Structural basis for EPC1-mediated recruitment of MBTD1 into the NuA4/TIP60 acetyltransferase complex

Heng Zhang[1, 5], Maëva Devoucoux[2, 5], Li Li[1], Gamze Ayaz[1, 3], Harry Cheng[1], Wolfram Tempel[1], Cheng Dong[1], Peter Loppnau[1], Jacques Côté[*2] and Jinrong Min[1, 4, 6*]

1. Structural Genomics Consortium, University of Toronto, Toronto, Ontario M5G 1L7, Canada.

2. St. Patrick Research Group in Basic Oncology, Laval University Cancer Research Center, Oncology division of CHU de Québec-Université Laval Research Center, Quebec City, QC G1R 3S3, Canada.

3. Department of Biology, Middle East Technical University, Ankara, Turkey.

4. Department of Physiology, University of Toronto, Toronto, Ontario M5S 1A8, Canada.

5. These authors contributed equally to this work.

6. Lead author

*Corresponding author: jacques.cote@crhdq.ulaval.ca or jr.min@utoronto.ca

Summary

MBTD1 has recently been identified to join the NuA4/TIP60 acetyltransferase complex to regulate gene expression and DNA repair. MBTD1 competes with 53BP1 for recognition of the H4K20me mark, while MBTD1-associated TIP60 carries out the acetylation of H2AK15 and thereby blocks the ubiquitylation-dependent association of 53BP1 to the sites of DNA double strand breaks. Interestingly, the non-catalytic subunit EPC1 has been shown to enlist MBTD1 into the dynamic NuA4/TIP60 acetyltransferase complex, but the detailed molecular mechanism remains unexplored. Here, we present the crystal structure of the MBTD1-EPC1 complex that sheds light on the molecular mechanism by which MBTD1 recognizes EPC1. Mutational analysis coupled with functional evaluations in human cell lines validate the physiological significance of the key residues involved in the MBTD1-EPC1 association. Overall, our study provides a structural framework for understanding the mechanism by which MBTD1 recruits the NuA4/TIP60 acetyltransferase complex to influence transcription and DNA repair pathway choice.

Introduction

Coordination of posttranslational modifications (PTMs) represents an important signalling mechanism to control often complex biological processes such as DNA repair (Chin, 2017; Dantuma and van Attikum, 2016; Woodsmith et al., 2013; Woodsmith and Stelzl, 2014). DNA double strand break (DSB) is one of the most severe forms of genomic lesions whose improper repairs can trigger genomic instability and lead to malignant transformation (Hustedt and Durocher, 2016; Pommier et al., 2016; Wilson and Durocher, 2017). Hence, defective DSB repair molecules have been associated with a variety of human syndromes including cancers, immunodeficiency and neurological disorders, highlighting the physiological significance of the integrity of the DSB repair network for human health (Canugovi et al., 2013; McKinnon and Caldecott, 2007; Rass et al., 2007; Sunada et al., 2018; Terabayashi and Hanada, 2018; Uryga et al., 2016). To properly repair these dangerous genetic lesions, mammalian cells have evolved complex yet highly orchestrated signalling cascades involving various PTMs such as ubiquitylation, methylation and acetylation for context-dependent signal propagation. The NuA4 (Nucleosome acetyltransferase of H4) /TIP60 acetyltransferase complex has been demonstrated to participate in the repair of DNA DSBs at multiple levels, underlining the pleiotropic contributions of acetylation as a PTM to the DNA DSB repair network (Dhar et al., 2017; Jacquet et al., 2016; Paquin and Howlett, 2018; Rossetto et al., 2010; Soria et al., 2012; Tang et al., 2013).

Two major pathways exist to repair DNA DSBs in mammalian cells, namely homologous recombination (HR) and non-homologous end joining (NHEJ) (Her and Bunting, 2018; Sunada et al., 2018). The choice between these two canonical DSB repair pathways frequently determines the cellular fate post DNA damage and can have significant implications for related

human syndromes as well as therapeutic strategies (Chapman et al., 2012; Liu et al., 2016; Sung, 2018). The recruitment and retention of 53BP1 has been suggested to potentiate NHEJ-mediated repair of DSBs primarily by safeguarding DSB ends from DNA end resection, a key requirement to initiate HR-mediated repair (Gupta et al., 2014; Liu and Huang, 2016). While 53BP1 can be recruited to the DSB sites in part via binding to the H4K20me mark, RNF168 mediated mono-ubiquitination on H2AK13/15 is also key to the 53BP1 appearance at the break, leading to the DSB being repaired through the error-prone NHEJ pathway (Fradet-Turcotte et al., 2013; Zimmermann and de Lange, 2014). Another H4K20me reader protein, MBTD1 (MBT domain-containing protein 1), has recently been demonstrated to be a stable subunit of the NuA4/TIP60 acetyltransferase complex (Jacquet et al., 2016). Notably, the MBTD1 knockout mice suffer skeletal malformations, impaired fetal liver hematopoiesis and premature death, highlighting the indispensable role of MBTD1 in mammalian physiology (Honda et al., 2011). Interestingly, MBTD1 affects the recruitment of the NuA4/TIP60 complex to a specific subset of genes to regulate their expressions (Jacquet et al., 2016). Furthermore, NuA4/TIP60-associated MBTD1 binding to the H4K20me mark at DSB sites competes with 53BP1 for association with the chromatin surrounding the lesion (Jacquet et al., 2016). Moreover, the MBTD1-linked Tip60 (Tat interacting protein 60) enzyme catalyzes acetylation of not only the histone H4 tail, destabilizing 53BP1 interaction (Tang et al., 2013), but also H2AK15 to prevent its ubiquitylation by RNF168, and thereby further blocking the recruitment and retention of 53BP1 to the DSB sites (Jacquet et al., 2016). Collectively, MBTD1 via its direct interaction with the EPC1 (Enhancer of polycomb homolog 1) subunit of the NuA4/TIP60 acetyltransferase complex negatively regulates 53BP1's bivalent associations with chromatin surrounding DSB sites through competition for binding to H4K20me and histone lysine acetylation (Jacquet et al., 2016). Therefore, this MBTD1-EPC1 interaction not only regulates certain gene expressions, but

also enables the recruitment of the DNA end resection machinery to the DSB site and commits the cellular response to the error-free HR pathway.

