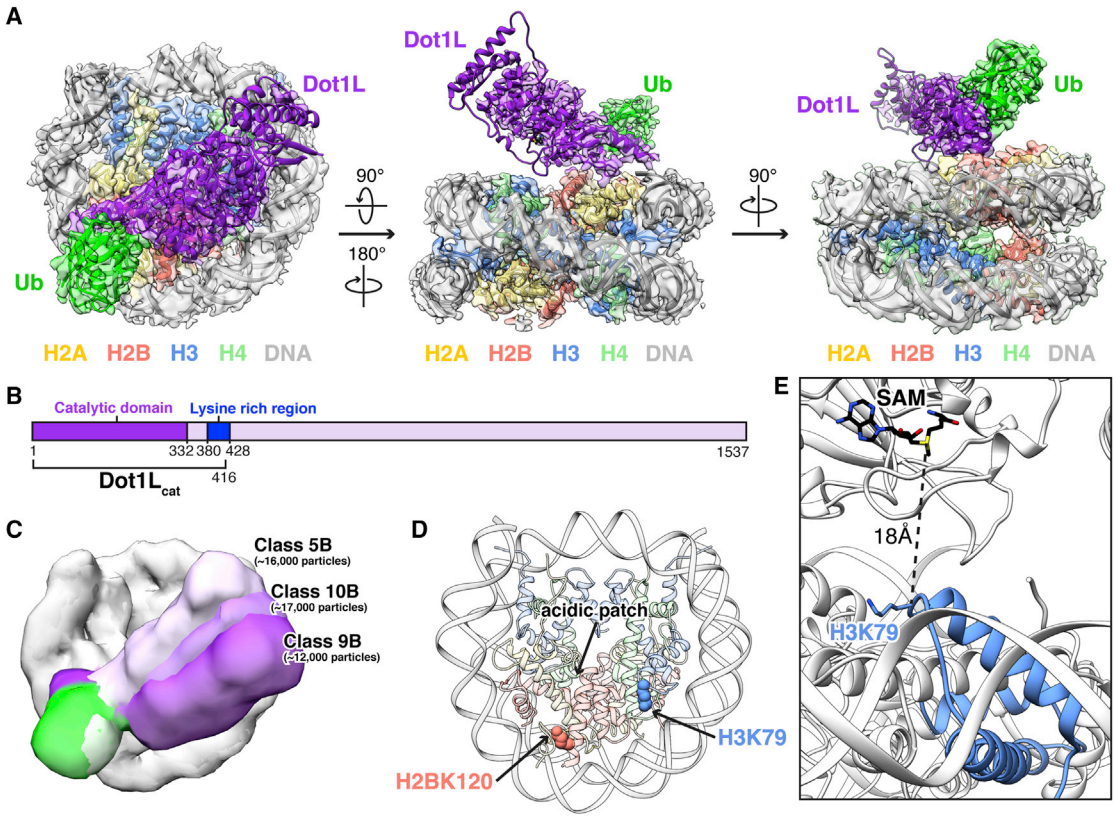


Figure 1. Cryo-EM Structure of Dot1L-H2BK120ub Nucleosome Complex

Figure360▷ For a Figure360 author presentation of Figure 1, see https://doi.org/10.1016/j.celrep.2019.01.058.

(A) Three orthogonal views of the complex with reconstructed map overlaying molecular model.

(B) Scheme of human Dot1L highlighting catalytic domain (dark purple), lysine rich region (blue), and minimal Dot1L$_{cat}$ fragment used in our studies.

(C) Overlay of 3D subclasses showing conformational heterogeneity of Dot1L (shades of purple) and ubiquitin (shades of green).

(D) Nucleosome structure with acidic patch and sites of ubiquitylation (H2BK120) and Dot1L-mediated methylation (H3K79) indicated.

(E) Zoomed view of Dot1L active site and targeted lysine in our model with distance between H3K79 Cα and SAM cofactor sulfur indicated.

See also Figures S1–S4 and Tables S1 and S2.

for leukemogenic transformation following MLL-fusion translocations (Bernt et al., 2011; Winters and Bernt, 2017). Yet key molecular details describing how Dot1L binds to the nucleosome and is activated by H2B ubiquitylation remain elusive.

Here, we report the 3.9 Å cryoelectron microscopy (cryo-EM) structure of the methyltransferase domain of human Dot1L bound to a designer nucleosome core particle assembled with a site-specifically modified H2BK120ub (Figure 1A). The structure captures Dot1L in a poised state in which Dot1L uses an arginine to anchor to the acidic patch on the nucleosome. We also observe critical hydrophobic interactions that mediate ubiquitin-dependent activity of Dot1L. In the structure, the catalytic site of Dot1L is separated from H3K79, indicating that Dot1L and/or the nucleosome must undergo conformational rearrangement from a poised to an active state for methylation.

RESULTS

Dot1L-H2BK120ub Nucleosome Complex Structure

In order to visualize the molecular details of Dot1L function on nucleosomes and activation by H2BK120 ubiquitylation, we first needed to prepare homogeneously site-specifically ubiquitylated nucleosomes. To this end, we generated and purified an H2BK120ub analog by crosslinking ubiquitin with a Gly76Cys mutation at its terminal residue to an H2BK120C mutant using dichloroacetone (DCA) as previously described (Figures S1A–S1C; Morgan et al., 2016). Importantly, we observe slightly increased Dot1L methyltransferase activity when using nucleosomes reconstituted with two copies of this purified DCA-crosslinked analog as compared to nucleosomes assembled with two copies of a more precise semisynthetic analog of H2BK120ub containing a single Gly76Ala point mutation in ubiquitin (McGinty et al., 2009). As this Gly76Ala analog was previously established to be indistinguishable from natively linked H2BK120ub in a Dot1L methyltransferase assay, we can extrapolate that Dot1L is activated by our DCA-crosslinked analog similarly to, if not to a greater extent than, native H2BK120ub (Figure S1D). We reconstituted our homogeneous DCA-crosslinked H2BK120ub nucleosome core particles with a minimal catalytic fragment of Dot1L (Dot1L$_{cat}$) (Figure S1C). This fragment spanning amino acids 2–416 includes the catalytic domain of Dot1L and an adjacent lysine-rich sequence that was previously identified to be

necessary and sufficient for nucleosome binding and activity (Figure 1B; Min et al., 2003). Although stable in solution, Dot1L$_{cat}$ dissociated from H2BK120ub nucleosomes during cryo-EM sample preparation under all conditions tested. To improve complex stability, we crosslinked the complex with glutaraldehyde prior to vitrification. This allowed for visualization of a Dot1L$_{cat}$-H2BK120ub nucleosome complex by cryo-EM (Table S1).

We initially resolved a 1:1 Dot1L$_{cat}$-H2BK120ub nucleosome complex from 3D reconstruction of ~100,000 particles at 3.5 Å resolution (Figure S2; classes 3 and 4). In this reconstruction, high-resolution features in the nucleosome region of the density map are consistent with the estimated resolution. However, conformational heterogeneity in Dot1L and ubiquitin limit resolution in their respective regions of the reconstructed map. To improve Dot1L and ubiquitin visualization, we reclassified the particles used for the 3.5 Å reconstruction into four subclasses, leading to the identification of two major conformational states (Figure S2; classes 2A and 4A). We used class 2A to generate a 3.9 Å reconstruction that, although having a lower overall resolution than our initial map, has improved density for Dot1L and ubiquitin generally and does not compromise high-resolution information at the Dot1L-nucleosome interface. The observed conformational heterogeneity in Dot1L and ubiquitin is further illustrated by finer classification into 10 subclasses, where a continuum of Dot1L and ubiquitin conformations is observed (Figures 1C and S2; classes 5B, 9B, and 10B). As these classes did not have sufficient particles to allow for high-resolution reconstruction, subsequent modeling was completed using the 3.9 Å map. It is worth noting that we anticipated a 2:1 complex (i.e., two copies of Dot1L per ubiquitylated nucleosome) due to the expected ability of the methyltransferase to bind to each face of the pseudosymmetric nucleosome core particle (McGinty et al., 2009; Zhou et al., 2016). However, we initially observed a 1:1 complex with only sparse density in the regions corresponding to Dot1L and ubiquitin on the opposite nucleosome face. This is likely due to loss of Dot1L on one nucleosome face during sample freezing and the resulting conformational heterogeneity of ubiquitin when unconstrained by Dot1L. This is consistent with a previous crystal structure of an unbound H2BK120ub nucleosome that lacks defined density for ubiquitin (Machida et al., 2016). Only in the finer classification described above did we identify minor class averages with definitive density for a 2:1 Dot1L:nucleosome complex (Figure S2; class 4B) or density for ubiquitylated nucleosomes only (Figure S2; class 7B), demonstrating that, despite crosslinking with glutaraldehyde, our particles include an ensemble of Dot1L:nucleosome stoichiometries.

With the optimal 3.9 Å resolution reconstructed map in hand, we generated a Dot1L$_{cat}$-H2BK120ub nucleosome model by docking high-resolution crystal structures of the nucleosome core particle, one Dot1L catalytic domain, and one ubiquitin into the map followed by iterative manual model building and real-space refinement (Figure 1A). Although present on the opposite nucleosome face, ubiquitin was not modeled on the nucleosome face not bound by Dot1L due to poorly resolved density. In our model, Dot1L anchors to the nucleosome on the edge of the nucleosome acidic patch—an emerging hotspot for nucleosome binding (Figure 1D). We do not observe density for the lysine rich region C-terminal to the catalytic domain, suggesting that it binds to the nucleosome in a heterogeneous manner. In fact, no density is unaccounted for by the model in this region or any other region of the map. Ubiquitin, which is covalently attached to H2BK120, occupies a position lifted off the nucleosome surface and packs against the nucleosome-binding region of Dot1L. Interestingly, Dot1L does not engage the nucleosome surface surrounding the targeted H3K79 residue. Rather, a gap exists between H3K79 and the Dot1L active site. Because we do not observe obvious density for the S-adenosylmethionine (SAM) cofactor in the Dot1L active site, we aligned our model with a Dot1L-SAM co-crystal structure, allowing us to determine the position of SAM. This alignment shows an 18 Å distance between the SAM sulfur atom and the H3K79 Cα atom (Figure 1E). We hypothesize that our model represents a poised state and that conformational rearrangement of the nucleosome and/or reorientation of Dot1L relative to the nucleosome is required for catalysis to occur. Interestingly, mass spectrometric analysis of our reconstituted Dot1L$_{cat}$-H2BK120ub(DCA) complex shows that the majority of H3 is monomethylated (Figure S1E). As we did not supplement our complex with SAM, this observation suggests that Dot1L$_{cat}$ purified with endogenous SAM bound and has subsequently methylated the nucleosome. Therefore, the majority of our reconstructed particles contain a bound Dot1L$_{cat}$ that is poised for H3K79 dimethylation.

It is worth noting that the poised conformation of Dot1L in our cryo-EM model is similar to a molecular replacement model of a non-glutaraldehyde crosslinked, Dot1L-H2BK120ub nucleosome complex that we solved using an 8 Å resolution X-ray diffraction dataset (Dot1L Cα root-mean-square deviation [RMSD] = 10.4 Å following alignment by histones; Figure S3A; Table S2). This crystal structure shows a 2:1 Dot1L-H2BK120ub nucleosome complex in which Dot1L is bound to the identical surface of the nucleosome in a nearly identical manner but held even farther away from H3K79 than in our cryo-EM model (Figure S3B). We initially considered this poised conformation in our crystallographic model to be an artifact of crystal packing due to the presence of a symmetry-related copy of Dot1L wedged between the Dot1L catalytic site and its targeted nucleosome (Figure S3A). However, the convergence of our models generated by cryo-EM and X-ray crystallography suggests a functional role for the observed conformation. Moreover, as our crystallographic complex was not glutaraldehyde crosslinked, it is unlikely that crosslinking has artificially forced Dot1L into the poised state observed in our cryo-EM structure. As the majority of particles in our cryo-EM dataset are classified into 3D classes with visible density for Dot1L in the poised conformation, we expect that this is the preferred equilibrium position of Dot1L on the nucleosome.

