Phase Separation of the Methyl-CpG-Binding Domain (MBD) Protein Family Underlies Heterochromatin Organization

by Hui Zhang

Darmstadt, 2022

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Phase Separation of the Methyl-CpG-Binding Domain (MBD) Protein Family Underlies Heterochromatin Organization

vom Fachbereich Biologie der Technischen Universität Darmstadt

zur Erlangung des Grades
Doctor rerum naturalium
(Dr. rer. nat.)

Dissertation by

Hui Zhang

Supervisor: Prof. Dr. M. Cristina Cardoso
Prof. Dr. Bodo Laube

Darmstadt. 27.01. 2022

Contents

1 Summary

1.1 Summary

In the last decade, phase separation has been proposed as a mechanism for the formation of a variety of subcellular compartments such as chromatin compartmentalization in eukaryotic cell nuclei. Yet, the requirements for chromatin compartmentalization have not been well characterized and several overlapping factors have been implied. This study aimed to gain insight into the function and underlying mechanisms of the MBD family proteins (MBDs) on chromatin organization.

Firstly, I characterized the phase separation properties, functions, and dynamics of the MBDs *in vitro* and *in vivo*. I found that the MBDs have different phase separation properties and functions in heterochromatin organization. *In vitro*, MBD1ΔCxxC3 formed only irregular aggregates, while MBD2, MBD3, MBD4, and MeCP2 formed liquid-like spherical droplets in certain conditions. *In vivo*, MBD1ΔCxxC3 and MBD3 did not influence heterochromatin compartment size, while MBD1, MBD2, MBD4, and MeCP2 promoted the heterochromatin clustering in a dose-dependent manner. Besides, MBD1, MBD1ΔCxxC3, and MBD4 were shown to decrease the roundness of heterochromatin compartments, while MBD2, MBD3, and MeCP2 did not. I also detected the molecule exchange dynamics of MBDs in between heterochromatin compartments and surrounding environments and found that all MBDs are highly dynamic. In conclusion, the LLPS properties were: MeCP2 > MBD2 > MBD3 > MBD4 ≥ MBD1ΔCxxC3. The functions in heterochromatin compartment size were: MeCP2 > MBD4 > MBD2 > MBD1ΔCxxC3 ≈ MBD3. The functions in heterochromatin compartment roundness were: (MBD3 >) MeCP2 > MBD2 > MBD4 > MBD1ΔCxxC3. The protein mobility was: MeCP2 < MBD2 < MBD1ΔCxxC3 < MBD4. The immobile fractions were: MeCP2 ≈ MBD2 > MBD4 ≈ MBD1ΔCxxC3. Thus, the heterochromatin organization could be regulated by the MBDs mediated phase separation. MBD2 and MeCP2 show the highest liquid-liquid phase separation (LLPS) properties *in vitro* and more significantly influence the heterochromatin compartment structure.

Secondly, I elucidated the functional difference of MBD2 isoforms *in vivo* and *in vitro*. MBD2 has three isoforms, MBD2a, MBD2b, and MBD2c. MBD2a is the longest isoform and contains N-terminus, MBD-TRD, and the C-terminus, while MBD2b and MBD2c contain the conserved MBD-TRD, but lack either the N-Terminus (MBD2b) or the C-Terminus (MBD2c). All MBD2 isoforms showed the ability of phase separation with distinct properties. MBD2b exhibited overall similar but weaker phase separation properties than the full-length MBD2a. MBD2b was less enriched in heterochromatin compartments and promoted the heterochromatin compartment clustering to a lesser extent than MBD2a. MBD2b showed a

faster molecule exchange between heterochromatin compartments and surrounding environments. Compared to MBD2a, MBD2c showed similar molecule dynamics and enrichment in the heterochromatin compartments but did not influence the heterochromatin compartment size. *In vitro*, the purified MBD2c formed only irregular aggregates. Besides, the N-terminus formed only irregular aggregates in all conditions tested and showed the ability to recruit DNA. The MBD-TRD alone did not form any condensates but formed irregular aggregates in the presence of DNA. The C-terminus could form either irregular aggregates or spherical droplets depending on buffer conditions that are not influenced by DNA. Thus, polymer-polymer phase separation (PPPS) also plays a role in MBD2 function, especially for MBD2c. Consistently, high levels of MBD2-MBD-TRD and MBD2c did not influence the total occupancy of heterochromatin while the high levels of MBD2b promoted the spreading of the heterochromatin region. In conclusion, the N-terminus is responsible for protein-DNA interaction, the MBD-TRD is responsible for PPPS and heterochromatin localization, and the C-terminus is responsible for the LLPS. The three isoforms of MBD2 exhibit distinct functional properties due to a combination of the three regions.

Thirdly, I demonstrated that MeCP2 alone could form liquid-like spherical droplets via self-interaction-induced LLPS. The fold enrichment of MeCP2 in the droplets *in vitro* and heterochromatin compartments *in vivo* was similar. MeCP2 LLPS could be promoted by various factors that somehow mimic the *in vivo* physiological conditions. DNA promoted the LLPS of MeCP2 in both DNA length and concentration-dependent manner by providing additional sites for multivalent weak protein-DNA interactions. Crowders promoted the MeCP2 LLPS by increasing the local protein concentration and enhancing the weak interactions via the “excluded volume” effect. The weak MeCP2-MeCP2 and MeCP2-DNA interactions were characterized by the emergence of a properly confined diffusion population in the droplets and heterochromatin compartments. Cytosine methylation restricted the size of MeCP2 condensates *in vivo* and *in vitro* with more confined overall mobility. The specific mC-DNA interaction contributed to the emergence of the static population with low to static diffusion in MeCP2 condensates *in vivo* and *in vitro*. The phase separation properties of MeCP2 were altered by the RTT-related nonsense mutations. *In vivo*, the MeCP2 promoted the heterochromatin compartment clustering, which was attenuated by RTT-related nonsense mutations. The phase separation properties and function of RTT-related nonsense mutations in heterochromatin compartment clustering are negatively correlated with the severity in RTT patients. RTT-related nonsense mutations influence the heterochromatin organization via impaired LLPS properties.

These results provide new insights into the function and underlying mechanism of the MBD protein family in heterochromatin organization and thus contribute to our general understanding of phase separation in nuclei architecture.

1.2 Zusammenfassung

In den letzten zehn Jahren wurde die Phasentrennung als Mechanismus für die Bildung verschiedener subzellulärer Kompartimente vorgeschlagen, z. B. für die Kompartimentierung des Chromatins in eukaryotischen Zellkernen. Die Voraussetzungen für die Chromatinkompartimentierung wurden jedoch nicht gut charakterisiert und es wurden mehrere sich überschneidende Faktoren diskutiert. Ziel dieser Studie war es, einen Einblick in die Funktion und die zugrunde liegenden Mechanismen der Methyl-CpG-bindende Domäne (MBD)-Proteinfamilie bei der Chromatinorganisation zu gewinnen.

Zunächst habe ich die Phasentrennungseigenschaften, Funktionen und Dynamik der Proteine der MBD-Familie *in vitro* und *in vivo* charakterisiert. Ich fand heraus, dass die Proteine der MBD-Familie unterschiedliche Phasentrennungseigenschaften und Funktionen bei der Organisation der Chromozentren haben. MBD1ΔCxxC3, eine Isoform, der die unmethylierte CpG-Bindungsdomäne CxxC3 fehlt, bildete unregelmäßige Aggregate, während MBD2, MBD3, MBD4 und MeCP2 unter bestimmten Bedingungen *in vitro* flüssigkeitsähnliche sphärische Tröpfchen bildeten. *In vivo* kolokalisierten MBD1, MBD2, MBD4 und MeCP2 mit den Chromozentren und förderten dosisabhängig die Chromozentren Clusterbildung. MBD1ΔCxxC3 und MBD3 hingegen hatten keinen Einfluss auf die Größe der Chromozentren. Außerdem zeigte sich, dass MBD1, MBD1ΔCxxC3 und MBD4 die Rundheit der Chromozentren verringerten, während die anderen MBD Proteine keinen Einfluss darauf hatten. Die Messung der Dynamik des Molekülaustauschs der MBDs zwischen den Chromozentren und der Umgebung ergab, dass alle MBDs sehr dynamisch sind. Zusammenfassend lässt sich festhalten, dass die MBD Proteine sich wie folgt in ihrer Fähigkeit zur Flüssig-Flüssig-Phasentrennung (LLPS) unterscheiden: MeCP2 > MBD2 > MBD3 > MBD4 ≥ MBD1ΔCxxC3. Ihre Auswirkung auf die Größe der Chromozentren war in absteigender Reihenfolge: MeCP2 > MBD4 > MBD2 > MBD1ΔCxxC3 ≈ MBD3. Der Einfluss der MBD Proteine auf die Rundheit der Chromozentren war: (MBD3 >) MeCP2 > MBD2 > MBD4 > MBD1ΔCxxC3. Bei der Proteinmobilität unterschieden sich die MBD Proteine wie folgt: MeCP2 < MBD2 < MBD1ΔCxxC3 < MBD4. Bei den immobilen Fraktionen waren: MeCP2 ≈ MBD2 > MBD4 ≈ MBD1ΔCxxC3. Die Organisation der Chromozentren könnte also durch die von den Proteinen der MBD-Familie vermittelte Phasentrennung reguliert werden. MBD2 und MeCP2 zeigen *in vitro* die besten Eigenschaften bei der LLPS und beeinflussen die Chromozentrenstruktur am stärksten.

Im zweiten Teil dieser Thesis habe ich die funktionellen Unterschiede der MBD2-Isoformen *in vivo* und *in vitro* aufgeklärt. MBD2 hat drei Isoformen: MBD2a, MBD2b und MBD2c. MBD2a umfasst den N-Terminus, die MBD-TRD und den C-Terminus, während MBD2b und MBD2c die konservierte MBD-TRD Domäne, aber entweder keinen N-Terminus oder keinen

C-Terminus aufweisen. Alle MBD2 Isoformen zeigten die Fähigkeit zur Phasentrennung, jedoch mit unterschiedlichen Eigenschaften. MBD2b zeigte *in vivo* und *in vitro* insgesamt ähnliche, aber schwächere Funktionen als die Isoform MBD2a. MBD2b war weniger stark in Chromozentren angereichert, förderte die Clusterbildung von Chromozentren in geringerem Maße und zeigte einen schnelleren Molekülaustausch zwischen den Chromozentren und der Umgebung. Im Vergleich zu MBD2a zeigte MBD2c eine ähnliche Moleküldynamik und Anreicherung in den Chromozentren, hatte aber keinen Einfluss auf die Größe der Chromozentren. *In vitro* bildete MBD2c nur unregelmäßige Aggregate. Der N-Terminus allein formte irreguläre Aggregate in allen getesteten Konditionen und zeigte die Fähigkeit, DNA zu rekrutieren. Die MBD-TRD allein formte keine Kondensate, sondern unregelmäßige Aggregate in Gegenwart von DNA. Der C-Terminus konnte abhängig von den Pufferkonditionen und unabhängig von anwesender DNA entweder unregelmäßige Aggregate oder sphärische Tröpfchen formen. Somit spielt die Polymer-Polymer-Phasentrennung (PPPS) eine Rolle bei der Funktion von MBD2, vor allem von MBD2c. Übereinstimmend mit diesen Beobachtungen, hatten auch hohe Level MBD2-MBD-TRD und MBD2c keinen Einfluss auf die Heterochromatinbeschäftigung, während hohe Level MBD2b die Ausweitung der Heterochromatinregion förderten. Zusammenfassend lässt sich festhalten, dass der N-Terminus für die Protein-DNA-Interaktion, die MBD-TRD für die PPPS und die Chromozentren Lokalisation und der C-Terminus für die LLPS verantwortlich ist. Die drei Isoformen von MBD2 weisen unterschiedliche funktionelle Eigenschaften auf, die auf eine Kombination der drei Regionen zurückzuführen sind.

Im dritten Kapitel habe ich gezeigt, dass MeCP2 allein flüssigkeitsähnliche sphärische Tröpfchen durch Selbstinteraktion-induzierte LLPS bilden kann. Die Anreicherung von MeCP2 in den Tröpfchen *in vitro* und den Chromozentren *in vivo* war ähnlich. Die LLPS von MeCP2 wurde durch verschiedene Faktoren gefördert, die die physiologischen Bedingungen *in vivo* nachahmen. Die Anwesenheit von DNA förderte die LLPS von MeCP2 sowohl in Abhängigkeit von der DNA-Länge als auch von der Konzentration, indem sie zusätzliche Stellen für multivalente schwache Protein-DNA-Wechselwirkungen bereitstellte. Crowder förderten die LLPS von MeCP2, indem sie die lokale Proteinkonzentration erhöhten und die schwachen Wechselwirkungen durch den "excluded volume"-Effekt verstärkten. Die MeCP2-MeCP2- und MeCP2-DNA-Wechselwirkungen wurden durch das Auftreten einer begrenzten Diffusionspopulation in den Tröpfchen und in den Chromozentren charakterisiert. Cytosin-Methylierung schränkte die Größe der MeCP2-Kondensate *in vivo* und *in vitro* ein und führte zu einer eingeschränkten Gesamtmobilität. Die spezifische mC-DNA-Wechselwirkung trug zum Entstehen der statischen Population mit geringer bis statischer Diffusion in MeCP2-Kondensaten *in vivo* und *in vitro* bei. RTT-bedingte Nonsense-Mutationen veränderten die Phasentrennungseigenschaften von MeCP2. *In vivo* förderte MeCP2 die Clusterbildung von

Chromozentren, die durch RTT-bezogene Nonsense-Mutationen stark abgeschwächt wurde. Die Phasentrennungseigenschaften und die Funktion von RTT-bedingten Nonsense-Mutationen bei der Chromozentren Clusterbildung sind negativ mit dem Schweregrad der Erkrankung von RTT-Patienten korreliert. RTT-bedingte Nonsense-Mutationen beeinflussen die Chromozentrum Organisation durch beeinträchtigte LLPS-Eigenschaften.

Diese Ergebnisse geben neue Einblicke in die Funktion und den zugrundeliegenden Mechanismus der MBD-Proteinfamilie in der Heterochromatin Organisation und tragen somit zum Gesamtverständnis der Phasentrennung im Zellkern bei.

2 Introduction

2.1 Chromatin organization and phase separation

2.1.1 Chromatin organization

In eukaryotic cells, chromatin is hierarchically organized into distinct domains with different epigenetic modifications, gene expression profiles, and chromatin dynamics (Botchkarev, Gdula et al. 2012). Based on the chromatin-chromatin interactions, chromatin is subcompartmentalized into topologically associating domains (TADs) (Dixon, Gorkin et al. 2016). Within TAD, loops are formed and maintained via loop extrusion (Fudenberg, Imakaev et al. 2016). Multiple TADs assemble to form A/B compartments (Fudenberg, Imakaev et al. 2016). A compartments are hallmarked by gene-rich, active transcription activities, and interior localization within nuclei, while B compartments are hallmarked by gene-poor and gene silencing lies on the nuclear periphery or nuclear center. Based on the compaction levels, the chromatin is divided into two components: euchromatin and heterochromatin. *In vivo*, the two components could be distinguished by differential DNA staining under the microscope (Fig. 1).

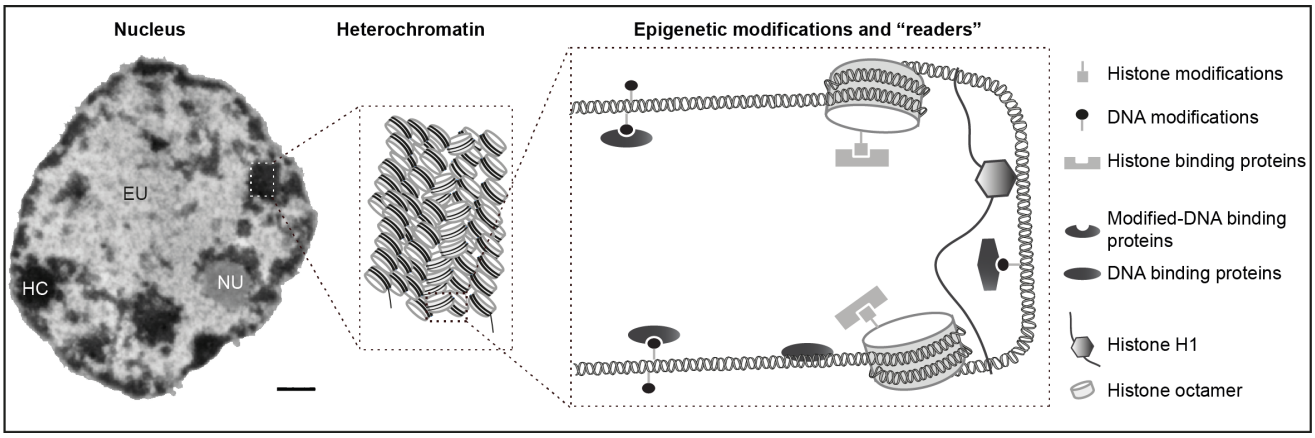


Figure 1 Heterochromatin recognition and epigenetic modifications.

Left: transmission electron microscopy (EM) image of a mouse liver cell nucleus with electron-dense regions corresponding to heterochromatin. Scale bar = 0.5 μm. Middle: graphical representation of densely packed nucleosomes within heterochromatin. Right: cartoon showing epigenetic modifications and relative "readers" within heterochromatin regions. NU: nucleolus; EU: euchromatin; HU: heterochromatin. Symbols are as indicated.

Euchromatin is composed of mainly transcriptional active genes hallmarked with low DNA methylation and high histone H3 lysine 4 acetylation (H3K4ac) (Grewal and Moazed 2003, Guillemette, Drogaris et al. 2011). Heterochromatin is the highly compacted form of chromatin with restricted DNA accessibility and includes the mainly inactive satellite sequences, such as major satellite DNA. Heterochromatin is hallmarked by high DNA methylation, trimethylation at histone 3 lysine 9 and 27 (H3K9me3 and H3K27me3) (Hennig

1999, Grewal and Moazed 2003, Allshire and Madhani 2018, Poleshko, Smith et al. 2019). Heterochromatin is located in the nuclear periphery and regions surrounding the nucleolus (Fig. 1).

Heterochromatin plays an essential role in nuclear architecture, genome stability, et al. (Peters, O'Carroll et al. 2001, Yan, Lim et al. 2011, Janssen, Colmenares et al. 2018, Greenstein and Al-Sady 2019). In mouse, Drosophila and plants, the pericentromeric heterochromatin aggregates to form large heterochromatin compartments (also called chromocenters, or heterochromatic foci), associated with an enrichment of heterochromatin related proteins. The composition, function, and dynamics of heterochromatin compartments have been investigated (Peters, O'Carroll et al. 2001, Cheutin, McNairn et al. 2003, Simon and Meyers 2011, Yan, Lim et al. 2011, Isaac, Sanulli et al. 2017, Johnson and Straight 2017, Johnson, Yewdell et al. 2017, Strom, Emelyanov et al. 2017, Erdel, Rademacher et al. 2020, Li, Coffey et al. 2020). Dysregulation of the heterochromatin compartments is associated with senescence and aging, which are hallmarked by deregulated lamin B1 (LMNB1) (Shah, Donahue et al. 2013). Low LMNB1 leads to loss of H3K9me3 modifications at the nuclear periphery and contributes to the reactivation of inflammatory genes (Chen, Zheng et al. 2014). Besides, disease-causing lamin A (LMNA) mutations lead to globally reduced levels of LMNB1, HP1α, H3K9me3, and H3K27me3 (McCord, Nazario-Toole et al. 2013, Fernandez, Scaffidi et al. 2014, Loi, Cenni et al. 2016).

