Molecular Insights into Methyllysine Recognition by the Methyl-CpG-Binding Domain of MeCP2

Jyotirmayee Padhan,[1] and Babu Sudhamalla*,[1]

[1]Department of Biological Sciences, Indian Institute of Science Education and Research Kolkata, Mohanpur Campus, Mohanpur, 741246 Nadia, West Bengal, India.

*Correspondence: s.babu@iiserkol.ac.in

Short title: Insights into MeCP2 Methyllysine Recognition

Keywords: MeCP2, MBD domain, Methyllysine, Histones, H3K27me3

ABSTRACT

Methyl-CpG binding protein 2 (MeCP2) is a reader of DNA methylation and is highly expressed in the central nervous system. Loss of function mutation of MeCP2 which impairs its DNA binding ability causes Rett syndrome, while gain of MeCP2 function causes MeCP2 duplication syndrome. In the context of gene regulation, MeCP2 act as both transcription activator and repressor. Despite over three decades of MeCP2 discovery, the function of this enigmatic protein is largely unknown. Recent study on MeCP2 reported that, in addition to binding DNA, its MBD domain also binds to methylated lysine residues on histone H3 tail. Here we aimed to characterize the methyllysine binding pocket of MeCP2-MBD. Our analysis of MeCP2-MBD crystal structure reveals the presence of an aromatic cage formed by five residues, namely, W104, F132, Y141, F142, and F155, where the trimethyllysine of H3K27me3 binds, as confirmed in docking studies. Mutation of the aforementioned aromatic residues to alanine abolished binding of MeCP2-MBD with its reported ligand H3K27me3 in ITC and pull-down experiments. Additionally, our genomic analysis of publicly available ChIP-seq data revealed that MeCP2 localizes with canonical histone methylation marks like H3K4me3, H3K9me3, H3K27me3 and H4K20me3 in the genome and regulate critical biological and cancer related pathways. Promoters of MeCP2 regulated genes, show signature of H3K4me3 and H3K27me3. In the genomic level, mutation of aromatic cage residue W104A, significantly reduced chromatin occupancy of MeCP2 as demonstrated by ChIP-qPCR experiments. Overall, our study establishes MeCP2 as a histone methylation reader and its histone binding ability is critical for MeCP2 to tether into chromatin.

INTRODUCTION

Methylation of DNA is a heritable epigenetic modification that involves the addition of a methyl group to the cytosine base in DNA.[1] This modification is catalyzed by a group of enzymes known as DNA methyltransferases (DNMTs), which transfer a methyl group from S-adenosylmethionine (SAM) to the C5 position of the cytosine, resulting in the formation of 5-methylcytosine (5mC).[2] The primary site of DNA methylation in mammals is the CpG dinucleotide, which is found in short stretches throughout the genome. These CpG sites often form CpG islands, regions that are typically unmethylated.[2] DNA methylation primarily has been associated with transcriptional repression. For example, transposable elements and viral sequences in the mammalian genome are heavily methylated to silence their expression, which could otherwise lead to genomic instability and mutations.[3, 4]

DNA methylation is recognized by specialized proteins, which are classified into three distinct families: MBD (methyl-CpG binding) proteins, UHRF (ubiquitin-like with PHD and RING finger) proteins, and zinc finger proteins.[5] The MBD proteins were the first to be characterized, with MeCP2 being the founding member of this family. MeCP2 was initially identified in 1992 due to its ability to bind to methylated DNA,[6] and was later found to be implicated in neurological disorder called Rett syndrome.[7, 8] MeCP2 is largely an intrinsically disordered protein, with only two domains exhibiting well-defined secondary structures: the MBD domain and the transcription repression domain (TRD), which are connected by an intermediary region[9, 10] **(**Figure 1A**)**. The core structure of MBD domain folds into a wedge-like shape consisting of two short α-helices and three anti-parallel β-strands which are connected by a loop region (Figure 1B).[11] This domain recognizes the methylated CpG dinucleotide through two conserved symmetrical arginine residues (R111 and R133), forming a stair-shaped motif that interacts with

DNA via hydrogen bonding, cation-π interactions, and nucleotide stacking.[12, 13] Conversely, the TRD of MeCP2 recruits histone deacetylases (HDACs) and Sin3A co-repressor complex to maintain a repressive state of the chromatin.[14, 15] Loss of DNA binding ability of MeCP2 manifests in Rett syndrome, where the critical DNA binding residue R133 is mutated to cystine.[16-18]

MeCP2 protein is abundantly expressed in nervous system and its transcription regulation is tightly controlled.[19-21] Loss of MeCP2 function results in Rett syndrome, while duplication of its locus leads to its overexpression, causing MeCP2 duplication syndrome.[22] Studies have indicated that MeCP2 functions not only as a transcriptional repressor but also plays a role in gene activation, depending on its interacting cofactors. In the mouse hypothalamus, for example, MeCP2 binds to the transcription activator CREB1 at the promoter of active genes.[23] More recently, it was demonstrated that MeCP2 directly interacts with RNA polymerase II (RNA Pol II) at the transcription start site (TSS) of many active genes in human neural cells.[24]

Additionally, various post-translational modifications (PTMs) also regulate MeCP2's function by influencing DNA binding and its interactions with chromatin remodelers. For instance, acetylation of lysine 171 of MeCP2, reduce its affinity for HDAC1 and ATRX.[25] Phosphorylation of MeCP2, triggered by calcium influx into neurons, facilitates the release of MeCP2 from the BDNF promoter.[26, 27] The functional versatility of MeCP2 contributed by PTMs regulate neuronal plasticity and morphology, cellular adaptation to stimuli, circuit formation, and the behavioral phenotype in mice.[28]

Despite over three decades of MeCP2's discovery, our understanding of this multifaceted protein's role in gene regulation remain incomplete. Emerging evidence suggests that the MBD domain of MeCP2 not only recognizes methylated DNA but also interacts with specific histone methylation marks. Recent studies have shown that the MBD domain of MeCP2 binds to

trimethyllysine marks on histone H3, such as H3K4me3, H3K9me3, H3K27me3, and H3K36me3.[29, 30] However, the molecular mechanism by which a canonical DNA binding domain interacts with methylated histones remain unexplored. In this study, we focused on characterizing the methyllysine binding site within the MBD domain of MeCP2. Our findings reveal the presence of an aromatic cage within the MBD domain of MeCP2, where methyllysines are typically known to bind. We demonstrate that mutations of this aromatic cage residues disrupt the binding of MeCP2-MBD to its methylated histone ligand, H3K27me3, and significantly impair MeCP2's localization to its target genes. These findings provide novel insights into the dual role of MeCP2 in chromatin recognition and regulation, thereby advancing our understanding of its function in gene expression.

METERIALS AND METHODS

General Materials, Methods, and Equipment. The plasmids for bacterial and mammalian expression were obtained from Addgene and Sino Biological, respectively. Mutagenesis primers and ChIP-qPCR primers were obtained from Sigma-Aldrich (Table S1 and S2). Commercially available competent bacterial cells were used for protein expression and mutagenesis. HEK293T cells were obtained from American Type Culture Collection and used following the manufacturer's protocol. All of the antibodies used in this study were purchased from established vendors and used following the manufacturer's protocol. The histone H3(15-36)K27me3 peptide was synthesized by S. Biochem Ltd and purified by HPLC to 98% purity (Figure S1). The concentrations of the peptides were determined on the basis of the observation that 1 mg/ml peptide generates an absorbance (A_{205}) of 30 at 205 nm.[31] The integrity of the purified peptides was confirmed by MALDI-MS (Figure S2).

Protein-Peptide Docking. Protein–peptide docking was performed using the HADDOCK2.4 web server.[32] The 3D structure of the MeCP2-MBD domain was obtained from the PDB (PDB code 6OGK), and the peptide representing H3(21-36)K27me3 was built in an extended conformation using PyMOL. Both the protein and peptide structures were uploaded to the server, and the active site residues within the protein's putative binding site were specified. The docking process was carried out in three stages: 1) rigid-body energy minimization (it0), 2) semi-flexible simulated annealing refinement (it1), and 3) final refinement in explicit solvent (water refinement). The resulting top clusters were ranked based on the HADDOCK score, a weighted sum of van der Waals interactions, electrostatics, desolvation energy, and restraint violations. The best-ranked cluster, characterized by the lowest average HADDOCK score, favorable interface RMSD, and a high buried surface area, was selected for further analysis. All models were visualized and assessed using PyMOL.

Mutagenesis, Expression and Purification of the MeCP2-MBD Domain. The C-terminal 6xHis-tagged MeCP2-MBD domain bacterial expression construct pNIC-CH was obtained from addgene (a gift from Cheryl Arrowsmith, addgene catalog no. 162259). MeCP2-MBD domain mutants W104A, F132A, Y141A, F142A, and F155A were generated using the QuikChange Lightning site-directed mutagenesis kit (Agilent catalog no. 210519), and the resulting mutations were confirmed by DNA sequencing. Protein expression and purification were performed using previously described methods with slight modifications.[33, 34] The wild-type MeCP2-MBD or its mutant plasmids were transformed into One Shot BL21 star (DE3) *E. coli* competent cells (Invitrogen, catalog no. C601003) using pNIC-CH kanamycin-resistant vector. A single colony was picked up and grown overnight at 37 °C in 10 mL of Luria-Bertani (LB) broth in the presence

of 50 μg mL^{-1} kanamycin. The culture was diluted 100-fold and allowed to grow at 37 °C to an optical density (OD_{600}) of 1.0. Protein expression was induced overnight at 17 °C with 0.5 mM isopropyl β-D-1-thiogalactopyranoside in an Innova 44 Incubator shaker (New Brunswick Scientific). Proteins were purified as follows. Harvested cells were resuspended in 15 mL of lysis buffer [50 mM HEPES (pH7.5), 500 mM NaCl, 5 mM β-mercaptoethanol, 5% glycerol, 25mM imidazole, lysozyme, DNase, and 1:200 (v/v) Protease Inhibitor Cocktail III (Calbiochem)]. The cells were lysed by pulse sonication and centrifuged at 13000 rpm for 40 min at 4°C. According to the manufacturer's instructions, the soluble extracts were subjected to Ni-NTA agarose resin (QIAGEN, catalog no. 30210). After 20 volumes of wash buffer [50 mM HEPES (pH 7.5), 500 mM NaCl, 5 mM β-mercaptoethanol, 5% glycerol, and 25 mM imidazole] had been passed through the column, proteins were eluted with a buffer containing 50 mM HEPES (pH 7.5), 500 mM NaCl, 5 mM β-mercaptoethanol, 5% glycerol, and 250 mM imidazole. Proteins were further purified by gel filtration chromatography (Superdex-75) using an AKTA pure FPLC system (GE Healthcare) in two different high and low salt concentration buffers; (low salt concentration, 50mM HEPES (pH 7.5), 200 mM NaCl, and 5% glycerol) and (high salt concentration, 50mM HEPES (pH 7.5), 500 mM NaCl, and 5% glycerol). Purified proteins were concentrated using an Amicon Ultra-10k centrifugal filter device (Merck Millipore Ltd.), and the concentration was determined using a Bradford assay kit (Bio-Rad Laboratories) with bovine serum albumin as a standard. The proteins were aliquoted and stored at −80 °C until use.