Significant variation of resolution exists across the reconstructed cryo-EM map (Figures S4A–S4H). The nucleosome dominates particle alignments and is thus well-resolved, allowing high-confidence modeling of most histone side chains (Figure S4E). Even though we initiated nucleosome modeling with a high-resolution nucleosome crystal structure, the final model closely resembles a published cryo-EM structure of an unbound nucleosome (RMSD = 1.2 Å across shared Cα atoms), which diverges from crystal structures most notably in the slight splaying

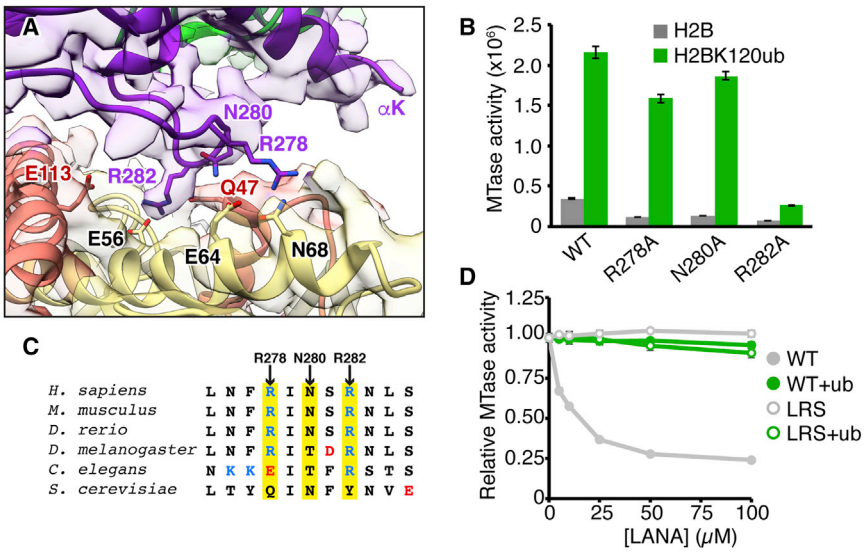


Figure 2. Dot1L Interacts with the Nucleosome Acidic Patch

(A) Zoomed view of Dot1L nucleosome-interaction loop engaged with nucleosome acidic patch surface. Model, reconstructed map, and relevant side chains are shown.

(B) Quantified Dot1L methyltransferase (MTase) assay using indicated mutants with unmodified (H2B) and ubiquitylated (H2BK120ub) nucleosomes.

(C) Sequence alignment of Dot1L nucleosome-interaction loop with positively and negatively charged residues colored blue and red, respectively, and key positions highlighted in yellow.

(D) Competitive MTase assay with varying concentrations of wild-type (WT) or nucleosome-binding-deficient (LRS = LRS 8–10 AAA) LANA fusion proteins on unmodified and H2BK120ub nucleosomes. Five replicates performed for all assays and means and SDs are shown.

See also Figure S5.

of the H2B αC helices away from the nucleosome surface (Bilokapic et al., 2018). In contrast to the nucleosome volume, the regions of the reconstructed map corresponding to Dot1L and ubiquitin exhibit more limited high-resolution features. The catalytic domain of Dot1L was easily docked using a well-resolved nucleosome binding loop and surrounding features, including the αK helix that closely borders the ubiquitin density. Secondary structure in nearby Dot1L regions, including the catalytic site, also aligns well to the map, but local resolution is limited due to the aforementioned conformational heterogeneity. More distant regions and large loops of Dot1L are less well resolved in our map. Ubiquitin was most challenging to model due to its smaller size and spheroidal shape. To improve rigid body docking of the ubiquitin structure, we masked the volumes of Dot1L and ubiquitin and continued refinement (Figure S2). Although this approach failed to generate side-chain resolution for Dot1L or ubiquitin, clear density for secondary structure elements allowed high-confidence docking of ubiquitin and further confirmed Dot1L placement in the nucleosome binding and active site regions (Figures S4I and S4J).

Dot1L-Nucleosome Acidic Patch Interaction

The catalytic domain of Dot1L binds to the nucleosome using a long pre-structured loop between Dot1L β9 and β10 sheets. We clearly resolve main-chain and select side-chain density for this Dot1L nucleosome-interaction loop (Figures 2A and S4F). The most striking feature within the loop is the Dot1L Arg282 side chain that clearly inserts into a cavity at the edge of the nucleosome acidic patch, in position to make a salt bridge and a hydrogen bond with the side chains of H2AE56 and H2BQ47, respectively. Two other residues in the loop, Arg278 and Asn280, are positioned to make direct nucleosome interactions, but their side chains are not clearly resolved in the map. To validate the importance of the nucleosome-interaction loop for Dot1L methyltransferase activity, we tested alanine mutants of key loop residues using an enzyme-coupled methyltransferase assay. We first established that the assay could detect $Dot1L_{cat}$ activity on unmodified nucleosomes and nucleosomes containing a near-native semisynthetic H2BK120ub analog with a Gly76Ala mutation in ubiquitin and that the time points and enzyme concentrations we used were in the linear range of the assay (Figure S1F). We observe a ~6-fold enhancement of $Dot1L_{cat}$ activity on H2BK120ub nucleosomes as compared to unmodified nucleosomes. This is a smaller enhancement than previously described using a direct assay with ^{3}H-SAM and may result from intrinsic differences between the assays, but the same overall trend is observed (McGinty et al., 2009). Because Dot1L activity could be detected on both ubiquitylated and unmodified nucleosomes, we were able to assess ubiquitin-dependent and independent effects of Dot1L mutations. As predicted by our model, Dot1L Arg282Ala leads to a 90% reduction in H3K79 methylation on ubiquitylated nucleosomes and nearly abolishes all activity on unmodified nucleosomes (Figure 2B). Asn280Ala and Arg278Ala mutations also result in losses of activity, especially on unmodified nucleosomes, although to a lesser degree than the Arg282Ala mutant Dot1L. Notably, the best resolved side chain in Dot1L at the nucleosome interface in our cryo-EM map is also the most important side chain for Dot1L activity, regardless of H2B ubiquitylation. The Dot1L nucleosome interaction loop is well conserved in higher eukaryotes but divergent in *S. cerevisiae* ortholog Dot1p, leaving open the possibility that Dot1p binds nucleosomes and is activated by ubiquitin through a distinct mechanism compared to human Dot1L (Figure 2C).

Although our structure shows that Dot1L binds to the acidic patch of ubiquitylated nucleosomes, an alternative binding mode has been suggested for Dot1L on unmodified nucleosomes (Zhou et al., 2016). To test whether Dot1L requires the acidic patch for activity on unmodified nucleosomes, we performed competitive methyltransferase assays by titrating a glutathione S-transferase (GST)-fusion of the nucleosome acidic-patch-binding 23 amino acid LANA (latency-associated

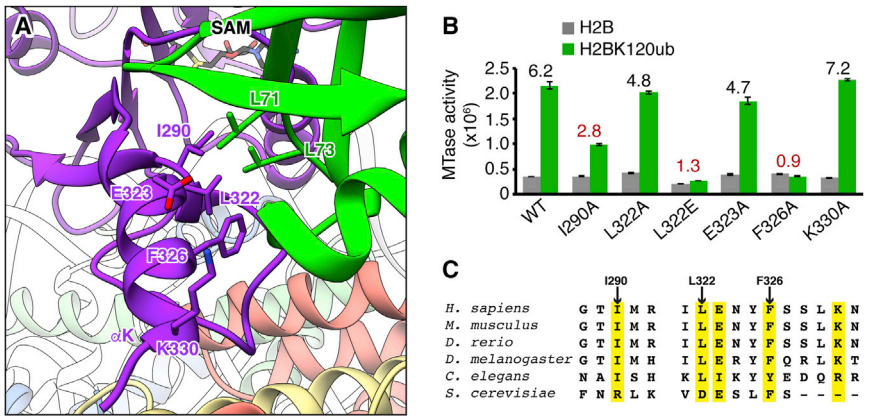


Figure 3. Ubiquitin Interacts with Dot1L to Enhance H3K79 Methylation

(A) Zoomed view of Dot1L-ubiquitin interaction with relevant side chains indicated.

(B) Quantified MTase assay of Dot1L mutants on unmodified (H2B) and ubiquitylated (H2BK120ub) nucleosomes. Numbers above bars indicate fold enhancement by ubiquitin for each Dot1L protein. A value of 1 indicates no difference in activity toward H2BK120ub relative to unmodified nucleosomes. Five replicates performed for all assays and means and SDs are shown.

(C) Sequence alignment of Dot1L ubiquitin-binding regions with key positions highlighted in yellow.

See also Figure S6.

nuclear antigen from Kaposi's sarcoma-associated herpesvirus) sequence into our assays (Figure 2D). Wild-type LANA, but not a nucleosome-binding-deficient mutant (residues 8–10, Leu-Arg-Ser to Ala-Ala-Ala triple mutation [LRS]), is able to inhibit Dot1L activity on unmodified nucleosomes in a dose-dependent manner. Although not ruling out alternative binding modes, this suggests that Dot1L activity on ubiquitylated and unmodified nucleosomes shares a similar acidic patch interaction dependence. To our surprise, LANA fails to substantially inhibit Dot1L enzymatic activity on ubiquitylated nucleosomes even at very high concentrations (Figure 2D). This is not due to the GST-fusion, as pre-assay cleavage of the GST tag leads to identical results (Figures S5A and S5B). This observation could be secondary to tighter Dot1L binding to H2BK120ub nucleosomes or the inability of LANA to bind to H2BK120ub nucleosomes altogether. Further investigation shows that GST-LANA forms a stable complex with unmodified nucleosomes but is unable to bind to H2BK120ub nucleosomes (Figure S5C). Importantly, this presents a mechanism in which H2BK120 ubiquitylation may activate Dot1L in isolation but may also further promote Dot1L activity in cells by acting as a gatekeeper to the acidic patch and neutralizing competition for nucleosome acidic patch binding.

Dot1L-Ubiquitin Hydrophobic Interaction

Ubiquitin occupies a position lifted off the nucleosome surface and rests against the αK helix of Dot1L that scaffolds the structure of the Dot1L nucleosome-interaction loop. Dot1L and ubiquitin exhibit conformational flexibility in our cryo-EM dataset. However, docking of Dot1L and ubiquitin into finely parsed 3D classes followed by alignment by Dot1L shows that Dot1L and ubiquitin move as a unit, suggesting a direct interaction (Figure S6A). Previously reported comprehensive mutagenesis of the ubiquitin surface identified only two ubiquitin side chains necessary for Dot1L activation, Leu71 and Leu73, both located near the C terminus of ubiquitin. Therefore, even though the ubiquitin-Dot1L interface appears large, we expected to find that only the small hydrophobic region of Dot1L adjacent to ubiquitin Leu71 and Leu73 would be responsible for ubiquitin-dependent activity. Although the ubiquitin Leu71 side chain is poorly resolved in our map, Leu71 is at the end of the β4 strand that is part of the globular-fold of ubiquitin, so we can confidently place its side chain in proximity to Dot1L αK residues Leu322 and Phe326 (Figures 3A and S4H). We have less confidence in the location of ubiquitin Leu73 in our model as it resides in the flexible ubiquitin tail. However, based on its importance for ubiquitin-dependent Dot1L activity, we hypothesized that Leu73 makes direct contact with a hydrophobic surface on Dot1L. A likely surface is the exposed Ile290 side chain of Dot1L. We tested the effects of alanine mutations of Ile290, Leu322, and Phe326, as well as Glu323 and Lys330 that are also positioned at the Dot1L-ubiquitin interface but more distant from ubiquitin Leu71 and Leu73 (Figure 3B). To assess H2BK120ub-specific effects, we compared fold increases in Dot1L activity due to H2B ubiquitylation for each of the Dot1L mutants. The Phe326Ala mutation effectively abolished all ubiquitin-dependent activity of Dot1L, with a calculated H2BK120ub fold increase of 0.9 (compared to 6.2 for wild-type Dot1L). An Ile290Ala mutation also preferentially disrupted H2BK120-dependent activity of Dot1L, with a calculated fold increase of 2.8. A Dot1L Leu322Ala mutant methylated both unmodified and ubiquitylated nucleosomes with near wild-type activity, so we prepared a more severe Leu322Glu mutant Dot1L. This mutation resulted in a more pronounced reduction in activity on H2BK120ub nucleosomes than unmodified nucleosomes, leading to a ubiquitin fold increase of 1.3. Glu323Ala and Lys330Ala mutations, which are not in the vicinity of ubiquitin Leu71 or Leu73 in our model, had minimal effects on Dot1L enzymatic activity on both nucleosomes, consistent with the published ubiquitin mutagenesis studies that found no other region of the ubiquitin surface to be necessary for Dot1L activation. Taken together, these results suggest that Dot1L Ile290, Leu322, and Phe326 hydrophobic interactions with ubiquitin are critical to ubiquitin-dependent activity. Similar to the nucleosome-interaction loop, sequences from Dot1L orthologs in higher eukaryotes align well with human Dot1L for positions at the ubiquitin interface (Figure 3C).