Consistent with the recently discovered pivotal role of MBTD1 in influencing DSB repair pathway choice, clinical studies have detected recurrent chromosomal translocations that result in the observed generations of MBTD1-associated oncogenic fusion proteins including ZMYND11-MBTD1 and MBTD1-CXorf67 fusions in acute myeloid leukemia and endometrial stromal sarcomas, respectively (Brunetti et al., 2018; De Braekeleer et al., 2014; de Rooij et al., 2016; Dewaele et al., 2014; Yamamoto et al., 2018). These clinical observations highlight the important functions of MBTD1 in gene regulation and DNA damage response, to maintain human genome stability and cancer suppression. MBTD1 consists of a N-terminal FCS zinc finger that binds to regulatory RNAs and four malignant brain tumor (MBT) domains that recognize H4K20me1/me2 mark (Eryilmaz et al., 2009; Lechtenberg et al., 2009). The first and second MBT domains are reported to accommodate transcription factor YY1 in vitro (Alfieri et al., 2013), whereas the fourth MBT domain harboring the semi-aromatic cage is responsible for histone mark readout (Eryilmaz et al., 2009). EPC1 contains EPcA, EpcB and EPcC domains at the N-terminus (Kee et al., 2007). The C-terminus of EPC1 was reported to interact with transcriptional repressor RFP (Tezel et al., 2002). The precise molecular mechanism by which part of MBTD1 associates with the EPC1 remains elusive. Through structural analysis coupled with biophysical and functional analysis, we have elucidated the interaction mechanism between MBTD1 and the EPC1-containing NuA4/TIP60 acetyltransferase complex. Hence, these data will provide valuable information for the development of inhibitors for dissociating TIP60 acetyltransferase from the DSB sites, and thereby will enable the manipulation of DSB repair pathway choice that can have clinical implications.

Results and Discussion

Characterization of the key EPC1 region responsible for recruiting MBTD1

MBTD1 has been recently revealed to associate with the NuA4/TIP60 acetyltransferase complex via interaction with the C-terminus of EPC1, a non-catalytic subunit of the core complex (Jacquet et al., 2016). The deletion of the last 255 residues of EPC1 leads to specific loss of MBTD1 from the NuA4/TIP60 acetyltransferase complex without affecting the association of the other subunits (Jacquet et al., 2016). To investigate the detailed molecular mechanism behind the MBTD1-mediated recruitment of the NuA4/TIP60 acetyltransferase complex to chromatin, we first aimed to narrow down the key region on the C-terminus of EPC1 (**Fig. 1A**). To map the minimal MBTD1-binding site on EPC1, we performed *in vitro* GST pull-down experiments using various purified GST-tagged EPC1 protein fragments and His-tagged MBTD1 proteins. Notably, GST-EPC1$^{644\text{-}799}$ efficiently bound MBTD1, whereas GST-EPC1$^{729\text{-}799}$ and GST-EPC1$^{673\text{-}776}$ lacking a hydrophobic region did not (**Fig. 1A and 1B**), suggesting that this hydrophobic fragment (residues 644-672) of EPC1 is probably responsible for binding of MBTD1. Interestingly, biophysical analysis by Isothermal Titration Calorimetry (ITC) confirmed that EPC1$^{644\text{-}672}$ directly interacts with MBTD1 with a dissociation constant (K_d) of ~67 nM (**Fig. 1C**), indicating that the hydrophobic C-terminal fragment of EPC1 (residues 644-672) enlists MBTD1 into the dynamic NuA4/TIP60 acetyltransferase complex (**Fig. 1A**).

Overall structure of the MBTD1-EPC1 complex

In order to elucidate how the NuA4/TIP60 acetyltransferase complex engages MBTD1 at the atomic level, we sought to determine the crystal structure of MBTD1 in complex with EPC1. Unfortunately, co-crystallization trials of purified MBTD1$^{140–562}$ and EPC1$^{644\text{-}672}$ proteins were not successful. Thus, we generated a fusion construct in which EPC1$^{644\text{-}672}$ was C-terminally

linked to MBTD1$^{140–562}$ using a (GSA)$_6$ linker. Remarkably, the MBTD1-EPC1 fusion protein yielded high-quality crystals that diffracted to a resolution of 1.9 Å and enabled us to solve the complex structure (**Table S1**). The well-defined electron density allowed us to build residues Lys649-Thr665 of EPC1 into the structural model with all four MBT domains (MBT1-4) of MBTD1 clearly visible in the complex structure (**Fig. 2A**). Intriguingly, the MBTD1-interacting region of EPC1 folds into a helical structure that locks into a hydrophobic binding groove formed by the MBT1 and MBT2 repeats of MBTD1 (**Fig. 2B**), indicating that hydrophobic intermolecular interactions dictate the recruitment of MBTD1 into the NuA4/TIP60 acetyltransferase complex. Specifically, two helices from MBT1 and MBT2 form a V-shaped groove (hereafter named: helical groove) to anchor EPC1. The overall structure of MBTD1 in the complex is highly similar to the previously determined apo structure with an RMSD value of approximately 0.36 Å (**Fig. S1A**) (Eryilmaz et al., 2009). However, subtle structural divergences are observed in the hydrophobic helical groove of MBTD1 between the apo and MBTD1-EPC1 complex structures. The binding of EPC1 triggers a mild yet significant displacement of the two helices of the hydrophobic helical groove of MBTD1 compared with the apo structure, especially residues I229-V243, resulting in a relatively more open hydrophobic binding groove upon EPC1 lock-in (**Fig. S1B**). Notably, the residues of the hydrophobic helical groove in the apo structure of MBTD1 have a high B-factor value, whereas they are well ordered in the complex structure, indicating that the EPC1-binding event can stabilize the hydrophobic helical groove of MBTD1 (**Fig. S1C**). Given the facts that 1. only the MBT4 repeat of MBTD1 contains the semi-aromatic cage to recognize H4K20me (Eryilmaz et al., 2009), 2. EPC1-binding does not involve the MBT4 repeat, 3. this binding event does not induce any significant conformational change on the MBT4 repeat structure, it is unlikely that the MBTD1-EPC1 interaction would affect MBTD1's binding ability towards H4K20me (**Fig. 2A**). In accordance with these structural insights, our

biophysical analysis by ITC demonstrated that MBTD1 displayed similar binding affinities towards H4K20me1 in the presence or absence of EPC1 (**Fig. 2C**). Collectively, the above data suggest that the recruitment of MBTD1 into the NuA4/TIP60 acetyltransferase complex does not impair H4K20me-binding, thereby conferring this histone mark binding capability to the NuA4/TIP60 acetyltransferase complex and influencing transcription and DNA repair pathway choice.