Nevertheless, the underlying mechanisms regulating the formation, maintenance, and/or dynamics of heterochromatin are still far from clear. Recent findings indicate a role of (liquid-liquid) phase separation on nuclei compartmentalization.

2.1.2 Liquid-liquid phase separation

The liquid-liquid phase separation (LLPS) is originally applied in membrane science to describe and characterize the demixing process during asymmetric membrane formation in polymer solution (Koenhen, Mulder et al. 1977). *In vitro*, the lysozyme was also shown to undergo demixing and form liquid droplets with high lysozyme concentrations, which promoted the crystallization process (a process called phase transition) (Muschol and Rosenberger 1997). Yet, the involvement of LLPS in cellular activities was not described until 2009, when Brangwynne et al. proposed that germline P granules are liquid-like spherical droplets with fast exchange dynamics, fusion, and fission properties (Brangwynne, Eckmann et al. 2009). This indicates that P granules are probably regulated by LLPS.

Now we are realizing that LLPS might be a general mechanism underlying the organization of various membraneless organelles with well-delineated physicochemical boundaries, but without phospholipid membrane barriers, such as germ granules, stress granules, nucleoli, centrosomes. Such membraneless organelles are essential for normal cellular activity,

whose disruption is increasingly linked to human disease as reviewed in recent papers (Fig. 2) (Shin and Brangwynne 2017, Babinchak and Surewicz 2020, Zbinden, Pérez-Berlanga et al. 2020, Lu, Qian et al. 2021, Pakravan, Orlando et al. 2021).

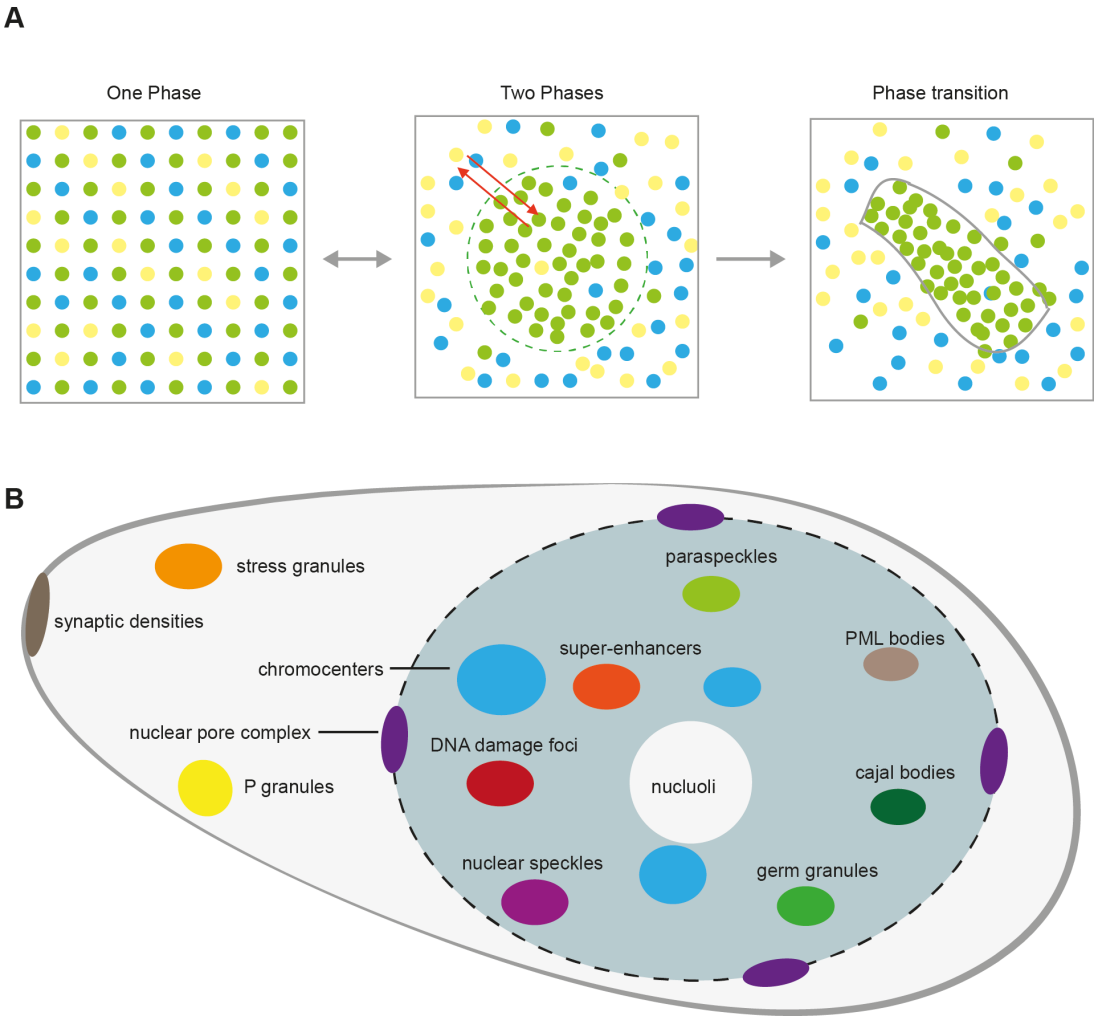


Figure 2 Liquid-liquid phase separation (LLPS) *in vitro* and *in vivo*.

(A) In a mixture containing different molecules, unstructured protein(s) demix to form two phases, the dense phase and the dilute phase. The dense phase (droplets) shows the properties of reversibility, spherical morphology, and dynamic molecule exchange between the two phases. At certain conditions (such as proteins with certain disease-related mutations), the liquid-like droplets further transmit into irregular aggregates with slow/no molecule dynamics and are irreversible.
(B) LLPS is now characterized as a common mechanism underlying the formation and dynamic regulation of membraneless organelles within the cytoplasm and nucleus.

Proteins that could undergo phase separation often contain intrinsically disordered regions (IDRs) or low complexity regions (LCRs) (Oldfield and Dunker 2014). LCRs have biased amino acid preferences and thus show distinct charge clusters. For instance, the LCR would be positively charged if enriched with positively charged amino acids, such as arginine, lysine, and histidine, and therefore promotes electrostatic interactions among molecules.

The LLPS is driven by multivalent weak hydrophobic interactions and/or electrostatic interactions including charge-charge, charge-π, π-π stacking interactions, and hydrogen

bonds (Li, Banjade et al. 2012, Brangwynne, Tompa et al. 2015, Molliex, Temirov et al. 2015, Krainer, Welsh et al. 2021).

The droplets formed by LLPS are always reversible, while they can undergo a further transition to less reversible gel-like and, at extreme conditions, solid condensates. This process is named phase transition. The driving force underlying phase transition is not clear, but likely sequence coded. Phase transition plays a role in cellular activity. For example, in yeast, poly(A)-binding protein (Pab1) undergoes LLPS under heat shock and forms grainy, amorphous particles with little or no dynamics at higher temperature and lower pH (Riback, Katanski et al. 2017). This suggests that the gel-like condensates are possibly an adaptive response during stress.

2.1.3 LLPS in chromatin compartmentalization and human disease

Recent studies focus on LLPS as a mechanism of chromatin compartmentalization.

Now, an increasing amount of chromatin-binding proteins are shown to bear the ability of LLPS, indicating that LLPS might be a general mechanism underlying chromatin compartmentalization.

Nucleosome arrays (NAs) were shown to undergo histone tail-dependent liquid-like phase separation in physiologic salt conditions (Gibson, Doolittle et al. 2019). This could be promoted by histone H1, controlled by linker DNA length, and disrupted by histone acetylation. Furthermore, NAs with acetylated histones could form a new liquid phase with multi-bromodomain proteins, and these droplets had distinct properties compared to droplets formed by unmodified histones. This indicates a role of LLPS in the segregation of euchromatin and heterochromatin.

In the euchromatin region, LLPS plays a role in local enrichment of certain factors that are essential for certain nuclear activities, including DNA replication, transcription, damage repair, and alternative splicing (Sabari, Dall'Agnese et al. 2018, Parker, Bell et al. 2019, Wang, Gao et al. 2019, Laflamme and Mekhail 2020, Levone, Lenzken et al. 2021, Pessina, Gioia et al. 2021, Reber, Jutzi et al. 2021). For example, the super-enhancer-associated transcription coactivators BRD4 and MED1 could form liquid-like condensates *in vivo* and *in vitro* in an IDR-dependent manner (Sabari, Dall'Agnese et al. 2018). The coactivators could recruit multiple transcriptional factors, such as OCT4, SOX2, and NANOG and transcriptional machinery (Boija, Klein et al. 2018, Sabari, Dall'Agnese et al. 2018). Besides, disrupting the phase separation properties of OCT4 with MED1 via acidic mutation (mutate negatively charged aspartic/glutamic acids to non-polar alanine) also decreased the ability of OCT4 in gene activation (Boija, Klein et al. 2018). This further links the LLPS with transcription activation.

In the heterochromatin region, heterochromatin protein 1α and β (HP1α and β), H3K9me3 readers, were shown to drive heterochromatin compartment formation via LLPS in a dimerization dependent manner and could be enhanced by phosphorylation and interactive partners like H3K9 trimethylated histones (Bannister, Zegerman et al. 2001, Larson, Elnatan et al. 2017, Strom, Emelyanov et al. 2017, Sanulli, Trnka et al. 2019, Qin, Stengl et al. 2021). Multivalent interactions of HP1β with H3K9me3-modified nucleosomes via its chromodomain enables phase separation and contribute to the formation of heterochromatin compartments *in vivo* (Qin, Stengl et al. 2021). Moreover, *in vitro* droplets formed by HP1β in the presence of H3K9me3 and SUV39H1, an HP1β binder and H3K9me3 enzyme, showed the ability to exclude active chromatin-related proteins, suggesting a role of LLPS in heterochromatin segregation from euchromatin (Wang, Gao et al. 2019). Besides HP1, LLPS of Polycomb complex PRC1 also plays an important role in polycomb-induced heterochromatin compaction and gene repression.

Abnormal phase transition is tightly correlated with human disease. Take Parkinson's disease as an example, which is hallmarked by α-Synuclein (α-Syn) aggregation and amyloid formation (Lücking and Brice 2000). Both *in vivo* and *in vitro* results show that α-Syn could form liquid-like droplets, which eventually transform into amyloid-hydrogel-containing oligomers and fibrillar species (Ray, Singh et al. 2020). In particular, familial Parkinson's disease mutations promote the LLPS and phase transition to aggregated states (Ray, Singh et al. 2020). Similar phenomena are also observed for tau, another component involved in Alzheimer's disease (Kanaan, Hamel et al. 2020).

2.1.4 Polymer-polymer phase separation (PPPS) and heterochromatin organization

Polymer-polymer phase separation (PPPS) promotes the formation of ordered collapsed globules in the presence of polymer scaffold via proteins that could bridge units within polymers in close spatial proximity and assemble cross-links (Erdel and Rippe 2018). Both LLPS and PPPS are driven by multivalent interactions but with different properties (Fig. 3) (Brangwynne, Tompa et al. 2015, Erdel and Rippe 2018, Wang, Cairns et al. 2019). In LLPS, protein-protein interactions contribute to the liquid-like droplet formation, and the protein-polymer interactions are not necessary. While in PPPS, the protein-polymer interactions promote the formation of ordered globules, and the protein-protein interactions are not necessary. Thus, upon removal of polymers, the PPPS mediated condensates will fall apart, while LLPS mediated droplets persist. Conceptually, the PPPS is similar to the coil-globule transition, a phenomenon that in solution, the polymers collapse into polymer-rich globules via polymer-polymer interaction (Moore 1977).

Although both LLPS and PPPS promote the formation of dynamic spherical condensates, LLPS and PPPS show distinct functional implications, such as size, density, and stability of

condensates (Erdel and Rippe 2018). Bodies formed by LLPS and PPPS show distinct responses to concentration changes. For PPPS, increasing the total concentration of bridging factors promotes the enrichment of the factors at binding sites, but the size of globules does not change till saturated concentration. Oppositely, for LLPS, the droplets become bigger at higher protein concentrations, while the protein concentrations remain invariant. Besides, condensates formed by PPPS and LLPS show distinct internal structures, molecular transport, coalescence, et al. For example, LLPS predicts a homogeneous molecule distribution while PPPS predicts a local enrichment of bridging factors if the binding sites within polymers are not homogeneously distributed.

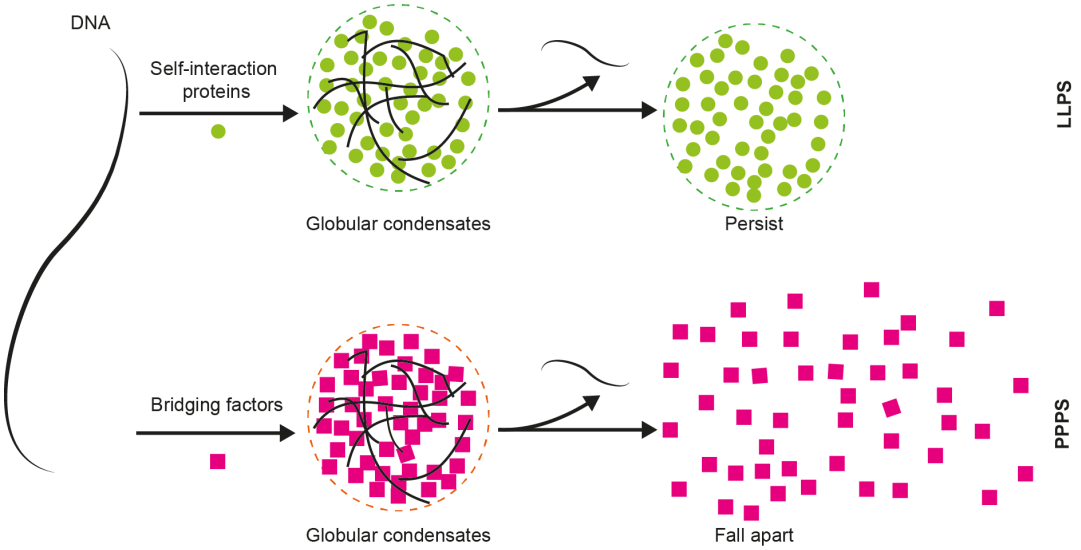


Figure 3 Comparison of liquid-liquid phase separation and polymer-polymer phase separation.

Unstructured proteins with the ability of self-interaction form liquid-like spherical droplets regardless of DNA (up). While proteins containing multiple DNA binding sites could condensate the DNA into globules. Upon removal of the DNA scaffold, the condensates fall apart (bottom).

LLPS and PPPS mechanisms are not mutually exclusive. Chromatin is the most common polymer in cell nuclei, which could be recognized by chromatin-binding proteins. Some of such proteins retain both bridging and protein-protein interaction functions, such as HP1 (heterochromatin protein 1) and MeCP2 (methyl CpG binding protein 2). HP1 exhibits both homodimerization and binding affinity with methylated H3 at lysine 9 (H3K9me3), a heterochromatin marker (Bannister, Zegerman et al. 2001). Indeed, several observations did not support the LLPS mechanism underlying HP1 function in heterochromatin compartments. Firstly, all three HP1 homologs showed no predominant enrichment (2~3 fold) in heterochromatin compared to chromatin compaction levels (~2 fold) in NIH-3T3 fibroblasts (Müller‐Ott, Erdel et al. 2014). Secondly, HP1α is less expressed *in vivo* compared to heterochromatin-associated proteins such as MeCP2 and SUV39H1(Müller‐Ott, Erdel et al. 2014). Thirdly, HP1α does not influence the heterochromatin clustering and

compaction (Müller‐Ott, Erdel et al. 2014). Lastly, the purified human HP1α could not form liquid-like droplets *in vitro* in the absence of DNA (Strom, Emelyanov et al. 2017, Erdel, Rademacher et al. 2020). The observation that HP1 could form bridges between nucleosomes suggests an alternative PPPS mechanism of HP1 function in heterochromatin compartalization. Indeed, Erdel et al. show that the size, global accessibility, and compaction of heterochromatin foci are independent of HP1, suggesting that HP1 might also adopt the PPPS mechanism in heterochromatin compartment formation (Erdel, Rademacher et al. 2020).

2.2 DNA methylation and readers

2.2.1 Overview of the epigenome

"An epigenetic trait" was defined as "a stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence" (Berger, Kouzarides et al. 2009). The epigenetic modifications involve DNA methylation, histone modifications, histone variants, nucleosome positioning, and non-coding RNAs, among which the covalent DNA methylation and histone modifications are the best-studied epigenetic modifications.

Epigenomics plays a crucial role in shaping the developmental decisions of multicellular organisms, human cancers, and other diseases. During development, the cells start from pluripotent stem cells, differentiate into various cell types with progressively restricted pluripotency, together with restricted gene expression patterns due to global epigenetic modifications (Reik, Dean et al. 2001, Reik 2007, Feng, Jacobsen et al. 2010). Genomic imprinting is an epigenetic phenomenon that leads to gene expression in a parent-of-origin-specific manner and is thus important during normal development. Loss of imprinting contributes to several inherited genetic diseases, such as Beckwith–Wiedemann syndrome (Walter and Paulsen , Robertson 2005, Kalish, Jiang et al. 2014). In cancers, genome-wide hypomethylation and gene-specific hypermethylation events are observed simultaneously in the same cell (Robertson 2005).

2.2.2 Cytosine methylation

DNA methylation was discovered as early as 1946 (McCarty and Avery 1946), and cytosine methylation was discovered in 1948 (Hotchkiss 1948). Subsequently, various studies demonstrated that cytosine methylation plays essential roles in multiple cellular processes including genome regulation, development, and disease (Walsh and Bestor 1999, Robertson and Wolffe 2000, Chen, Akbarian et al. 2001, Vaillant and Paszkowski 2007, Lister and Ecker 2009, Zhang, Kimatu et al. 2010, Severin, Zou et al. 2011, Schubeler 2015, Angeloni and Bogdanovic 2019). Besides, adenine was reported to be methylated in *E. coli* (Dunn and Smith 1955), but the underlying mechanism is still poorly understood. DNA methylation

exists in both prokaryotic and eukaryotic cells and occurs predominantly at cytosine sites and rarely at adenine sites in mammals (Vanyushin 2005, Iyer, Zhang et al. 2016, Wu, Wang et al. 2016, Koh, Goh et al. 2018, Xiao, Zhu et al. 2018, Beaulaurier, Schadt et al. 2019, Woodcock, Horton et al. 2020). Here I will review the observations on cytosine methylation.