Dot1L Active State Model

We believe our structure represents a poised conformation of Dot1L on the nucleosome. Our biochemical analysis shows that key interactions observed in the poised state, especially those mediated by Dot1L Arg282, are also required for activity and are thus likely to interact during catalysis. Therefore, we

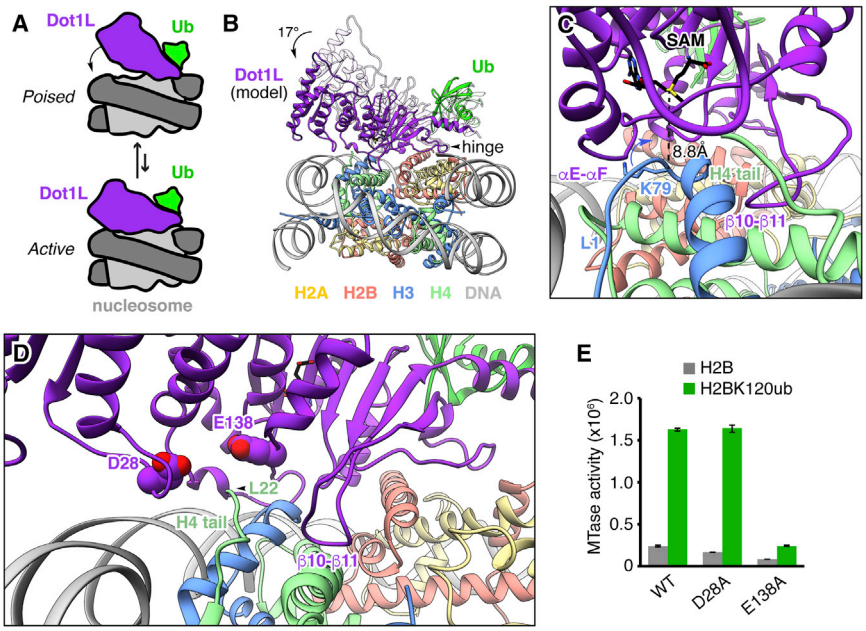


Figure 4. Model of Active Conformation of Dot1L on the Nucleosome

(A) Scheme of proposed poised to active conformational rearrangement of Dot1L and ubiquitin on the nucleosome.

(B) Front view of model of Dot1L and ubiquitin in an active conformation made by a 17° rotation of Dot1L about hinge Arg282 Cα atom of Dot1L. Ubiquitin was adjusted with Dot1L, keeping the orientation of ubiquitin relative to Dot1L constant. Cryo-EM model shown in white for comparison.

(C) Zoomed side view of active model with SAM sulfur-H3K79 Cα distance labeled. Loops with anticipated conformational changes are indicated.

(D) Front view of active model with key residues of acidic cleft indicated.

(E) Quantified MTase assay of Dot1L acidic cleft mutants on unmodified (H2B) and ubiquitylated (H2BK120ub) nucleosomes. Five replicates performed for all assays and means and SDs are shown.

modeled the active state of the Dot1L-H2BK120ub complex by fixing the Dot1L Arg282 Cα as a hinge position and rotating Dot1L and ubiquitin toward H3K79 on the nucleosome surface (Figures 4A and 4B). As no non-SET domain methyltransferase structures exist in complex with a cofactor and a lysine-containing substrate, we used the distance between the *S*-adenosylhomocysteine (SAH) cofactor and the target lysine Cα in a ternary complex of the SET domain containing SET8 methyltransferase to guide active state modeling (Couture et al., 2005). This SET8 structure shows an 8.8 Å distance between the Cα of the target lysine and the sulfur atom of the cofactor. A 17° rotation of Dot1L about Arg282 was required to replicate this distance in the Dot1L-nucleosome complex while placing the Cα of H3K79 at the entrance to the putative Dot1L substrate channel. This active state results in only minor steric clashes between Dot1L flexible loops and the nucleosome surface. Based on these clashes, we anticipate that the flexible β10-β11 and αE-αF loops in Dot1L will need to move from their locations in the Dot1L crystal structure to allow Dot1L to bind in an active conformation (Figure 4C). Although our model places the Dot1L active site at an optimal distance relative to the H3K79 Cα atom, the H3K79 side chain is oriented along the nucleosome surface in the poised state (Figure 1E). Therefore, our modeling also suggests that the H3 L1 loop containing H3K79 will need to undergo minor conformational rearrangement to direct the H3K79 side chain away from the nucleosome surface and toward the active site (Figure 4C).

It has been demonstrated that a basic sequence in the H4 N-terminal tail, specifically H4R17 and R19, is critical for Dot1L activity (Fingerman et al., 2007; McGinty et al., 2009). We were only able to model residues starting with H4L22 in our poised structure. Because there is an acidic cleft on the surface of Dot1L adjacent to the location from which the H4 N-terminal tail emerges from the nucleosome, we hypothesized that the H4 tail basic patch can bind to this Dot1L acidic cleft to promote formation of a catalytically active conformation of Dot1L on the nucleosome. To explore this hypothesis, we mutated two prominent residues in the Dot1L acidic cleft (Figure 4D). Interestingly, the Dot1L Glu138Ala mutation, but not a nearby Asp28Ala mutation, led to a substantial loss in Dot1L activity on both unmodified and ubiquitylated nucleosomes (Figure 4E). This result, together with our poised structure, is suggestive of a specific interaction between the H4 tail basic patch and the Dot1L acidic cleft. However, due to the proximity of Glu138 to the Dot1L active site, we cannot rule out other explanations. Future structural characterization of an active complex will illuminate further molecular details.

DISCUSSION

We used cryo-EM to solve the structure of Dot1L bound to an H2BK120ub nucleosome in a poised state. In our structure, Dot1L binds to the nucleosome acidic patch using a nucleosome-interaction loop. The acidic patch is emerging as a hotspot for nucleosome binding (Kalashnikova et al., 2013; McGinty and Tan, 2016). Examination of structures of proteins and peptides bound to the nucleosome show that a common feature is the use of two arginines to bind to several locations within the acidic patch (Figures 5A–5C). The first example was observed in a crystal structure of the LANA sequence bound to the nucleosome (Barbera et al., 2006). In this structure, LANA Arg9 inserts into a cavity surrounded by H2AE61, D90, and E92 side chains. Other structures, including those with RCC1, Sir3, 53BP1, RING1B, and CENP-C, have demonstrated that many more proteins use arginines with nearly identical conformations to bind in this same acidic patch cavity, leading to the introduction of the term arginine anchor (Armache et al., 2011; Kato et al., 2013; Makde et al., 2010; McGinty et al., 2014; Wilson et al., 2016). All of these proteins also engage the nucleosome with one or more additional arginines that bind to the acidic patch in other locations. Other structures, including our Dot1L-nucleosome

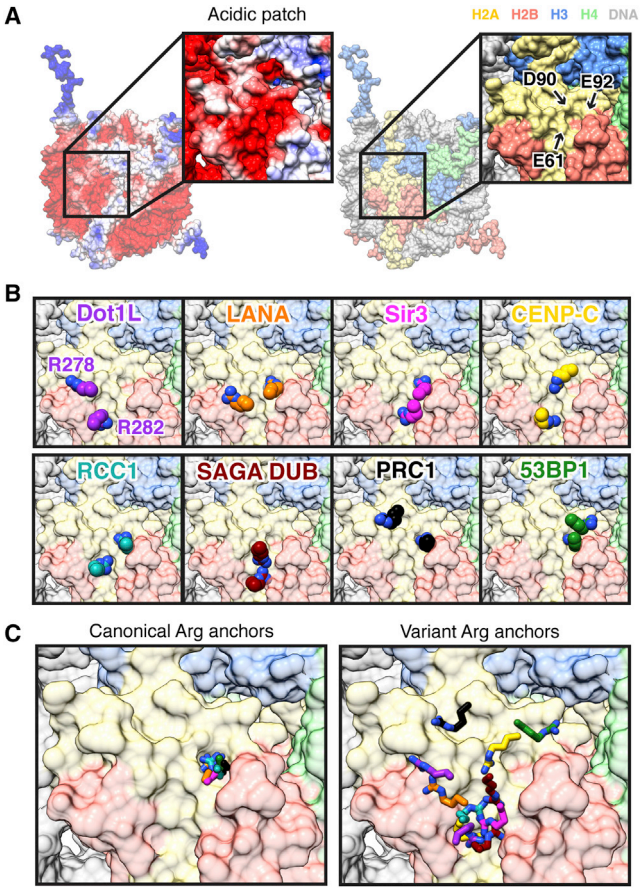


Figure 5. Canonical and Variant Arginine Anchors Bind the Nucleosome Acidic Patch

(A) Electrostatic (generated with Adaptive Poisson-Boltzmann Solver [APBS] using AMBER) and histone-colored surfaces of nucleosome.

(B) Zoomed views of arginine anchors of Dot1L, LANA (PDB: 1ZLA), Sir3 (PDB: 3TU4), CENP-C (PDB: 4X23), RCC1 (PDB: 3MVD), SAGA DUB module (PDB: 4ZUX), PRC1/RING1B (PDB: 4RP8), and 53BP1 (PDB: 5KGF).

(C) Overlay of all canonical (left) and variant (right) arginine anchors.

structure presented here and a structure of the SAGA deubiquitinating (DUB) module (Morgan et al., 2016), have nucleosome interactions with multiple arginines and no binding in the H2AE61/D90/E92 cavity. In general, the arginines that bind these other acidic patch surfaces exhibit more diversity in both binding sites and arginine side chain conformations than those that bind to the H2AE61/D90/E92 cavity (Figure 5C). We suggest the use of the terms canonical arginine anchor and variant arginine anchor to describe these two classes of arginines. The Dot1L variant arginine anchor Arg282 most closely resembles the CENP-C variant arginine anchor Arg714. Although we do not observe density for Dot1L Arg278 in our cryo-EM map, the position of its Cα atom is similar to that of the LANA variant arginine anchor Arg12. Overall, the analysis of diverse acidic-patch-binding proteins demonstrates both shared and protein-specific molecular mechanisms for nucleosome binding in this region.

Interestingly, we found that the LANA sequence is unable to bind to H2BK120ub nucleosomes under conditions where binding to unmodified nucleosomes was observed. This suggests that, in addition to promoting Dot1L activity in a purified system, H2BK120ub may also prevent other proteins from binding the nucleosome in a physiologic cellular environment. This would limit competition, further promoting intranucleosomal methylation, and may also increase local concentrations of Dot1L in areas of the nucleus containing H2BK120ub-modified chromatin. Macromolecular crystals of H2BK120ub nucleosomes do not alter the nucleosome structure and lack density for ubiquitin, suggesting ubiquitin has conformational flexibility in the absence of other bound factors (Machida et al., 2016). Therefore, it is unlikely that ubiquitin competes with LANA by binding directly to the acidic patch. Rather, we favor the explanation that a conformationally flexible ubiquitin in the vicinity of the acidic patch can prevent acidic patch binding of other proteins. Some proteins for which H2BK120ub inhibits chromatin binding may have been identified in nucleosome affinity proteomics experiments with ubiquitylated nucleosomes, including RCC1, another acidic patch binder (Shema-Yaacoby et al., 2013). H2BK120ub is one of many acidic patch proximal histone modifications, and several of these modifications also tune the nucleosome binding and/or activity of chromatin effector proteins (Dann et al., 2017; Fujiki et al., 2011; Jbara et al., 2016). Given the emerging trend of acidic-patch-dependent chromatin binding, this is likely to be a major mechanism in the regulation of chromatin processes.