Recognition mechanism of EPC1 by MBTD1

Analysis of the complex structure revealed that the interaction of EPC1 with MBTD1 buries a total surface area of 723 Å^2. In the complex structure, EPC1 exhibits an extended conformation and is sandwiched between two helices of MBTD1 (**Fig. 2B**). Two hydrophobic leucine residues (Leu651 and Leu663) of EPC1 fit snugly into two separate hydrophobic pockets in MBTD1, respectively (**Fig. 3A**). Specifically, Leu663 located at the C-terminus of EPC1 occupies the hydrophobic cleft formed by Trp193, Pro244, Trp255 and Leu259 of MBTD1 (**Fig. 3A**). Leu651 from the N-terminus of EPC1 forms hydrophobic interactions with Val232 and Leu263 of MBTD1. Similarly, at the center of the binding interface, Ala660 is accommodated by hydrophobic patch created by Leu259, Val260 and Leu263 of MBTD1. Ala656, Phe658and Ala662 of EPC1 also made hydrophobic interactions with their respective neighboring residues of MBTD1 (**Fig. 3A**). In addition to the hydrophobic contacts, hydrogen bonds mediate the intermolecular interactions of Leu651, Ala662 and Val664 of EPC1 with MBTD1, further strengthening their association (**Fig. 3A**). Intriguingly, these residues involved in the binding of EPC1 with MBTD1 are largely conserved across different species (**Fig. S2**), suggesting an evolutionarily conserved recognition mechanism of EPC1 by MBTD1. Overall, the above data

indicate that hydrophobic intermolecular interactions act as the major driving forces to enlist MBTD1 via EPC1 in the dynamic NuA4/TIP60 complex.

Given that the transcription factor YY1 has been previously demonstrated to directly bind to MBTD1 *in vitro* (Alfieri et al., 2013), we sought to compare the recognition mechanisms of YY1 and EPC1 by MBTD1. Interestingly, MBTD1 employs the same helical groove to capture both EPC1 and YY1, suggesting a conserved scaffold role of the helical groove in the assembly of MBTD1-associated protein complexes (**Fig. S1C**). Structural comparison analysis indicates that whereas EPC1 folds into a short helix in the MBTD1-EPC1 complex, YY1 adopts a two-stranded β-sheet in the MBTD1-YY1 complex. Several EPC1-binding residues of MBTD1 are not involved in binding YY1, leading to relatively weaker intermolecular interactions (**Fig. S1D**). Consistently, a much weaker binding affinity has been reported between YY1 and MBTD1 in the micro-molar affinity range as supported by the observation of high B-factors of the helical groove residues in the MBTD1-YY1 complex (**Fig. S1C**), and no detectable binding is observed between YY1 and MBTD1 *in vivo* (Alfieri et al., 2013; Jacquet et al., 2016).

Mutational analysis of MBTD1-EPC1 interaction

To validate the importance of the key residues involved in the MBTD1-EPC1 interaction, we introduced point mutations of some key residues into MBTD1 and EPC1 and tested their interactions by ITC. Notably, Leu651, Ala660 andLeu663 of EPC1 involve the hydrophobic pockets on MBTD1, suggesting an important role of these residues in the complex formation. Consistently, the triple mutant EPC1[L651D/A660/L663D] abolished MBTD1-binding (**Fig. 3B**). Similarly, the aspartic acid substitutions of the key residues of MBTD1 (A236D/L259D/L263D), identified from the MBTD1-EPC1 complex structure, impaired the MBTD1-EPC1 association (**Fig. 3B**). Furthermore, GST pull-down assays showed that the aspartic acid substitutions of

some other key residues in the MBTD1-EPC1 interface also significantly dampened the complex formation (Fig. S3A and S3B). Additionally, as the side-chain of Val664 of EPC1 is not involved in the interaction with MBTD1, the EPC1^{V664D} mutant displayed a comparable binding ability with MBTD1 to the wild type (WT) EPC1 (**Fig. S3A**). Collectively, the mutagenesis analysis confirmed the importance of the key residues involved in the hydrophobic interactions between MBTD1 and EPC1 in leashing MBTD1 to the NuA4/TIP60 complex.

To further examine the impacts of the key EPC1 mutations on their abilities to associate with MBTD1 *in vivo*, we generated isogenic K562 cell lines with a single copy of the epitope-tagged EPC1 gene inserted at the *AAVS1* safe harbor locus as described previously (Jacquet et al., 2016). As expected, a C-terminal deletion mutant of EPC1 (EPC1$^{1-464}$) completely lost its ability to associate with MBTD1 in cells (**Fig. 3C**), further confirming that the C-terminal portion of EPC1 is responsible for enlisting MBTD1 into the NuA4/TIP60 acetyltransferase complex *in vivo*. In accordance with the above ITC results, the triple mutant EPC1$^{L651D/A660D/L663D}$ showed complete loss of co-purification with MBTD1 but retained association with Tip60 *in vivo* (**Fig. 3C**). Similar experiments with K562 cell lines expressing epitope-tagged MBTD1 confirmed that residues Ala236, Leu259, and Leu263 of MBTD1 are fundamental to MBTD1's ability to associate with the NuA4/TIP60 complex *in vivo* (**Fig. 3C**). Thus, the above results validated the physiological significance of these key residues of MBTD1 and EPC1 for their association in human cell lines.