Isothermal Titration Calorimetry (ITC) Measurements. ITC binding experiments were performed as previously described.[33, 34] Briefly, experiments were carried out on a MicroCal PEAQ-ITC instrument (Malvern). Experiments were conducted at 15 °C while the samples were

being stirred at 750rpm. Buffers of protein and peptides were matched to 50 mM HEPES (pH 7.5), 200 mM NaCl, and 5% glycerol. Each titration comprised one initial injection of 0.4 μL lasting 0.8 s, followed by 18 injections of 2 μL lasting 4 s each at 2.5 min intervals. The initial injection was discarded during data analysis. The microsyringe (40 μL) was loaded with a peptide sample solution, and the peptide concentration varied from 1.5 to 2 mM. It was injected into the cell (200 μL), occupied by a protein concentration of 80-100 μM. All of the data were fitted to a single-binding site model using the MicroCal PEAQ-ITC analysis software to calculate the stoichiometry (N), binding constant (K_D), enthalpy (ΔH), and entropy (ΔS) of the interaction. The final titration figures were prepared using OriginPro 2020 software (OriginLab).

Mammalian Cell Culture and Transfection. HEK293T cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and a penicillin/streptomycin cocktail, and maintained in a humidified incubator at 37°C with 5% CO2. Once the cells reached approximately 90% confluency, they were removed from the incubator, the medium was discarded, and the cells were washed with cold PBS. The collected cells were then rewashed with cold PBS, pelleted, and stored as a dry pellet at −80°C for subsequent histone extraction.

The mammalian expression construct pCMV3-C-HA-MeCP2 was obtained from Sino Biological (Catalog no. HG12052-CY). The following point mutations W104A and R133C were introduced in pCMV3-C-MeCP2 plasmid as described above. Cells were transfected at 80% confluency with a ratio of 1 μg DNA per 1×10^6 cells for pCMV3-C-HA-MeCP2, and its mutant plasmids in different cell culture dishes. A 2:1 ratio of Lipofectamine 2000 (Invitrogen, catalog no. 11668019) to DNA was used. First DNA and lipofectamine were incubated separately for

5 min at RT in Opti-MEM (Gibco, catalog no. 31985070). Then the DNA and lipofectamine were mixed together and incubated at RT for 20 min and then added to the cells. Six hours post transfection, media was replaced. Cells were collected after 24 to 36 h post transfection and frozen as a dry pellet at −80 °C before use.

Acid Extraction of Histones. Histones were extracted from HEK293T cells using the acid-extraction protocol as previously described.[35] Briefly, frozen cell pellets from 100 mm dish of HEK293T cells were resuspended in 1 ml of hypotonic lysis buffer (10 mM Tris-HCl pH 8.0, 1 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT, 1 mM PMSF, and protease inhibitor cocktail) and incubated for 30 min on nutator at 4 °C. Samples were then centrifuged (10,000*g*, 10 min at 4 °C), and the supernatant was discarded entirely. The nuclei was resuspended in 800 µl of 0.4 N H_2SO_4, vortexed intermittently for 5 min, and further incubated at 4 °C on a nutator for overnight. The nuclear debris was pelleted by centrifugation (16000*g*, 10 min at 4 °C), and the supernatant containing histones were collected. The histones were precipitated by adding 264 µl trichloroacetic acid drop by drop to histone solution and the tube was inverted several times to mix the solutions and then the sample was incubated on ice for 30 min. Finally, histones were pelleted by centrifugation (16000*g*, 10 min at 4 °C), and the supernatant was discarded. The histones pellet was washed twice with ice-cold acetone, followed by centrifugation (16000*g*, 5 min at 4 °C), and the supernatant was carefully removed. The histone pellet was air-dried for 20 min at RT and subsequently dissolved in an appropriate volume of ddH2O and transferred into a fresh tube. The aliquoted histones were stored at −80 °C before use.

Ni-NTA Pull-Down Assays and Western Blotting. Pull-down experiments were performed by following the method previously described.[33, 34] First, 15-20 µg of acid-extracted histones isolated from HEK293T cells were mixed with 50 µM recombinant 6xHis-tagged-MeCP2-MBD wild-type and its mutant proteins (in high salt buffer) were incubated at room temperature on a nutator for 2 hours. Then, 50 µL of Ni-NTA agarose (QIAGEN, catalog no. 30210) was added to the reaction mixture and incubated for 45 min at room temperature on a nutator. The beads were then washed three times with 1 mL of wash buffer containing 50 mM HEPES (pH 7.5), 500 mM NaCl, and 25mM imidazole. Bound histones were eluted by incubating the beads in 20 µL of elution buffer [50 mM HEPES (pH 7.5) and 500 mM imidazole]. Equal volumes of eluted samples were separated on a 12% SDS-PAGE gel and transferred onto a 0.45 µm PVDF membrane at a constant voltage of 80 V for 1 h at 4 °C. The membrane was rinsed in TBST buffer [50 mM Tris (pH 7.4), 200 mM NaCl, and 0.1% Tween 20] and blocked for 1h at room temperature in 5% milk buffer prepared in TBST. Immunoblotting was performed overnight at 4 °C with the following primary antibodies: H3K27me3 (Cell Signaling, catalog no. 9733S) and 6x-His tag (Invitrogen, catalog no. MA121315). The membranes were washed with TBST buffer thrice at room temperature for 5 min each. The blots were then incubated with the HRP-conjugated secondary antibodies goat anti-rabbit IgG (Invitrogen, catalog no. 31466) or goat anti-mouse IgG (Invitrogen, catalog no. 31431) with 5% non-fat dry milk, diluted 1:20000 in TBST. The membranes were rewashed with TBST buffer thrice at room temperature for 5 min each. Protein bands were visualized by chemiluminescence using Super Signal West Femto substrate (Invitrogen, catalog no. 34094) following the manufacturer's protocol.

Public ChIP-seq Data Sets. MECP2 (SRR882078),[36] input DNA (SRR037637) and histone marks ChIP-seq raw data files of H3K4me3 (SRR639066),[37] H3K9me3 (SRR1617639),[38] H3K27me3 (SRR1617636),[38] H3K36me3 (SRR2877566),[39] and H4K20me3 (SRR1511591)[40] in IMR90 cells were obtained from Sequence Read Archive (SRA).

Preprocessing and Analysis of ChIP-seq Data. ChIP-seq data were analyzed using the usegalaxy server.[41] Raw reads were quality-filtered and adapter-trimmed using the Trimmomatic trimming tool.[42] Next, the trimmed reads were aligned in single-end mode with the human reference genome (hg19) using Bowtie2 (version 2.4.2).[43] Alignment followed by peak calling was performed with MACS2 (version 2.1.0) over input.[44] As recommended for IDR (irreproducibility discovery rate) analysis, the final list of peaks was selected with an IDR threshold of 0.001 to obtain a conservative set of peaks. Peaks overlapping with blacklisted regions as defined by the ENCODE project were discarded. Heat maps and correlation plot were prepared using the deepTools2 (version 3.5.4).[45] Further analysis was performed in the R (version 3.4.0)/Bioconductor environment with plots generated relying on the ChIPseeker library.[46] We used the Bioconductor package ChIPseeker.[47] to display the overlaps between the different data sets and IGV for ChIP-seq signal visualization in a genomic context.

ChIP-qPCR. ChIP-qPCR experiments were performed as previously reported with slight modifications.[48] Briefly, the HEK293T cells were transfected with HA-tagged MeCP2 wild-type and its mutants (W104A and R133C) plasmids in different cell culture dishes. Thirty-six hours post-transfection, cells were collected and fixed in 1% formaldehyde for 20 min, and quenched with 125 mM Glycine for 5 min at RT. The cells were thoroughly washed with PBS supplemented

with protease inhibitors. The pellets were resuspended in a high salt lysis/sonication buffer (25 mM Tris–HCl pH 7.5, 0.5% sodium deoxycholate, 0.1% SDS, 1% Triton X-100, 5 mM EDTA, 800 mM NaCl and protease inhibitors) and sonicated using Qsonica Q700 (pulsed 5 s on, 30 s off, for 60 s processing time). Insoluble cell debris was pelleted *via* centrifugation at 13,600*g* for 30 min at 4 °C. Dynabeads Protein G were incubated with 2.5 μg anti-HA antibody (Invitrogen, catalog no. 26183), or control IgG antibody (CST, catalog no. 5415) for 30 min at RT with rotation in Ab binding and washing buffer. The supernatant containing the soluble chromatin was incubated with Ab-conjugated magnetic beads at 4 °C for overnight, with rotation. Next, the beads were washed with wash buffer A (50 mM HEPES pH 7.9, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% SDS and 0.1% sodium deoxycholate), wash buffer B (50 mM HEPES pH 7.9, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% SDS and 0.1% sodium deoxycholate), and wash buffer C (25 mM Tris–HCl pH 7.5, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40 alternative and 0.5% sodium deoxycholate). Finally, the beads were washed with tris-EDTA buffer for two times and eluted with elution buffer (10 mM Tris, pH 7.5, 1 mM EDTA, and 1% SDS) and incubated at 65 °C for 30 min. The eluted samples were treated with RNase A (Invitrogen, catalog no. EN0531) to a final concentration of 20 μg/ml and incubated at 65 °C for 8 h to digest any contaminated RNA. Next, the samples were treated with Proteinase K (Invitrogen, catalog no. 25530049) to a final concentration of 200 μg/ml and incubated at 45 °C for 2 h to digest proteins. Finally, the DNA was purified using a QIAquick PCR Purification Kit (QIAGEN, catalog. no. 28104) and used for qPCR analysis.