In our structure, ubiquitin is lifted off the nucleosome surface and is in a position to interact with the catalytic domain of Dot1L. Computational sorting of particles with different conformational states of Dot1L with respect to the nucleosome demonstrates that Dot1L and ubiquitin move together through a continuum of positions. We also identify mutants in Dot1L that preferentially disrupt activity on H2BK120ub nucleosomes over unmodified nucleosomes. Taken together, our data suggest that ubiquitin enhances Dot1L activity through direct interactions with the catalytic domain of Dot1L. Contrary to this model, photocrosslinking studies place ubiquitin Leu71 and Leu73 in proximity to the H2A N-terminal tail (Zhou et al., 2016). We have not observed any density for ubiquitin near the emergence of the H2A tail from the nucleosome disk or for the first 15 residues of the H2A tail either interacting with ubiquitin or other parts of the complex. As our cryo-EM map predicts conformational flexibility in ubiquitin, we anticipate that crosslinking to H2A may occur only in minor ubiquitin conformations. Previous results show that moving ubiquitin from H2BK120 to H2BK125 leads to a modest decrease in Dot1L activity, and moving ubiquitin to H2BK116 or K108 dramatically or entirely inhibits Dot1L, respectively (Figure S6B; Chatterjee et al., 2010). Although we do not see clear density for the ubiquitin-H2BK120 linkage in our cryo-EM map, we hypothesize that the ubiquitin C-terminal tail is flexible enough to allow ubiquitin to reach its observed position in our structure from an H2BK125 attachment site. However, H2BK108 and H2BK116 attachment places ubiquitin between Dot1L and the nucleosome and likely prevents productive binding of Dot1L as a result.

In summary, we solved the cryo-EM structure of Dot1L bound to an H2BK120ub nucleosome. Dot1L binds to the nucleosome acidic patch using a nucleosome-interaction loop and interacts

directly with ubiquitin for activation. Dot1L-mediated methylation of H3K79 is essential to the pathogenesis of MLL-rearranged leukemias, prompting the development of Dot1L inhibitors that are now in clinical trials. Our structure and subsequent biochemical validation identify mutations to interrogate the nucleosome-specific and ubiquitin-dependent functions of Dot1L. This will enable precision hypothesis-driven analysis of the functions of Dot1L in normal and leukemia model cells. Further, our studies suggest that H2BK120 ubiquitylation may function as a gatekeeper of the acidic patch. The tuning of acidic patch binding by adjacent post-translational modifications is likely to be a fundamental mechanism in the regulation of chromatin signaling.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Preparation of proteins and nucleosomes
  - Reconstitution of Dot1L-nucleosome complexes
  - Cryo-EM sample preparation
  - Cryo-EM data collection and analysis
  - Crystallization and X-ray crystallography
  - Lysine methyltransferase assays
  - Gel filtration-based binding assay
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes two tables and six figures and can be found with this article online at https://doi.org/10.1016/j.celrep.2019.01.058.

ACKNOWLEDGMENTS

We thank Andrew Thieme and Erin Blanding for reagent preparation. Brittany Ford and Fabiola Jaramillo assisted us with cryo-EM sample preparation and screening. Cryo-EM data were collected at the Shared Materials and Instrumentation Facility at Duke University, within the collaborative framework of the Molecular Microscopy Consortium. X-ray data were collected at SER-CAT 22-ID beamline at the Advanced Photon Source (supported by NIH S10_RR25528 and S10_RR028976 and the US DOE contract no. W-31-109-Eng-38). Leonard Collins at the UNC Biomarker Mass Spectrometry Facility collected mass spectrometry data. We thank the McGinty lab, Brian Strahl, and Song Tan for project discussions and comments on the manuscript. This work was supported by Searle Scholars Program, Pew-Stewart Scholars in Cancer Research, and Damon Runyon Dale F. Frey awards to R.K.M.; NIH T32GM008570 to C.J.A.; and by the NIH Intramural Research Program/US National Institute of Environmental Health Sciences (NIEHS) ZIC ES103326 to M.J.B.

AUTHOR CONTRIBUTIONS

C.J.A. and R.K.M. planned experiments, collected and analyzed data, and prepared the manuscript. M.R.B. and M.J.B. planned experiments, collected and analyzed data, and edited the manuscript. A.H., Y.K., and E.H.B. collected and analyzed data. All authors have commented on and agreed to the content of this manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 12, 2018
Revised: January 9, 2019
Accepted: January 15, 2019
Published: February 12, 2019

REFERENCES

Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr. D Struct. Biol. *74*, 531–544.

Armache, K.J., Garlick, J.D., Canzio, D., Narlikar, G.J., and Kingston, R.E. (2011). Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. Science *334*, 977–982.

Barbera, A.J., Chodaparambil, J.V., Kelley-Clarke, B., Joukov, V., Walter, J.C., Luger, K., and Kaye, K.M. (2006). The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. Science *311*, 856–861.

Bernt, K.M., Zhu, N., Sinha, A.U., Vempati, S., Faber, J., Krivtsov, A.V., Feng, Z., Punt, N., Daigle, A., Bullinger, L., et al. (2011). MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. Cancer Cell *20*, 66–78.

Bilokapic, S., Strauss, M., and Halic, M. (2018). Histone octamer rearranges to adapt to DNA unwrapping. Nat. Struct. Mol. Biol. *25*, 101–108.

Briggs, S.D., Xiao, T., Sun, Z.-W., Caldwell, J.A., Shabanowitz, J., Hunt, D.F., Allis, C.D., and Strahl, B.D. (2002). Gene silencing: trans-histone regulatory pathway in chromatin. Nature *418*, 498.

Chatterjee, C., McGinty, R.K., Fierz, B., and Muir, T.W. (2010). Disulfide-directed histone ubiquitylation reveals plasticity in hDot1L activation. Nat. Chem. Biol. *6*, 267–269.

Couture, J.-F., Collazo, E., Brunzelle, J.S., and Trievel, R.C. (2005). Structural and functional analysis of SET8, a histone H4 Lys-20 methyltransferase. Genes Dev. *19*, 1455–1465.

D'Arcy, A., Sweeney, A.M., and Haber, A. (2004). Practical aspects of using the microbatch method in screening conditions for protein crystallization. Methods *34*, 323–328.

Dann, G.P., Liszczak, G.P., Bagert, J.D., Müller, M.M., Nguyen, U.T.T., Wojcik, F., Brown, Z.Z., Bos, J., Panchenko, T., Pihl, R., et al. (2017). ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference. Nature *548*, 607–611.

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., and Richmond, T.J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. J. Mol. Biol. *319*, 1097–1113.

Dillon, S.C., Zhang, X., Trievel, R.C., and Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol. *6*, 227.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Feng, Q., Wang, H., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Struhl, K., and Zhang, Y. (2002). Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. Curr. Biol. *12*, 1052–1058.

Fingerman, I.M., Li, H.-C., and Briggs, S.D. (2007). A charge-based interaction between histone H4 and Dot1 is required for H3K79 methylation and telomere silencing: identification of a new trans-histone pathway. Genes Dev. *21*, 2018–2029.

Fujiki, R., Hashiba, W., Sekine, H., Yokoyama, A., Chikanishi, T., Ito, S., Imai, Y., Kim, J., He, H.H., Igarashi, K., et al. (2011). GlcNAcylation of histone H2B facilitates its monoubiquitination. Nature *480*, 557–560.

Greer, E.L., and Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. Nat. Rev. Genet. *13*, 343–357.

Holt, M.T., David, Y., Pollock, S., Tang, Z., Jeon, J., Kim, J., Roeder, R.G., and Muir, T.W. (2015). Identification of a functional hotspot on ubiquitin required for stimulation of methyltransferase activity on chromatin. Proc. Natl. Acad. Sci. USA *112*, 10365–10370.

Hyun, K., Jeon, J., Park, K., and Kim, J. (2017). Writing, erasing and reading histone lysine methylations. Exp. Mol. Med. *49*, e324.

Jbara, M., Maity, S.K., Morgan, M., Wolberger, C., and Brik, A. (2016). Chemical synthesis of phosphorylated histone H2A at Tyr57 reveals insight into the inhibition mode of the SAGA deubiquitinating module. Angew. Chem. Int. Ed. Engl. *55*, 4972–4976.

Kalashnikova, A.A., Porter-Goff, M.E., Muthurajan, U.M., Luger, K., and Hansen, J.C. (2013). The role of the nucleosome acidic patch in modulating higher order chromatin structure. J. R. Soc. Interface *10*, 20121022.

Kato, H., Jiang, J., Zhou, B.-R., Rozendaal, M., Feng, H., Ghirlando, R., Xiao, T.S., Straight, A.F., and Bai, Y. (2013). A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. Science *340*, 1110–1113.

Kim, J., Hake, S.B., and Roeder, R.G. (2005). The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. Mol. Cell *20*, 759–770.

Kim, S.-A., Chatterjee, N., Jennings, M.J., Bartholomew, B., and Tan, S. (2015). Extranucleosomal DNA enhances the activity of the LSD1/CoREST histone demethylase complex. Nucleic Acids Res. *43*, 4868–4880.

Kimanius, D., Forsberg, B.O., Scheres, S.H., and Lindahl, E. (2016). Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. eLife *5*, e18722.

Kucukelbir, A., Sigworth, F.J., and Tagare, H.D. (2014). Quantifying the local resolution of cryo-EM density maps. Nat. Methods *11*, 63–65.

Lacoste, N., Utley, R.T., Hunter, J.M., Poirier, G.G., and Côte, J. (2002). Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. J. Biol. Chem. *277*, 30421–30424.

Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J. Mol. Biol. *276*, 19–42.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997a). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251–260.

Luger, K., Rechsteiner, T.J., Flaus, A.J., Waye, M.M., and Richmond, T.J. (1997b). Characterization of nucleosome core particles containing histone proteins made in bacteria. J. Mol. Biol. *272*, 301–311.

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Preparation of nucleosome core particle from recombinant histones. Methods Enzymol. *304*, 3–19.

Machida, S., Sekine, S., Nishiyama, Y., Horikoshi, N., and Kurumizaka, H. (2016). Structural and biochemical analyses of monoubiquitinated human histones H2B and H4. Open Biol. *6*, 160090.

Makde, R.D., England, J.R., Yennawar, H.P., and Tan, S. (2010). Structure of RCC1 chromatin factor bound to the nucleosome core particle. Nature *467*, 562–566.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Cryst. *40*, 658–674.

McGinty, R.K., and Tan, S. (2016). Recognition of the nucleosome by chromatin factors and enzymes. Curr. Opin. Struct. Biol. *37*, 54–61.

McGinty, R.K., Kim, J., Chatterjee, C., Roeder, R.G., and Muir, T.W. (2008). Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. Nature *453*, 812–816.

McGinty, R.K., Köhn, M., Chatterjee, C., Chiang, K.P., Pratt, M.R., and Muir, T.W. (2009). Structure-activity analysis of semisynthetic nucleosomes: mechanistic insights into the stimulation of Dot1L by ubiquitylated histone H2B. ACS Chem. Biol. *4*, 958–968.

McGinty, R.K., Henrici, R.C., and Tan, S. (2014). Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. Nature *514*, 591–596.

McGinty, R.K., Makde, R.D., and Tan, S. (2016). Preparation, crystallization, and structure determination of chromatin enzyme/nucleosome complexes. Methods Enzymol. *573*, 43–65.

Min, J., Feng, Q., Li, Z., Zhang, Y., and Xu, R.-M. (2003). Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. Cell *112*, 711–723.

Morgan, M.T., Haj-Yahya, M., Ringel, A.E., Bandi, P., Brik, A., and Wolberger, C. (2016). Structural basis for histone H2B deubiquitination by the SAGA DUB module. Science *351*, 725–728.

Ng, H.H., Feng, Q., Wang, H., Erdjument-Bromage, H., Tempst, P., Zhang, Y., and Struhl, K. (2002a). Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. Genes Dev. *16*, 1518–1527.

Ng, H.H., Xu, R.-M., Zhang, Y., and Struhl, K. (2002b). Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79. J. Biol. Chem. *277*, 34655–34657.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Sawada, K., Yang, Z., Horton, J.R., Collins, R.E., Zhang, X., and Cheng, X. (2004). Structure of the conserved core of the yeast Dot1p, a nucleosomal histone H3 lysine 79 methyltransferase. J. Biol. Chem. *279*, 43296–43306.