Functional implications of uncoupling MBTD1 from the NuA4/TIP60 acetyltransferase complex

MBTD1 has been found to co-localize with other NuA4/TIP60 complex subunits at the transcription start sites of active genes, affecting their recruitment and thereby regulating gene

transcription (Jacquet et al., 2016). We next examined the functional significance of the interaction between MBTD1 and EPC1 on gene transcription. The MBTD1 or EPC1 mutants that impair the complex formation exhibited significantly reduced binding to the *RPSA* gene promoter, a *bone fide* target of the NuA4 /TIP60 complex, as measured by chromatin immunoprecipitation (ChIP) assay (**Fig. 4A**). While a portion of EPC1 is retained, the MBTD1 recruitment is completely lost, in agreement with its dissociation from the complex. Furthermore, MBTD1 and EPC1 have been suggested to function as putative tumor suppressors (Jacquet et al., 2016). Clonogenic survival analysis revealed that transfected cells expressing EPC1 or MBTD1 mutants displayed fewer colony growth compared to the empty vector control (**Fig. 4B** and **S4**). Notably, significantly more colonies were recovered with the MBTD1 mutant versus its WT version, indicating that MBTD1-dependent growth suppression requires interaction with the NuA4/TIP60 complex at least in part. Collectively, these results confirm the physiological significance of the key residues identified in the complex structure of MBTD1-EPC1, highlighting that the hydrophobic interactions between EPC1 and MBTD1 determine MBTD1's functional association with the NuA4/TIP60 acetyltransferase complex in mammalian cells.

The MBTD1-associated NuA4/TIP60 complex recognizes the H4K20me mark at DSB sites and catalyzes the acetylation of the histone H4 tail and H2AK15 surrounding the sites of broken chromatins (Tang et al., 2013; Jacquet et al., 2016). These DNA damage-signaling events act in concert to favor downstream recruitment of the HR-mediated repair machinery as the choice of DSB repair pathways, blocking at multiple levels the recruitment and retention of the key NHEJ factor 53BP1 to the DSB sites (Jacquet et al., 2016). Given that we have identified the key residues of MBTD1 responsible for its enlistment into the NuA4/TIP60 complex, we aimed to further validate these key residues' functional importance in mammalian cells. It was shown

that overexpression of WT MBTD1 led to mild increase of H2AK15ac signal in MBTD1 knockout cells, whereas overexpression of the MBTD1 mutant failed to produce this effect (**Fig. S5A**). Furthermore, post-gamma irradiation clonogenic survival analysis demonstrated that complementation of MBTD1 knockout cells (2 different clones (Jacquet et al. 2016)) by transfection of a WT MBTD1 expression vector rescues the radiosensitivity of MBTD1 knockout cells (**Fig. 4C**). In contrast, complementation of the MBTD1 knockout cells with the MBTD1 mutant failed to rescue the radiosensitivity of MBTD1 knockout cells (**Fig. 4C**). Interestingly, ectopic expression of MBTD1 seems to also protect WT cells (containing endogenous MBTD1 protein), while the mutant again fails to do so. These results highlight the functional significance of the MBTD1-EPC1 interactions in repairing damaged chromatin in mammalian cells.

Since MBT1 is important for proper expression of specific genes, directly bound by the NuA4/TIP60 complex or indirectly, we next investigated whether WT MBTD1 but not the mutant lacking the EPC1 binding ability rescues the gene regulation defects in MBTD1 knockout cells. The expression levels of *HIST1H4C* and *CCDN3* have been shown to be dysregulated in opposite manners in the MBTD1 knockout cells, *HIST1H4C* being overexpressed while *CCDN3* being underexpressed (Jacquet et al., 2016). We measured the expression levels of these two genes in the MBTD1 knockout cells transfected with empty vector, WT MBTD1 or mutant vectors by RT-qPCR. The *HOXB2* expression is measured as a negative control as it is not affected by the loss of MBTD1. Complementation of the MBTD1 knockout cells with a WT MBTD1 vector rescued their transcription defects, significantly decreasing the expression of *HIST1H4C* while increasing the expression of *CCDN3* (**Fig. 4D**). In contrast, complementation of the MBTD1 knockout cells with the MBTD1 mutant failed to rescue the misregulation of *HIST1H4C* and *CCDN3*, while being expressed at similar levels compared to the WT vector (**Fig.**

4D and **S5B**). Thus, the above findings validated the key role of the MBTD1-EPC1 interactions in NuA4/TIP60 acetyltransferase complex-mediated transcriptional regulation.

In brief, our study has elucidated the molecular mechanism by which MBTD1 enrolls into the dynamic NuA4/TIP60 acetyltransferase complex to regulate transcription and DNA repair pathway choice. These findings will provide valuable information for the development of inhibitors for selectively dissociating MBTD1 from the NuA4/TIP60 acetyltransferase complex, which can have important clinical implications. Drug-induced disruption of the MBTD1 interaction with NuA4/TIP60 could be used to sensitize cancer cells to radiotherapy, alone or in combination with PARP inhibitors. This approach may be better than directly inhibiting Tip60 HAT activity which can lead to much larger side effects on other/healthy cells

Acknowledgements

We are grateful to Amelie Fradet-Turcotte for important advice on radiosensitivity clonogenic assays and the use of their irradiator. Diffraction data were collected from Structural Biology Center (SBC)-CAT beam lines at the Advanced Photon Source of the Argonne National Laboratory. SBC-CAT is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada through Ontario Genomics Institute [OGI-055], Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck KGaA, Darmstadt, Germany, MSD, Novartis Pharma AG, Ontario Ministry of Research, Innovation and Science (MRIS), Pfizer, S ão Paulo Research Foundation-FAPESP, Takeda, and Wellcome. This

work was also supported by the Canadian Institutes of Health Research (CIHR) (FDN-143314 to J.C.) and the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number RGPIN-2016-06300 to J.M. J.C. holds the Canada Research Chair in Chromatin Biology and Molecular Epigenetics.