RESULTS AND DISCUSSION

Recognition of Histone H3K27me3 Mark by MeCP2 via an Aromatic Cage in the MBD Domain. Histone lysine methylation is a critical epigenetic mark that is dynamically regulated by lysine methyltransferases (KMTs) and lysine demethylases (KDMs).[49] The readers of histone lysine methylation are found in diverse protein families; however, the mechanism of methyllysine recognition remains largely conserved across these readers.[50] Histone methylation readers typically feature an aromatic cage consisting of two to four aromatic residues, where the methyllysine binds.[50-52] This binding is driven by cation-π interaction between the electron rich aromatic residues in the reader and the positively charged methylammonium group of lysine, along with hydrophobic and van der waals interactions.[50-52] Recent studies have shown that the MBD domain of MeCP2, in addition to its well-known role in methyl-CpG binding, also interacts with the histone H3 trimethylated at lysine 27 (H3K27me3) mark.[29] However, the precise binding site for this interaction remains unclear.

To identify the region of trimethyllysine binding within the MBD domain of MeCP2, we carefully examined its crystal structure (PDB code 6OGK) to search for the presence of such aromatic cage. We found that the MBD domain of MeCP2 contains five aromatic residues, viz, W104, F132, Y141, F142, and F155 that are arranged in a cage-like structure, which we hypothesized to be the possible site of methyllysine binding. To investigate the potential binding mode of MeCP2-MBD with H3K27me3 peptide, we performed protein-peptide docking using the HADDOCK2.4 web server. The five aforementioned aromatic residues were kept as active site residue and the H3(21-36)K27me3 was modeled in an extended conformation. HADDOCK generated 1000 docking models, which were clustered based on interface RMSD. The top-ranked cluster showed a favorable HADDOCK score of –53.76 ± 2.01, with the Kme3 group consistently occupying a hydrophobic pocket formed by aromatic cage residues (Figure 1C). The interaction

was stabilized by cation–π interactions, hydrophobic, and salt bridge interaction between the trimethyl lysine and residues in the binding pocket of the protein (Figure S3). These results suggest that the aromatic cage of MeCP2 MBD is a plausible Kme3-binding pocket.

Mutation of Aromatic Cage Residue Disrupts Binding of MeCP2-MBD with H3K27me3. To validate our docking results *in vitro*, we individually mutated these aromatic residues to alanine, overexpressed the MeCP2-MBD domain in *E. coli* cells, and purified the proteins to high homogeneity (Figure S4). We then performed ITC binding assays to assess the effect of these mutations on the binding affinity of MeCP2-MBD toward the H3K27me3 peptide. The wild-type MeCP2-MBD domain exhibited strong binding affinity for the H3K27me3 peptide, with a dissociation constant (K_D) of 5.2 ± 0.4 μM (Figure 2A), which is consistent with previously reported values.[30] In contrast, mutation of the aromatic cage residues, W104, F132, Y141, F142 and F155 to alanine drastically reduced the binding affinity for the H3K27me3 peptide, with a K_D ranging from >600 to >1000 μM (Figures 2B-F).

Next, we assessed the binding of MeCP2-MBD wild-type and its mutants to endogenous histones, which serve as more physiologically relevant substrates than peptides. Histones were isolated from HEK293T cells and incubated with 6xHis-tagged MeCP2-MBD wild-type and its mutant proteins (Figure 3A). Subsequent enrichment with Ni-NTA beads followed by western blotting with H3K27me3 antibody reveals that the wild-type MeCP2-MBD efficiently bound to H3K27me3 bearing histones, while mutants W104A, F132A, Y141A and F155A showed a complete loss of binding (Figure 3B). Notably, the F142A mutant exhibited a significant reduction in binding to H3K27me3 histones compared to wild-type protein (Figure 3B). Our findings from ITC binding and histone pull-down experiments align with previous reports on methyllysine

readers. For example, the BAH domain of BAHD1, which also recognizes the H3K27me3 mark, features an aromatic cage consisting of tryptophan (W667) and two tyrosines (Y645 and Y669).[53] Mutations of these aromatic residues to alanine not only abolished the binding of BAH domain with H3K27me3 in ITC binding assays but also reduced the localization of BAHD1 with H3K27me3 in cells.[53] Similarly, mutations of aromatic residues in other methyllysine readers, such as PHD, chromodomain, tudor domain, and MBT domain, have been shown to reduce ligand binding affinity.[54-59]

Several reports suggest that some epigenetic readers also exhibit DNA binding capabilities, which promotes their engagement with chromatin and their affinity for DNA could be higher than for histones.[60] For instance, the chromodomain of CBX8, a reader of the H3K27me3 mark and a component of polycomb repressive complex 1 (PRC1), binds to the H3K27me3 peptide with weak affinity (K_D = 0.7 mM) *in vitro*, while showing no interaction with its unmodified counterpart.[61] However, electrophoretic mobility shift assays (EMSA) reveal that the CBX8 chromodomain binds both H3K27me3-modified and unmodified nucleosomes at similar concentration, suggesting that interactions with DNA primarily drive this binding.[61] Given that, the MBD domain is also primarily a DNA interacting module, we propose that for MeCP2, the interaction of its MBD domain with DNA might be the main driving force behind its methyllysine recognition. Upon DNA binding, an induced-fit mechanism could occur, enabling the aromatic cage to open and accommodate the trimethyllysine. Our presumption aligns with observations that the binding pockets of methyllysine readers can undergo flexible structural changes to bind their cognate ligands.[62, 63]

MeCP2 Localizes with Canonical Histone Methylation Marks in the Genome. Histone lysine methylation plays a crucial role in diverse nuclear functions, including heterochromatin formation, chromatin phase separation, and transcriptional regulation.[49, 64] The N-terminal tail of histones contains five canonical lysines that are subjected to methylation: H3K4, H3K9, H3K27, H3K36 and H4K20.[65, 66] The occurrence of different methylation marks with in the genome is indicative of distinct biological functions. For instance, the presence of H3K4me3 mark at gene promoters is associated with active transcription,[67, 68] while H3K9me3 is a hallmark of constitutive heterochromatin, and H3K27me3 mark associated with facultative heterochromatin.[69-71] These histone methylation marks are recognized by epigenetic reader proteins, which facilitate the recruitment of transcriptional regulatory complexes to the chromatin.[52]

A recent study revealed that MeCP2 also binds to various histone methylation marks, including H3K4me3, H3K9me3, and H3K36me3, with dissociation constants (K_D) of 0.12 μM, 0.075 μM, and 3.1 μM, respectively.[30] This prompted us to investigate the colocalization of MeCP2 with these histone marks across the genome. We utilized publicly available ChIP-seq data to understand the localization of MeCP2 with the canonical histone methylation marks in the genome. Pearson correlation analysis showed that more than 50% of MeCP2 peaks overlap with peaks of H3K4me3, H3K9me3, H3K27me3, and H4K20me3 (Figure 4A). Among these marks, H3K4me3 exhibited the strongest positive correlation, with 66% of MeCP2 peaks coinciding with H3K4me3 peaks (Figure 4A). In contrast, despite MeCP2 binding to H3K36me3,[30] we observed a negative correlation between their genomic localizations (Figure 4A). Next, the significant peaks of MeCP2 and histone methylation marks were used to plot Venn diagram, in order to understand their overlap in the genome. Our Venn diagram analysis reveals that a total of 846 peaks (11.7%) out of 7216 peaks of MeCP2 overlaps with H3K4me3, H3K9me3, H3K27me3 and H4K20me3

peaks in the genome (Figure 4B). To further understand the colocalization of these canonical histone methylation marks in the MeCP2 occupied genes, we plotted heat maps. Heat map visualization demonstrate that transcription start site (TSS) of MeCP2 bound genes has abundant enrichment of H3K4me3 (Figure 4C). Additionally, the histone marks H3K9me3, H3K27me3, and H4K20me3 are distributed across the 5kb region from TSS of a significant subset of MeCP2 associated genes (Figure 4C). Consistent with the Pearson correlation analysis, the H3K36me3 marks are notably absent in MeCP2 occupied regions (Figure 4C).

The localization of an epigenetic reader with specific histone modification signifies the readers role in regulating gene expression and chromatin dynamics.[72] Initially identified as a DNA methylation reader, MeCP2 was thought to primarily act as transcription repressor by recruiting histone deacetylases (HDACs) and Sin3A repressor complex.[14, 15] However, subsequent studies have shown that, MeCP2 could also activate transcription by forming complex with the transcription factor CREB1 at the promoter of active genes.[23] Our analysis supports this dual functionality of MeCP2, as it associates with both transcriptionally active marks, such as H3K4me3, and repressive marks, including H3K9me3, H3K27me3, and H4K20me3. The localization of MeCP2 with H3K9me3 and H4K20me3 can be attributed to its role in the formation and maintenance of pericentromeric heterochromatin (PHC) by deposition of H3K9me3 and H4K20me3, potentially by recruitment of specific histone methyltransferases.[73] Consistent with this, earlier report also indicated that MeCP2 recruits histone methyltransferase specific for H3K9 methylation.[74]

MECP2 and Histone Methylation Marks Play Crucial Roles in Regulating Cellular Pathways and Cancer Development. The binding of epigenetic readers to specific histone marks

facilitates the recruitment of regulatory machinery, leading to the expression of relevant outputs that are implicated in various cellular functions. Since MeCP2 interacts with and localizes alongside several histone methylation marks, we sought to understand how this engagement influences biological pathways. To explore this, we conducted KEGG pathway enrichment analysis to assess the significance of histone marks recognized by MeCP2. The results, illustrated in the bubble map of KEGG enrichment analysis, demonstrate that MeCP2, in conjunction with canonical histone methylation marks, regulates critical cellular pathways such as Apelin, Wnt, Rap1, insulin, calcium, Hippo, cAMP, Ras, and mTOR signaling, as well as melanogenesis, autophagy, and cellular senescence (Figure 5). As MeCP2 is predominantly expressed in the nervous system, we also observed its involvement in neuronal pathways, including circadian rhythm regulation, glutamatergic and cholinergic synapse signaling, serotonergic synapse function, axon guidance, and neurodegenerative processes (Figure 5). Additionally, our analysis reveals that MeCP2 and histone methylation marks are implicated in the progression of various cancers, such as hepatocellular carcinoma, gastric cancer, breast cancer, colorectal cancer, chronic myeloid leukemia, and melanoma (Figure 5).