Schübeler, D., MacAlpine, D.M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D.E., O'Neill, L.P., Turner, B.M., Delrow, J., et al. (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. Genes Dev. *18*, 1263–1271.

Shah, N.H., Dann, G.P., Vila-Perelló, M., Liu, Z., and Muir, T.W. (2012). Ultrafast protein splicing is common among cyanobacterial split inteins: implications for protein engineering. J. Am. Chem. Soc. *134*, 11338–11341.

Shahbazian, M.D., Zhang, K., and Grunstein, M. (2005). Histone H2B ubiquitylation controls processive methylation but not monomethylation by Dot1 and Set1. Mol. Cell *19*, 271–277.

Shema-Yaacoby, E., Nikolov, M., Haj-Yahya, M., Siman, P., Allemand, E., Yamaguchi, Y., Muchardt, C., Urlaub, H., Brik, A., Oren, M., and Fischle, W. (2013). Systematic identification of proteins binding to chromatin-embedded ubiquitylated H2B reveals recruitment of SWI/SNF to regulate transcription. Cell Rep. *4*, 601–608.

van Leeuwen, F., Gafken, P.R., and Gottschling, D.E. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell *109*, 745–756.

Vasudevan, D., Chua, E.Y.D., and Davey, C.A. (2010). Crystal structures of nucleosome core particles containing the '601' strong positioning sequence. J. Mol. Biol. *403*, 1–10.

Vijay-Kumar, S., Bugg, C.E., and Cook, W.J. (1987). Structure of ubiquitin refined at 1.8 A resolution. J. Mol. Biol. *194*, 531–544.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M.Q., and Zhao, K. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat. Genet. *40*, 897–903.

Wilson, M.D., Benlekbir, S., Fradet-Turcotte, A., Sherker, A., Julien, J.-P., McEwan, A., Noordermeer, S.M., Sicheri, F., Rubinstein, J.L., and Durocher, D. (2016). The structural basis of modified nucleosome recognition by 53BP1. Nature *536*, 100–103.

Winters, A.C., and Bernt, K.M. (2017). MLL-rearranged leukemias-an update on science and clinical approaches. Front Pediatr. *5*, 4.

Wood, K., Tellier, M., and Murphy, S. (2018). DOT1L and H3K79 methylation in transcription and genomic stability. Biomolecules *8*, 11.

Zhang, K. (2016). Gctf: real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12.

Zhao, Y., and Garcia, B.A. (2015). Comprehensive catalog of currently documented histone modifications. Cold Spring Harb. Perspect. Biol. *7*, a025064.

Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.

Zhou, L., Holt, M.T., Ohashi, N., Zhao, A., Müller, M.M., Wang, B., and Muir, T.W. (2016). Evidence that ubiquitylated H2B corrals hDot1L on the nucleosomal surface to induce H3K79 methylation. Nat. Commun. *7*, 10589.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
E. coli BL21(DE3)pLysS	Novagen	Cat.#69388-3
E. coli HB101	ATCC	Cat.#67593
Chemicals, Peptides, and Recombinant Proteins		
TALON Metal Affinity Resin	Clontech	Cat.#635669
Source S resin	GE Healthcare	Cat.#17094405
Superdex 200 increase 10/300 column	GE Healthcare	Cat.#28990944
Source Q resin	GE Healthcare	Cat.#17094705
High density cobalt resin	Agarose Bead Technologies	Cat.#6BCL-QHCo
1,7-Dichloroacetone	Sigma-Aldrich	Cat.#384372
Vydac 218TP 5um C18 Column 150x4.6mm	Mac-Mod Analytical	Cat.#18TP5415
Vydac 218TP 5um C18 Column 250x10mm	Mac-Mod Analytical	Cat.#18TP510
Vydac 218TP 10um C18 Column 250x22mm	Mac-Mod Analytical	Cat.#218TP1022
2-Mercaptoethanesulfonic acid sodium salt	Sigma-Aldrich	Cat.#M1511
Zeba spin desalting column	Thermo Fisher	Cat.#PI87764
Glutaraldehyde	Sigma-Aldrich	Cat.#G5882
Bis-Tris	Hampton Research	Cat.#HR2-783
Ammonium phosphate dibasic	Hampton Research	Cat.#HR2-629
Poly(ethylene glycol) methyl ether (PEG2000-MME)	Sigma-Aldrich	Cat.#202509
Silicon fluid	Clearco	Cat.#PSF-1cSt
Paraffin oil	Fisher	Cat.#O1214
Poly(ethylene glycol) (PEG400)	Sigma-Aldrich	Cat.#202398
Critical Commercial Assays		
MTase-Glo Methyltransferase Assay	Promega	Cat.#V7601
Deposited Data		
hDot1L-H2BK120ub nucleosome poised structure	This paper	PDB: 6NN6
hDot1L-H2BK120ub nucleosome 3.9 Å map	This paper	EMD-0458
hDot1L-H2BK120ub nucleosome 3.5 Å map	This paper	EMD-0459
hDot1L-H2BK120ub nucleosome 7.6 Å masked map	This paper	EMD-0460
Xenopus nucleosome crystal structure with 601 DNA	Vasudevan et al., 2010	PDB: 3LZ0
Human Dot1L catalytic domain crystal structure	Min et al., 2003	PDB: 1NW3
Human ubiquitin crystal structure	Vijay-Kumar et al., 1987	PDB: 1UBQ
Xenopus nucleosome crystal structure with alpha-satellite DNA	Davey et al., 2002	PDB: 1KX5
LANA peptide/nucleosome crystal structure	Barbera et al., 2006	PDB: 1ZLA
Sir3 BAH domain/nucleosome crystal structure	Armache et al., 2011	PDB: 3TU4
CENP-C/nucleosome crystal structure	Kato et al., 2013	PDB: 4X23
RCC1/nucleosome crystal structure	Makde et al., 2010	PDB: 3MVD
SAGA DUB module/nucleosome crystal structure	Morgan et al., 2016	PDB: 4ZUX
PRC1 ubiquitylation module/nucleosome crystal structure	McGinty et al., 2014	PDB: 4RP8
53BP1/nucleosome cryo-EM structure	Wilson et al., 2016	PDB: 5KGF
Recombinant DNA		
Plasmid expressing hDot1L(2-416)	gift from Song Tan	pST50Tr-STRaHISNhDot1Lt1x1
Plasmid expressing hH2A.D	This paper	pST50Trc3-hH2A.D
Plasmid expressing hH2B.C	This paper	pST50Trc4-hH2B.C
Plasmid expressing hH3.2	This paper	pST50Tr-hH3.2

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Plasmid expressing xH4	Luger et al., 1997b	pET3a-xH4
Plasmid containing multiple copies of 147 bp Widom 601 DNA	McGinty et al., 2016	pST55-16xNCP601a
Plasmid containing multiple copies of 155 bp Widom 601 DNA	Kim et al., 2015	pST55-16xNCP601a155M
Plasmid expressing LANA	gift from Song Tan and Jiehuan Huang	pST50Tr-AVTGSTHISNLANA
Plasmid expressing LANA(LRS > AAA) mutant	gift from Song Tan and Jiehuan Huang	pST50Tr-AVTGSTHISNLANAx1
Plasmid expressing ExoX(2-167; D6A)	gift from Song Tan	pST50Tr-STRaHSTNGS3ExoXt1x1
Plasmid expressing Ubiquitin(G76C)	gift from Song Tan and Ryan Henrici	pST50Tr-STRaHSTNGS2hUbqx7
Plasmid expressing xH2B(K120C)	This paper	pST50Trc4-xH2Bx27
Plasmid expressing xH2A	Luger et al., 1997b	pET3a-xH2A
Plasmid expressing xH2B	Luger et al., 1997b	pET3a-xH2B
Plasmid expressing xH3	Luger et al., 1997b	pET3a-xH3
Plasmid expressing NpuC	This paper	pST50Tr-MBPNNpuC-HIS
Plasmid expressing xH2B(1-116)-NpuN	This paper	pST50Tr-xH2Bt6NpuNSMOHST
Plasmid expressing Ubiquitin(1-75)-Npu	This paper	pST50Tr-hUbqt1NpuSMOHST
Software and Algorithms		
Microsoft Excel	Microsoft	https://products.office.com/en-us/excel
EPU	FEI	https://www.fei.com/software/epu-automated-single-particles-software-for-life-sciences/
MotionCor2	Zheng et al., 2017	http://msg.ucsf.edu/em/software/motioncor2.html
GCTF	Zhang, 2016	https://www.mrc-lmb.cam.ac.uk/kzhang/Gctf/
RELION 2.1	Kimanius et al., 2016	http://www2.mrc-lmb.cam.ac.uk/relion
Chimera	Pettersen et al., 2004	https://www.cgl.ucsf.edu/chimera
Phenix	Afonine et al., 2018	https://www.phenix-online.org/download/
Coot	Emsley et al., 2010	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
ResMap	Kucukelbir et al., 2014	http://resmap.sourceforge.net/
CCP4 Phaser MR	McCoy et al., 2007	http://www.ccp4.ac.uk/html/phaser.html
Other		
Holey carbon grid, Cu 300 mesh	Quantifoil Micro Tools GmbH	Quantifoil R 1.2/1.3

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Robert K. McGinty (rmcginty@email.unc.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

BL21(DE3)pLysS and HB101 *E. coli* cells were used for recombinant expression of proteins and growth of nucleosomal DNA repeat plasmids, respectively. Competent cells of these strains were prepared from cells restreaked from glycerol stocks originally obtained by Song Tan from Novagen and ATCC, respectively.

METHOD DETAILS

Preparation of proteins and nucleosomes

An unpublished plasmid with the gene fragment encoding human Dot1L(2-416) (Dot1L$_{cat}$) in the pST50Tr vector with an N-terminal Strep peptide (STR)-hexahistidine (His$_6$) affinity tag was obtained as a gift from Song Tan. STR-His$_6$-hDot1L$_{cat}$ was expressed in *E. coli* BL21(DE3)pLysS cells at 18°C. The tagged protein was enriched by metal-affinity chromatography using Talon resin (Clontech). The affinity tag was removed using the tobacco etch virus (TEV) protease prior to further purification by cation exchange chromatography using a Source S resin (GE Healthcare). For structural biology, an additional gel filtration chromatography step with a Superdex 200 increase 10/300 column (GE Healthcare) was performed. All Dot1L$_{cat}$ point mutants were cloned by site-directed mutagenesis and expressed and purified identically to the wild-type protein fragment. All Dot1L mutants were verified by intact mass spectrometry using an Agilent 6520 Accurate Mass Quadrupole Time-of-Flight mass spectrometer. Unpublished pST50Tr plasmids containing the sequence encoding LANA peptide residues 1-23 or a triple mutant of residues 8-10 (LRS to AAA) with an N-terminal AviTag (AVT)-glutathione-S-transferase (GST)-His$_6$ tag were obtained as a gift from Jiehuan Huang and Song Tan. AVT-GST-His$_6$-LANA and -LANA(LRS > AAA) were expressed in *E. coli* BL21(DE3)pLysS cells at 37°C. The fusion proteins were purified by Talon metal-affinity chromatography and anion exchange chromatography using a Source Q resin (GE Healthcare). An unpublished pST50Tr plasmid containing STR-His$_{10}$-tagged *E. coli* exodeoxyribonucleaseX (ExoX) residues 2-167 with an Asp6Ala inactivating mutant was a gift from Song Tan. This inactive ExoX was expressed and purified as described above for Dot1L$_{cat}$. TEV cleavage leaves a 3x Gly-Ser sequence at the N terminus of ExoX.