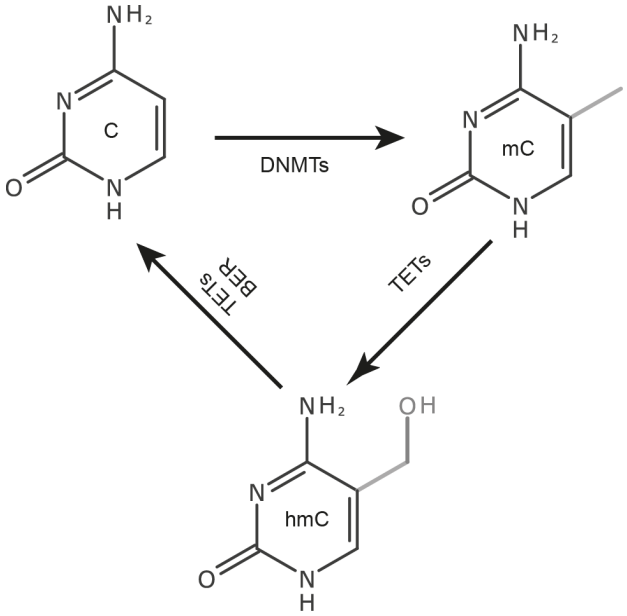


Figure 4 Graphic summary of cytosine methylation and demethylation.

The cytosine methylation at the 5′ carbon (5mC) is catalyzed by DNMT enzymes and could be oxidized to 5hmC by TET enzymes. 5hmC undergoes further oxidization and base-excision repair (BER) to convert back to an unmodified cytosine (C).

Cytosine methylation is also one of the best-studied and most mechanistically elucidated epigenetic modifications and is well conserved among most plant, animal, and fungal models. The cytosine methylation by adding a methyl group to the carbon 5 position of cytosine (5mC) is catalyzed by the DNA methyltransferase Dnmts (Dnmt1, Dnmt3a/b) (Fig. 4) (Ludwig, Zhang et al. 2016). Cytosine methylation has been involved in various pathways, such as gene transcription (Dolinoy, Weidman et al. 2007, Reik 2007), DNA damage repair (Rossetto, Truman et al. 2010, Williamson, Zhu et al. 2018), chromatin remodeling (Williamson, Zhu et al. 2018, Li, Xia et al. 2019), chromatin accessibility (Klemm, Shipony et al. 2019, Li, Wu et al. 2021) and chromatin organization (Noma, Allis et al. 2001, Richards and Elgin 2002, Grewal and Moazed 2003, Ng, Robert et al. 2003, Huisinga, Brower-Toland et al. 2006). Cytosine methylation plays an essential role during normal development. Transgenic mice expressing decreased but not completely absent DNA methyltransferase (MTase) activity showed abnormal development and embryonic lethality (Li, Bestor et al. 1992, Yoder, Walsh et al. 1997). Others showed that during the early development, cells undergo removal and reestablishment of genome-wide methylome, which as a result influences the gene imprinting, cell fate decision, gene expression et al. (Razin, Webb et al.

1984, Brandeis, Ariel et al. 1993, Smith and Meissner 2013). Dysregulation of cytosine methylation is correlated with developmental disorders and diseases including cancer (Gopalakrishnan, Van Emburgh et al. 2008, Smith and Meissner 2013, Vukic and Daxinger 2019, Yuan and Huang 2019, Chen, Miao et al. 2020).

2.2.3 CpG methylation and function

Cytosine methylation is primarily restricted to the CpG dinucleotides (mCpG) and occupies around 4% of all cytosine (~ 1% of all nucleic acid bases) and 60%-80% of all CpG dinucleotides in most vertebrates (Ehrlich, Gama-Sosa et al. 1982, Lister, Mukamel et al. 2013, Smith and Meissner 2013, Bachman, Uribe-Lewis et al. 2014).

In vitro, mCpG stabilizes the DNA helix, increases its melting temperature, and reduces DNA helicase and RNA/DNA polymerase speed (Rausch, Zhang et al. 2021). This suggests a role of mCpG in regulating replication and transcription speed. At the single molecular level, Cassina et al. found that methylation produces measurable incremental changes in persistence lengths, together with unchanged DNA contour lengths (Cassina, Manghi et al. 2016). This indicates that DNA methylation-induced DNA conformational modification may play a role in the binding affinities of protein-DNA interactions. Cytosine methylation also promotes the condensation of chromatin fibers in the presence of linker histone (Karymov, Tomschik et al. 2001), indicating a role of DNA methylation on heterochromatin compaction *in vivo*. However, *in vivo* studies by Gilbert, et al. showed that DNA methylation affects the nuclear organization and nucleosome structure but not chromatin compaction. Yet, we showed before that DNA methylation binding protein MeCP2 promotes the chromocenter (heterochromatin compartment) fusion (clustering) in mouse cell nuclei (Brero, Easwaran et al. 2005, Bertulat, De Bonis et al. 2012, Casas-Delucchi, Becker et al. 2012). These studies indicate that DNA might, directly and indirectly, regulate chromatin dynamics by itself and multiple methylation binding proteins (methylation “readers”).

In vivo, mCpG is tightly correlated with gene repression and is mainly located at the promoters of transposable elements, imprinting genes, and tissue-specific genes with lower CpG densities. Transposable elements occupy around 45% of the mammalian genome but their replication may lead to gene disruption and DNA mutation due to insertion (Ukai, Ishii‐Oba et al. 2003). Thus, the transposable elements are commonly silenced by bulk methylation (Schulz, Steinhoff et al. 2006). Imprinted genes are expressed from only one of the two inherited parental chromosomes which are regulated by DNA methylation and histone modifications. For example, during male gametogenesis, the methylation at the maternal allele of the H19 locus was delayed, thus the maternal allele retains transcriptional activity (Lee, Singh et al. 2010). Yet, methylation-independent imprinting was also reported (Inoue, Jiang et al. 2017).

mCpG also occurs in gene bodies, but the functional consequences are still fuzzy or may vary. The housekeeping genes are generally more highly methylated at the gene bodies than the tissue-specific genes, indicating that the methylation at the gene bodies correlates with transcription activity (Aran, Toperoff et al. 2011, Muyle and Gaut 2019). One hypothesis is that DNA methylation at the gene body enhances splicing accuracy, reduces the accumulation of histone variant H2A.Z, and/or establishes constitutive expression patterns within housekeeping genes (Entrambasaguas, Ruocco et al. 2021). But contrary results showed that loss of gene body methylation does not affect the gene expression, the enrichment of histone variant H2A.Z, and histone modification profiles in *E. salsugineum* and *Arabidopsis thaliana* epigenetic recombinant inbred lines. Thus, if/how gene body DNA methylation functions is still an open question (Bewick, Ji et al. 2016).

mCpG is lacking from the CpG dense regions (CpG islands). CpG island consists of around one thousand base pairs with enrichment for CpG dinucleotides and is always located at the promoters of many housekeeping genes (Bird, Taggart et al. 1985). The hypomethylation at CpG islands is maintained by several possible mechanisms. Firstly, the binding of transcription factors could exclude the binding and enrichment of DNA methyltransferase and thus prevent the *de novo* methylation at the CpG islands. This is confirmed by the result that truncating or deleting the binding sites of transcription factors promoted the methylation at CpG islands (Lienert, Wirbelauer et al. 2011). Secondly, histone modifications could influence the hypomethylation state of CpG islands. H3K4 methylation inhibits the recognition and binding between Dnmt3 and histone H3 (Ooi, Qiu et al. 2007). Thirdly, other factors might participate in the maintenance of CpG island hypomethylation, such as TET dioxygenases and base excision repair (BER) enzymes (Fig. 4) (Tahiliani, Koh et al. 2009, Bellacosa and Drohat 2015, Ko, An et al. 2015, Nissar, Kadla et al. 2021). Although CpG islands at the promoters are rarely methylated, the CpG islands at the intragenic and gene body regions show tissue-specific methylation patterns. During tumorigenesis and aging, these CpG islands are frequently hypermethylated due to methylation-induced rapid and terminal heterochromatin formation.

2.2.4 Non-CpG methylation and function

The non-CpG methylation (mCH, H = A, T, C) was firstly discovered in 1970 in newborn mice and mouse embryo cell cultures (Salomon and Kaye 1970). Subsequent research using the same method confirmed the existence of non-CpG methylation in mouse fibroblasts and human spleen (Grafstrom, Yuan et al. 1985, Woodcock, Crowther et al. 1987). Woodcock et al. also indicated that 54.5% genome-wide cytosine methylation was present in mCpA, mCpT, and mCpC (Woodcock, Crowther et al. 1987). But more precise technologies are required for further characterization of such non-CpG methylation. Applying

a combination of dual-labeling nearest neighbor technique and bisulfite genomic sequencing methods, Ramsahoye et al. found that, in embryonic stem cells, non-CpG methylation was significantly enriched at CpA and, to a lesser extent, at CpT, which was catalyzed by Dnmt3a. In somatic tissues, non-CpG methylation was less detectable (Ramsahoye, Biniszkiewicz et al. 2000, Ziller, Muller et al. 2011). Yet, the first genome-wide single-base-resolution map of methylated cytosines in a mammalian genome was not presented until 2009 (Lister, Pelizzola et al. 2009). Lister et al. also indicated that ~ ¼ of cytosine methylation was identified in non-CpG sites in human embryonic stem cells, which disappeared upon differentiation. These data suggest that embryonic stem cells may use different methylation mechanisms to regulate gene expression. The non-CpG methylation was mainly enriched in gene bodies and depleted from protein binding sites and enhancers (Lister, Pelizzola et al. 2009). It is clear now that the non-CpG methylation is universally distributed in the mammalian genome and enriched in oocytes, ES cells, neurons, and glial cells (Lister, Pelizzola et al. 2009, Xie, Barr et al. 2012, Lister, Mukamel et al. 2013, Guo, Su et al. 2014).

The underlying mechanism of non-CpG methylation is still unclear. One hypothesis is that the non-CpG methylation is a by-product of the hyperactivity of non-specific *de novo* methyltransferases Dnmt3a and Dnmt3b (Ziller, Muller et al. 2011, Lee, Park et al. 2017). Indeed, Dnmt3a and Dnmt3b could preferentially methylate cytosines in CAC and CAG contexts, respectively (Lee, Park et al. 2017).

Non-CpG methylation is recognized as a key epigenetic mark in mammalian embryonic stem cells (ESCs) and neurons, regulating cell type-specific functions. In human and mouse neurons, the non-CpG methylation occupies around 53% and 38% of all cytosine methylation, respectively (Lister, Mukamel et al. 2013). During human synaptogenesis, highly conserved non-CG methylation accumulates in neurons and is enriched in genes escaping X-chromosome inactivation (Lister, Mukamel et al. 2013). These indicate an essential role of non-CpG methylation in brain development. Non-CpG methylation could promote and repress gene expression depending on the binding partners. MeCP2 could recognize non-CpG methylated sites, recruit corepressors such as SIN3 transcription regulator (Sin3a) and histone deacetylases (HDACs) and induce gene repression (Guo, Su et al. 2014). In ESCs, histone H3 tri-methylated at lysine 36 (H3K36me3) recruits Dnmt3b to the actively transcribed genes, resulting in hyper-methylation at non-CpG sites. These transcriptional active genes with high non-CpG methylation are involved in embryo development (Lee, Park et al. 2017). Besides, the mitochondrial genomes are also highly methylated at the non-CpG sites and the methylation patterns display notable differences in mitochondrial genomes between normal and cancer cells. This suggests a function of non-CpG methylation in cancer (Patil, Cuenin et al. 2019).

2.2.5 DNA methylation readers

In mammals, DNA methylation is recognized by various methyl-CpG binding proteins (MBPs) (Ludwig, Zhang et al. 2016). Among various MBPs with different structures (Ludwig, Zhang et al. 2016), the methyl-CpG binding domain (MBD) containing protein family is the first discovered and well-characterized (Lewis, Meehan et al. 1992, Tate and Bird 1993, Cross, Meehan et al. 1997, Hendrich and Bird 1998). The MBD family proteins (MBDs) are characterized by the shared 70-85 amino acid-long methyl-CpG binding domain (MBD) with specific mCpG binding affinity for all MBDs except MBD3. The MBDs are highly conserved inside the MBD, but share little similarities outside the MBD, indicating distinct and possibly complementary functions among MBDs (Hendrich and Bird 1998). The MBDs are now thought to function as a scaffold protein bridging DNA methylation and binding partner such as histone deacetylase, indicating a role of MBDs in gene expression, epigenetic modifications, and chromatin organization.

Eleven members of MBDs have been discovered and characterized, including MeCP2, MBD1-6, SETDB1/2, and BAZ2A/B (Fig. 5), and will be introduced in detail as follows.

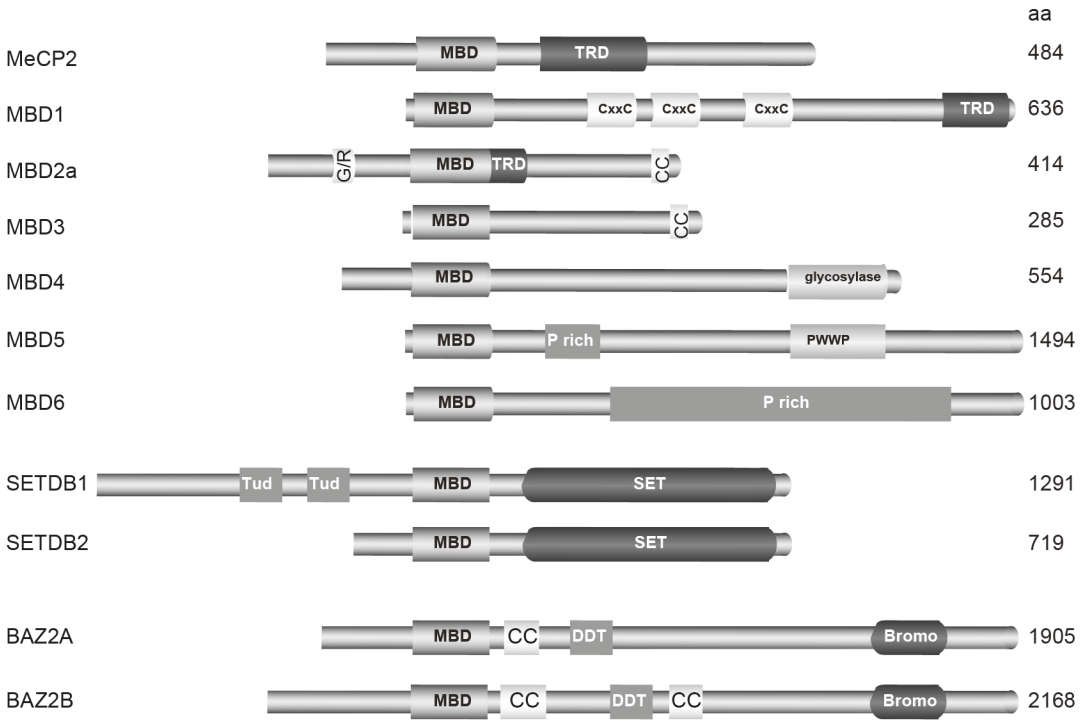


Figure 5 Schematic graph showing the structures of methyl-CpG binding domain-containing protein family.

MBD: methyl-CpG binding domain; NTD: N-terminal domain; TRD: transcriptional repression domain. CxxC: unmethylated CpG binding zinc finger; G/R: glycine/arginine; CC: coiled-coil domain; Glycosylase: DNA glycosylase function; P rich: Proline-rich domain; PWWP: Proline-tryptophan-tryptophan-proline; Tud: Tudor domain; SET: Suvar3-9, Enhancer-of-zeste, Trithorax domain; DDT: DNA binding domain; Bromo: Bromodomain.

2.3 MeCP2

2.3.1 MeCP2 isoforms and structure

The Methyl-CpG Binding Protein 2 (MeCP2) gene is highly conserved in Euteleostomi (bony vertebrates) and humans and is located on the X chromosome. The MeCP2 protein has two isoforms (MeCP2 e1 (exon 1) and MeCP2 e2 (exon 2)) with different amino termini due to alternative splicing and different translational start sites (Fig. 6). MeCP2-e1 is 498 amino acid-long with 21 unique N-terminal residues. MeCP2-e2 is 486 amino acids long with 9 unique N-terminal residues leading to a partially different N-terminal domain (NTD). The remaining protein sequence keeps the same to both isoforms with two functionally characterized domains: the methyl-CpG binding domain (MBD) and the transcriptional repression domain (TRD). The MBD specifically recognizes and binds mCpG, while the TRD was found to bind multiple transcriptional repressors, thus silencing gene expression (Jones, Veenstra et al. 1998, Nan, Ng et al. 1998, Kokura, Kaul et al. 2001, Lunyak, Burgess et al. 2002, Suzuki, Yamada et al. 2003, Forlani, Giarda et al. 2010). However, the TRD was also shown to bind to multiple transcriptional activators and activate gene expression (Chahrour, Jung et al. 2008, Gonzales, Adams et al. 2012, Leoh, van Heertum et al. 2012). Thus, the function of MeCP2 in gene expression is context-dependent. More recently, the TRD has been narrowed down to the N-CoR/SMRT interacting domain (NID) (Lyst, Ekiert et al. 2013). The MBD and TRD (or NID) are linked by the intervening domain (ID). Both isoforms end with a long C-terminal domain (CTD).

The two isoforms of MeCP2 are abundantly expressed in the central nervous system, but with different expression levels and distributions in developing and postnatal mouse brains. MeCP2 e1 is the predominant isoform in the brain and has an earlier expression onset than MeCP2 e2 (Olson, Zachariah et al. 2014). The two isoforms are commonly considered functionally equivalent, yet recent evidence shows that MeCP2 e1 plays a role in neuronal maturation (Li, Wang et al. 2013) and is more relevant for RTT (Fichou, Nectoux et al. 2009, Saunders, Minassian et al. 2009, Sheikh, de Paz et al. 2017).

As MeCP2 e2 was the first isoform to be known and commonly explored during the past, the coding sequence from MeCP2 e2 was applied throughout the work.