Mutation in the Aromatic Cage of the MBD Domain Reduces MeCP2 Localization to Chromatin. Our ITC binding experiments and histone pull-down assays demonstrated that the MBD domain of MeCP2 recognizes trimethylated lysine 27 on histone H3 through an aromatic cage. Moreover, mutations in the aromatic cage of the MBD domain disrupt the binding of MeCP2 to both histone peptides and endogenous histones *in vitro*. To investigate how MeCP2 interacts with chromatin in the genomic context, we mutated the aromatic cage residue W104 to alanine and transiently expressed HA-tagged wild-type MeCP2 and its mutant W104A in HEK293T cells (Figure S5). Given that MeCP2 is also a DNA methylation reader, we included the R133C

mutation in our study. To validate the interaction between MeCP2 and histones at the cellular level, we performed immunoprecipitation assays. Our results show that HA-MeCP2 wild-type efficiently pulls down histone H3, while its mutant W104A shows a complete loss of binding to histone H3 in cells (Figure 6). In contrast, the R133C mutant retains its ability to bind histones (Figure 6), suggesting that residue R133 does not contribute to histone recognition.

Next, we immunoprecipitated DNA from these cells and performed quantitative ChIP to assess MeCP2 localization at target genes. We selected ten genes previously reported as MeCP2 target; DNMT1, ICAM1, HIPK3, DKK1, ILF3, ICAM3, KDM1A, ZNF713 and HDAC1.[75] First, we examined the localization of MeCP2 and histone marks in the bodies of these genes. Analysis of publicly available ChIP-seq data reveals that MeCP2 and the histone marks H3K4me3 and H3K27me3 are enriched in the promoter of these aforementioned genes (Figure 7A and S6). We validated this observation by performing ChIP-qPCR (Figure S7). Next, we performed ChIP-qPCR from immunoprecipitated DNA from cells expressing either WT or mutated MeCP2 plasmid. Our analysis of ChIP-qPCR experiment shows that mutation of aromatic cage residue W104 leads to more than 50% reduction in enrichment of MeCP2 in its target genes and this reduction is comparable to that observed for R133C mutation (Figure 7B). Together, these finding highlight the importance of aromatic cage of MeCP2 in histone recognition and chromatin localization. While R133C mutation is well established to impair DNA binding, the loss of histone interaction in case of W104A mutant suggest that histone binding also contributes to MeCP2's genomic localization independent of DNA methylation. This dual mode of chromatin binding may be crucial for MeCP2's function in the context of gene regulation.

Our analysis also suggests that the promoter of some MeCP2 target genes contain both transcription activating H3K4me3 and transcription repressing H3K27me3 marks. The presence

of both these histone modifications in a gene promoter is a signature of bivalent domain. Bivalent domains have traditionally been associated with expression of developmentally regulated genes in embryonic stem cells (ESC) keeping them in poised state.[76, 77] However, more recently it has been suggested that bivalency is a mechanism commonly used by cells to tightly regulate the expression of tissue specific genes.[78] Hence, based on our observation, we propose that MeCP2 rather than functioning solely as a transcription activator or repressor, plays a broader transcriptional regulatory role.

CONCLUSION

In conclusion, our study identifies MeCP2 as a histone methylation reader, highlighting its ability to bind methylated lysine residues in histones, specifically through an aromatic cage in its MBD domain. Mutation of key aromatic residues disrupts MeCP2's binding to histone marks, such as H3K27me3, and impairs its chromatin localization. Genomic analysis further demonstrates that MeCP2 interacts with histone marks like H3K4me3 and H3K27me3 at gene promoters, regulating critical biological and cancer-related pathways. These findings underscore the importance of MeCP2's histone-binding ability for its chromatin association and gene regulatory functions.

ASSOCIATED CONTENT

Supporting Information

This article contains supporting information. The supporting information contains list of primers, peptide characterization, western blot data, ChIP-seq data, and ChIP-qPCR data. Table S1-S2 and Figure S1-S7.

Accession Codes

MeCP2, UniProt P51608

AUTHOR INFORMATION

Corresponding Author: *E-mail: s.babu@iiserkol.ac.in

ORCID: Babu Sudhamalla: 0000-0002-6610-1424.

Author Contributions

B.S. conceived the ideas. J.P., and B.S. designed the experiments. J.P. performed the experiments. J.P., and B.S. analyzed data. J.P., and B.S. prepared original draft. B. S. writing-reviewing and editing. B.S. supervision, B.S. resources. B.S. project administration. B.S. funding acquisition.

Funding

This work was supported by grants from SERB (CRG/2022/005242), DBT Ramalingaswami Fellowship (BT/RLF/Re-entry/56/2018), and Intramural grant from IISER Kolkata (Ministry of Education), Govt. of India to B.S.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank infrastructural facilities supported by IISER Kolkata, and DST-FIST (SR/FST/LS-II/2017/93). J.P. acknowledges, PMRF for her doctoral fellowship, from Govt. of India.

ABBREVIATIONS

MeCP2, Methyl-CpG binding protein 2; MBD, methyl-CpG-binding domain; SAM, S-adenosylmethionine; 5mC, 5-methylcytosine; UHRF, ubiquitin-like with PHD and RING finger; TRD, transcription repression domain.

REFERENCES

[1] Kim, M., and Costello, J. (2017) DNA methylation: an epigenetic mark of cellular memory, *Experimental & Molecular Medicine 49*, e322-e322.

[2] Moore, L. D., Le, T., and Fan, G. (2013) DNA Methylation and Its Basic Function, *Neuropsychopharmacology 38*, 23-38.

[3] Bhat, A., Ghatage, T., Bhan, S., Lahane, G. P., Dhar, A., Kumar, R., Pandita, R. K., Bhat, K. M., Ramos, K. S., and Pandita, T. K. (2022) Role of Transposable Elements in Genome Stability: Implications for Health and Disease, *Int J Mol Sci 23*.

[4] Deniz, Ö., Frost, J. M., and Branco, M. R. (2019) Regulation of transposable elements by DNA modifications, *Nat Rev Genet 20*, 417-431.

[5] Mahmood, N., and Rabbani, S. A. (2019) DNA Methylation Readers and Cancer: Mechanistic and Therapeutic Applications, *Front Oncol 9*, 489.

[6] Meehan, R. R., Lewis, J. D., and Bird, A. P. (1992) Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA, *Nucleic Acids Res 20*, 5085-5092.

[7] Good, K. V., Vincent, J. B., and Ausió, J. (2021) MeCP2: The Genetic Driver of Rett Syndrome Epigenetics, *Front Genet 12*, 620859.

[8] Bienvenu, T., Carrié, A., de Roux, N., Vinet, M. C., Jonveaux, P., Couvert, P., Villard, L., Arzimanoglou, A., Beldjord, C., Fontes, M., Tardieu, M., and Chelly, J. (2000) MECP2 mutations account for most cases of typical forms of Rett syndrome, *Hum Mol Genet 9*, 1377-1384.

[9] Hite, K. C., Adams, V. H., and Hansen, J. C. (2009) Recent advances in MeCP2 structure and function, *Biochem Cell Biol 87*, 219-227.

[10] Singh, J., Saxena, A., Christodoulou, J., and Ravine, D. (2008) MECP2 genomic structure and function: insights from ENCODE, *Nucleic Acids Res 36*, 6035-6047.

[11] Wakefield, R. I., Smith, B. O., Nan, X., Free, A., Soteriou, A., Uhrin, D., Bird, A. P., and Barlow, P. N. (1999) The solution structure of the domain from MeCP2 that binds to methylated DNA, *J Mol Biol 291*, 1055-1065.

[12] Free, A., Wakefield, R. I., Smith, B. O., Dryden, D. T., Barlow, P. N., and Bird, A. P. (2001) DNA recognition by the methyl-CpG binding domain of MeCP2, *J Biol Chem 276*, 3353-3360.

[13] Zou, X., Ma, W., Solov'yov, I. A., Chipot, C., and Schulten, K. (2012) Recognition of methylated DNA through methyl-CpG binding domain proteins, *Nucleic Acids Res 40*, 2747-2758.

[14] Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex, *Nature 393*, 386-389.

[15] Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J., and Wolffe, A. P. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription, *Nat Genet 19*, 187-191.

[16] Kucukkal, T. G., and Alexov, E. (2015) Structural, Dynamical, and Energetical Consequences of Rett Syndrome Mutation R133C in MeCP2, *Comput Math Methods Med 2015*, 746157.

[17] Ghosh, R. P., Horowitz-Scherer, R. A., Nikitina, T., Gierasch, L. M., and Woodcock, C. L. (2008) Rett syndrome-causing mutations in human MeCP2 result in diverse structural changes that impact folding and DNA interactions, *J Biol Chem 283*, 20523-20534.

[18] Masuyama, T., Matsuo, M., Jing, J. J., Tabara, Y., Kitsuki, K., Yamagata, H., Kan, Y., Miki, T., Ishii, K., and Kondo, I. (2005) Classic Rett syndrome in a boy with R133C mutation of MECP2, *Brain Dev 27*, 439-442.

[19] Sharma, K., Singh, J., Frost, E. E., and Pillai, P. P. (2018) MeCP2 in central nervous system glial cells: current updates, *Acta Neurobiol Exp (Wars) 78*, 30-40.

[20] Gulmez Karaca, K., Brito, D. V. C., and Oliveira, A. M. M. (2019) MeCP2: A Critical Regulator of Chromatin in Neurodevelopment and Adult Brain Function, *Int J Mol Sci 20*.

[21] Good, K., Kalani, L., Vincent, J., and Ausió, J. (2025) Multifaceted roles of MeCP2 in cellular regulation and phase separation: implications for neurodevelopmental disorders, depression, and oxidative stress, *Biochem Cell Biol 103*, 1-12.

[22] Ta, D., Downs, J., Baynam, G., Wilson, A., Richmond, P., and Leonard, H. (2022) A brief history of MECP2 duplication syndrome: 20-years of clinical understanding, *Orphanet J Rare Dis 17*, 131.