AUTHOR CONTRIBUTIONS

H.Z., M.D., J.C. and J.M. designed the research. H.Z., M.D., G.A., H.C., W.T., C.D. and P.L. performed experiments. H.Z., M.D., L.L., J.C. and J.M. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

METHOD DETAILS

Protein expression and purification

The MBTD1$^{140–562}$-(GSA)$_6$-EPC1$^{644–672}$ fusion protein construct was cloned into the pET28-MHL vector. Recombinant proteins were expressed in *E. coli* strain BL21 (DE3) in TB medium overnight at 16 °C. Cells were collected by centrifugation at 4 °C and resuspended in 25 mM Tris–HCl pH 8.0, 500 mM NaCl, 5 mM β-mercaptoethanol, 1 mM PMSF. Subsequently, cells were lysed by sonication and debris was removed by centrifugation at 4 °C. The supernatants were loaded onto a Ni-NTA resin (QIAGEN) gravity column and eluted with buffer containing 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 250 mM imidazole, 5 mM β-mercaptoethanol. The proteins were further purified by anion-exchange chromatography and subsequent gel-filtration chromatography in a buffer containing 25 mM Tris-HCl pH 8.0, 150 mM NaCl. Peak fractions were concentrated to 16 mg/mL.

Crystallization, data collection and structure determination

Crystal screening was performed by sitting-drop vapor-diffusion at 18 °C. Crystals of MBTD1-EPC1 were obtained using a reservoir solution composed of 0.1 M sodium citrate tribasic dihydrate pH 5.0, 10% w/v polyethylene glycol 6,000. The crystals were soaked in a cryoprotectant of crystallization solution supplemented with 20% glycerol before being flash-frozen. Diffraction data were collected at beam line 19-ID of the Advanced Photon Source and reduced with XDS (Kabsch, 2010) and AIMLESS (Evans and Murshudov, 2013). The structure was solved by molecular replacement with coordinates from PDB entry 3FEO (Eryilmaz et al., 2009). The model was iteratively refined with REFMAC (Murshudov et al., 2011) and rebuilt with COOT (Emsley et al., 2010). Model geometry was evaluated with MOLPROBITY (Chen et al., 2010).

GST pull-down

EPC1 fragments with the indicated boundaries were cloned into the pET28GST-LIC vector that encodes an N-terminal GST-tag. All point mutations were generated using the Quick Change Site–directed mutagenesis system (Stratagene, La Jolla, CA). All GST-tagged proteins were expressed in BL21(DE3) cells and purified by GST affinity. Proteins were then dialyzed overnight against dialysis buffer (25 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM β-mercaptoethanol). The GST-tagged proteins were first incubated with glutathione resin for 1 hour at 4 ℃ in dialysis buffer. His-tagged MBTD1 proteins were incubated with the beads for 30 min at 4 ℃. The samples were washed four times with 1 mL dialysis buffer to remove excess unbound protein, and then analyzed through SDS-PAGE.

Isothermal titration calorimetry

Proteins were expressed and purified as described above. All ITC experiments were performed on a Microcal VP-ITC (Microcal, Amherst, MA) at 25 °C. All measurements were carried out at a stirring speed of 307 rpm, in 25 mM Tris-HCl, pH 8.0, 150 mM NaCl. The ITC experiments were repeated at least two times. The binding data were subsequently analyzed using Origin (MicroCal).

Generation of isogenic cell lines and affinity purification of complexes

K562 cells were used to express various 3xFlag-tagged EPC1 proteins including EPC1^{WT}, EPC1$^{1-464}$, EPC1$^{L651D/L663D}$, and EPC1$^{L651D/A660D/L663D}$ as well as 3xFlag-tagged MBTD1^{WT} and MBTD1$^{A236D/L259D/L263D}$ from the AAVS1 safe harbor using ZFN-mediated insertion as described previously (Dalvai et al., 2015). After whole cell extraction, anti-Flag immunoprecipitation was performed with anti-FLAG agarose affinity gel (Sigma M2), followed by elution with 3xFLAG peptide (200 µg/ml from Sigma in the following buffer: 20 mM HEPES pH7.9, 100 mM KCl,

0.1% Triton X-100, 20% glycerol, 100 μM ZnCl2, 1 mM DTT and supplemented with proteases, deacetylases and phosphatases inhibitors).

Flag eluates from the above preparations were normalized following SDS-PAGE and transfer onto a nitrocellulose membrane. The following antibodies were used at the indicated dilution: anti-FLAG-HRP conjugate (Sigma M2, 1:5000), anti-Tip60 (Abcam 137518, 1:1000); anti-MBTD1 (Abcam ab116361, 1:1000); anti-GAPDH (ThermoFisher, 39-8600, 1:10000).

ChIP-qPCR Assay

For anti-FLAG ChIP, 1 mg of cross-linked chromatin from K562 cells was incubated with 10 μg of anti-FLAG antibody (Sigma, M2) or anti-IgG antibody (Millipore- PP64) pre-bound on 300 μl of Dynabeads Prot-G (Invitrogen) overnight at 4 ℃. The beads were washed extensively and eluted in 0.1% SDS, 0.1 M NaHCO3. Crosslink was reversed with 0.2 M NaCl and incubation overnight at 65 ℃. Samples were treated with RNase and Proteinase K for 2 h and recovered by phenol chloroform extraction and ethanol precipitation. Quantitative real-time PCRs were performed on a LightCycler 480 (Roche) with SYBR Green I (Roche) to confirm specific enrichment at the RPSA locus. Ratios of immunoprecipitated signal versus input are based on duplicate experiments and errors bars represent the range. Primers used: AGAAAGCGGGCTAACATCCT; GCTGTGCTGTCACCACTTGT.

Clonogenic growth assay

EPC1^{WT}, EPC1$^{L651D/A660D/L663D}$, MBTD1^{WT}, MBTD1$^{A236D/L259D/L263D}$ or empty vectors were transfected into U2OS cells which were then seeded onto dishes in duplicates. Subsequently, puromycin [1 μg/ μl] was added to the culture media at 48hours post-transfection for selection. After ten days of culture, the colonies formed on the dishes were fixed with methanol, followed

by staining with crystal violet (Sigma C6158) and washing with distilled water. Images of the dishes were then taken.