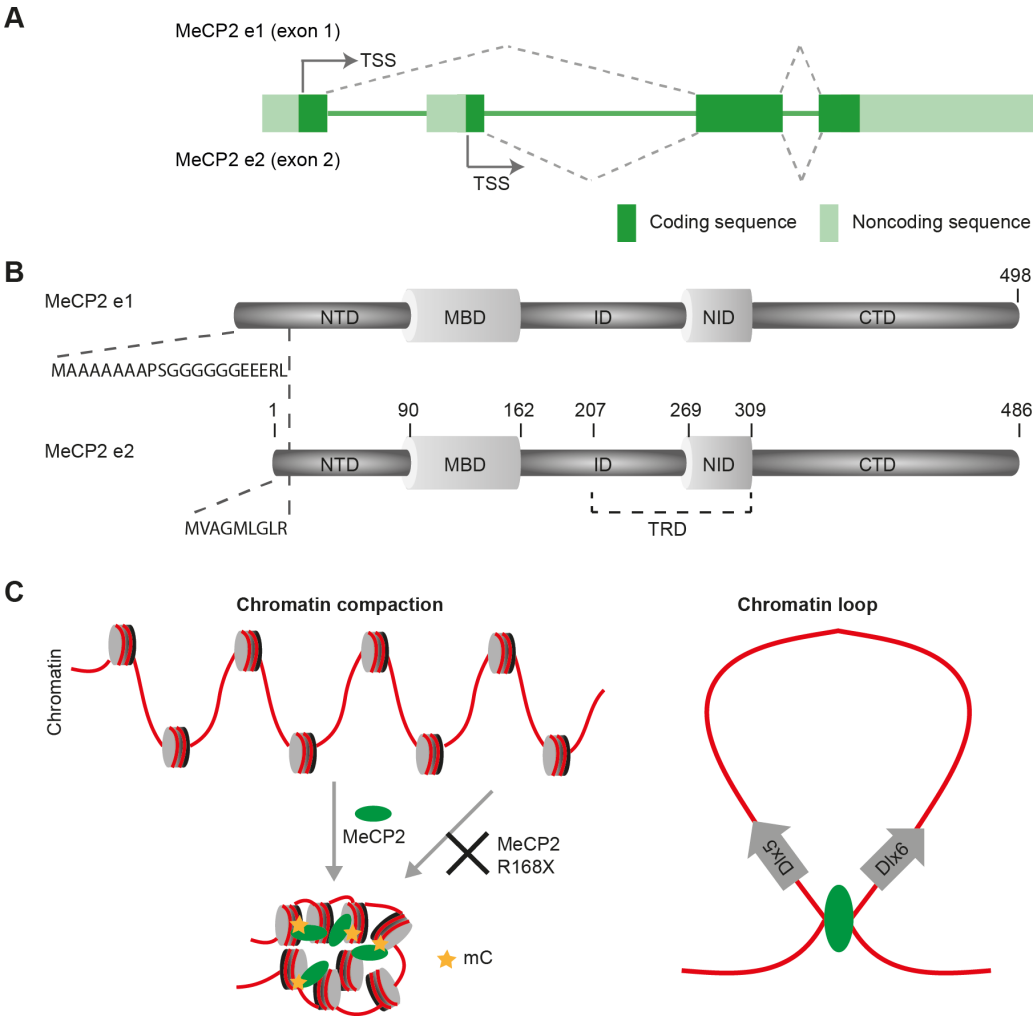


Figure 6 Schematic representation of MECP2 structure and function in chromatin organization.

(A) MeCP2 gene organization. TSS: translational start site.
(B) Protein structure of MeCP2 two isoforms: MeCP2 e1 (exon 1) and MeCP2 e2 (exon 2). NTD: N-terminal domain; MBD: methyl binding domain; ID: intervening domain; NID: N-CoR/SMRT interacting domain; CTD: C-terminal domain; TRD: transcriptional repression domain.
(C) MeCP2 binding to 5mC induces nucleosome arrays (NAs) compaction. While the nonsense mutation MeCP2 R168X failed to compact NAs.
(D) MeCP2 promotes silent chromatin loop formation in the imprinted Dlx5-Dlx6 locus.

2.3.2 DNA binding and genome-wide distribution

MeCP2 was originally recognized to bind to symmetrically methylated CpG sites (mCpG) via the MBD domain (Lewis, Meehan et al. 1992, Nan, Meehan et al. 1993). The mCpG binding affinity could be enhanced by the flanking regions of MBD (Ghosh, Nikitina et al. 2010). The specific MBD-mCpG interaction is mediated by three amino acid residues: D121, R111, and R133 (Ho, McNae et al. 2008). Further point mutations at D121, R111, and R133 significantly reduced MeCP2-mCpG binding affinity (Chahrour, Jung et al. 2008, Agarwal, Becker et al. 2011, Casas-Delucchi, van Bemmel et al. 2012). *In vivo*, using genome-wide chromatin immunoprecipitation-sequencing (ChIP-seq) analysis, MeCP2 was found to bind globally across the genome and tracked the mCpG density (Skene, Illingworth et al. 2010).

The MBD was also reported to bind to asymmetric methylated CpG sites (Buchmuller, Kosel et al. 2020) and non-CpG methylation sites (mCpH, H = A, T, C) although with a higher preference for mCAC over all other forms (Chen, Chen et al. 2015, Gabel, Kinde et al. 2015, Lagger, Connelly et al. 2017, Sperlazza, Bilinovich et al. 2017, Liu, Xu et al. 2018). Distinct from other mammalian somatic cell types, post-mitotic neurons accumulate high levels of mCpH, especially mCAC, suggesting an essential function of mCAC in neuron maturation and functioning (Xie, Barr et al. 2012, Varley, Gertz et al. 2013, Guo, Su et al. 2014). Recent research by Tillotson and collaborators showed that such non-CpG methylation is an essential target of MeCP2 function in neurons, whose lack would lead to severe phenotypic features that largely mirrored those seen in mouse models of RTT (Lagger, Connelly et al. 2017, Tillotson, Cholewa-Waclaw et al. 2020). RTT (Rett syndrome) is a neurological disorder caused by MeCP2 dysfunction (Amir, Van den Veyver et al. 1999, Chen, Akbarian et al. 2001, Guy, Hendrich et al. 2001, Collins, Levenson et al. 2004, Van Esch, Bauters et al. 2005, Samaco, Fryer et al. 2008).

The MeCP2 could also bind to hydroxymethyl-cytosine (5hmC), an oxidative intermediate of 5mC catalyzed by the Ten-eleven translocation dioxygenase (Tet). In the brain, 5hmC accounts for ~40% of all cytosine modification in the brain (Kriaucionis and Heintz 2009, Szulwach, Li et al. 2011). The binding affinity of MeCP2 to 5hmC is lower than 5mC but higher than unmodified DNA (Buchmuller, Kosel et al. 2020). Further, the R133 was reported to directly bind to 5hmC, as mutations at the R133 site (R133C) abolished the specific binding with 5hmC (Mellén, Ayata et al. 2012). Besides, the binding of MeCP2 with 5hmC was reported to function in gene activation (Mellén, Ayata et al. 2012).

The MeCP2 was also reported to bind to unmethylated DNA, which is mediated by the intervening domain (ID), TRD, and C-terminal domain (CTD) alpha (aa:310-355 (Nikitina, Shi et al. 2007, Ghosh, Nikitina et al. 2010). Furthermore, three AT-hook-like domains were identified within the ID, TRD, and CTD domains (AT-hook 1, aa 184–195; AT-hook 2, aa 264–273; AT-hook 3, aa 341–364). The AT-hook motif is a short motif binding to the minor groove of AT-rich DNA via the core consensus amino acid sequence RGRP (Lyst, Connelly et al. 2016). In the cell nuclei, the MeCP2 AT-hook domains were shown to play important roles in MeCP2 mobility (Piccolo, Liu et al. 2019).

These methylation-dependent and independent DNA binding capabilities allow MeCP2 to bind to different sites on the DNA under different contexts and thus, possibly contribute to genome-wide chromatin organization. *In vivo*, MeCP2 binding across the genome might reduce the transcriptional noise. MeCP2 was reported to inhibit the oxidation of 5mC. The MeCP2-mC binding protects the methylated sites from TET binding, prevents the oxidation of mC to 5hmC, and therefore restricts transcriptional noise (Skene, Illingworth et al. 2010, Ludwig, Zhang et al. 2017).

To summarize, MeCP2 could bind to DNA of various contexts with various functions in nuclei metabolism and chromatin organization.

2.3.3 MeCP2 on chromatin architecture

MeCP2 is a multifunctional epigenetic reader and is regulated at multiple levels including specific isoforms, interacting factors, post-translational modifications, and their interplay within the chromatin context (Schmidt, Zhang et al. 2020). Yet, it is not well understood how MeCP2 orchestrates nuclei structure.

In neurons, MeCP2 expression is nearly as abundant as histone octamers (Skene, Illingworth et al. 2010). In MeCP2 deficient neurons, the level of histone H1 doubled (Skene, Illingworth et al. 2010) whereas, in wild-type neurons, the H1 level was half of the amount of H1 in other cells (Pearson, Bates et al. 1984), indicating that MeCP2 may act as a histone H1-like chromatin linker. Accordingly, MeCP2 was shown: to accelerate histone H1 exchange *in vivo* and decrease histone H1 dwell time in chromatin (Ghosh, Horowitz-Scherer et al. 2010); to have a similar mobility to H1 *in vivo*; and to share an overlapping binding site on nucleosomes with H1 *in vitro* (Misteli, Gunjan et al. 2000, Kumar, Kamboj et al. 2008, Ghosh, Horowitz-Scherer et al. 2010, Skene, Illingworth et al. 2010). In fact, by *in vitro* fluorescence anisotropy assays, it was observed that MeCP2 could replace histone H1 from chromatin (Nan, Campoy et al. 1997, Ghosh, Horowitz-Scherer et al. 2010) and globally alter the chromatin state.

MeCP2 regulates global epigenetic modifications. MeCP2 deficiency was reported to affect global chromatin composition and state by increasing H3 acetylation (Skene, Illingworth et al. 2010). MeCP2 was also shown to increase H3K9me2 at the promoter of the SIRT1 (Wang, Wang et al. 2018), indicating a role of MeCP2 in facultative heterochromatin regulation. Hence, MeCP2 could probably dampen transcriptional noise from repetitive DNA elements including satellite DNA in a DNA methylation-dependent manner (Skene, Illingworth et al. 2010). On the other hand, MeCP2 was also reported to activate gene expression by binding and recruiting the transcription activator CREB1 in euchromatin (Chahrour, Jung et al. 2008).

MeCP2 is enriched at pericentric heterochromatin (chromocenter) and proposed to participate in the chromocenter organization (Lewis, Meehan et al. 1992). Indeed, Brero et al. showed that, during myogenic differentiation, the number of chromocenters decreased concomitantly with increased MeCP2 level and genome methylation (Brero, Easwaran et al. 2005). Of note, ectopic MeCP2-YFP could promote pericentric heterochromatin clustering even in the absence of cellular differentiation.

Further, the role of MeCP2 during neuronal differentiation was analyzed comparing wild type and MeCP2 deficient mouse embryonic stem cells (mESCs) (Bertulat, De Bonis et al. 2012).

An increased MeCP2 level and enrichment at chromocenters were measured during neuronal differentiation, together with significant chromocenter clustering. Accordingly, the chromocenter clustering function was impaired during the differentiation of MeCP2 deficient mESCs. Furthermore, ectopic expression of MeCP2 with RTT mutations showed impaired heterochromatin accumulation and decreased chromatin clustering function (Agarwal, Becker et al. 2011), suggesting a role of heterochromatin organization in RTT.

MeCP2 was described to compact nucleosomal arrays (Georgel, Horowitz-Scherer et al. 2003) *in vitro* (Nikitina, Ghosh et al. 2007). Georgel et al. in 2003 observed using electron microscopy that NAs formed both extensively condensed ellipsoidal particles and oligomeric superstructures upon addition of MeCP2. This was independent of DNA methylation and relied upon regions downstream of MBD, as R168X nonsense mutation failed to assemble oligomeric superstructures (Georgel, Horowitz-Scherer et al. 2003, Nikitina, Ghosh et al. 2007). This was further confirmed by the observation that the ID, TRD, and C-terminal domain alpha (CTD alpha) could bind and compact NAs and that R270X and R273X, truncated within the TRD and missing the whole CTD, could not compact and oligomerize NAs (Ghosh, Nikitina et al. 2010, Baker, Chen et al. 2013). These facts could in part explain how nonsense mutations of MeCP2 lead to severe symptoms of RTT.

Besides, MeCP2 was proposed to be involved in the formation of a silent chromatin loop at the imprinted Dlx5-Dlx6 locus which was lost in RTT (Horike, Cai et al. 2005). Current models though propose that the chromatin loops are promoted by ‘loop extrusion’, where cohesin extrudes chromatin until it encounters boundaries created by CTCF (CCCTC-binding factor) binding (Sanborn, Rao et al. 2015, Fudenberg, Imakaev et al. 2016), albeit the underlying mechanism is unclear. MeCP2 has been reported to interact with ATRX and cohesin subunits SMC1 and SMC3 using coimmunoprecipitation experiments in the mouse forebrain (Kernohan, Jiang et al. 2010). ATRX was proposed to create an extended DNA linker region for CTCF binding (Kernohan, Vernimmen et al. 2014), and CTCF was reported to promote loop formation (Lewis and Murrell 2004, Nativio, Wendt et al. 2009). Of note, the interaction of MeCP2 with cohesin subunit SMC3 was found to be induced by S229 phosphorylation and inhibited by the S80 phosphorylation of MeCP2 (Gonzales, Adams et al. 2012), indicating a role of MeCP2 and its modifications on chromatin looping.

MeCP2 was described to form loops involving undersaturated (DNA partially occupied by nucleosomes) nucleosomal arrays *in vitro* (Nikitina, Shi et al. 2007). Wild type MeCP2 was shown to form nucleosome-MeCP2-nucleosome ‘sandwich’ structures bringing two nucleosomes closely together. While the RTT truncation mutant R294X was shown to form DNA-MeCP2-DNA ‘stem’ motifs, bringing nucleosome entry and exit site in close proximity (Nikitina, Shi et al. 2007). Interestingly, the RTT mutation R106W, which does not bind to methylated DNA (Table 4), did not induce any chromatin conformations. Thus, MeCP2 loop

formation was proposed to proceed in a two-step process involving methylation-dependent DNA binding followed by methylation-independent interactions between MeCP2 CTD and nucleosomes (Nikitina, Shi et al. 2007). Of note, MeCP2 was also shown to bind to four-way junction DNA, which has a similar conformation as the 'stem' motif (Galvao and Thomas 2005, Nikitina, Shi et al. 2007).

It is still far from clear how MeCP2 organizes heterochromatin structure, but emerging evidence suggests a role of phase separation in heterochromatin condensation.

2.3.4 Dysfunction

MeCP2 was shown to be associated with the neurological disorder Rett syndrome (RTT), as mutations in this gene were found in about 80% of Rett patients (Amir, Van den Veyver et al. 1999). RTT affects mostly young girls and is characterized by normal development until 7-18 months of age, followed by developmental stagnation and decline of higher brain functions (Hagberg, Aicardi et al. 1983). Mutations causing RTT and related neurological disorders have been identified along the entire MeCP2 locus, but effects vary depending on the mutation type and location. Missense and nonsense mutations are the most commonly found and relatively well studied. A collection of all RTT-related mutations can be found in the online RettBASE: RettSyndrome.org (http://mecp2.chw.edu.au/cgi-bin/mecp2/search/printGraph.cgi).

In the following, we will concentrate on RTT mutations impacting MeCP2 DNA binding and chromatin organization.

MeCP2 RTT-related missense mutations are largely found in the MBD, and a large proportion of these mutations reduce the 5mC binding affinity and, consequently, lead to impaired heterochromatin organization and function in cells (Agarwal, Becker et al. 2011).

MeCP2 R133 and R111 residues located within the MDB directly contact 5mC, and mutations at either site decrease MeCP2 localization at heterochromatin *in vivo* albeit to different extents. MeCP2 R111G is a rare RTT mutation found only in one patient, which abolishes MeCP2 localization to heterochromatin (Agarwal, Becker et al. 2011). MeCP2 R133 mutation influences the pericentric heterochromatin localization depending on the amino acid substitution. MeCP2 R133C and R133L decrease the enrichment at heterochromatin, whereas R133H promotes it (Agarwal, Becker et al. 2011, Casas-Delucchi, Becker et al. 2012). Furthermore, artificially targeting MeCP2 R111G and R133L mutants to pericentric heterochromatin rescued their ability to cluster heterochromatin (Casas-Delucchi, Becker et al. 2012).

In addition to missense mutations, several nonsense RTT mutations have been described within the ID or the TRD and account for around 15% of RTT-causing mutations. In general,

these truncations showed decreased protein stability *in vivo* and DNA binding affinity *in vitro* (Table 1) (Yusufzai and Wolffe 2000).

Table 1: Summary of high-frequency RTT-related MeCP2 nonsense mutations and phenotypes

Mutation	Frequency	Effect on: mice, cell, protein	References
R168X	364	**Mice**: Breathing dysfunction, hind limb clasping, and atrophy, hypoactivity. Decreased life span of ~12 weeks. **Male mice**: Impaired motor and cognitive function and reduced anxiety, abnormal hypoxic and hypercapnic responses, apnea incidence, irregular breath cycle and decreased breathing rate, enriched outside heterochromatin. **Protein**: Decreased chromatin compaction ability, decreased methyl-DNA binding.	Lawson-Yuen et al. 2007 Schaevitz et al. 2013 Bissonnette et al. 2014 Georgel et al. 2003 Yusufzai et al. 2000
R255X	313	**Mice**: Decreased brain weight, increased breathing, incidence of arrhythmia, anxiety, motor and learning impairments. **Cell**: mTORC1 pathway abnormalities, decreased nucleolin level, increased phosphorylation of mTORC2 (S2481) and mTORC1 (S2448). **Protein**: Decreased methyl-DNA binding.	Pitcher et al. 2015 Olson et al. 2018 Yusufzai et al. 2000
R270X	274	**Male**: Severe neonatal encephalopathy and death before 4 years of age. **Mice**: Median life span of 85 days, increased body weight, decreased brain weight. **Cell**: Less athalassemia/mental retardation syndrome X linked (ATRX) foci. **Protein**: Decreased methyl-DNA binding, failed to form a higher order structure with nucleic acids, and reduced activity to oligomerize nucleic acids.	Villard et al. 2007 Baker et al. 2013 Yusufzai et al. 2000
G273X	1	**Mice:** Longer life span than R270X mice, bigger brain than that in R270X mice but smaller than that in WT. **Cell:** Impaired function in heterochromatin compartment clustering **Protein**: Impaired binding activity with NAs (better than R270X)	Baker et al. 2013
R294X	237	**Cell**: Induce caspase-mediated apoptosis, rescued by FoxG1. **Protein**: Decreased methyl-DNA binding; decreased stability.	Lundvall et al. 2006 Yusufzai et al. 2000

X means point mutation generating a truncated protein. Mutation numbering according to human MeCP2 isoform starting in exon 2.

MeCP2 R168X generates a truncated protein with a deletion of the complete TRD and C-terminal region. Male and female mice with R168X expression showed typical RTT phenotype, but little is known about the underlying mechanism. Although the entire MBD is retained, MeCP2 R168X has an impaired ability to form higher-order structures as tested by *in vitro* nucleosomal array (NA) assays (Georgel, Horowitz-Scherer et al. 2003).

MeCP2 R255X generates a truncated protein retaining partial ID domain, which was shown to promote the binding affinity of MBD to mCpG. RTT patients with R255X showed a 100% skewed XCI pattern in leukocytes (Nielsen, Henriksen et al. 2001). Mice with R255X showed typical RTT phenotypes, including decreased brain weight, increased breathing, the incidence of arrhythmia, anxiety, motor and learning impairments (Pitcher, Herrera et al. 2015). Molecularly, MeCP2 R255X increased the phosphorylation levels of mTORC2

(S2481) and mTORC1 (S2448), leading to mTORC1 pathway abnormalities (Pitcher, Herrera et al. 2015, Olson, Pejhan et al. 2018).