Consensus sequences from human H2A and H2B and the sequence for H3.2 were cloned into pST50Tr vectors. Recombinant *Xenopus* and human histones were expressed, purified and reconstituted into nucleosomes as previously described (Luger et al., 1997b, 1999). Nucleosomes were assembled with 147 or 155 bp of the 601 DNA sequence (Lowary and Widom, 1998) with the nucleosome positioning sequence centered in each DNA fragment (Kim et al., 2015; McGinty et al., 2016). H2BK120ub(DCA) was prepared essentially as previously described (Morgan et al., 2016). Briefly, *Xenopus* H2BK120C was cloned by site-directed mutagenesis and the pST50Tr-STR-His$_{10}$-ubiquitin(Gly76Cys) plasmid was obtained as a gift from Ryan Henrici and Song Tan. Purified *Xenopus* H2A/H2BK120C dimer and purified STR-His$_{10}$-ubiquitin(Gly76Cys) were combined at 0.5 and 1 μM concentrations, respectively, in crosslinking buffer (6 M urea, 33.3 mM sodium borate pH 8.5, 5 mM TCEP) and cooled on ice. Crosslinking was performed by the addition of dichloroacetone (DCA, Sigma-Aldrich) to a final concentration of 2.5 μM. After 60 min on ice, DCA was quenched with the addition of 50 mM β-mercaptoethanol. Histone dimer was refolded from the mixture as previously reported and purified with high density cobalt resin (Agarose Bead Technologies) to remove any H2A/H2BK120C lacking DCA-crosslinked ubiquitin. TEV protease was used to remove the STR-His$_{10}$ tag on ubiquitin, and H2BK120ub(DCA)-containing histone dimer was purified away from crosslinked ubiquitin dimers by cation exchange chromatography. A Gly-Ser-Gly-Ser sequence remains on the N terminus of ubiquitin after protease cleavage.

Xenopus H2BK120ub(Gly76Ala) was prepared essentially as previously described (McGinty et al., 2009) with the following exceptions. The ubiquitin(1-75)-MES thioester was prepared by cloning ubiquitin residues 1-75 in frame with a fused Npu intein ending in the AAFN sequence (Shah et al., 2012) followed by a SUMO-His$_{10}$ tag in the pST50Tr vector. The ubiquitin intein fusion protein was expressed in *E. coli* BL21(DE3)pLysS cells at 37°C, purified using metal-affinity chromatography using Talon resin (Clontech), and dialyzed into 100 mM Na_2HPO_4, pH 7.2, 150 mM NaCl, 1 mM EDTA. The ubiquitin(1-75)-MES thioester was formed by the addition of 100 mM 2-mercaptoethanesulfonic acid sodium salt (MESNa, Sigma-Aldrich) and 5 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP, VWR) and incubation overnight at 4°C, followed by purification with cation exchange chromatography using a Source S resin. The *Xenopus* H2B(1-116)-Npu fusion exhibited premature cleavage during expression, so we opted to use a split version (Shah et al., 2012). To this end, the H2B fragment was cloned in frame with the N-terminal Npu fragment (NpuN) with a C-terminal SUMO-His$_{10}$ tag in a pST50Tr vector. The C-terminal Npu fragment (NpuC) was cloned with a TEV-cleavable N-terminal MBP tag and a C-terminal His$_6$ tag. MBP-NpuC-His$_6$ was expressed in *E. coli* BL21(DE3)pLysS cells at 37°C and purified using metal-affinity chromatography using Talon resin (Clontech), cleaved with TEV protease, and further purified by cation exchange chromatography using a Source S resin. H2B(1-116)-NpuN-SUMO-His$_6$ was also expressed in *E. coli* BL21(DE3)pLysS cells at 37°C and extracted from inclusion bodies using standard histone inclusion body purification procedures (Luger et al., 1999). H2B(1-116)-NpuN-SUMO-His$_6$ and NpuC-His$_6$ were dialyzed into 1.6 M guanidine, 100 mM Na_2HPO_4, 140 mM NaCl, and 1 mM EDTA and combined in roughly equimolar quantities. The H2B(1-116)-MES thioester was formed by addition of 200 mM MESNa and 5 mM TCEP and incubation overnight at 30°C prior to purification by cation exchange chromatography using a Source S resin. Peptide synthesis, expressed protein ligation, and desulfurization were performed exactly as previously described (McGinty et al., 2009). Semisynthetic intermediates and the final H2BK120ub(Gly76Ala) product were purified using Vydac C18 columns (Mac-Mod Analytical) and mass spectrometric characterization was performed using an Agilent 6520 Accurate Mass Quadrupole Time-of-Flight mass spectrometer. H2BK120ub analogs were assembled into nucleosomes as described above.

Reconstitution of Dot1L-nucleosome complexes

The Dot1L$_{cat}$-H2BK120ub(DCA) nucleosome complex used for cryo-EM experiments was reconstituted as previously reported for other nucleosome complexes (McGinty et al., 2016). Briefly, 2.5 molar equivalents of Dot1L$_{cat}$ were added in four steps to H2BK120ub(DCA) nucleosomes containing all *Xenopus* histones and 147 bp 601 DNA at 5 minute intervals in reconstitution buffer

(25 mM Tris-Cl pH 7.6, 150 mM NaCl, 1 mM DTT, 0.1 mM PMSF). The complex was purified by gel filtration using a Superdex 200 increase column equilibrated in reconstitution buffer. For crystallography experiments, Dot1L$_{cat}$ was reconstituted using DCA-crosslinked H2BK120ub nucleosomes containing 155 bp 601 DNA and the inactive ExoX as a crystallization chaperone. Reconstitution of the complex was performed as described above, with the following changes: 1) 2.5 molar equivalents of ExoX was added in four steps after addition of Dot1L and 2) reconstitution and purification were performed in reconstitution buffer containing 50 mM NaCl.

Cryo-EM sample preparation

Dot1L$_{cat}$-H2BK120ub(DCA) nucleosome complex at ~10 mg/ml was buffer exchanged into crosslinking buffer (10 mM HEPES pH 7.5, 150 mM NaCl) using a Zeba spin desalting column (Thermo Fisher). Glutaraldehyde (Sigma-Aldrich) was added to a final concentration of 0.1% and crosslinking was allowed to proceed for 5 min at room temperature. Glutaraldehyde was quenched by the addition of Tris-Cl at pH 7.5 to a final concentration of 20 mM. After 15 min at room temperature, samples were buffer exchanged into crosslinking buffer. Complex was diluted to 0.93 mg/ml in freezing buffer (10 mM HEPES pH 7.5, 48 mM NaCl) prior to spotting on a Quantifoil R1.2/1.3, 300 mesh grid for vitrification after blotting for 4 s using a Vitrobot Mark IV (FEI) cryoplunger at 100% humidity and 4°C. An uncrosslinked complex was analyzed by liquid chromatography coupled mass spectrometry using an Agilent 6520 Accurate Mass Quadrupole Time-of-Flight mass spectrometer to assess the methylation state of histone H3.

Cryo-EM data collection and analysis

Cryo-EM data was collected on a Titan Krios (FEI) operating at 300 kV. The images were recorded using the automated data acquisition software EPU (FEI) using a Falcon III direct electron detector (FEI) operated in electron counting mode at a nominal magnification of 75,000, corresponding to a pixel size of 1.08 Å, and at a defocus range of −1.25 μm to −2.75 μm. Each exposure lasted 60 s and was collected as a 30-frame movie at a dose rate of ~0.8e^-/pix/s, resulting in a total electron dose of ~42 e^-/Å^2. Whole-frame movie alignment was performed with MotionCor2 (Zheng et al., 2017), CTF parameters were estimated with GCTF (Zhang, 2016), and Relion v2.1 (Kimanius et al., 2016) was used for subsequent data analysis. Initially, ~1,000 particles were manually picked to generate 2D class averages to use as templates for iterative rounds of automatic particle picking. In total, 408,526 automatically selected particles were subjected to reference-free 2D classification. Iterative rounds of 2D classification were used to narrow the set to 156,709 high quality particles. 3D classification using a 50 Å low pass filtered nucleosome structure (PDB: 6ESF) generated in Relion as reference identified two classes (3 and 4) containing clearly defined Dot1L-nucleosome complexes. Movie refinement, particle polishing, 3D refinement, and post-processing of particles in classes 3 and 4 resulted in a 3.5 Å reconstruction showing significant conformational heterogeneity in Dot1L and ubiquitin. Polished particles from classes 3 and 4 were subjected to further 3D classification into 4 new subclasses (designated with A). Refinement and post-processing of the resultant class 2A yielded a 3.9 Å reconstruction used for modeling. Masking of the Dot1L and ubiquitin densities for class 2A and continuing refinement resulted in a 7.6 Å map with improved density for Dot1L and ubiquitin. Finally, the original class 3 and class 4 polished particles were reclassified into 10 3D subclasses (designated with B), allowing further parsing of stoichiometry and conformational heterogeneity. Crystal structures for the nucleosome (PDB: 3LZ0) (Vasudevan et al., 2010), the catalytic domain of Dot1L (PDB: 1NW3) (Min et al., 2003), and ubiquitin (PDBID 1UBQ) (Vijay-Kumar et al., 1987) were docked into the 3.9 Å reconstructed map using Chimera (Pettersen et al., 2004) and refined by iterative real-space refinement in Phenix (Afonine et al., 2018) and manual model building in Coot (Emsley et al., 2010). Local resolution was calculated with ResMap (Kucukelbir et al., 2014). Resolution of reconstructions was calculated using the gold standard FSC criterion of 0.143. All molecular graphics were prepared with Chimera.

Crystallization and X-ray crystallography

The Dot1L$_{cat}$-ExoX-H2BK120ub(DCA) nucleosome complex was concentrated to 10 mg/ml and crystallized by mixing 1:1 with 25 mM Bis-Tris pH 6 (Hampton Research), 25 mM $(NH_4)_2HPO_4$ (Hampton Research), and 1.5% PEG2000-MME (Sigma-Aldrich) at room temperature using modified microbatch under oil with Al's oil (equal volumes of silicon fluid (Clearco) and paraffin oil (Fisher)) (D'Arcy et al., 2004). Notably, this complex was not crosslinked with glutaraldehyde. Inactive ExoX was used as a crystallization chaperone (R.K.M. and S. Tan, unpublished data). Crystals were collected after 20-35 days and dehydrated and cryoprotected by soaking at room temperature in 25 mM Bis-Tris pH 6, 25 mM $(NH_4)_2HPO_4$, and 3.75% PEG2000-MME in drops containing increasing concentrations of PEG400 (final concentration 24%, increment 4%, soak interval 15 min). Upon completion of soaking, crystals were flash cooled in liquid nitrogen. Diffraction data was collected on the Advanced Photon Source SER-CAT 22-ID beamline at 1 Å and 100 K. Data was processed using HKL2000 (Otwinowski and Minor, 1997) in the C2221 space group. A molecular replacement model was generated with CCP4 Phaser MR (McCoy et al., 2007) using a polyalanine half nucleosome model generated from PDB: 1KX5 (Davey et al., 2002) and a polyalanine Dot1L catalytic domain model generated from PDB: 1NW3, with loop residues 296-313 deleted.

Lysine methyltransferase assays

Lysine methyltransferase assays were performed using the MTase-Glo enzyme-methyltransferase kit (Promega) using the manufacturer's protocol. Briefly, 1 μM nucleosomes were combined with 15 nM wild-type or mutant Dot1L$_{cat}$ in methyltransferase buffer (20 mM Tris-Cl pH 8.0, 50 mM NaCl, 1 mM EDTA, 0.1 mg/ml BSA, 1 mM DTT, and 20 μM SAM) in a total volume of 8 μl. The concentration of 15 nM Dot1L was selected as it was very close to the linear range of activity on H2BK120ub nucleosomes, while allowing

activity on unmodified nucleosomes to be measured robustly. Using the same concentration allowed the direct comparison of activity on ubiquitylated and unmodified nucleosomes. The assay was allowed to proceed at 30°C for 30 min prior to developing the assay and quantifying luminescence using an Envision 2103 Multilabel plate reader (Perkin Elmer). Assays were performed with 5 replicates. All assays were performed with human histones except for *Xenopus* unmodified or ubiquitylated H2B. No significant difference in Dot1L activity was observed between nucleosomes containing *Xenopus* H2B and those containing canonical human H2B, validating the use of chimeric nucleosomes. For LANA competition methyltransferase activity assays, AVT-GST-His_6-LANA or AVT-GST-His_6-LANA(LRS > AAA) were added to the assay at 0, 5, 10, 25, 50, and 100 μM final concentrations. An additional assay was performed after incubation of LANA fusion proteins with 0.025 molar equivalents of TEV protease overnight at room temperature. Cleavage of LANA affinity tags was verified by SDS-PAGE.