Gamma-irradiation and clonogenic survival assay

MBTD1-3xFlag, MBTD1-A236D/L259D/L263D-3xFlag mutant or an empty vector were transfected into U2OS cells or MBTD1 KO cell lines previously described (Jacquet et al. 2016). Cells were then seeded in 10cm dishes. Puromycin was added to the culture media [1 μg/ml] 48hrs post-transfection for two days of selection. One day before irradiation, the cells were counted to plate 500 cells per condition in 6-well plates. Irradiation with 0 or 0.5 gray was performed using a Cellrad Faxitron irradiator. After 12 days of culture, the colonies formed were fixed with methanol, followed by staining with crystal violet (Sigma C6158) and washes with distilled water. Images were then taken, and colonies counted.

Western blot analysis of whole cell extracts from the above transfections were done to verify equivalent expression of WT and mutant proteins. The following antibodies were used at the indicated dilution: anti-FLAG-HRP conjugate (Sigma M2, 1:5000) and anti-GAPDH (ThermoFisher, 39-8600, 1:10000). The error bars represent the range based on independent experiments.

Chromatin extracts

Chromatin-enriched extracts were prepared as described before (Gwak et al., 2016). Briefly, U2OS cells were transfected with WT MBTD1-3xFlag, MBTD1-A236D/L259D/L263D-3xFlag mutant or an empty vector (4 μg) using Lipofectamine 2000. 2 days post-transfection, puromycin was added [1μg/ml]. 48hrs later cells were harvested, washed once with PBS 1X and resuspended in buffer 1 (50mM Tris-HCl pH7.5, 100mM NaCl, 1mM EDTA, 1mM DTT, 1X protease inhibitor cocktail (Sigma- P8340), 1mM PMSF, 20mM N-ethylmaleimide, 10mM NaB,

10mM NaF). Cell suspension was incubated 5 min on ice, then centrifuged at 1000g for 15 minutes. Supernatant was discarded. Nuclei were then resuspended in buffer 2 (50mM Tris-HCl pH7.5, 300mM NaCl, 1mM EDTA, 5mM $CaCl_2$, 1mM DTT, 1X protease inhibitor cocktail, 1mM PMSF), incubated 10 min on ice and sonicated 3x30 seconds. Chromatin enriched extracts were finally clarified by centrifugation at 1000g for 20 min. The following antibodies were used for Western blotting at the indicated dilution: anti-H2AK15ac (Abcam ab101447, 1:1000); anti-H4penta-Acetyl (Upstate 06-946, 1:3000); anti-H3 (Abcam 1791, 1:10 000); anti-H4 (Abcam 7311, 1:5000).

Reverse Transcription-qPCR

WT MBTD1-3xFlag, MBTD1-A236D/L259D/L263D-3xFlag mutant or an empty vector were transfected in a MBTD1 KO U2OS cell line (#15 (Jacquet et al. 2016). Puromycin [1 µg/ml] was added 48hrs post-transfection and 48hrs later total RNA was extracted with the RNeasy Plus Mini kit (Qiagen). 500 ng of RNA was reverse transcribed by oligo-dT and random priming into cDNA with a qScript cDNA SuperMix kit (QuantaBio-VWR), according to the manufacturer's instructions.

Quantification of the amount of cDNA was done with SYBR Green I (Roche) on a LightCycler 480 (Roche) for real-time PCR, using primers specific for *HIST1H4C* and *CCDN3* coding regions, two genes that are affected in MBTD1 KO cells (Jacquet et al. 2016). The *HOXB2* gene was used as negative control as it is not affected in MBTD1 KO cells. The *36B4* gene was used as a housekeeping internal control for normalisation. The error bars represent the range based on independent experiments. The oligonucleotide sequences used for expression analysis by RT-qPCR are listed below:

Primer name	For	Rev

36B4	CGACCTGGAAGTCCAACTAC	ATCTGCTGCATCTGCTTG
HIST1H4C	ATCCAGGGCATTACAAAACCG	AACCTTAAGCACACCTCGAGT
CCDN3	GGAAGATGCTGGCTTACTGGA	GACAGGTAGCGATCCAGGTAG
HOXB2	TAATAAGTACCTGTGCCGGCC	CTTCATGCGCCGGTTCTGAAA

Reference List

Alfieri, C., Gambetta, M.C., Matos, R., Glatt, S., Sehr, P., Fraterman, S., Wilm, M., Muller, J., and Muller, C.W. (2013). Structural basis for targeting the chromatin repressor Sfmbt to Polycomb response elements. Genes Dev *27*, 2367-2379.

Brunetti, M., Gorunova, L., Davidson, B., Heim, S., Panagopoulos, I., and Micci, F. (2018). Identification of an EPC2-PHF1 fusion transcript in low-grade endometrial stromal sarcoma. Oncotarget *9*, 19203-19208.

Canugovi, C., Misiak, M., Ferrarelli, L.K., Croteau, D.L., and Bohr, V.A. (2013). The role of DNA repair in brain related disease pathology. DNA Repair (Amst) *12*, 578-587.

Chapman, J.R., Taylor, M.R., and Boulton, S.J. (2012). Playing the end game: DNA double-strand break repair pathway choice. Mol Cell *47*, 497-510.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallographica Section D: Biological Crystallography *66*, 12-21.

Chin, J.W. (2017). Expanding and reprogramming the genetic code. Nature *550*, 53-60.

Dalvai, M., Loehr, J., Jacquet, K., Huard, C.C., Roques, C., Herst, P., Cote, J., and Doyon, Y. (2015). A Scalable Genome-Editing-Based Approach for Mapping Multiprotein Complexes in Human Cells. Cell Rep *13*, 621-633.

Dantuma, N.P., and van Attikum, H. (2016). Spatiotemporal regulation of posttranslational modifications in the DNA damage response. EMBO J *35*, 6-23.

De Braekeleer, E., Auffret, R., Douet-Guilbert, N., Basinko, A., Le Bris, M.J., Morel, F., and De Braekeleer, M. (2014). Recurrent translocation (10;17)(p15;q21) in acute poorly differentiated myeloid leukemia likely results in ZMYND11-MBTD1 fusion. Leuk Lymphoma *55*, 1189-1190.

de Rooij, J.D., van den Heuvel-Eibrink, M.M., Kollen, W.J., Sonneveld, E., Kaspers, G.J., Beverloo, H.B., Fornerod, M., Pieters, R., and Zwaan, C.M. (2016). Recurrent translocation t(10;17)(p15;q21) in minimally differentiated acute myeloid leukemia results in ZMYND11/MBTD1 fusion. Genes Chromosomes Cancer *55*, 237-241.