[23] Chahrour, M., Jung, S. Y., Shaw, C., Zhou, X., Wong, S. T., Qin, J., and Zoghbi, H. Y. (2008) MeCP2, a key contributor to neurological disease, activates and represses transcription, *Science 320*, 1224-1229.

[24] Liu, Y., Flamier, A., Bell, G. W., Diao, A. J., Whitfield, T. W., Wang, H. C., Wu, Y., Schulte, F., Friesen, M., Guo, R., Mitalipova, M., Liu, X. S., Vos, S. M., Young, R. A., and Jaenisch, R. (2024) MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, *Neuron 112*, 1943-1958.e1910.

[25] Pandey, S., Simmons, G. E., Jr., Malyarchuk, S., Calhoun, T. N., and Pruitt, K. (2015) A novel MeCP2 acetylation site regulates interaction with ATRX and HDAC1, *Genes Cancer 6*, 408-421.

[26] Chen, W. G., Chang, Q., Lin, Y., Meissner, A., West, A. E., Griffith, E. C., Jaenisch, R., and Greenberg, M. E. (2003) Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2, *Science 302*, 885-889.

[27] Martinowich, K., Hattori, D., Wu, H., Fouse, S., He, F., Hu, Y., Fan, G., and Sun, Y. E. (2003) DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation, *Science 302*, 890-893.

[28] Bellini, E., Pavesi, G., Barbiero, I., Bergo, A., Chandola, C., Nawaz, M. S., Rusconi, L., Stefanelli, G., Strollo, M., Valente, M. M., Kilstrup-Nielsen, C., and Landsberger, N. (2014) MeCP2 post-translational modifications: a mechanism to control its involvement in synaptic plasticity and homeostasis?, *Front Cell Neurosci 8*, 236.

[29] Lee, W., Kim, J., Yun, J.-M., Ohn, T., and Gong, Q. (2020) MeCP2 regulates gene expression through recognition of H3K27me3, *Nature Communications 11*, 3140.

[30] Ortega-Alarcon, D., Claveria-Gimeno, R., Vega, S., Kalani, L., Jorge-Torres, O. C., Esteller, M., Ausio, J., Abian, O., and Velazquez-Campoy, A. (2024) Extending MeCP2 interactome: canonical nucleosomal histones interact with MeCP2, *Nucleic Acids Res 52*, 3636-3653.

[31] Anthis, N. J., and Clore, G. M. (2013) Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm, *Protein Sci 22*, 851-858.

[32] Honorato, R. V., Trellet, M. E., Jiménez-García, B., Schaarschmidt, J. J., Giulini, M., Reys, V., Koukos, P. I., Rodrigues, J. P. G. L. M., Karaca, E., van Zundert, G. C. P., Roel-Touris, J., van Noort, C. W., Jandová, Z., Melquiond, A. S. J., and Bonvin, A. M. J. J. (2024) The HADDOCK2.4 web server for integrative modeling of biomolecular complexes, *Nature Protocols 19*, 3219-3241.

[33] Barman, S., Roy, A., Padhan, J., and Sudhamalla, B. (2022) Molecular Insights into the Recognition of Acetylated Histone Modifications by the BRPF2 Bromodomain, *Biochemistry 61*, 1774-1789.

[34] Bardhan, I., Barman, S., Roy, A., and Sudhamalla, B. (2023) Novel insights into the recognition of acetylated histone H4 tail by the TRIM24 PHD-Bromo module, *Biochem J 480*, 629-647.

[35] Shechter, D., Dormann, H. L., Allis, C. D., and Hake, S. B. (2007) Extraction, purification and analysis of histones, *Nature Protocols 2*, 1445-1457.

[36] Maunakea, A. K., Chepelev, I., Cui, K., and Zhao, K. (2013) Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition, *Cell Research 23*, 1256-1269.

[37] Jin, F., Li, Y., Dixon, J. R., Selvaraj, S., Ye, Z., Lee, A. Y., Yen, C.-A., Schmitt, A. D., Espinoza, C. A., and Ren, B. (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells, *Nature 503*, 290-294.

[38] Vallot, C., Ouimette, J. F., Makhlouf, M., Féraud, O., Pontis, J., Côme, J., Martinat, C., Bennaceur-Griscelli, A., Lalande, M., and Rougeulle, C. (2015) Erosion of X Chromosome Inactivation in Human Pluripotent Cells Initiates with XACT Coating and Depends on a Specific Heterochromatin Landscape, *Cell Stem Cell 16*, 533-546.

[39] Kirschner, K., Samarajiwa, S. A., Cairns, J. M., Menon, S., Pérez-Mancera, P. A., Tomimatsu, K., Bermejo-Rodriguez, C., Ito, Y., Chandra, T., Narita, M., Lyons, S. K., Lynch, A. G., Kimura, H., Ohbayashi, T., Tavaré, S., and Narita, M. (2015) Phenotype Specific Analyses Reveal Distinct Regulatory Mechanism for Chronically Activated p53, *PLOS Genetics 11*, e1005053.

[40] Nelson, D. M., Jaber-Hijazi, F., Cole, J. J., Robertson, N. A., Pawlikowski, J. S., Norris, K. T., Criscione, S. W., Pchelintsev, N. A., Piscitello, D., Stong, N., Rai, T. S., McBryan, T., Otte, G. L., Nixon, C., Clark, W., Riethman, H., Wu, H., Schotta, G., Garcia, B. A., Neretti, N., Baird, D. M., Berger, S. L., and Adams, P. D. (2016) Mapping H4K20me3 onto the chromatin landscape of senescent cells indicates a function in control of cell senescence and tumor suppression through preservation of genetic and epigenetic stability, *Genome Biol 17*, 158.

[41] Community, T. G. (2022) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update, *Nucleic Acids Research 50*, W345-W351.

[42] Bolger, A. M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics 30*, 2114-2120.

[43] Langmead, B., and Salzberg, S. L. (2012) Fast gapped-read alignment with Bowtie 2, *Nat Methods 9*, 357-359.
[44] Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W., and Liu, X. S. (2008) Model-based Analysis of ChIP-Seq (MACS), *Genome Biology 9*, R137.
[45] Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F., and Manke, T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis, *Nucleic Acids Res 44*, W160-165.
[46] Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., Bravo, H. C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K. D., Irizarry, R. A., Lawrence, M., Love, M. I., MacDonald, J., Obenchain, V., Oleś, A. K., Pagès, H., Reyes, A., Shannon, P., Smyth, G. K., Tenenbaum, D., Waldron, L., and Morgan, M. (2015) Orchestrating high-throughput genomic analysis with Bioconductor, *Nature Methods 12*, 115-121.
[47] Yu, G., Wang, L. G., and He, Q. Y. (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization, *Bioinformatics 31*, 2382-2383.
[48] Sullivan, A. E., and Santos, S. D. M. (2020) An Optimized Protocol for ChIP-Seq from Human Embryonic Stem Cell Cultures, *STAR Protoc 1*, 100062.
[49] Black, J. C., Van Rechem, C., and Whetstine, J. R. (2012) Histone lysine methylation dynamics: establishment, regulation, and biological impact, *Mol Cell 48*, 491-507.
[50] Vann, K. R., Vishweshwaraiah, Y. L., Dokholyan, N. V., and Kutateladze, T. G. (2021) Searching for methyllysine-binding aromatic cages, *Biochem J 478*, 3613-3619.
[51] Huang, X., Chen, Y., Xiao, Q., Shang, X., and Liu, Y. (2024) Chemical inhibitors targeting histone methylation readers, *Pharmacol Ther 256*, 108614.
[52] Hyun, K., Jeon, J., Park, K., and Kim, J. (2017) Writing, erasing and reading histone lysine methylations, *Experimental & Molecular Medicine 49*, e324-e324.
[53] Zhao, D., Zhang, X., Guan, H., Xiong, X., Shi, X., Deng, H., and Li, H. (2016) The BAH domain of BAHD1 is a histone H3K27me3 reader, *Protein Cell 7*, 222-226.
[54] Jacobs, S. A., and Khorasanizadeh, S. (2002) Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail, *Science 295*, 2080-2083.
[55] Peña, P. V., Hom, R. A., Hung, T., Lin, H., Kuo, A. J., Wong, R. P., Subach, O. M., Champagne, K. S., Zhao, R., Verkhusha, V. V., Li, G., Gozani, O., and Kutateladze, T. G. (2008) Histone H3K4me3 binding is required for the DNA repair and apoptotic activities of ING1 tumor suppressor, *J Mol Biol 380*, 303-312.
[56] Li, H., Fischle, W., Wang, W., Duncan, E. M., Liang, L., Murakami-Ishibe, S., Allis, C. D., and Patel, D. J. (2007) Structural basis for lower lysine methylation state-specific readout by MBT repeats of L3MBTL1 and an engineered PHD finger, *Mol Cell 28*, 677-691.
[57] Fischle, W., Wang, Y., Jacobs, S. A., Kim, Y., Allis, C. D., and Khorasanizadeh, S. (2003) Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains, *Genes Dev 17*, 1870-1881.
[58] Botuyan, M. V., Lee, J., Ward, I. M., Kim, J. E., Thompson, J. R., Chen, J., and Mer, G. (2006) Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair, *Cell 127*, 1361-1373.
[59] Hung, T., Binda, O., Champagne, K. S., Kuo, A. J., Johnson, K., Chang, H. Y., Simon, M. D., Kutateladze, T. G., and Gozani, O. (2009) ING4 mediates crosstalk between histone