Besides, other nonsense mutations like R270X and R294X are frequently found in RTT patients with severe phenotypes. Moreover, Baker and coworkers generated two mouse models bearing the R270X and G273X, which mimic the RTT-causing R279fs and G273fs, separately (Baker, Chen et al. 2013). They found that the G273X had a delayed disease onset and longer lifespan compared to the R270X mice. In comparison, the G273X bears the AT-hook2 (aa: 264–273) which is impaired in R270X mice. Thus, the AT-hook2 seems essential for the MeCP2 function. Besides, the MECP2-R270X mice exhibited a reduction in heterochromatic foci ellipticity in CA1 pyramidal cells at age 8-9 weeks, while the MECP2-G273X mice showed no significantly reduced heterochromatic foci ellipticity in CA1 pyramidal cells until 5 months old (Ito-Ishida, Baker et al. 2020).

Of note, the MeCP2 nonsense mutations leading to larger deletions cause a more severe phenotype than those nonsense mutations causing smaller C-terminal truncations (Vashi and Justice 2019). Molecularly, the larger C-terminal deletions caused weaker methyl-DNA binding properties (Yusufzai and Wolffe 2000), suggesting that the function of MeCP2 is, at least partially, dependent on specific binding to methylated cytosine.

2.4 MBD2

2.4.1 Isoforms and structure feature

Methyl-CpG-binding domain protein 2 (MBD2) was originally identified via searching the EST (expressed sequence tag) database using the amino acid sequence of the MeCP2-MBD region (Hendrich and Bird 1998, Jongeneel 2000). In humans, the MBD2 gene is located at chromosome 18 (18q21.2), encoding three isoforms of MBD2 due to alternative splicing or alternative translational start sites (TSS) (Fig. 7). The MBD2a encodes the full-length long isoform with a glycine and arginine-rich region (G/R) at N-terminus, coiled-coil domain (CC) containing region at the C-terminus and the conserved MBD and TRD in between. MBD2b encodes the short isoform with a deletion of the N-terminus due to alternative TSS. The MBD2c encodes a premature version of MBD2 lacking the CC containing C-terminus due to the inclusion of the alternative third exon. All three isoforms contain the MBD and TRD.

The MBD2 MBD is responsible for mCpG binding (Hendrich and Bird 1998, Buchmuller, Kosel et al. 2020), and the TRD is responsible for interactions with multiple transcription repression factors, such as Sin3A (Boeke, Ammerpohl et al. 2000). Distinct from MeCP2, in which the MBD and TRD are linked by the ID domain, the MBD and TRD of MBD2 overlap to a large extent, indicating a close relationship between DNA methylation and transcriptional repression of the MBD2 (Boeke, Ammerpohl et al. 2000). Besides, the MBD2 MBD shows a

much weaker (or no) binding affinity to mCpG than the MeCP2 MBD (Buchmuller, Kosel et al. 2020). The arginines in the G/R are reported to be posttranslationally methylated which reduced the mCpG binding properties of the MBD. The MBD2 CC domain promotes the assembly of the nucleosome remodeling and histone deacetylase (NuRD) complex via binding with the CC domain from p66a (Gnanapragasam, Scarsdale et al. 2011), yet the CC of MBD2 fails to self-interaction (Becker, Allmann et al. 2013, Walavalkar, Gordon et al. 2013). The highly disordered region in between TRD and CC was shown to be responsible for binding with histone deacetylase core sub-complex (HDCC, a subcomplex of NuRD), as modifying two consecutive residues (R286E/L287A) disrupts binding of MBD2 to the HDCC and therefore blocks methylation-dependent gene silencing by NuRD (Desai, Webb et al. 2015, Yu, Azzo et al. 2019).

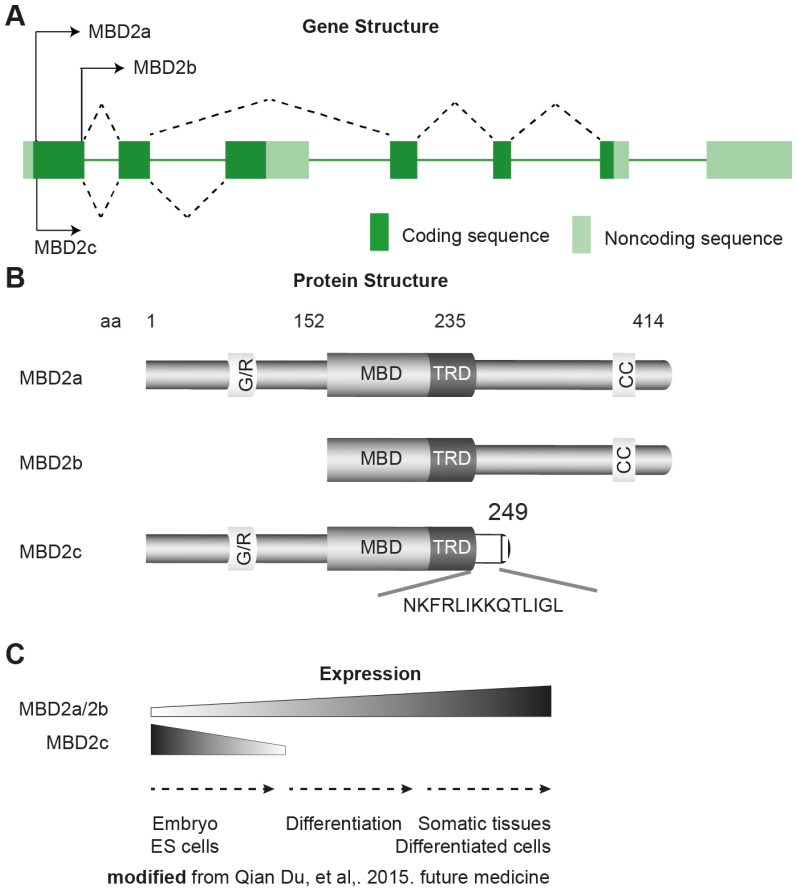


Figure 7 Schematic representation of MBD2 isoforms and expression profile during differentiation.

(A-B) The MBD2b arises from the second alternative translational start site in exon1, generating a truncated isoform lacking the N-terminus. The MBD2c is a premature form lacking the C-terminus due to the alternative third exon. G/R: glycine and arginine-rich region, MBD2 has 11 G/R repeats at the N-terminus, also called $(G/R)_{11}$ in some papers; MBD: methyl binding domain; TRD: transcriptional repression domain; CC: coiled-coil domain. The CC domain is a helical peptide that performs as a monomeric helix in solution but forms stable interactions with the CC from p66a due to charge distribution.
(C) MBD2a and MBD2b are highly expressed in somatic cells in tissues and differentiated cells in culture, while MBD2c is highly expressed in embryos and embryonic stem (ES) cells and promotes the maintenance of pluripotency.

The three structurally distinct isoforms of MBD2 exhibit different expression profiles and functions. MBD2a and MBD2b is highly expressed in somatic tissues (Hendrich and Bird 1998, Wood and Zhou 2016) while MBD2c is highly expressed in embryonic stem cells (Hendrich and Bird 1998, Lu, Loh et al. 2014). *In vivo* studies using MBD2 knockout (KO) mice showed that MBD2 KO was not embryonically lethal yet it depicted a maternal nurturing defect (Hendrich, Guy et al. 2001). This suggests a role of MBD2 in mouse brain development. Moreover, MBD2a and MBD2c were shown to perform opposite roles in cell fate decisions (Lu, Loh et al. 2014). MBD2a was shown to repress the expression of OCT4, NANOG, and SOX2 via recruiting the NuRD complex and therefore promote the ES cell differentiation. MBD2c, the isoform lacking the C-terminus, binds to the promoters of OCT4 and SOX2 but does not recruit the NuRD complex and thus did not influence the expression of those genes. Moreover, either MBD2a repression or MBD2c overexpression enhanced the reprogramming efficiency of the somatic cells (Lee, Prasain et al. 2013) (Lu, Loh et al. 2014).

2.4.2 DNA binding and genome-wide distribution

The MBD2 MBD specifically binds to symmetrically methylated CpG sites, which could be enhanced by the C-terminus (Hendrich and Bird 1998, Desai, Webb et al. 2015). *In vivo* chromatin immunoprecipitation assays showed that MBD2 mainly binds to highly methylated CpG sites with no detectable sequence specificity. Distinct from MeCP2, which tracks the mCpG density (Skene, Illingworth et al. 2010), MBD2 is absent from low density methylated CpG sites (Chatagnon, Perriaud et al. 2011, Baubec, Ivanek et al. 2013, Gunther, Rust et al. 2013, Menafra, Brinkman et al. 2014, Devailly, Grandin et al. 2015).

In vivo, MBD2 (MBD2a) promotes the de novo and maintenance methylation via interaction with Dnmt1 and Dnmt3a and thereby functions in CpG islands hypermethylation in cancer cells (Stirzaker, Song et al. 2017). Meanwhile, it was reported that the arginine methylation at the G/R region reduced the MBD-mCpG interaction, together with decreased MBD2-NuRD complex formation at the mCpG sites and impaired transcriptional repression function (Tan and Nakielny 2006).

Although the MBD2 shows only a weak interaction with hemimethylated CpG and hydroxymethylcytosine (hmC) and no binding with unmodified cytosine *in vitro* (Hendrich and Bird 1998, Hashimoto, Liu et al. 2012, Buchmuller, Kosel et al. 2020), the MBD2a and MBD2b were also detected in unmethylated sites *in vivo* (Baubec, Ivanek et al. 2013). MBD2c, the isoform lacking the C-terminal NuRD interaction region, was lost from unmethylated sites but retained at highly methylated CpG sites (Baubec, Ivanek et al. 2013). Thus, the interaction with NuRD could probably direct the localization of MBD2 at unmethylated sites (Baubec, Ivanek et al. 2013). Besides, the G/R is highly positively

charged, likely contributing to nonspecific MBD2-DNA binding via weak electrostatic interactions.

2.4.3 MBD2 on chromatin architecture

MBD2 was shown to be enriched into the highly methylated pericentromeric heterochromatin and promote chromocenter clustering in mouse cells (Hendrich and Bird 1998, Brero, Easwaran et al. 2005). Using the LacO/LacI system, researchers found that LacI-MBD2a and LacI-MBD2b dramatically converted the euchromatic LacO array into heterochromatin with repressed transcription by recruiting the NuRD complex (Gunther, Rust et al. 2013).

The MBD2/NuRD complex is involved in transcription regulation and chromatin architecture. The MBD2/NuRD complex was isolated as early as 1989, whose components and functions were further characterized by subsequent research.

On the one hand, two specific methyl-CpG binding proteins (complex) were identified, the multiprotein complex MeCP1 and a single peptide MeCP2 (Meehan, Lewis et al. 1989). MBD2 was recognized as a component of the MeCP1 complex via the EST searching based on the MBD from MeCP2 (Lewis, Meehan et al. 1992, Hendrich and Bird 1998). Further, the MBD2 was shown to be involved in the histone deacetylase (HDAC) complex including the Mi2, a member of the SWI2/SNF2 helicase/ATPase family (Zhang, Ng et al. 1999). On the other hand, two histone deacetylase complexes were identified: Sin3a/HDAC complex (Zhang, Iratni et al. 1997) and NuRD complex (Zhang, LeRoy et al. 1998). Sin3A/HDAC could bind with MeCP2 and thus be located at the mCpG sites (Zhang, Ng et al. 1999). The NuRD complex contains histone deacetylase and Mi2 (Wade, Jones et al. 1998, Zhang, LeRoy et al. 1998, Wade, Gegonne et al. 1999), thereby the NuRD showed both histone deacetylase and chromatin remodeling function. The MBD2 and MBD3 were found to be the core components of the HDAC subunit within the NuRD complex (Ng, Zhang et al. 1999). The NuRD complex could be divided into two stable distinct subunits, the histone deacetylase core sub-complex (HDCC) and the chromatin remodeling sub-complex which are bridged by the MBD2 or MBD3 (Zhang, Ng et al. 1999, Gnanapragasam, Scarsdale et al. 2011, Yu, Azzo et al. 2019). Subsequent research revealed that the NuRD complex contains at least 7 proteins with variants, including metastasis Tumor-Associated 1, 2, or 3 (MTA1, 2, or 3), Histone Deacetylase 1 or 2 (HDAC1 or 2), Retinoblastoma Binding Protein 4 or 7 (RBBP4 or 7), GATAD2A or B, Chromodomain Helicase DNA Binding Protein 3, 4, or 5 (CHD3, 4, or 5),

Cyclin-Dependent Kinase 2 Associated Protein 1 (CDK2AP1), and MBD2 or 3 (Leighton and Williams Jr 2020).

Taken together, the NuRD complex likely represents the original MeCP1 complex, and the MBD2 bridges the three epigenetic regulatory elements: DNA methylation, histone

deacetylation, and ATP-dependent chromatin remodeling. Firstly, the MBD2 recruits the chromatin remodeling sub-complex via the specific coiled-coil (CC) interaction, which is located at the C-terminus of MBD2 and the N-terminus of p66a (homolog in human, GATAD2A) (Gnanapragasam, Scarsdale et al. 2011). Secondly, the MBD2 recruits the HDCC sub-complex via the interaction between the disordered region linking the TRD and CC of MBD2 and HDCC. This is further confirmed by the observation that MBD2 point mutations R286E/L287A disrupted the binding to HDCC (Desai, Webb et al. 2015). Lastly, MBD2 recruits the NuRD complex to the highly methylated sites. As a result, MBD2 decreased the chromatin accessibility at the highly methylated CpG sites via chromatin remodeling and chromatin deacetylation.

The MBD2-NuRD complex could also be influenced by other MBDs, especially MBD3, which also participates in the NuRD complex assembly. The MBD2-NuRD and MBD3-NuRD are mutually exclusive subunits with different biochemical and functional properties (Hendrich, Guy et al. 2001, Le Guezennec, Vermeulen et al. 2006).

The MBD2 isoforms also perform differently to NuRD complex assembly due to post-translational modification or interactions. MBD2/NuRD could interact with the protein arginine methyltransferase (PRMT) 1 and/or 5, which catalyzed the arginine methylation at the G/R region. Such arginine methylation was shown to weaken the interaction with NuRD and methylated DNA, thus decreasing the transcription repression ability (Tan and Nakielny 2006). This is not the case for the MBD2b as it lacks the G/R containing N-terminus. MBD2c lacks a C-terminus and therefore does not participate in the assembly of the NuRD complex, yet it retains the binding ability with mCpG. Thus, the MBD2c might compete with MBD2a/b for certain mCpG sites and regulate the distribution of the NuRD complex.

MBD2 could also promote histone arginine methylation by recruiting the PRMT1 and PRMT5 *in vivo*. MBD2-PRMT1 interaction catalyzes the asymmetrical dimethylation of histone H4 at arginine 3 (H4R3me2) and promotes gene expression (Wang, Huang et al. 2001). MBD2-hSWI/SNF-PRMT5 interaction catalyzes the methylation of histone H3 at arginine 8 and thereby represses the expression of ST7 and NM23 tumor-suppressor genes (Pal, Vishwanath et al. 2004).

MBD2 participates in the maintenance and spreading of cytosine methylation via recruiting the Dnmts (Stirzaker, Song et al. 2017). Besides, MBD2 occupancy inhibits the binding of Tet, thus inhibiting the Tet-mediated oxidative demethylation (Ludwig, Zhang et al. 2017).

MBD2 was also shown to promote gene transcription by interacting with multiple factors that are independent of the NuRD complex. MBD2a, but not MBD2b, was shown to selectively interact with different transcriptional activators such as RNA helicase A (RHA) and NGFI-A (egr-1) and promote the gene activation. Interestingly, the specific binding of MBD2a to RHA inhibits the binding to histone deacetylase (Fujita, Fujii et al. 2003, Weaver, Hellstrom et al.

2014). These studies further highlight the importance of G/R containing N-terminus in regulating the MBD2 function. The MBD2 could also promote gene expression via promoter demethylation, but the underlying mechanism is still controversial (Cervoni and Szyf 2001, Wang, Liu et al. 2013). Besides, MBD2 was shown to indirectly interact with the histone acetyltransferase pCAF via the TACC3 to methylated promoters and activate transcription of repressed genes (Angrisano, Lembo et al. 2006). The highly disordered regions between TRD and CC direct the interaction with TACC3 (Angrisano, Lembo et al. 2006).

2.4.4 Dysfunction

Although closely correlated with neuronal functions in both mouse models and humans, loss of MBD2 in mice only causes mild phenotypes with maternal nurturing phenotypes including little body size and impaired nurturing behavior (Hendrich, Guy et al. 2001). Latter research by Wood, K. H and collaborators showed that the MBD2 knockout mice showed no significant alterations in locomotor activity, sociability, anxiety-related behaviors, learning, and memory compared to the wildtype and heterozygous littermates (Wood, Johnson et al. 2016). In cells, MBD2 was shown to regulate the specific transition stage during olfactory receptor neuron (ORN) differentiation. Mbd2 null progenitors display enhanced proliferation, while the Mbd2 null ORNs display a decreased lifespan (Macdonald, Verster et al. 2010). Taken together, the MBD2 seems dispensable during brain development. Yet considering the high heterogeneity of the brain, one possible explanation of this might be that MBD2 is indispensable during the differentiation and functioning of certain cell types, which is too small a population to be identified when the whole brain and/or all neurons were taken.

MBD2 is also involved in human cancers (Berger and Bird 2005, May, Owen et al. 2018, Li, Li et al. 2020). Overexpression of MBD2 is associated with aberrant hypermethylation at the promoters of tumor-suppressor genes, such as GSTP1 and 14–3–3σ and thus inhibits their expression in human prostate cancer cells (Lodygin, Diebold et al. 2004, Pulukuri and Rao 2006, Chatagnon, Bougel et al. 2009, Stirzaker, Song et al. 2017). Besides, MBD2 recruits the PRMT5 to the $p16^{INK4a}$ and $p14^{ARF}$ CpG islands and inhibits the expression of $P16^{INK4a}$ via H4R3 methylation in human colon cancer cells (Magdinier and Wolffe 2001, Le Guezennec, Vermeulen et al. 2006, Martin, Jorgensen et al. 2008). During gliomagenesis, high levels of MBD2 suppress the antiangiogenic activity of a key tumor suppressor BAI1, and thus drive tumor growth (Zhu, Hunter et al. 2011). Besides, a high frequency of MBD2 mutations was detected in human colon and lung cancers (Bader, Walker et al. 2003) but the exact relationships between mutations and cancers are to be elucidated.

MBD2 is also involved in immunity (Szyf 2010, Wood and Zhou 2016, May, Owen et al. 2018). MBD2 promotes the Dendritic cells (DCs) directed $CD4^+$ T-cell differentiation into diverse helper (Th) cells (Cook, Owen et al. 2015). MBD2 absence during the differentiation

of naive CD8 T cells into effector and memory cells leads to impaired memory CD8 T cells (Kersh 2006). Primary immune thrombocytopenia (ITP) is hallmarked with genome-wide hypomethylation, which is positively correlated with decreased levels of MBD2 and MBD4, indicating that the MBD2 is involved in the pathogenesis of ITP (Chen, Gu et al. 2011).