Dewaele, B., Przybyl, J., Quattrone, A., Finalet Ferreiro, J., Vanspauwen, V., Geerdens, E., Gianfelici, V., Kalender, Z., Wozniak, A., Moerman, P*., et al.* (2014). Identification of a novel, recurrent MBTD1-CXorf67 fusion in low-grade endometrial stromal sarcoma. Int J Cancer *134*, 1112-1122.

Dhar, S., Gursoy-Yuzugullu, O., Parasuram, R., and Price, B.D. (2017). The tale of a tail: histone H4 acetylation and the repair of DNA breaks. Philos Trans R Soc Lond B Biol Sci *372*.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallographica Section D: Biological Crystallography *66*, 486-501.

Eryilmaz, J., Pan, P., Amaya, M.F., Allali-Hassani, A., Dong, A., Adams-Cioaba, M.A., Mackenzie, F., Vedadi, M., and Min, J. (2009). Structural studies of a four-MBT repeat protein MBTD1. PLoS One *4*, e7274.

Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? Acta Crystallographica Section D: Biological Crystallography *69*, 1204-1214.

Fontana, G.A., Reinert, J.K., Thoma, N.H., and Rass, U. (2018). Shepherding DNA ends: Rif1 protects telomeres and chromosome breaks. Microb Cell *5*, 327-343.

Fradet-Turcotte, A., Canny, M.D., Escribano-Diaz, C., Orthwein, A., Leung, C.C.Y., Huang, H., Landry, M.C., Kitevski-LeBlanc, J., Noordermeer, S.M., Sicheri, F.*, et al.* (2013). 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark. Nature *499*, 50-+.

Gupta, A., Hunt, C.R., Chakraborty, S., Pandita, R.K., Yordy, J., Ramnarain, D.B., Horikoshi, N., and Pandita, T.K. (2014). Role of 53BP1 in the regulation of DNA double-strand break repair pathway choice. Radiat Res *181*, 1-8.

Gwak, J., Shin, J.Y., Lee, K., Hong, S.K., Oh, S., Goh, S.H., Kim, W.S., and Ju, B.G. (2016). SFMBT2 (Scm-like with four mbt domains 2) negatively regulates cell migration and invasion in prostate cancer cells. Oncotarget *7*, 48250-48264.

Her, J., and Bunting, S.F. (2018). How cells ensure correct repair of DNA double-strand breaks. J Biol Chem *293*, 10502-10511.

Honda, H., Takubo, K., Oda, H., Kosaki, K., Tazaki, T., Yamasaki, N., Miyazaki, K., Moore, K.A., Honda, Z., Suda, T.*, et al.* (2011). Hemp, an mbt domain-containing protein, plays essential roles in hematopoietic stem cell function and skeletal formation. Proc Natl Acad Sci U S A *108*, 2468-2473.

Hustedt, N., and Durocher, D. (2016). The control of DNA repair by the cell cycle. Nat Cell Biol *19*, 1-9.

Jacquet, K., Fradet-Turcotte, A., Avvakumov, N., Lambert, J.P., Roques, C., Pandita, R.K., Paquet, E., Herst, P., Gingras, A.C., Pandita, T.K.*, et al.* (2016). The TIP60 Complex Regulates Bivalent Chromatin Recognition by 53BP1 through Direct H4K20me Binding and H2AK15 Acetylation. Mol Cell *62*, 409-421.

Kabsch, W. (2010). Xds. Acta Crystallographica Section D: Biological Crystallography *66*, 125-132.

Kee, H.J., Kim, J.-R., Nam, K.-I., Park, H.Y., Shin, S., Kim, J.C., Shimono, Y., Takahashi, M., Jeong, M.H., and Kim, N. (2007). Enhancer of polycomb1, a novel homeodomain only protein-binding partner, induces skeletal muscle differentiation. Journal of biological chemistry *282*, 7700-7709.

Lechtenberg, B.C., Allen, M.D., Rutherford, T.J., Freund, S.M., and Bycroft, M. (2009). Solution structure of the FCS zinc finger domain of the human polycomb group protein L (3) mbt‑like 2. Protein Science *18*, 657-661.

Liu, M., Wang, H., Lee, S., Liu, B., Dong, L., and Wang, Y. (2016). DNA repair pathway choice at various conditions immediately post irradiation. Int J Radiat Biol *92*, 819-822.

Liu, T., and Huang, J. (2016). DNA End Resection: Facts and Mechanisms. Genomics Proteomics Bioinformatics *14*, 126-130.

McKinnon, P.J., and Caldecott, K.W. (2007). DNA strand break repair and human genetic disease. Annu Rev Genomics Hum Genet *8*, 37-55.

Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallographica Section D: Biological Crystallography *67*, 355-367.

Paquin, K.L., and Howlett, N.G. (2018). Understanding the Histone DNA Repair Code: H4K20me2 Makes Its Mark. Mol Cancer Res.

Pommier, Y., Sun, Y., Huang, S.N., and Nitiss, J.L. (2016). Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. Nat Rev Mol Cell Biol *17*, 703-721.

Rass, U., Ahel, I., and West, S.C. (2007). Defective DNA repair and neurodegenerative disease. Cell *130*, 991-1004.

Rossetto, D., Truman, A.W., Kron, S.J., and Cote, J. (2010). Epigenetic modifications in double-strand break DNA damage signaling and repair. Clin Cancer Res *16*, 4543-4552.

Soria, G., Polo, S.E., and Almouzni, G. (2012). Prime, repair, restore: the active role of chromatin in the DNA damage response. Mol Cell *46*, 722-734.

Sunada, S., Nakanishi, A., and Miki, Y. (2018). Crosstalk of DNA double-strand break repair pathways in poly(ADP-ribose) polymerase inhibitor treatment of breast cancer susceptibility gene 1/2-mutated cancer. Cancer Sci *109*, 893-899.

Sung, P. (2018). Introduction to the Thematic Minireview Series: DNA double-strand break repair and pathway choice. J Biol Chem *293*, 10500-10501.