Gel filtration-based binding assay

AVT-GST-His_6-LANA was reconstituted with H2BK120ub(Gly76Ala) or unmodified nucleosomes as described above for Dot1L-nucleosome reconstitutions at a 4:1 molar ratio (~4.5 μM nucleosome) in 10 mM Tris-Cl pH 8.0, 50 mM NaCl, and 1 mM DTT. Interaction was evaluated by separation on a Superdex 200 increase gel filtration column equilibrated in the same buffer. Chromatography fractions were analyzed by SDS-PAGE.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were performed with Microsoft Excel. All methyltransferase data are reported as the mean ± standard deviation of five replicates. All samples in a given figure panel were performed simultaneously. Quantification of resolution for and statistical validation of cryo-EM density maps were carried out in RELION 2.1.

DATA AND SOFTWARE AVAILABILITY

The cryo-EM maps and atomic coordinates of the Dot1L-H2BK120ub nucleosome complex have been deposited in the EMDB and PDB under ID codes EMD-0458 (3.9 Å map used for poised complex structure), EMD-0459 (3.5 Å map), EMD-0460 (7.6 Å map from Dot1L and ubiquitin masking), and PDB: 6NN6 (poised structure). Coordinates for X-ray crystallography molecular replacement model will be distributed upon request.

Cell Reports, Volume *26*

Supplemental Information

Structural Basis for Recognition of Ubiquitylated Nucleosome by Dot1L Methyltransferase

Cathy J. Anderson, Matthew R. Baird, Allen Hsu, Emily H. Barbour, Yuka Koyama, Mario J. Borgnia, and Robert K. McGinty

Table S1. Cryo-EM data collection and processing, related to Figure 1.

	Dot1L -H2BK120ub nucleosome (Classes 3 and 4)	Dot1L -H2BK120ub nucleosome (Class 2A)	Dot1L -H2BK120ub nucleosome (Class 2A masked)
Data collection			
Microscope	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios
Detector	FEI Falcon III	FEI Falcon III	FEI Falcon III
Magnification	75,000	75,000	75,000
Voltage (kV)	300	300	300
Electron exposure ($e^-/Å^2$)	42	42	42
Defocus range (μm)	-1.25 to -2.75	-1.25 to -2.75	-1.25 to -2.75
Pixel size (Å)	1.08	1.08	1.08
Symmetry imposed	C1	C1	C1
Data processing			
Initial particle images	408,526	408,526	408,526
Final particle images	101,430	39,558	39,558
Map resolution (Å)	3.5	3.9	7.6
FSC threshold	0.143	0.143	0.143

Table S2. X-ray crystallography data collection, related to Figure 1.

	Dot1L-ExoX-H2BK120ub nucleosome
Data collection	
Space group	C2221
Cell dimensions	
a, *b*, *c* (Å)	104.97, 200.05, 214.46
α, β, γ (°)	90.0, 90.0, 90.0
Reflections	29,554
Resolution (Å)	50.00-8.00 (8.14-8.00)
R_{merge}	0.215 (0.765)
R_{pim}	0.068 (0.238)
$I / \sigma I$	19.8 (2.0)
Completeness (%)	97.5 (90.6)
Redundancy	11.7 (9.6)

*Values in parentheses are for highest-resolution shell.

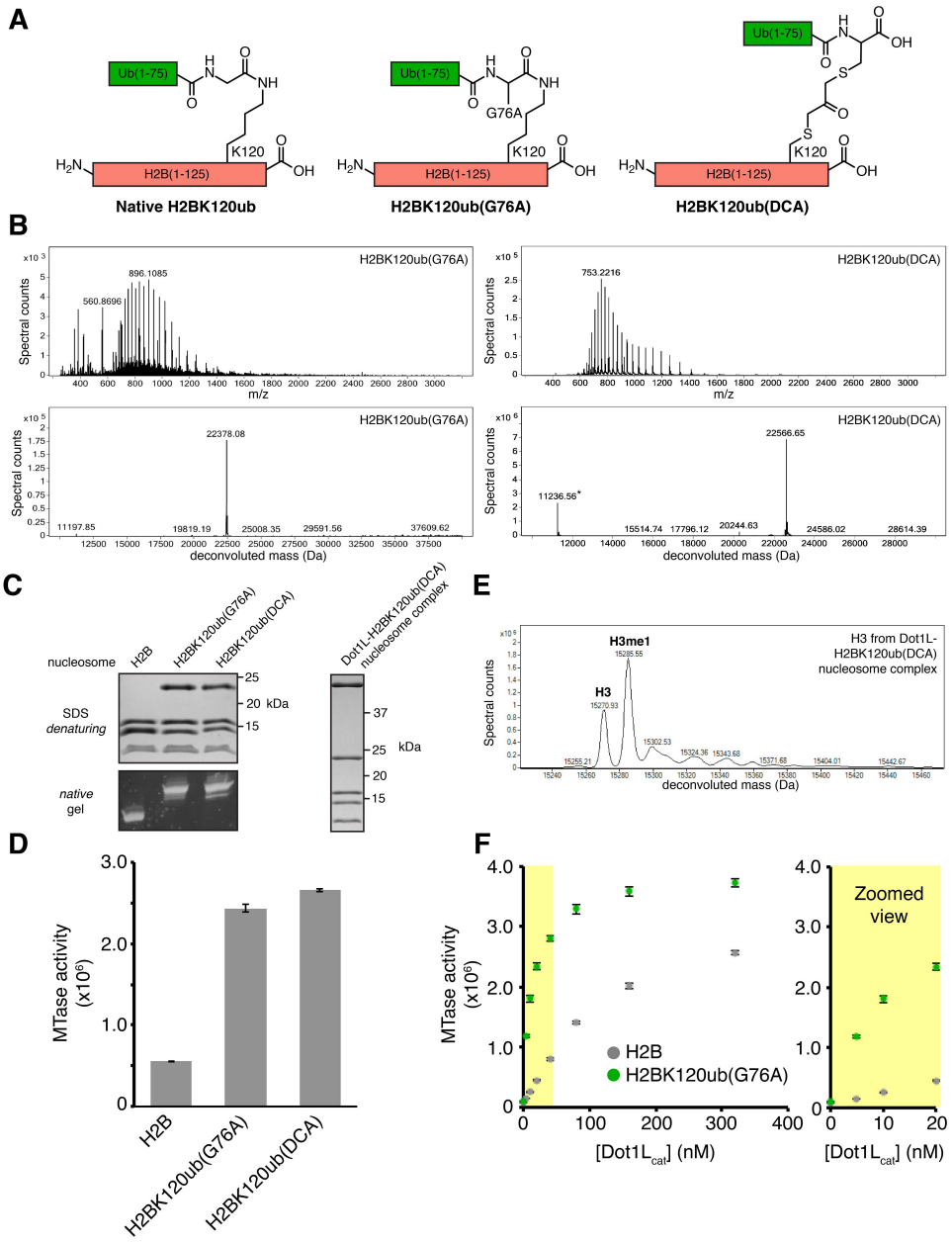


Figure S1, related to Figure 1. Preparation and characterization of H2BK120ub analogs. A, Scheme showing differences between native, G76A, and DCA-crosslinked linkages at the H2B-ubiquitin junction. **B,** Left, ESI-mass spectrum of semisynthetic H2BK120(Gly76Ala = G76A) (top) and deconvoluted mass spectrum (bottom, expected mass 22,378 Da); Right, equivalent spectra for H2BK120(DCA) from Dot1L-H2BK120(DCA) nucleosome complex (expected mass 22,566 Da; asterisk marks mass of coeluting histone H4). **C,** SDS denaturing and 10% native acrylamide gels of nucleosomes containing unmodified or ubiquitylated H2BK120 (G76A or DCA crosslinked) (left) and of reconstituted Dot1L-H2BK120ub(DCA) complex used for cryo-EM (right). **D,** Quantified methyltransferase assay using nucleosomes reconstituted with indicated unmodified or ubiquitylated H2B. **E,** Deconvoluted ESI-mass spectrum of

histone H3 from Dot1L-H2BK120ub(DCA) nucleosome complex (expected mass unmethylated 15,271 Da, monomethylated 15,285 Da). **F,** Quantified methyltransferase assay with designated Dot1L$_{cat}$ concentrations. Zoomed view of indicated region (yellow) at right. Five replicates performed for all assays and means and standard deviations are shown.

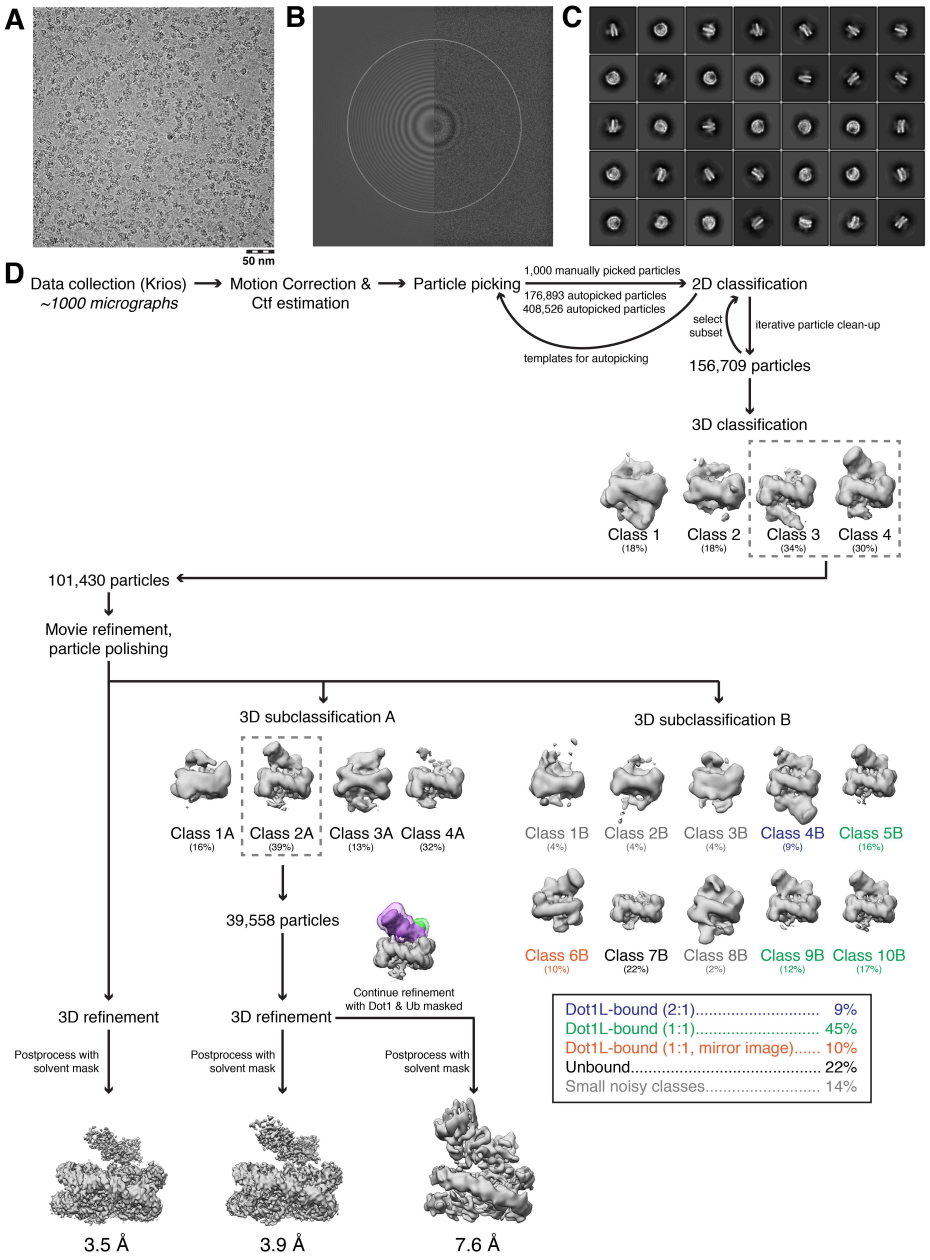


Figure S2, related to Figure 1. Workflow for cryo-EM data processing. Representative **A,** micrograph of complex with 50 nm scale bar, **B,** CTF estimation by GCTF, and **C,** reference-free 2D classifications. **D,** Data analysis scheme showing multiple classification strategies and reconstructed maps discussed in the main text. Classes 3 and 4 of initial 3D classification gave rise to a 3.5 Å map with poorly resolved density for Dot1L and ubiquitin (left). Dot1L and ubiquitin densities were improved by further classification into 4 subclasses leading to a 3.9 Å map from class 2A (middle). The Dot1L and ubiquitin volumes were masked to improve these regions of the map to allow for high confidence docking of high resolution crystal structures into the map.