2.5 Other MBD family proteins

MBD1, MBD3, and MBD4 were originally identified together with MBD2 in 1998 (Hendrich and Bird 1998) by homology searching using the MBD from the MeCP2.

2.5.1 MBD1

In humans, MBD1 (also known as PCM1) is located at chromosome 18, encoding the largest member of the MBD protein family with many variants (https://www.ncbi.nlm.nih.gov/gene/4152). The longest isoform of MBD1 consists of 605 amino acids (643 amino acids in mus musculus). Similar to MeCP2 and MBD2, MBD1 contains the MBD and TRD with mCpG binding and transcriptional repression function, respectively. Besides, depending on variants, MBD1 contains two or three CxxC domains, which are called CxxC1, CxxC2, and CxxC3 (if the variants contain the third one) from the N-terminus to the C-terminus. CxxC is a small zinc finger motif with ~50–70 amino acids and can specifically bind to unmethylated CpG via the highly conserved KFFG motif. MBD1 CxxC3 has a positively charged surface and specifically binds to unmethylated CpG, whereas the CxxC1 and CxxC2 show a negatively charged surface and do not bind to unmethylated CpG (Jørgensen, Ben-Porath et al. 2004, Xu, Liu et al. 2018). MBD1 is involved in DNA methylation-dependent and independent transcriptional repression at the heterochromatin and euchromatin, respectively (Fujita, Takebayashi et al. 1999, Ng, Jeppesen et al. 2000). Still, the precise contribution of MBD and CxxC3 on the MBD1 genome-wide distribution is not clear.

MBD1 could modulate the DNA methylation levels. MBD1 could recruit the Tet1 to the highly methylated heterochromatic regions in a CxxC3 dependent manner and promote the 5mC to 5hmC conversion. While the MBD1 variant without the CxxC3 was shown to block the localization of Tet1 at highly methylated CpG sites and thus inhibit the 5mC to 5hmC conversion (Zhang, Rausch et al. 2017).

MBD1 participates in heterochromatin architecture. *In vivo*, MBD1 is enriched into the highly methylated pericentromeric heterochromatin foci and maintains the heterochromatin state by recruiting the histone H3 methylase Suv39h1, HP1, and histone deacetylase (Hendrich and Bird 1998, Fujita, Fujii et al. 2003). Besides, MBD1 forms complex with MBD1-containing chromatin-associated factor (MCAF) 1 and SET domain bifurcated histone lysine

methyltransferase 1 (SETDB1) and therefore facilitates the formation of heterochromatic domains (Ichimura, Watanabe et al. 2005).

In mouse models, MBD1 KO is not embryonically lethal but shows deficits in adult neurogenesis and hippocampal function together with increased genomic instability (Zhao, Ueba et al. 2003). In adult neural stem/progenitor cells (NSPCs), MBD1 knock-out was shown to be accompanied by increased Fgf-2 level. Further, both acute knockdown of Mbd1 and overexpression of Fgf-2 in adult NSPCs inhibited their neuronal differentiation. Thus, MBD1 plays an essential role in NSPCs differentiation via regulating the Fgf-2 expression (Li, Barkho et al. 2008). In NSPCs, MBD1 also regulates the expression of miR-184, which further regulates the expression of Numblike (Numbl). Numbl is a key regulator and induces the NSPCs differentiation during brain development (Liu, Teng et al. 2010).

In the primary sensory neurons of dorsal root ganglion (DRG), MBD1-deficient mice exhibit reduced responses to acute mechanical, heat, cold, and capsaicin stimuli and the blunted nerve injury-induced pain hypersensitivities. Molecularly, MBD1 inhibits the expression of Oprm1 and Kcna2 by recruiting DNA methyltransferase DNMT3a and promoting the methylation at the promoters (Mo, Wu et al. 2018).

2.5.2 MBD3

In humans, the MBD3 gene is located at chromosome 19 (19p13.3) and encodes the smallest member of the MBD protein family. The MBD3 protein shows high sequence similarity to MBD2 (71.1%) and is thought to have arisen from an ancient gene duplication during the evolution of the vertebrate lineage (Hendrich and Bird 1998, Hendrich and Tweedie 2003). MBD3 has two isoforms due to an alternate in-frame splice site in the 5' coding region in exon 1. Different from MBD2a while similar to MBD2b, MBD3 is coded from the MBD domain, lacking the G/R containing N-terminus (Hendrich and Bird 1998). MBD3 is expressed in both ESCs and somatic tissues (Hendrich and Bird 1998, Roloff, Ropers et al. 2003).

MBD3 is the only MBD-containing protein that does not bind to symmetrically methylated CpG sites but retains the binding affinity to unmethylated DNA and hydroxymethylated DNA (Hendrich and Bird 1998). This is because the two amino acids in the MBD domain that are responsible for the direct interaction with DNA were mutated (K30H and Y34F) (Buchmuller, Kosel et al. 2020).

MBD3 could regulate the DNA methylation levels. On the one hand, MBD3 maintains the DNA methylation and silencing state of the paternal H19 allele(Reese, Lin et al. 2007). On the other hand, MBD3 maintains the demethylated and active state of rRNA genes (Brown, Suderman et al. 2008). These suggest that MBD3 might function in maintaining the DNA modification state, preventing a conversion between unmethylated and methylated states.

Recently, MBD3 was reported to bind to the 5hmC (Yildirim, Li et al. 2011, Iurlaro, Ficz et al. 2013), which was questioned by other publications (Buchmuller, Kosel et al. 2020). Genome-wide, MBD3 localizes at promoters, gene bodies, and enhancers of active genes (Shimbo, Du et al. 2013). MBD3 was also reported to be enriched at the 5hmC sites in a Tet-dependent manner (Baubec, Ivanek et al. 2013). Indeed, Mbd3 knockdown affects the expression of 5hmC-marked genes in ESCs (Baubec, Ivanek et al. 2013), suggesting that MBD3 might perform 5hmC specific functions.

Like MBD2, MBD3 is a scaffold subunit of the NuRD complex and plays an essential role in ES cell pluripotency maintenance and differentiation. Mbd3-deficient embryonic stem (ES) cells showed LIF-independent self-renewal and defects in differentiation as they failed to commit to developmental lineages (Kaji, Caballero et al. 2006, Kaji, Nichols et al. 2007). Disrupting the Mbd3/NuRD complex by inhibiting GATAD2A facilitates deterministic murine iPSC reprogramming of MEF cells induced by the four Yamanaka factors (Oct4, Sox2, Klf4, and Myc) (Mor, Rais et al. 2018). More recently, Schmolka N and collaborators observed that the ectopic expression of MBD2a and MBD2b but not MBD2c could partially rescue the differentiation block in MBD3 KO ESCs during neuron differentiation (Schmolka, Bhaskaran et al. 2021). This indicates that the MBD2/NuRD and MBD3/NuRD are functionally redundant.

Contradictory, Mbd3-deficient neural stem cells, EpiSCs, and preiPSCs displayed a severe defect in reprogramming to iPSCs, while MBD3 overexpression could facilitate the reprogramming progress in a context-dependent manner (dos Santos, Tosti et al. 2014). In primary human fibroblasts, inhibiting MBD3 reduced the efficiency of reprogramming and the final iPSCs were incapable of trilineage differentiation. While overexpression of MBD3 increased the number of fibroblast-derived iPSC colonies with trilineage differentiation capability (Jaffer, Goh et al. 2018). Similarly, MBD3 overexpression improved the reprogramming of cloned pig embryos (Wang, Shi et al. 2019). Besides, MBD3 restricted ES cells from differentiating towards the trophectoderm lineage (Zhu, Fang et al. 2009). Taken together, MBD3 seems essential to maintain the pluripotent stem cells in the native state with trilineage differentiation capability. Loss of MBD3 induces the native ES cells to a less pluripotent stage with restricted differentiation ability. Consistently, Mbd3-deficient blastocysts in the inner cell mass (ICM) failed to develop into mature epiblast after implantation (Kaji, Nichols et al. 2007). Thus, MBD3 KO mice are embryonically lethal (Hendrich, Guy et al. 2001, Kaji, Caballero et al. 2006).

2.5.3 MBD4

MBD4 (also known as MED1) gene is located in chromosome 3 (3q21.3) and encodes a protein with 5 variants due to alternate in-frame splice sites and alternative exon inclusion

and exclusion. The longest isoform consists of 580 amino acid residues with a conserved MBD domain and a C-terminal glycosylase domain (Hendrich and Bird 1998), a homolog to bacterial DNA damage specific endonucleases with glycosylase activity during base excision repair (BER) (McCullough, Dodson et al. 1999). BER is initiated with DNA glycosylase mediated excision of damaged DNA base pairs, generating an abasic site and followed by the endonuclease catalytic backbone incision and DNA polymerase catalytic gap-filling (Bosshard, Markkanen et al. 2012, Sjolund, Senejani et al. 2013).

The deamination of 5-methylcytosine (5-meC) and cytosine (C) to thymine (T) and uracil (U), respectively, generates G:T and G:U mismatch pairs. These occur at a rate of 2–300 lesions/cell/day and lead to C:G to T:A transitions and genome instability if not repaired before replication (Duncan and Miller 1980). Besides, such transitions frequently occur in human cancer and other diseases (Cooper and Youssoufian 1988). MBD4 is responsible for the removal of deaminated C (U), 5-meC (T), and 5-hmC (hmU) within CpG sites independent of cytosine methylation (Hendrich, Abbott et al. 1999, Moréra, Grin et al. 2012, Sjolund, Senejani et al. 2013, Yakovlev, Kuznetsova et al. 2017, Pidugu, Bright et al. 2021).

Emerging evidence shows that MBD4 promotes cytosine demethylation (Rai, Huggins et al. 2008). In zebrafish, deaminase/glycosylase pair AID/Mbd4 (human) overexpression was shown to promote widespread DNA demethylation, while AID or Mbd4 knockdown caused remethylation of a set of common genes (Rai, Huggins et al. 2008). However, contradictory studies using the zebrafish embryos showed that no aberrant methylation was observed upon AID or MBD4 (zebrafish) knockdown (Shimoda, Hirose et al. 2014). One possible explanation lies that the zebrafish MBD4 contains the DNA glycosylase domain but lacks the MBD domain, thus zebrafish MBD4 might not be able to bind to the methylated CpG site and convert the mCpG to CpG. These studies also indicate an MBD- and species-dependent demethylation of MBD4, but more work should be done.

Although lacking a transcriptional repression domain, MBD4 represses gene transcription at mCpG sites via promoting DNA methylation and histone modification. (Kondo, Gu et al. 2005, Laget, Miotto et al. 2014, Liao, Li et al. 2017). MBD4 recruited HDAC and Sin3A at mCpG sites and repressed gene expression in an HDAC dependent manner (Kondo, Gu et al. 2005). In mammalian cells, MBD4 cooperates with DNMT1 to mediate methyl-DNA repression and protect mammalian cells from oxidative stress (Laget, Miotto et al. 2014). More recently, downregulation of MBD4 contributed to overexpression and hypomethylation of the CD70 gene in SLE CD4+ T cells, suggesting a positive role of MBD4 in DNA methylation (Liao, Li et al. 2017).

MBD4 is enriched in the highly methylated pericentromeric heterochromatin foci in mouse cells and functions in heterochromatin organization (Hendrich and Bird 1998). Meng et al. showed that MBD4 recruited both the UHRF1, an E3 ubiquitin ligase, and USP7, a

deubiquitinating enzyme to the heterochromatin compartments, and therefore promoted the stability of Dnmt1(Meng, Harrison et al. 2015).

MBD4 KO mice are viable and fertile but have increased CG to TA mutations at CpG sites (Millar, Guy et al. 2002, Wong, Yang et al. 2002). Moreover, heterozygous mice with deficient MBD4 and adenomatous polyposis coli Min (Apc^{Min}) mutation, a mutation with intestinal tumor formation (Su, Kinzler et al. 1992), showed increased gastrointestinal tumor formation.

In conclusion, MBD4 is associated with various nuclear pathways including DNA repair, DNA (de)methylation, transcriptional repression, and heterochromatin organization. However, the precise contribution of MBD4 to these processes remains unclear.

2.5.4 MBD5/6, SETDB1/2, and BAZ2A/B

Besides, other MBD-containing proteins were also reported, including MBD5, MBD6, SETDB1/2, and BAZ2A/B but their functions are uncharacterized.

MBD5 has two isoforms due to alternative splicing, with the short isoform lacking a PWWP-containing C-terminus (Laget, Joulie et al. 2010). PWWP is a ~70 amino acid long domain centered by proline-tryptophan-tryptophan-proline (P-W-W-P) and is responsible for protein-protein interaction. MBD5 is enriched at heterochromatin compartments in both PWWP and MBD dependent manner, as both mutations at the MBD (P39A) and PWWP (W1399A, P1400A) disrupted MBD5 distribution at heterochromatin compartments (Laget, Joulie et al. 2010). MBD5 long isoform is highly expressed in the brain, while the short isoform is highly expressed in oocytes, suggesting an isoform-specific role in the development and especially brain development. Indeed, deletion, duplication, and mutation of MBD5 lead to MBD5-associated neurodevelopmental disorder (MAND), a 2q23.1 microdeletion syndrome characterized by microcephaly, intellectual disabilities, severe speech impairment, and seizures (Williams, Mullegama et al. 2010, Talkowski, Mullegama et al. 2011, Mullegama and Elsea 2016, Myers, Marini et al. 2021). Further studies showed that in the heterozygous hypomorphic $Mbd5^{+/GT}$ mouse, the haploinsufficiency of MBD5 changed the gene expression in the cortex region, suggesting context-dependent effects (Camarena, Cao et al. 2014, Seabra, Aneichyk et al. 2020).

MBD6 has only one isoform, is highly expressed in testis, and is localized to heterochromatin compartments in an MBD-dependent manner (Laget, Joulie et al. 2010). Surprisingly, both MBD5 and MBD6 do not bind methylated DNA (Laget, Joulie et al. 2010).

MBD5 and MBD6 also share some functions *in vivo*. Both MBD5 and MBD6 promote gene silencing via recruiting a J-domain protein, SILENZIO, which serves as a co-chaperone with HSP70s (Ichino, Boone et al. 2021). The MBD of MBD5 and MBD6 mediate the interaction with mammalian PR-DUB Polycomb protein complexes in a mutually exclusive manner,

which catalyzes deubiquitination of H2AK119. While MBD6-PRDUB is specifically enriched at the DNA damage sites (Baymaz, Fournier et al. 2014). Besides, MBD6 protein level is downregulated in uterine leiomyoma (UL) tissues and MBD6 is mutated and differentially expressed in gastric (GC) and colorectal (CRC) cancers. These indicate a role of MBD6 in cancer (Choi, Yoo et al. 2015, Liang, Zeng et al. 2019).

Distinct from the core MBD proteins, SETDB1/2 contains the SET domain that confers histone H3K9 lysine methyltransferase activity. SETDB1 promotes chromatin compaction and gene repression by H3K9 methylation (Schultz, Ayyanathan et al. 2002) and is involved in normal physiology and brain development (Markouli, Strepkos et al. 2021). SETDB1 is involved in cancer, such as non-small cell lung cancer (NSCLC) (Sun, Ding et al. 2015) tumorigenesis, and breast cancer Metastasis (Batham, Lim et al. 2019). During NSCLC tumorigenesis, SETDB1 was shown to stimulate the WNT/β-catenin pathway and diminish P53 expression (Sun, Ding et al. 2015).

SETDB2 is also involved in the carcinogenesis of human gastric cancers (GCs). Nishikawaji et al. noticed that SETDB2 was highly expressed in GC tissues and that SETDB2 knockdown significantly decreased cell proliferation, migration, and invasion together with lower global H3K9 trimethylation (H3K9me3) levels. Molecularly, SETDB2 inhibited the expression of tumor suppressor genes WWOX and CADM1 (Nishikawaji, Akiyama et al. 2016). SETDB2 is also involved in the immune response. During influenza virus infection, SETDB2 regulates the crosstalk between the type-I interferon and NF-κB pathway. The NF-κB pathway represents an important mechanism for virus-induced susceptibility to bacterial superinfection. More recently, COVID-19 infection was shown to repress the expression of SETDB2, followed by decreased H3K9me3 levels at NFkB binding sites on inflammatory gene promoters. As a result, the inflammatory response is increased upon coronavirus infection.

Bromodomain adjacent to zinc finger domain protein 2 (BAZ2) contains an MBD and a bromodomain that confers acetylated histone binding capability. BAZ2 is a subunit of the nucleolar remodeling complex NoRC and controls rDNA accessibility and transcription by interacting with acetylated lysine sites on histone H3 (Peña-Hernández, Aprigliano et al. 2021) or H4 (Zhou and Grummt 2005, Anosova, Melnik et al. 2015).

BAZ2A overexpression is involved in prostate cancer (PCa) (Peña-Hernández, Aprigliano et al. 2021). High levels of BAZ2A are frequently detected in metastatic tumors with cell proliferation, viability, invasion, and repression of genes silenced in aggressive PCa (Wood, Johnson et al. 2016). Molecularly, the inactive enhancers are marked by acetylation, which is catalyzed by the EP300, a histone acetyltransferase that acetylates H3K14ac (Jin, Yu et al. 2011, Peña-Hernández, Aprigliano et al. 2021). BAZ2A binds to inactive enhancers via bromodomain-H3K14ac interaction (Peña-Hernández, Aprigliano et al. 2021). Besides, RNA-

mediated BAZ2A interaction with TOP2A and KDM1A was also shown to regulate the gene expression in PCa.

The molecular function of BAZ2B is still not fully defined. *In vivo*, BAZ2B could bind to H3K14ac (Charlop-Powers, Zeng et al. 2010) and might function in neurodevelopment (Guo, Bettella et al. 2019). BAZ2B significantly promotes reprogramming of human hematopoietic committed progenitors to multipotent state and enhances self-renewal and engraftment of human CD34+ cells (Arumugam, Shin et al. 2020). This suggests a potential role of BAZ2B in human B cell development.

3 Aims of this work

It is increasingly clear that the phase separation could promote the assembly of distinct microscale condensates. In cell nuclei, heterochromatin plays an essential role in genome integrity, gene expression, and nuclear architecture. Many factors have been well characterized for their functions in the formation, maintenance, and dynamics of heterochromatin, yet the underlying mechanism is controversially discussed. In the last years, it has been shown that HP1 can promote heterochromatin formation via phase separation, but whether this is the case for the other factors is unclear, especially for DNA methylation “readers”.

In this work, I focus on the core members of the MBD protein family (MBD1-4, MeCP2) and their abilities regarding phase separation *in vivo* and *in vitro* (Fig. 8). The work is divided in three main aims as followed:

i) The function of the MBD family proteins in the heterochromatin compartments and their phase separation properties.

Previous work showed that MeCP2 could promote heterochromatin compartment fusion. I investigated whether this it is the same for the other MBD family proteins. *In vitro*, I purified the MBD family proteins and checked their abilities to phase separate and the morphology of condensates.

ii) The functions of MBD2 isoforms in the heterochromatin compartments.