H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation, *Mol Cell 33*, 248-256.
[60] Weaver, T. M., Morrison, E. A., and Musselman, C. A. (2018) Reading More than Histones: The Prevalence of Nucleic Acid Binding among Reader Domains, *Molecules 23*.
[61] Connelly, K. E., Weaver, T. M., Alpsoy, A., Gu, B. X., Musselman, C. A., and Dykhuizen, E. C. (2019) Engagement of DNA and H3K27me3 by the CBX8 chromodomain drives chromatin association, *Nucleic Acids Res 47*, 2289-2305.
[62] Luise, C., Robaa, D., and Sippl, W. (2021) Exploring aromatic cage flexibility of the histone methyllysine reader protein Spindlin1 and its impact on binding mode prediction: an in silico study, *J Comput Aided Mol Des 35*, 695-706.
[63] Ruthenburg, A. J., Allis, C. D., and Wysocka, J. (2007) Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark, *Mol Cell 25*, 15-30.
[64] Martin, C., and Zhang, Y. (2005) The diverse functions of histone lysine methylation, *Nat Rev Mol Cell Biol 6*, 838-849.
[65] Greer, E. L., and Shi, Y. (2012) Histone methylation: a dynamic mark in health, disease and inheritance, *Nature Reviews Genetics 13*, 343-357.
[66] Husmann, D., and Gozani, O. (2019) Histone lysine methyltransferases in biology and disease, *Nat Struct Mol Biol 26*, 880-889.
[67] Liang, G., Lin, J. C., Wei, V., Yoo, C., Cheng, J. C., Nguyen, C. T., Weisenberger, D. J., Egger, G., Takai, D., Gonzales, F. A., and Jones, P. A. (2004) Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome, *Proc Natl Acad Sci U S A 101*, 7357-7362.
[68] Yan, C., and Boyd, D. D. (2006) Histone H3 acetylation and H3 K4 methylation define distinct chromatin regions permissive for transgene expression, *Mol Cell Biol 26*, 6357-6371.
[69] Matsuwaka, M., Kumon, M., and Inoue, A. (2025) H3K27 dimethylation dynamics reveal stepwise establishment of facultative heterochromatin in early mouse embryos, *Nature Cell Biology 27*, 28-38.
[70] Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins, *Nature 410*, 116-120.
[71] Saksouk, N., Simboeck, E., and Déjardin, J. (2015) Constitutive heterochromatin formation and transcription in mammals, *Epigenetics Chromatin 8*, 3.
[72] Musselman, C. A., Lalonde, M. E., Côté, J., and Kutateladze, T. G. (2012) Perceiving the epigenetic landscape through histone readers, *Nat Struct Mol Biol 19*, 1218-1227.
[73] Fioriniello, S., Csukonyi, E., Marano, D., Brancaccio, A., Madonna, M., Zarrillo, C., Romano, A., Marracino, F., Matarazzo, M. R., D'Esposito, M., and Della Ragione, F. (2020) MeCP2 and Major Satellite Forward RNA Cooperate for Pericentric Heterochromatin Organization, *Stem Cell Reports 15*, 1317-1332.
[74] Fuks, F., Hurd, P. J., Wolf, D., Nan, X., Bird, A. P., and Kouzarides, T. (2003) The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation, *J Biol Chem 278*, 4035-4040.
[75] Castro-Piedras, I., Vartak, D., Sharma, M., Pandey, S., Casas, L., Molehin, D., Rasha, F., Fokar, M., Nichols, J., Almodovar, S., Rahman, R. L., and Pruitt, K. (2020) Identification of Novel MeCP2 Cancer-Associated Target Genes and Post-Translational Modifications, *Front Oncol 10*, 576362.

[76] Macrae, T. A., Fothergill-Robinson, J., and Ramalho-Santos, M. (2023) Regulation, functions and transmission of bivalent chromatin during mammalian development, *Nat Rev Mol Cell Biol 24*, 6-26.

[77] Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L., and Lander, E. S. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells, *Cell 125*, 315-326.

[78] Blanco, E., González-Ramírez, M., Alcaine-Colet, A., Aranda, S., and Di Croce, L. (2020) The Bivalent Genome: Characterization, Structure, and Regulation, *Trends Genet 36*, 118-131.

Figure legends

Figure 1. Domain organization of MeCP2 and docked structure of its MBD domain with H3K27me3 peptide. (A) Schematic representation of MeCP2 showing methyl-CpG binding domain (MBD) and a transcription repressor domain (TRD) connected by an intermediary domain (ID). (B) Cartoon representation of the X-ray crystal structure of MeCP2-MBD domain (PDB code 6OGK). (C) Close-up view of the docked complex of MeCP2-MBD domain and H3K27me3 peptide. The aromatic cage residues within the binding pocket are highlighted.

Figure 2. Exothermic ITC plots showing binding of MeCP2-MBD domain WT and mutants to H3K27me3 peptide. (A) MeCP2-MBD WT, (B) MeCP2-MBD W104A, (C) MeCP2-MBD F132A, (D) MeCP2-MBD Y141A, (E) MeCP2-MBD F142A, and (F) MeCP2-MBD F155A. The calculated binding constants are indicated.

Figure 3. Interaction of MeCP2-MBD WT and mutants with endogenous histone H3. (A) Schematic depicting the preparation of endogenous histone from HEK293T cells followed by enrichment with MeCP2-MBT WT and mutant proteins. (B) Western blot analysis of histone H3K27me3 enrichment by MeCP2-MBT WT and mutant proteins.

Figure 4. ChIP-seq analysis of MeCP2 localization and its association with canonical histone methylation marks in the genome. (A) Pearson correlation heat map depicting the relationship between MeCP2 binding and canonical histone methylation marks (B) Venn diagram illustrating the overlap of genome binding site between MeCP2 and histone marks H3K4me3, H3K9me3, H3K27me3 and H4K20me3. (C) Heat maps of ChIP-seq signals for MeCP2, H3K4me3, H3K9me3, H3K27me3, H3K36me3 and H4K20me3 within ±5 Kb of the TSS of MeCP2 bound genes (bottom). Line graphs showing the average ChIP-seq signal profiles around the TSS of MeCP2 bound genes (top). Genes were sorted from high to low based on MeCP2 peak intensity.

Figure 5. Bubble map of KEGG pathway enrichment analysis for MeCP2 and associated histone methylation marks. The Y-axis shows the names of the enriched pathways, while the size of each node indicates the gene ratio corresponding to MeCP2 and histone methylation marks H3K4me3, H3K9me3, H3K27me3, H3K36me3 and H4K20me3 marks. The P-value is represented by a color scale, with statistical significance increasing from blue (low significance) to red (high significance).

Figure 6. Immunoprecipitation (IP) of MeCP2 from HEK293T cells demonstrates that MeCP2 interacts with histone H3. Mutation of tryptophan at position 104 (W104A) disrupts the MeCP2–histone H3 interaction, whereas mutation of arginine at position 133 (R133A) does not impair histone H3 binding

Figure 7. ChIP-seq analysis of MeCP2 binding to the human genome. (A) IGV snapshots showing the co-localization of MeCP2 and different histone methylation marks at the promoters of MeCP2 target genes. (B) ChIP-qPCR analysis showing reduced MeCP2 localization at target sites upon mutation of the aromatic cage residue (W104A) and the DNA-binding residue (R133C). Data are presented as mean ± SD (n = 3). Statistical significance was determined using Student's t-test (*$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$; ns, not significant).

Figure 1

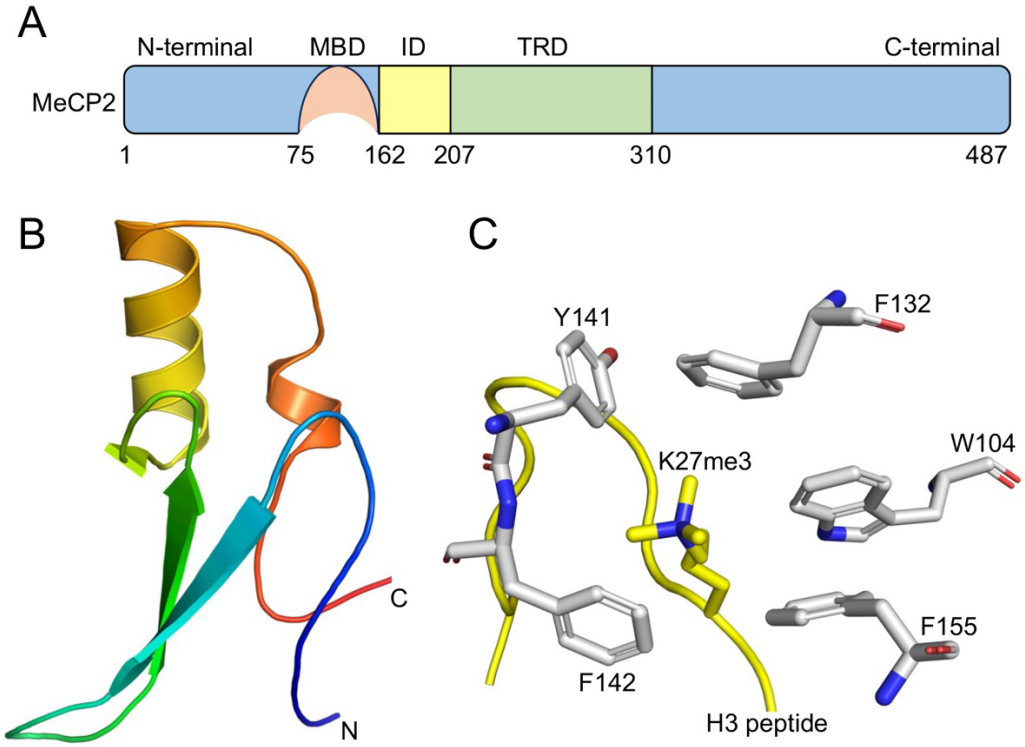

A
N-terminal
MBD
ID
TRD
C-terminal
MeCP2
1
75
162
207
310
487
B
C
N
C
Y141
F132
K27me3
W104
F155
F142
H3 peptide