Finally, finer classification of the particles from the initial classes 3 and 4 into 10 subclasses was used to parse stoichiometric and conformational heterogeneity (right).

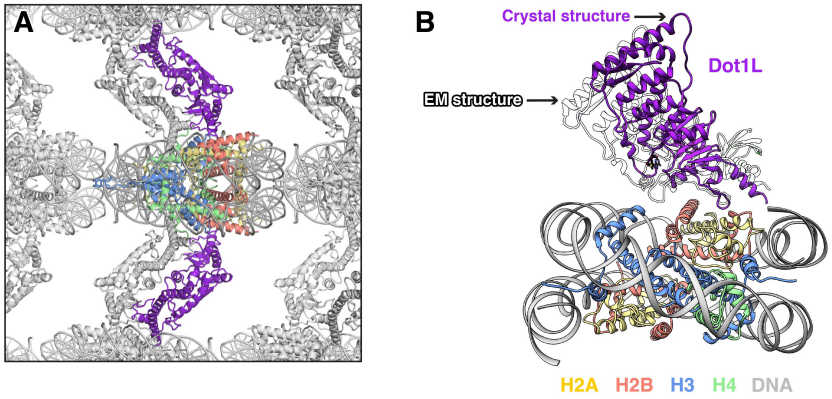


Figure S3, related to Figure 1. Comparison of X-ray and cryo-EM models of Dot1L-nucleosome complex. A, Symmetry-related complexes in crystal used for molecular replacement model with one nucleosome and two Dot1Ls colored. Another Dot1L from an adjacent unit cell wedges between Dot1L and the nucleosome to which it is bound on each nucleosome face. **B,** Alignment of cryo-EM and X-ray molecular replacement models using histones for alignment of structures. Dot1L from crystal model (purple) is lifted slightly away from the nucleosome as compared with Dot1L from the cryo-EM structure (white). The nucleosome is only shown for the cryo-EM structure for simplicity.

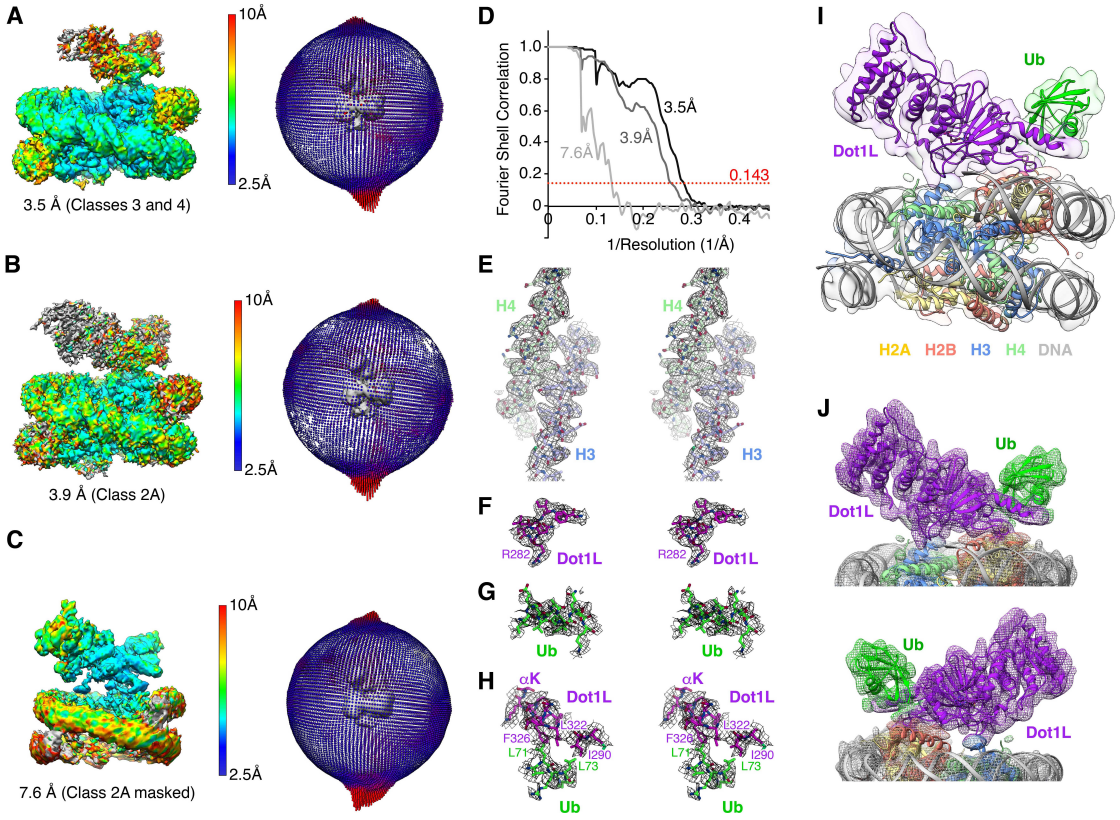


Figure S4, related to Figure 1. Local and overall resolution of reconstructed maps. **A,** Reconstructed map (3.5 Å) colored with local resolution (left), and orientational views (right). **B** and **C,** As in A for 3.9 Å reconstructed map and 7.6 Å reconstructed map after masking of Dot1L and ubiquitin volumes. **D,** FSC curves for post-processed reconstructions. **E-H,** Stereo views of local EM density maps from 3.9 Å reconstruction of **E**, H3 residues 89-112 and H4 residues 51-76, **F**, the Dot1L nucleosome interaction loop residues 277-284, **G**, ubiquitin residues 23-33, and **H**, Dot1L-ubiquitin interface including Dot1L residues 289-291+322-331 and ubiquitin residues 70-74. Masked .mrc maps in E-H contoured at 8 with 2.5 Å carve in PyMOL. **I,** Final model overlaid with 7.6 Å reconstruction from masking of Dot1L and ubiquitin volumes in Class 2A. **J,** Front and back zoomed views of overlay in panel I.

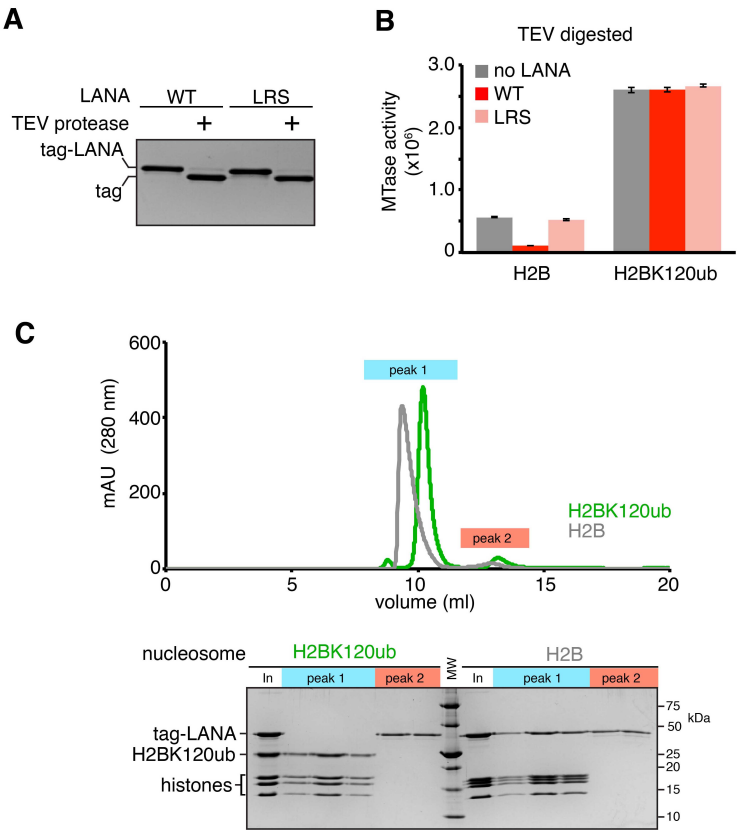


Figure S5, related to Figure 2. LANA cannot bind H2BK120ub nucleosomes. A, Denaturing gel showing TEV cleavage of LANA fusions prior to assay in B. Cleaved LANA is not observed on gel due to its small size. **B,** Quantified methyltransferase assay using no LANA or TEV cleaved wild-type or nucleosome binding-deficient (LRS = LRS 8-10 AAA) LANA fusion proteins on unmodified and H2BK120ub nucleosomes. **C,** Overlaid gel filtration chromatograms of reconstituted LANA-unmodified nucleosome complex and failed LANA-H2BK120ub nucleosome reconstitution (top). Peak 1 contains nucleosome or nucleosome-LANA complex and peak 2 contains free LANA fusion protein. Gel of representative fractions from each peak is shown (bottom). Five replicates performed for all assays and means and standard deviations are shown. In = Input.

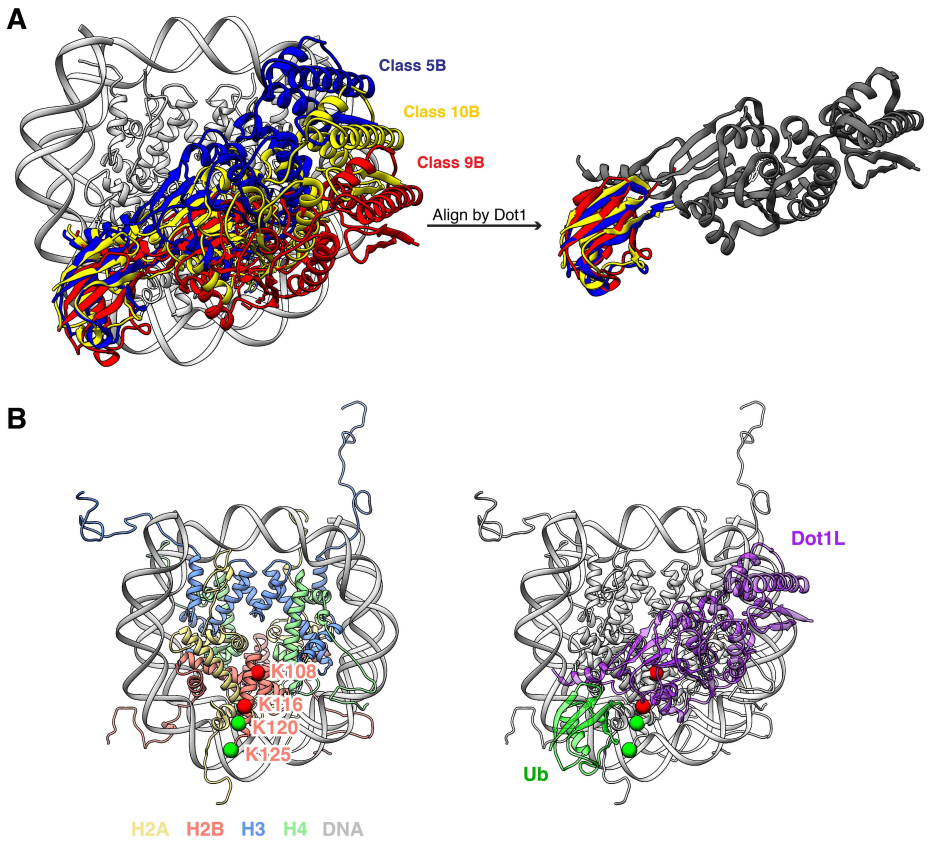


Figure S6, related to Figure 3. Analysis of ubiquitin conformational heterogeneity and attachment site relative to Dot1L. A, Overlay of docked structures for subclasses 5B, 9B, and 10B with Dot1L and ubiquitin colored (left). Alignment of Dot1L in these classes show that the ubiquitin maintains a similar orientation relative to Dot1L (right). **B,** Left, top view of nucleosome (PDBID 1KX5) with positions of Cα atoms from H2BK108, K116, K120, and K125 indicated. Dot1L is activated by ubiquitin attachment at green but not red sites. 1KX5 was used because H2BK125 is not observed in poised Dot1L-bound structure. Right, same view with Dot1L and ubiquitin from cryo-EM structure overlaid and displayed with transparency to allow visualization of underlying histone positions.