MBD2 has three isoforms, one lacking the N-terminus, one lacking the C-terminus, and one with both termini. Although all contain the MBD-TRD, N-terminus and C-terminus contain different functional domains, suggesting that the three isoforms might perform N- or C-terminus dependent functions in heterochromatin compartments. Thus, I tested their phase separation properties, functions in heterochromatin compartment clustering, and molecular dynamics inside heterochromatin compartments.

iii) The phase separation properties of MeCP2 and the influence of RTT-related nonsense mutations.

Molecularly, I analyzed how self-interaction contributes to MeCP2 phase separation and how nonspecific and specific protein-DNA interactions influence the phase separation properties of MeCP2. I examined how RTT-related nonsense mutations influence the function of MeCP2 in LLPS induced heterochromatin organization.

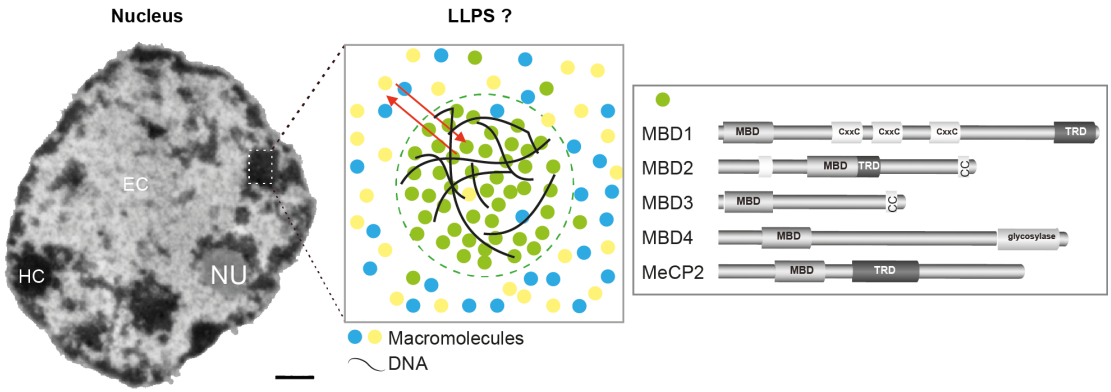


Figure 8 Schematic graph showing the work of this thesis

The heterochromatin is hallmarked with high cytosine methylation, which could be "read" by the methyl-CpG binding protein family except for MBD3. In this work, I focus on the core members of the MBD protein family (MBD1-4, MeCP2) and checked their possibilities of phase separation *in vitro* and their influence on heterochromatin dynamics *in vivo*.

4 Materials and Methods

4.1 Plasmids

All plasmids used in this study and the source references are listed in Table 2.

4.1.1 Bacterial expression plasmids

The pTYB1 vector was used for fusion protein purification because it utilizes the inducible self-cleavage activity of intein to separate the target proteins (MBDs) from the affinity tag (chitin-binding domain, CBD).

The pTYB1-MeCP2wt (pc1294) plasmid (Georgel, Horowitz-Scherer et al. 2003) was modified to generate various MBD expression plasmids pTYB1-GFP-MBDs. The GFP-MeCP2, MBD1ΔCxxC3, MBD2a, MBD2b, MBD3, and MBD4 coding sequence was amplified by PCR from pEG-MeCP2 (pc1208) (Kudo, Nomura et al. 2003), pGmMBD1.3 (pc 2899), pFB-mMBD2aG (pc1565), pFB-mMBD2.2G (pc2083), pFB-mMBD3G (pc1566), and pFB-mMBD4G (pc1567), respectively using primer pairs containing NdeI (NEB, R0111S) for the forward primer and EcoRI (NEB, R0101S) for the reverse primer (Table 3). Then, the amplified coding sequence was digested with NdeI/EcoRI and ligated into the NdeI/EcoRI digested pTYB1-MeCP2wt vector to finally generate the pTYB1-GFP-MeCP2, MBD1ΔCxxC3, MBD2a, MBD2b, MBD3, and MBD4, separately.

Using either pTYB1-MeCP2wt or pTYB1-GFP-MeCP2 as templates, the Q5 site-directed mutagenesis strategy was adopted following the standard protocol (NEB, E0554S) using the primer pairs listed in Table 3 (from IDT) to generate the truncated versions R168X, R255X, R269X, G273X, R294X, and ΔAT-hook2 without or with GFP tag respectively, followed by ligation, transformation of *E. coli* cells Top 10 (Table 4), and DNA sequencing.

Using the pTYB1-MBD2a as a template, the pTYB1-MBD2-N (aa: 1-152), MBD2a-MBD2a-1-235, MBD2a-C (236-414) were generated by Q5 site-directed mutagenesis strategy following the standard protocol (NEB, E0554S). Using pTYB1-MBD2b (1-235) as a template, the pTYB1-MBD-TRD were generated by Q5 site-directed mutagenesis strategy.

4.1.2 Mammalian expression plasmids

pEG-MeCP2-R168X, -R255X, -R269X, -G273X, -R294X, and ΔAT-hook2 were engineered using the Q5 site-directed mutagenesis with the pEG-MeCP2 (pc1208) vector as the PCR template as described above. pEG-MBD2-1-235 was engineered by inserting the MBD2a-1-235 coding sequence into the pEGFP-N3 (pc714) plasmid. In brief, the MBD2a-1-235 coding sequence was amplified using a primer pair containing HindIII (NEB, R0104S) for the forward primer and SalI (NEB, R0138S) for the reverse primer (Table 3). Then, the amplified

coding sequence was digested with HindIII/SalI and ligated into the HindIII/SalI digested pEGFP-N3 vector to finally generate the pEG-MBD2a-ΔC.

Using the pEG-MBD2a as template, the pEG-MBD2-N (aa: 1-152), -MBD2a-1-235, -MBD2a-C (236-414) were generated by Q5 site-directed mutagenesis strategy following the standard protocol (NEB, E0554S). Using pEG-MBD2b (1-235) as template, the pEG-MBD-TRD were generated by Q5 site-directed mutagenesis strategy.

pEG-MBD2c was generated by inserting the synthesized short double-strand short C-terminus of MBD2c into the C-terminus of MBD2a-ΔC within pEG-MBD2a-ΔC-MeCP2. In brief, the forward and reverse strands of the C-terminus of MBD2c with EcoNI and SalI overhangs separately were synthesized and annealed by 95 °C for 3 min and slow cooling down till room temperature. The dsDNA with EcoNI and SalI overhangs was then inserted into the EcoNI/SalI digested pEGFP-MBD2a-ΔC vector to finally generate the pEG-MBD2c vector.

4.2 Protein purification and analysis

4.2.1 Protein purification

(GFP tagged) MeCP2 full length and truncations, MBD1ΔCxxC3, MBD2b, MBD3, MBD4 carrying the C-terminal intein-CBD were expressed in BL21(DE3) *E. coli* cells. MBD2a carrying the C-terminal intein-CBD were expressed in BL21(DE3) pLysS *E. coli* cells. MBD2a-N, MBD2a-ΔC (MBD2c), MBD2a-MBD-TRD, and MBD2a-C carrying the C-terminal intein-CBD were expressed in BL21(DE3) Star *E. coli* cells (Table 4).

Expression was induced with 0.5 mM IPTG (Sigma-Aldrich, I6758-10G) at room temperature (RT) overnight. Subsequently, the cell lysates were prepared by pelleting and resuspending the bacteria in lysis buffer (20 mM Tris-HCl pH 8.5, 500 mM NaCl, 0.25% Triton X-100 and protease inhibitors PMSF, AEBSF, E64, and pepstatin A), followed by sonication on ice and centrifugation at 15000 rpm for 30 min. The cleared lysates were incubated with 2 ml chitin beads (NEB, S6651S) at 4 °C with rotation for 3 h to allow CBD-chitin binding. Then beads were washed and treated in benzonase buffer (20 mM Tris-HCl, pH 8.5, 2 mM MgCl2, 20 mM NaCl, 0.1 mM PMSF) with benzonase (MERCK, 70746-3, 1:2000 dilution) at 37 °C for 4 h, followed by washing and treatment in DNase buffer (20 mM Tris-HCl pH 8.5, 50 mM KCl, 2 mM MgCl2) with DNase I (4 µg/ml), RNase A (0.2 µg/µl) at 37 °C for 25 min to remove DNA and RNA contaminants. Finally, proteins were eluted by cleavage at 4 °C for two days in cleavage buffer (20 mM Tris-HCl pH 8.5, 500 mM NaCl) with 50 mM DTT (Sigma-Aldrich, D9779-5G). The eluted fraction was concentrated using Amicon® Ultra Centrifugal Filters (MERCK), aliquoted, flash frozen, and stored at -80 °C in storage buffer (20 mM Tris-HCl, pH 8.5, 300 mM NaCl).

4.2.2 Protein quantification

Protein concentrations were determined using Pierce™ 660 nm Protein Assay Kit (Thermo Fisher Scientific, 22660) following the manufacturer's instruction. In brief, 10 µl of BSA standard (Thermo Fisher Scientific, 23208) proteins and storage buffer (blank) were mixed with 150 µl Protein Assay Reagent in the 96-well micro test plate (SARSTEDT, 82.1581.001) and incubated at room temperature for 5 min. Three replicates for each condition were performed. The absorbance at 660 nm was measured using a plate reader Infinite 200 (TECAN). The blank-corrected absorbance was calculated by subtracting the average absorbance of the blank. The standard curve was generated by plotting the average blank-corrected absorbance for each BSA standard versus the relative concentrations (µg/µl). The protein concentrations were calculated according to the standard curve using the blank-corrected measurements.

4.2.3 Protein purity analysis

2 µg and 10 µg protein were loaded separately onto an SDS-PAGE gel and 15% Tris-borate EDTA polyacrylamide gel. The gels were stained with coomassie (to detect the proteins) and ethidium bromide (EtBr; to detect the potential contamination with nucleic acids) separately after electrophoresis. The SDS-PAGE gels after coomassie staining were captured by colorimetric trans-illumination imaging using the Amersham Imager 600 (Table 6) equipped with white light trans-illumination following the manufacturer's instruction. The Tris-borate EDTA PAGE gels after EtBr staining were imaged using the VWR genosmart (Table 6) UV trans-illumination system.

4.3 DNA templates and methylation

4.3.1 Synthesis of short DNA templates for binding assay

42 bp dsDNA was synthesized by primer extension using the large (Klenow) fragment of *E. coli* DNA polymerase I (NEB, M0210L) as described before (Rothbauer, Zolghadr et al. 2008, Zhang, Hastert et al. 2017). In brief, the longer CG-up/MG-up and shorter Fill-in-647N (Table 3) were annealed by slowly cooling down to 37 °C from 95 °C in NEB buffer 2 (50 mM NaCl, 10 mM Tris–HCl, 10 mM MgCl2, 1 mM dithiothreitol; New England Biolabs). The short CG/MG-down were extended by adding 1 mM dATP, dGTP, dTTP, and 1mM dCTP (Carl Roth) or 0.1 mM dmCTP (Jena Bioscience, NU-1125S) and Klenow fragment polymerase, followed by incubating for 1 h at 37 °C to generate the double-strand oligos with or without CpG methylation.

20 bp dsDNA with or without cytosine methylation was generated by resuspending and mixing the Cy5-MG/CG-up and MG/CG-down (Table 3) to a final concentration of 10 mM in

a solution containing 20 mM Tris pH 8.5 and 150 mM NaCl, followed by heating at 95 °C for 2 min and gradual cooling by switching off the thermomixer.

4.3.2 Synthesis of long DNA templates for phase separation assay

The DNA used for the phase separation assay with different lengths and methylation levels were synthesized by PCR (polymerase chain reaction) using Q5 polymerase (NEB, M0491S) as described before (Ludwig, Zhang et al. 2017). In brief, pUC18-MINX plasmid (Table 2) was applied as a template and different reverse (Rev) primers (Table 3) were used to amplify DNA of different lengths.

4.3.3 Synthesis of methylated DNA

The DNA with various methylation levels at cytosine residues randomly were generated by partially replacing the dCTP with dmCTP, followed by Q5 polymerase directed PCR as aforementioned.

The 800 bp DNA with CpG methylation was obtained with the CpG methyltransferase M.SssI (NEB, M0226S) after PCR and followed by DNA purification from agarose gel according to the manufacturer's instructions. Briefly, 1 µg of purified 800 bp DNA product was mixed with 160 µM SAM (S-adenosyl-methionine; NEB, B9003S), methylated by 4 units M.SssI for 4 h at 37 °C in the 1 x NEB buffer 2.

4.3.4 DNA methylation assay

DNA methylation levels were determined with methylation-sensitive restriction enzymes HpaII (NEB, R0171S) and its methylation insensitive isoschizomers MspI (NEB, R0106S). In brief, 150 ng DNA was incubated with MspI or HpaII (control with no enzyme) at 37 °C for 2 h before being loaded to 15% Tris-borate EDTA PAGE. After electrophoresis, the gel was stained with EtBr for 4 min and washed with water. The images were captured by the Amersham Imager 600 equipped with RGB fluorescence (Cy5-20 bp DNA) and the VWR genosmart UV trans-illumination system (Table 6).

4.4 *In vitro* phase separation and detection

4.4.1 *In vitro* phase separation

Proteins were firstly thawed on ice, centrifuged at 14000 rpm, 4 °C for 10 min to remove all aggregates. Phase separation in solutions (20 mM Tris-HCl pH 8.5) with various concentrations of salt, protein, crowding agents, and DNA was achieved by incubating for 45 min at RT (Fig. 9A).

4.4.2 Microscopy analysis

To check the droplet morphology, phase separation samples were loaded onto chambers made of double-sided tapes and sealed with coverslips (Fig. 9B). Fluorescence and differential interference contrast (DIC) images were taken using a Nikon Eclipse TiE2 microscope equipped with a Plan Apo λ 40x objective or a Nikon Ti-E microscope equipped with a CFI Planapochromat VC 20x objective (Table 6). All images were processed and analyzed using ImageJ.

The biophysical properties of droplets were quantified using ImageJ according to the fluorescent images from the phase separation assay. Droplets were identified and segmented using an FFT/bandpass filter to decrease the background and fluorescence intensity-based thresholding which was manually adjusted. Droplets with an area > 0.1 μm^2 were considered. The aspect ratio of droplets was calculated by the ratio of the minimal Feret diameter to the maximal Feret diameter.

4.4.3 Droplet sedimentation assay

The phase separation samples were sedimented by centrifugation at 14000 rpm for 15 min at RT. The top half of the supernatants were applied to a 12% SDS-PAGE gel, which was stained with Coomassie for 1-2 h after electrophoresis and subsequently washed with destaining buffer (100 ml acetic acid, 100 ml ethanol, and 500 ml H_2O) overnight. The image was taken using an Amersham Imager machine (Table 6), quantitatively analyzed, and plotted (Fig. 9C).

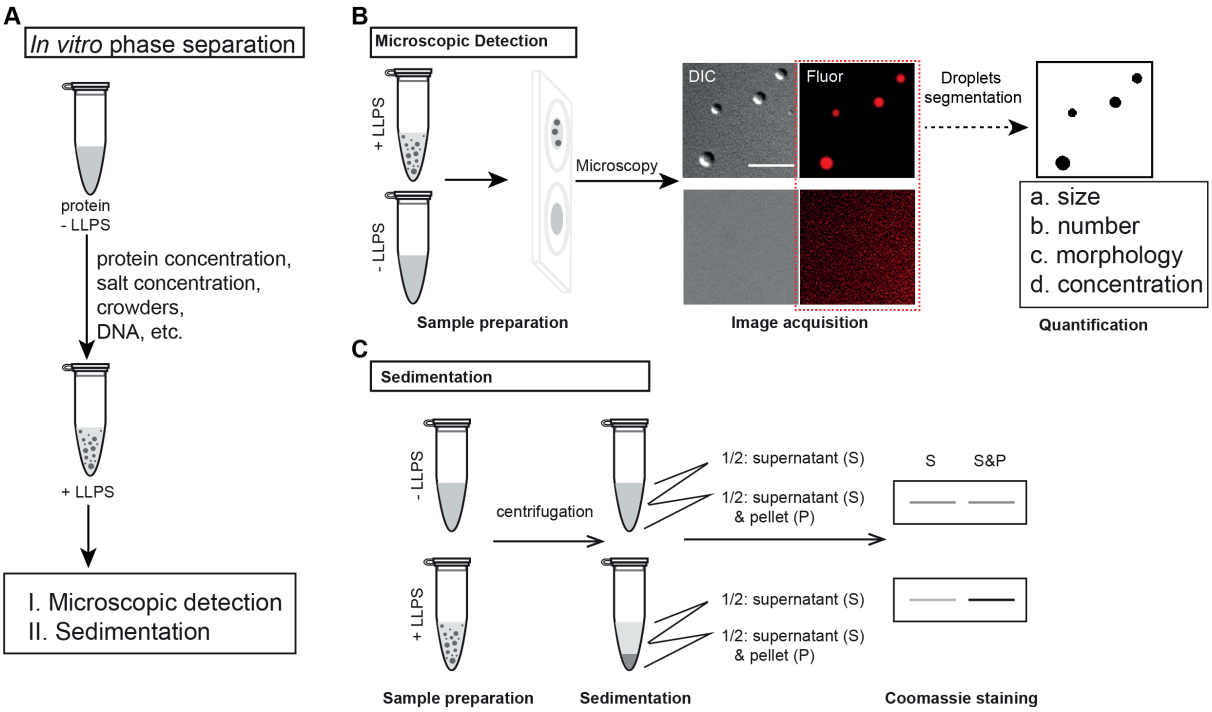


Figure 9 Graph elucidating the *in vitro* phase separation and detection

(A) *In vitro* phase separation assay. Purified proteins were incubated to form protein droplets at different conditions (Protein concentration, salt concentration, etc.) at room temperature for 45 min.
(B) Microscopic detection of droplets. The mixtures after in vitro phase separation assay were transferred to chambers made of double-sided tapes and covered with coverslips. Droplets were visualized under the microscope. With proper fluorescent labeling, the droplets were segmented followed by quantitative analysis.
(C) Sedimentation assay. The droplets after in vitro phase separation assay were pelleted by centrifuge. The clear supernatants were collected and loaded into SDS-PAGE, followed by coomassie staining and quantitative analysis.

4.4.4 Turbidity assay

20 µl phase separation solutions with various conditions were prepared as above and transferred to the 384-well plate with an optically clear bottom (PerkinElmer, 6007550). The phase separation was done by incubating at RT for 45 min, followed by absorbance measurement at 340 nm at room temperature using a plate reader Infinite 200 (TECAN).

4.5 In-droplet and in-solution MeCP2 quantification

4.5.1 Standard curve

Solutions containing gradient concentrations of GFP-MeCP2 were loaded onto chambers made of double-sided tapes and sealed with coverslips. The images were taken using a Nikon Eclipse TiE2 microscope equipped with Plan Apo λ 40X objective (Table 6). The mean fluorescence intensities of free GFP-MeCP2 were measured using ImageJ and plotted versus the corresponding known GFP-MeCP2 concentrations to generate the standard curves.