Figure 2

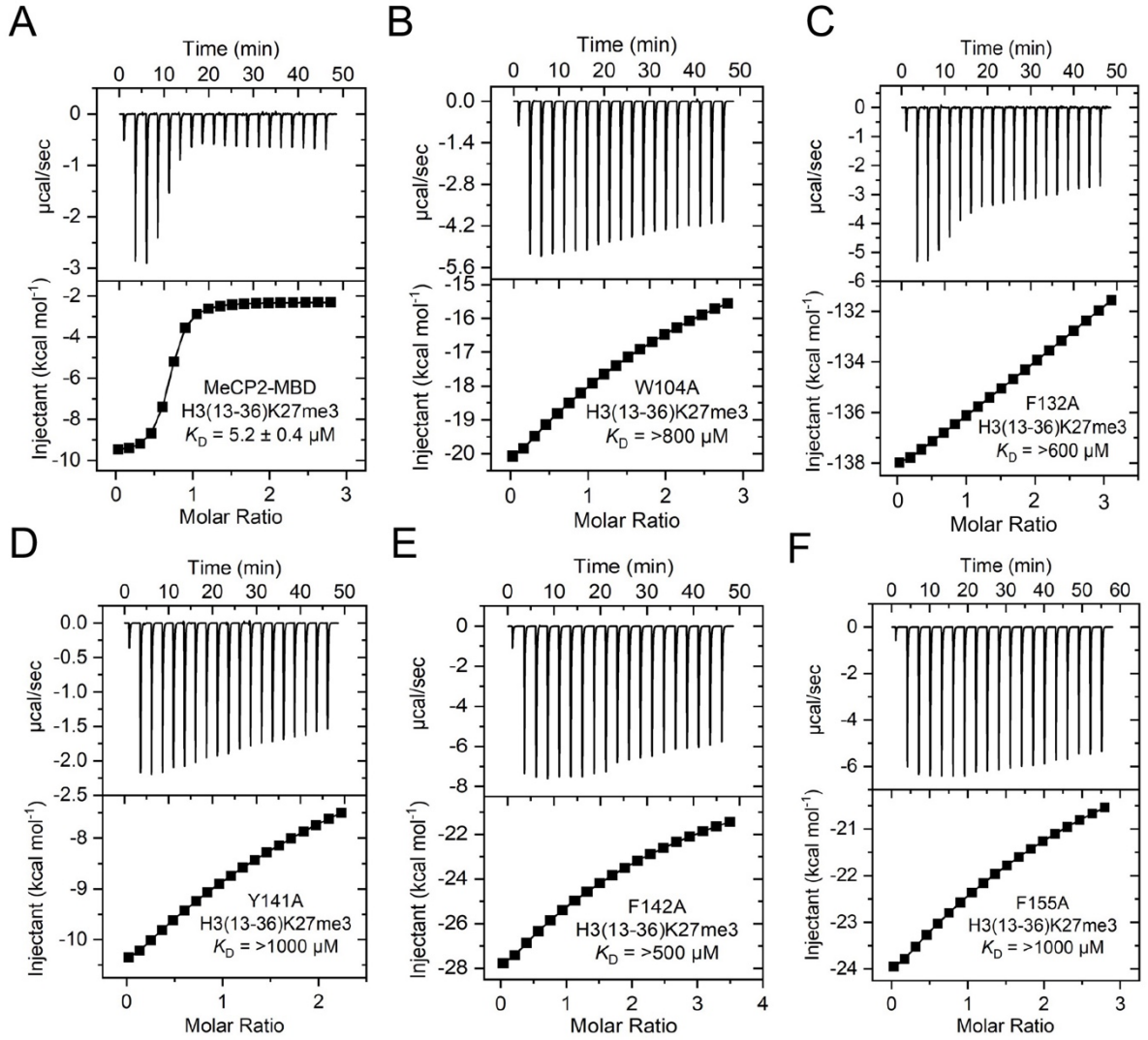

A
Time (min)
μcal/sec
Injectant (kcal mol-1)
MeCP2-MBD
H3(13-36)K27me3
KD = 5.2 ± 0.4 μM
Molar Ratio
B
W104A
H3(13-36)K27me3
KD = >800 μM
C
F132A
H3(13-36)K27me3
KD = >600 μM
D
Y141A
H3(13-36)K27me3
KD = >1000 μM
E
F142A
H3(13-36)K27me3
KD = >500 μM
F
F155A
H3(13-36)K27me3
KD = >1000 μM

Figure 3

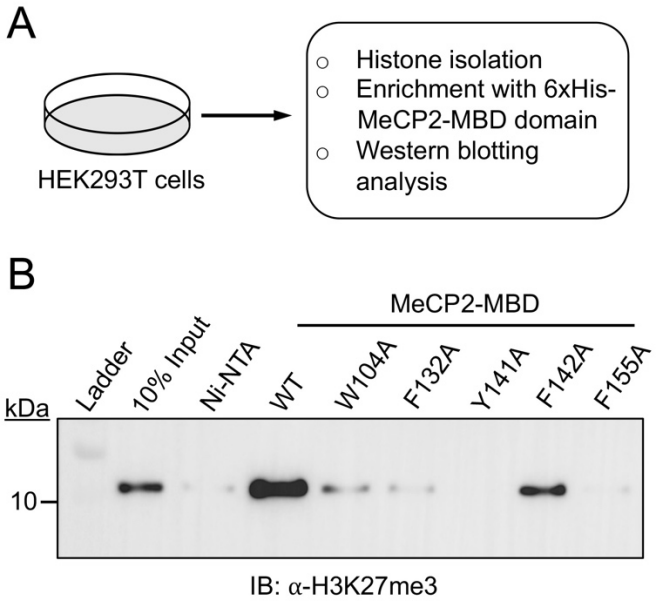

A
HEK293T cells
Histone isolation
Enrichment with 6xHis-MeCP2-MBD domain
Western blotting analysis
B
MeCP2-MBD
Ladder
10% Input
Ni-NTA
WT
W104A
F132A
Y141A
F142A
F155A
kDa
10
IB: α-H3K27me3

Figure 4

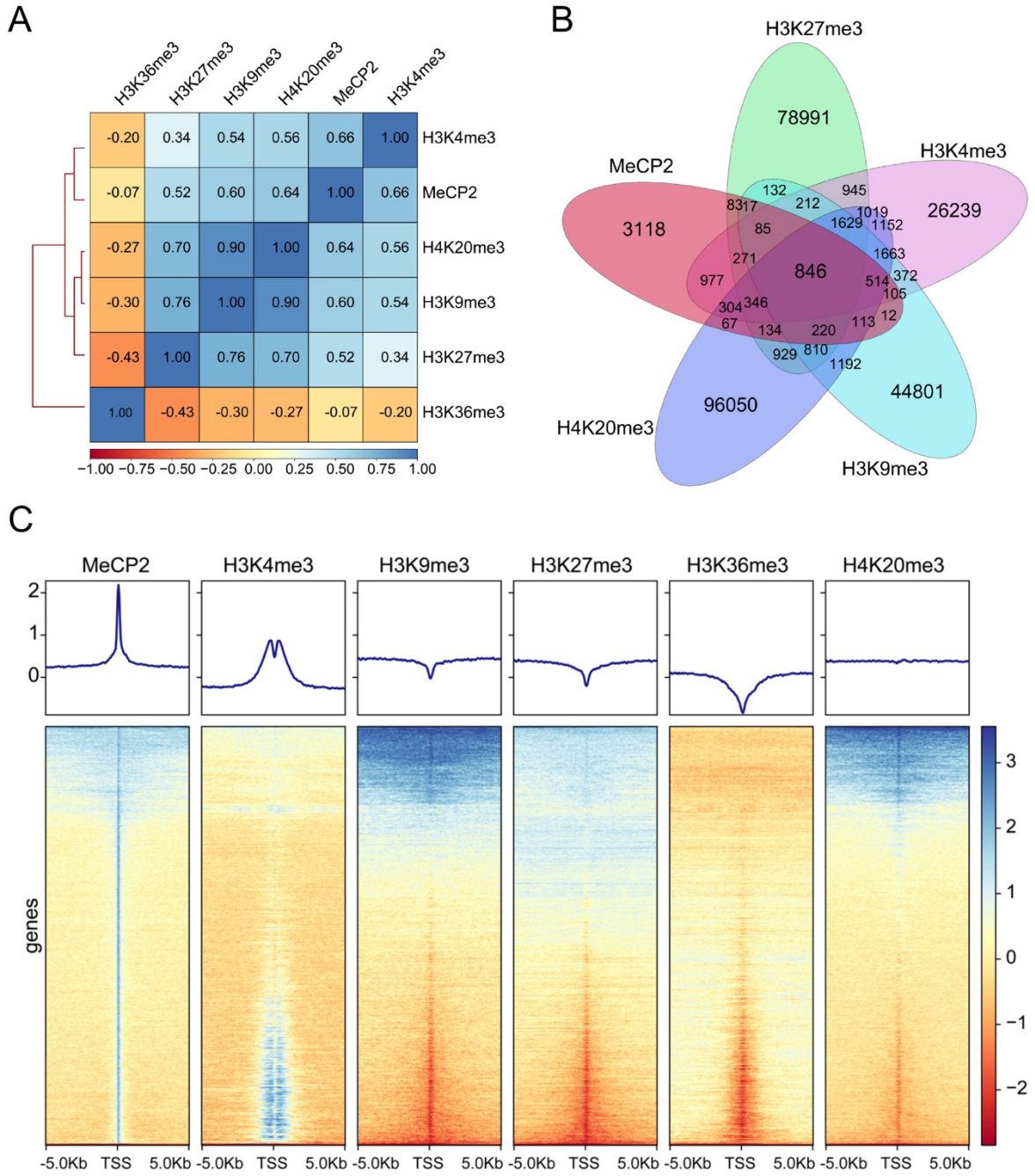

A
H3K36me3
H3K27me3
H3K9me3
H4K20me3
MeCP2
H3K4me3
-0.20 0.34 0.54 0.56 0.66 1.00 H3K4me3
-0.07 0.52 0.60 0.64 1.00 0.66 MeCP2
-0.27 0.70 0.90 1.00 0.64 0.56 H4K20me3
-0.30 0.76 1.00 0.90 0.60 0.54 H3K9me3
-0.43 1.00 0.76 0.70 0.52 0.34 H3K27me3
1.00 -0.43 -0.30 -0.27 -0.07 -0.20 H3K36me3
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
B
H3K27me3
78991
MeCP2
3118
H3K4me3
26239
8317
132
212
945
1019
1629
1152
85
271
1663
977
846
514
372
105
304
346
67
134
220
113
12
929
810
1192
96050
44801
H4K20me3
H3K9me3
C
MeCP2
H3K4me3
H3K9me3
H3K27me3
H3K36me3
H4K20me3
2
1
0
genes
-5.0Kb TSS 5.0Kb
3
2
1
0
−1
−2