Tang, J., Cho, N.W., Cui, G., Manion, E.M., Shanbhag, N.M., Botuyan, M.V., Mer, G., and Greenberg, R.A. (2013). Acetylation limits 53BP1 association with damaged chromatin to promote homologous recombination. Nat Struct Mol Biol *20*, 317-325.

Terabayashi, T., and Hanada, K. (2018). Genome instability syndromes caused by impaired DNA repair and aberrant DNA damage responses. Cell Biol Toxicol.

Tezel, G., Shimono, Y., Murakumo, Y., Kawai, K., Fukuda, T., Iwahashi, N., and Takahashi, M. (2002). Role for O-glycosylation of RFP in the interaction with enhancer of polycomb. Biochemical and biophysical research communications *290*, 409-414.

Uryga, A., Gray, K., and Bennett, M. (2016). DNA Damage and Repair in Vascular Disease. Annu Rev Physiol *78*, 45-66.

Wilson, M.D., and Durocher, D. (2017). Reading chromatin signatures after DNA double-strand breaks. Philos Trans R Soc Lond B Biol Sci *372*.

Woodsmith, J., Kamburov, A., and Stelzl, U. (2013). Dual coordination of post translational modifications in human protein networks. PLoS Comput Biol *9*, e1002933.

Woodsmith, J., and Stelzl, U. (2014). Studying post-translational modifications with protein interaction networks. Curr Opin Struct Biol *24*, 34-44.

Yamamoto, K., Yakushijin, K., Ichikawa, H., Kakiuchi, S., Kawamoto, S., Matsumoto, H., Nakamachi, Y., Saegusa, J., Matsuoka, H., and Minami, H. (2018). Expression of a novel ZMYND11/MBTD1 fusion transcript in CD7(+)CD56(+) acute myeloid leukemia with t(10;17)(p15;q21). Leuk Lymphoma, 1-5.

Zimmermann, M., and de Lange, T. (2014). 53BP1: pro choice in DNA repair. Trends Cell Biol *24*, 108-117.

FIGURE LEGENDS

Figure 1. Characterization of the MBTD1-EPC1 interaction.

(A) Schematic domain organization of MBTD1 and EPC1. EPc, Epl1 enhancer of polycomb; HR, hydrophobic region; MBT, malignant brain tumor; ZnF, zinc finger. Interactions between MBTD1 and EPC1 are indicated by a two-way arrow.

(B) Pull-down of MBTD1 by GST-tagged EPC1 fragments. Orange box denote the input protein.

(C) The binding of EPC1 to MBTD1 was measured with isothermal titration calorimetry (ITC). The N value is ~ 1.3. The GST-EPC1 protein was titrated against the MBTD1 protein.

Figure 2. Crystal structure of EPC1 bound to MBTD1.

(A) Overall structure of MBTD1 in complex with EPC1. The color scheme is the same as in Fig.1A.

(B) A close-up view of the electrostatic potential surface of MBTD1 showing the hydrophobic pockets for binding of EPC1.

(C) ITC titration curves of H4K20me1 (sequence: AKRHRKmeVLRDN) to MBTD1 (black curve) or the MBTD1-EPC1 complex (red curve).

Figure 3. Recognition mechanism of EPC1 by MBTD1.

(A) Detailed interactions of MBTD1 with EPC1.

(B) Thermodynamic analysis of the interaction between EPC1 triple mutation and MBTD1 (left), and EPC1 and MBTD1 triple mutation (right).

(C) K562 isogenic cell lines stably expressing 3xFlag tagged-EPC1 or 3xFlag tagged-MBTD1 with the indicated mutations from the AAVS1 locus. Complexes were immunoprecipitated with anti-Flag antibody, eluted with Flag peptides and analyzed by western blotting with the indicated antibodies.

Figure 4. Interaction of MBTD1 and EPC1 impacts gene regulation, DNA damage response and cell growth.

(A) MBTD1 or EPC1 mutants affect recruitment to a target gene *in vivo*. Chromatin immunoprecipitations was performed using K562 cells stably expressing MBTD1 or EPC1 mutant proteins. qPCR for the promoter region of the *RPSA* gene was performed on the recovered DNA and is presented as a ratio versus input signal. Error bars represent the range of two independent experiments.

(B) Clonogenic assay using U2OS cells transfected with the indicated expression vectors or empty vectors. Bar chart depicts percentages of clonogenic growth (CFU, colony forming units) compared to empty vector controls. WT and mutant proteins are ectopically expressed to similar levels (Fig. S4B). Error bars are range of two independent experiments.

(C) WT MBTD1 but not the EPC1-binding mutant suppresses the DNA damage sensitivity of MBTD1 KO cells. Clonogenic survival assay showing the percentages of viable colonies for each condition post a single dose of gamma-irradiation (0.5Gy) relative to their respective untreated controls. WT and MBTD1 KO U2OS cells (Jacquet et al. 2016) were transfected with empty vector, MBTD1-WT or MBTD1-A236D/L259D/L263D and then exposed to 0 or 0.5Gy

of gamma radiation. Cell survival was measured by counting colony forming units (CFU) after staining. Error bars represent the range based on two independent experiments.

(D) WT MBTD1 but not the EPC1-binding mutant complements the gene regulation defects in MBTD1 KO cells. Expression levels of genes that are dysregulated in MBTD1 KO cells. (Jacquet et al. 2016) as measured by RT-qPCR 48h post-transfection after transfection with empty vector, MBTD1-WT or MBTD1-A236D/L259D/L263D. The loss of MBTD1 increases the expression of *HIST1H4C* but decreases the expression of *CCDN3*. *HOXB2* is a negative control as it is not affected by the loss of MBTD1. Bar chart depicts the relative expression of the indicated genes relative to the cells complemented with WT MBTD1. The *36B4* gene was used as a housekeeping internal control for normalization between samples. WT and mutant MBTD1 are expressed to similar levels in this experiment (Fig. S5B). Error bars represent the range based on two independent experiments.

Fig1

A

B

Input   GST Pull-down

C

EPC1(644-672)+MBTD1(140-562)

A

H4K20me1/2 binding pocket

MBT4

MBT1

MBT3

MBT2

EPC1

B

EPC1

C

Fig3

Fig4

A

B

C

D