Figure 5

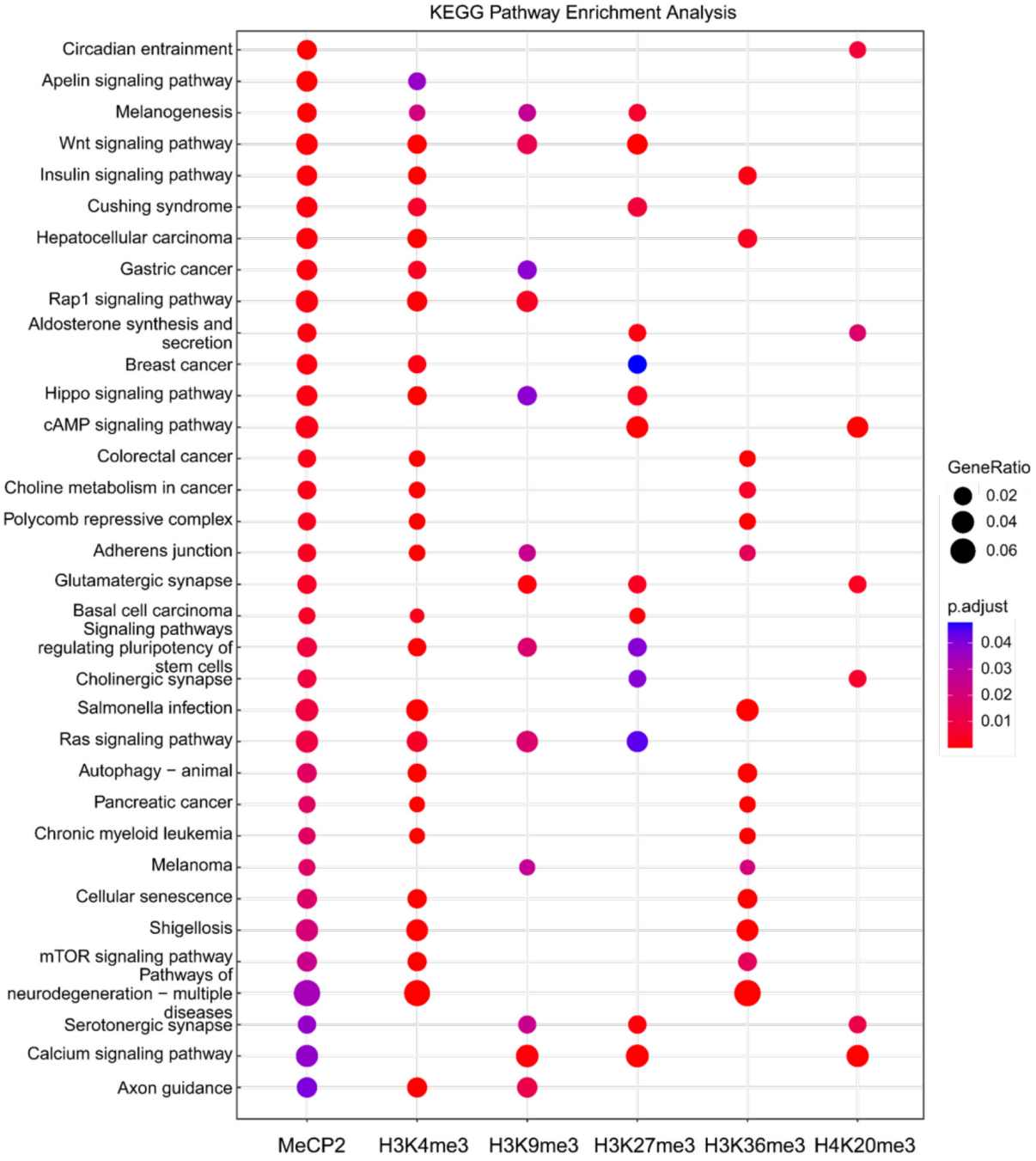

KEGG Pathway Enrichment Analysis
Circadian entrainment
Apelin signaling pathway
Melanogenesis
Wnt signaling pathway
Insulin signaling pathway
Cushing syndrome
Hepatocellular carcinoma
Gastric cancer
Rap1 signaling pathway
Aldosterone synthesis and secretion
Breast cancer
Hippo signaling pathway
cAMP signaling pathway
Colorectal cancer
Choline metabolism in cancer
Polycomb repressive complex
Adherens junction
Glutamatergic synapse
Basal cell carcinoma
Signaling pathways regulating pluripotency of stem cells
Cholinergic synapse
Salmonella infection
Ras signaling pathway
Autophagy − animal
Pancreatic cancer
Chronic myeloid leukemia
Melanoma
Cellular senescence
Shigellosis
mTOR signaling pathway
Pathways of neurodegeneration − multiple diseases
Serotonergic synapse
Calcium signaling pathway
Axon guidance
MeCP2
H3K4me3
H3K9me3
H3K27me3
H3K36me3
H4K20me3
GeneRatio
0.02
0.04
0.06
p.adjust
0.04
0.03
0.02
0.01

Figure 6

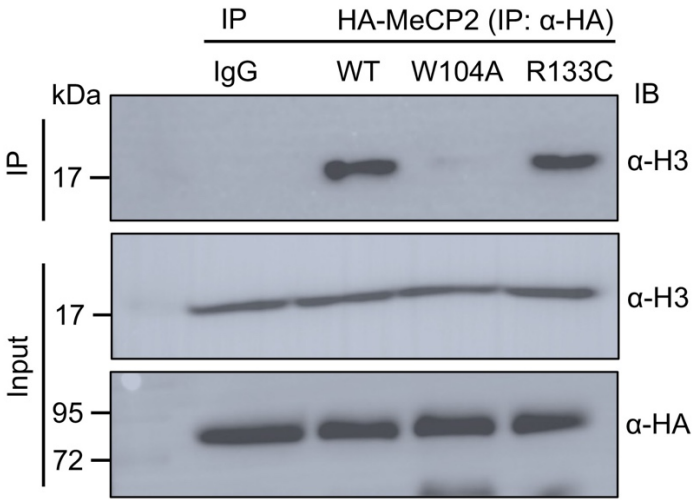
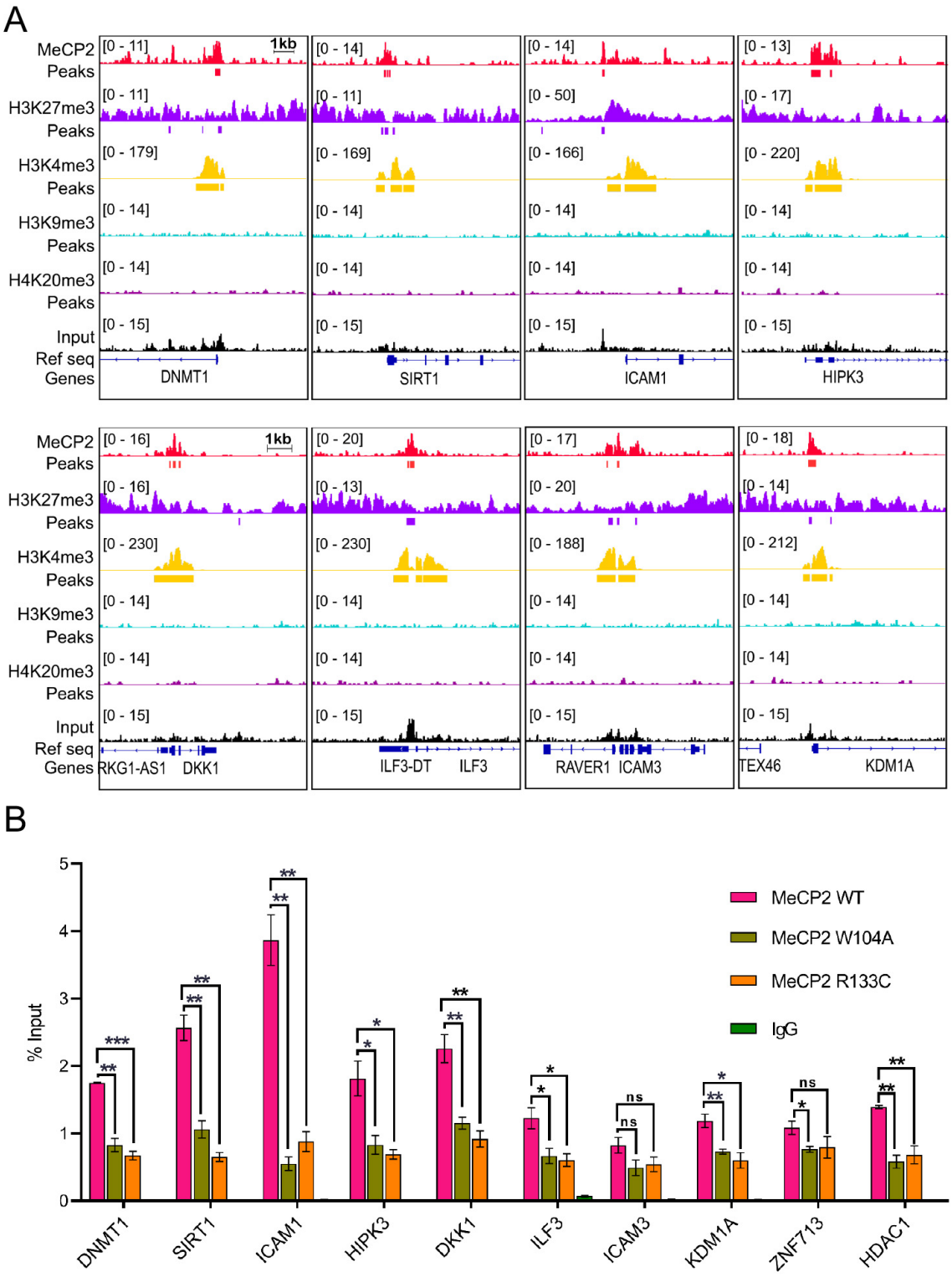

IP
HA-MeCP2 (IP: α-HA)
IgG
WT
W104A
R133C
kDa
IB
IP
17
α-H3
Input
17
α-H3
95
72
α-HA

Figure 7


A
MeCP2
Peaks
H3K27me3
Peaks
H3K4me3
Peaks
H3K9me3
Peaks
H4K20me3
Peaks
Input
Ref seq
Genes
1kb
DNMT1
SIRT1
ICAM1
HIPK3
RKG1-AS1
DKK1
ILF3-DT
ILF3
RAVER1
ICAM3
TEX46
KDM1A
B
% Input
MeCP2 WT
MeCP2 W104A
MeCP2 R133C
IgG
DNMT1
SIRT1
ICAM1
HIPK3
DKK1
ILF3
ICAM3
KDM1A
ZNF713
HDAC1