4.5.2 In-droplet and in-solution MeCP2 quantification

Purified GFP-MeCP2 was mixed with unlabeled MeCP2 in a molar ratio of 1:99 and diluted to a final concentration of 80 µM in buffer (20 mM Tris-HCl, pH 8.5, 300 mM NaCl). *In vitro* phase separation was done by incubating at RT for 45 min at various conditions. To quantify the protein concentration in droplets, droplet mixtures were then moved to chambers made of double-sided tapes and sealed with coverslips. To quantify the protein concentration in solution, droplets were sedimented by centrifuge at 14000 rpm for 15 min at RT, the upper clear supernatants were transferred to new PCR tubes and mixed by pipetting before being transferred to chambers. The images were taken by a Nikon Eclipse TiE2 microscope equipped with a Plan Apo λ 40x objective (Table 6). Mean fluorophore intensity in droplets or solutions was measured using ImageJ, the protein concentrations were calculated relative to the standard curves.

4.6 Pull-down interaction assay

The immobilization of untagged human MeCP2 and truncations with intein-CBD were generated as described above. Briefly, the human MeCP2 or truncations fused with intein-

CBD were bound to the chitin beads by incubating the bacteria extracts with chitin beads. Then the beads were treated with nuclease to remove nucleic acid contaminants as described above. 25 µl clean beads (control) and beads with immobilized MeCP2 or truncations were transferred to cold 1.5 ml tubes, washed with PBS supplemented with 125 mM NaCl, 0.05% NP-40, and protease inhibitors PMSF, AEBSF, E64, and pepstatin A for twice, and incubated with 300 µl 0.1 µg/µl corresponding GFP tagged full-length MeCP2 or truncations in PBS supplemented with 125 mM NaCl, 0.05% NP-40, and protease inhibitors for 90 min with rotation at 4 °C. Finally, the beads were collected by centrifugation, washed three times using the PBS supplemented with 125 mM NaCl, 0.05% NP-40, and protease inhibitors and followed by adding 50 µl 1X protein loading buffer (1% SDS, 25 mM Tris pH 6.8, 5% glycerol, 50 mM DTT and 0.005% bromophenol blue), and boiling at 95 °C for 5 min. Samples were then loaded to SDS-PAGE gel, electrophoresis was performed at 90 V for 100 min followed by semi-dry transfer at 25 V for 35 min. Then the membranes were blocked with 3% low-fat milk in PBS for 30 min at RT, incubated with rabbit anti-GFP antibody (Table 7) at 4 °C overnight on a rotary shaker, and followed by anti-rabbit IgG conjugated with Cy3 for 1 h at RT. The fluorescence signals were detected using an Amersham Imager (Table 6).

4.7 Microscale thermophoresis assay

The microscale thermophoresis (MST) assay was conducted using the Monolith NT.115 (NanoTemper) following the manufacturer's instructions. To check if MeCP2 is capable of self-interaction, 10 µl 200 or 300 µM GFP tagged MeCP2 or truncations were mixed with equal volumes of untagged MeCP2 gradients or truncations in buffer containing 20 mM Tris-HCl and 150 mM NaCl. To check the binding affinity of MeCP2 with DNA, 10 µl 100 µM short oligos labeled with ATTO-647N and Cy5 with or without methylation (Table 3) were mixed with equal volumes of untagged MeCP2 gradients or truncations in buffer containing 20 mM Tris-HCl and 150 mM NaCl. Then the mixtures were transferred into Monolith NT^{TM} capillaries and the fluorescent changes to microscopic temperature gradients were measured by Monolith NT.115 using 20% (for self-interaction assay) or 40-50% (for protein-DNA interaction) excitation power. The dissociation constants (Kd) were calculated using MO. Affinity Analysis software.

4.8 Mammalian cell culture and transfection

C2C12 mouse myoblast cells (Table 5) were cultured in Dulbecco’s modified Eagle medium (DMEM) high glucose (Sigma-Aldrich Chemie GmbH, D6429) supplemented with 20% fetal calf serum (FCS), 1x L-glutamine (Sigma-Aldrich Chemie GmbH, G7513), and 1 µM gentamicin (Sigma- Aldrich Chemie GmbH, G1397).

MEF-P (P53-/-, methylation proficient) and MEF-PM (P53-/- DNMT1n/m, methylation deficient) mouse embryonic fibroblast (Table 5) were cultured in DMEM high glucose (Sigma-Aldrich Chemie GmbH, D6429) supplemented with 15% FCS, 1x L-glutamine (Sigma-Aldrich Chemie GmbH, G7513), and 1 µM gentamicin (Sigma- Aldrich Chemie GmbH, G1397).

The transfection was performed using a Neon Transfection System (Thermofisher) following the manufacturer's instruction. Cells were seeded into 35 mm plates with a glass bottom for live cell experiments (Fig. 10). Cells were seeded into plates with coverslips for fixed cell experiments (Fig. 10).

4.9 Fluorescence and image analysis

36 h after transfection, C2C12 cells were fixed with 3.7% formaldehyde in phosphate buffered saline (PBS) for 15 min, washed 3 times using 0.02% PBST (0.02% Tween 20 in PBS), permeabilized using 0.5% Triton X-100 in PBS for 10 min, and washed 3 times using 0.02% PBST. For DNA visualization, cells were incubated with 4,6-Diamidine2-phenylindole dihydrochloride (DAPI, 1 g/ml) for 6 min, followed by 3 times washing using 0.02% PBST. Cells were finally kept in mounting media and stored at -20 °C till use.

GFP and DAPI signals were imaged using Leica TCS SP5 II confocal microscope with a HCX PL APO 100x / 1.44 oil Corr CS objective (Table 6). 405 nm Diode Laser (50 mW) and 488 nm argon ion laser (~20 mW) were applied for DAPI and GFP signal excitation, respectively. The settings for scanning were scanning at 400 Hz, 2.5x zoom, image format 1024 × 1024 pixels, pinhole 95.55 µm.

The images were further quantitatively analyzed with ImageJ (Fig. 10). The cell nuclei and heterochromatin compartments were identified by Auto Threshold based on the DAPI intensities (nuclei segmentation: "Auto threshold" "Li white"; heterochromatin compartment segmentation: "Gaussian Blur..." "sigma=4", "Subtract Background..." "rolling=15"). Cells were subgrouped by GFP intensities. Nuclei and heterochromatin compartment parameters (size, number, GFP intensity, Ferret diameters, etc.) in each cell were measured.

The mean heterochromatin compartment size per nuclei was calculated and plotted.

The GFP fold enrichment was calculated as the ratio of mean GFP intensity within heterochromatin compartments to the mean GFP intensity outside heterochromatin compartments per nuclei. The equation is: $E = I_{CC} / [(I_N - I_{CC} \times S_{CC} \times N) / (S_N - S_{CC} \times N)$. E was the fold enrichment; I_{CC} and I_N were mean GFP intensity inside heterochromatin compartments per nuclei and mean GFP intensity per nuclei, respectively; S_{CC} and S_N were the mean heterochromatin compartment area per nuclei and nuclei area; N was the number of heterochromatin compartments per nuclei.

The Feret diameter (or caliper diameter) is defined as the distance between the two parallel planes restricting the object perpendicular to that direction. Minimal and maximal Feret diameters (MinFeret and MaxFeret separately) per heterochromatin compartment were measured. Heterochromatin compartments with mean GFP intensity > 0.01 were applied. Aspect ratio (AR) was calculated using the equation AR = MinFeret/MaxFeret and plotted.

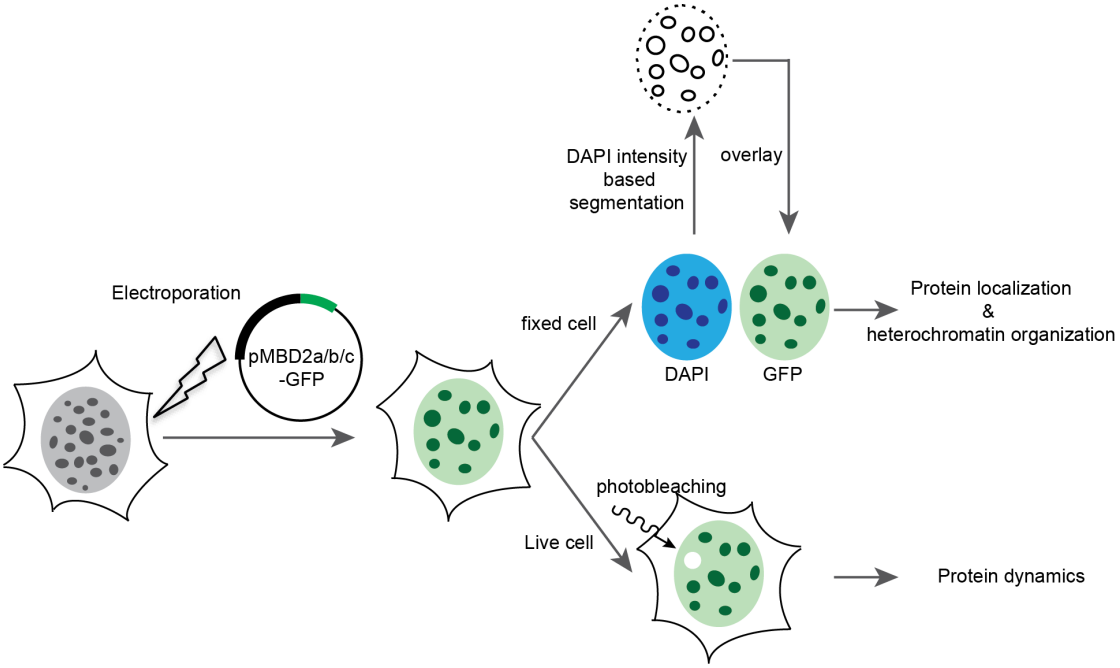


Figure 10 Schematic graph examining the functions and dynamics of MBDs *in vivo*.

C2C12 cells were transfected with plasmids expressing GFP tagged MBDs. After incubation, the cells were fixed and stained with DAPI (top). The nuclei and heterochromatin compartments were segmented based on the DAPI intensities. Alternative, the live cells were directed applied for fluorescence recovery after photobleaching (FRAP) 24 hours after transfection (bottom).

4.10 Fluorescence recovery after photobleaching (FRAP)

24 h after transfection, live cells were transferred onto a prewarmed Leica TCS SP5 II confocal microscope with a HCX PL APO 100x / 1.44 oil Corr CS objective (Fig. 10, Table 6). The FRAP assay was performed at 37 °C. For excitation of GFP, the 13% 488 nm argon ion laser was applied. The emission at 490-560 nm was detected. The settings for scanning were scanning at 400 Hz, 10x zoom, image format 256 × 256 pixels, pinhole 95.55 µm. The photobleaching was obtained by 100% 488 nm argon ion laser power. The fluorescent changes were tracked with 9 images taken before and 130 taken after photobleaching with an interval of ~1.3 s.

FRAP recovery plots were generated from the fluorescence intensities in the bleached region of interest (ROI), nucleoplasm (NP), and background (BG) at each time step. The fluorescence recovery in the ROI was double normalized using the equation $F_t = [(I_{ROI(t)} - I_{BG(t)})/(I_{ROI(0)} - I_{BG(0)})]/[(I_{NP(t)} - I_{BG(t)})/(I_{NP(0)} - I_{BG(0)})]$ (Phair, Gorski et al. 2003).F_t was the

normalized fluorescence recovery at each time point (t). $I_{ROI(t)}$ and $I_{ROI(0)}$ were the mean fluorescence intensity of the bleached ROI at time point (t) and prebleached point, separately. $I_{BG(t)}$ and $I_{BG(0)}$ were the mean fluorescence intensity of the background (BG) at time point (t) and prebleached point, separately. $I_{NP(t)}$ and $I_{NP(0)}$ were the mean fluorescence intensity of the unbleached NP at time point (t) and prebleached point, separately. Half recovery time (t1/2) and immobile fractions were extracted from the mean exponential fitting. Plots were generated using the GraphPad Prism software (https://www.graphpad.com/scientific-software/prism/).

Table 2: Plasmid characteristics.

Name	pc number*	Fluorophore	Gene species	Promoter	Reference
pEGFP-MeCP2	pc1208	EGFP	*Homo sapiens*	CMV	Kudo et al., 2003
pEGFP-MeCP2_R111G	pc1441	EGFP	*Homo sapiens*	CMV	Kudo et al, 2003
pEGFP-MeCP2_R168X	pc4748	EGFP	*Homo sapiens*	CMV	This study
pEGFP-MeCP2_R255X	pc4749	EGFP	*Homo sapiens*	CMV	This study
pEGFP-MeCP2_R270X	pc4774	EGFP	*Homo sapiens*	CMV	This study
pEGFP-MeCP2_G273X	pc4775	EGFP	*Homo sapiens*	CMV	This study
pEGFP-MeCP2_R294X	pc4776	EGFP	*Homo sapiens*	CMV	This study
pEGFP-MeCP2ΔAT-hook2	pc4777	EGFP	*Homo sapiens*	CMV	This study
pMBD1-GFP	pc1191	EGFP	*Mus musculus*	CMV	Hendrich and Bird, 1998
pGmMBD1.3	pc2899	EGFP	*Mus musculus*	CMV	Zhang et al. 2017
pGFP-MBD3	pc1193	EGFP	*Mus musculus*	CMV	Hendrich and Bird, 1998
pMBD4-GFP	pc1194	EGFP	*Mus musculus*	CMV	Hendrich and Bird, 1998
peMBD2G	pc2399	EGFP	*Mus musculus*	CMV	Cardoso Lab
pMBD2.2-GFP	pc2068	EGFP	*Mus musculus*	CMV	Cardoso Lab
pmMBD2.4G	pc2841	EGFP	*Mus musculus*	CMV	Zhang et al. 2017
pMBD2.1G-GFP	pc2067	EGFP	*Mus musculus*	CMV	Cardoso Lab
pmMBD2.6G	pc2843	EGFP	*Mus musculus*	CMV	Zhang et al. 2017
pEGFP-MBD2aΔC	pc4794	EGFP	*Mus musculus*	CMV	This study
pEGFP-MBD2c	pc4795	EGFP	*Mus musculus*	CMV	This study
pTYB1-MeCP2	pc1294	-	*Homo sapiens*	T7	Georgel, et al. 2003
pTYB1-GFP-MeCP2	pc4741	GFP	*Homo sapiens*	T7	This study
pTYB1-MeCP2-R168X	pc4734	-	*Homo sapiens*	T7	This study
pTYB1-GFP-MeCP2-R168X	pc4742	GFP	*Homo sapiens*	T7	This study
pTYB1-MeCP2-R255X	pc4735	-	*Homo sapiens*	T7	This study
pTYB1-GFP-MeCP2-R255X	pc4743	GFP	*Homo sapiens*	T7	This study
pTYB1-MeCP2-R270X	pc4778	-	*Homo sapiens*	T7	This study
pTYB1-MeCP2-G273X	pc4779	-	*Homo sapiens*	T7	This study

pTYB1-MeCP2-R294X	pc4781	-	*Homo sapiens*	T7	This study
pTYB1-MeCP2ΔAT-hook2	pc4782	-	*Homo sapiens*	T7	This study
pTYB1-MBD1ΔCxxC3	pc4783	-	*Mus musculus*	T7	This study
pTYB1-MBD3	pc4784	-	*Mus musculus*	T7	This study
pTYB1-MBD4	pc4785	-	*Mus musculus*	T7	This study
pTYB1-MBD2a	pc4786	-	*Mus musculus*	T7	This study
pTYB1-MBD2b	pc4787	-	*Mus musculus*	T7	This study
pTYB1-MBD2c	pc4788	-	*Mus musculus*	T7	This study
pTYB1-MBD2a-N	pc4789	-	*Mus musculus*	T7	This study
pTYB1-MBD2a-MBDTRD	pc4791	-	*Mus musculus*	T7	This study
pTYB1-MBD2ΔC	pc4792	-	*Mus musculus*	T7	This study
pTYB1-MBD2-C	pc4793	-	*Mus musculus*	T7	This study
pUC18-MINX-M3	pc3902	-	Synthetic	AmpR	Zhang, et al. 2017; Zillmann M, et al. 1988

*pc: plasmid collection. "-": no

Table 3: Oligonucleotide characteristics.

Name	Sequence [5' - 3']	Application	Reference
NdeI-GMeCP2-F	ACATATGATGGTGAGCAAGGGCGAG	pTYB1-GFP-MeCP2	This study
EcoRI-GMecp2-R	CGAATTCGCTAACTCTCTCGGTCACG	pTYB1-GFP-MeCP2	This study
pTYB1-BB PCR-F-1	TCGAGGGCTCTTCCTGCTTTGCC	TYB1-hMeCP2-1-167, 1-255	This study
pTYB1-BB PCR-F-2	GAGGGCTCTTCCTGCTTTGC	pTYB1-MBD2a-N, MBDTRD, MBD2aΔC,	This study
hMECP2-1_167-R	CCGGGAGGGGCTCCCTCT	pTYB1- hMECP2-1_167; pEG-hMECP2-1_167	This study
hMECP2-1_254-R	CTTCCTGCCGGGGCGTTT	pTYB1- hMECP2-1_254; pEG-hMECP2-1_254	This study
hMECP2-1_269-R	GCCCCGTTTCTTGGGAAT	pTYB1- hMECP2-1_269; pEG-hMECP2-1_269	This study
hMECP2-1_272-R	CGGCTTTCGGCCCCGTTTC	pTYB1- hMECP2-1_272; pEG-hMECP2-1_272	This study
hMECP2-1_309-R	CCGGGTCTTGCGCTTCTT	pTYB1- hMECP2-1_309; pEG-hMECP2-1_309	This study

hMECP2-1_293-R	GATAGAAGACTCCTTCACG	pTYB1- hMECP2-1_293; pEG-hMECP2-1_293	This study
pEG-hMECP2-BB-F-2	TGACTTTACACGGAGGATCCACCG	pEG-hMeCP2-1-167, 1-254	This study
MeCP2-AT-hook2-F	AGTGTGGTGGCAGCCGCTG	pTYB1-MeCp2ΔAT-hook2, pEG-MeCp2ΔAT-hook2	This study
MeCP2-AT-hook2-R	GGCCTGAGGGTCGGCCTCA		This study
NdeI-MBD1-F	AAGAAGGAGATATACATATGATGGCTGAGTCCTGGCAG	pTYB1-MBD1	This study
EcoRI-MBD1-R	GAAGAGCCCTCGAGGAATTCCAAAACTTCCTCCTTCAA	pTYB1-MBD1	This study
NdeI-MBD2-F	AAGAAGGAGATATACATATGATGCGCGCGCACCCGGGG	pTYB1-MBD2a	This study
EcoRI-MBD2-R	GAAGAGCCCTCGAGGAATTCCGCCTCATCTCCATCGTC	pTYB1-MBD2a	This study
NdeI-MBD3-F	AAGAAGGAGATATACATATGATGGAGCGGAAGAGGTGG	pTYB1-MBD3	This study
EcoRI-MBD3-R	GAAGAGCCCTCGAGGAATTCCACTCGCTCTGGCTCCG	pTYB1-MBD3	This study
NdeI-MBD4-F	AAGAAGGAGATATACATATGATGGAGAGCCCAAACCTT	pTYB1-MBD4	This study
EcoRI-MBD4-R	GAAGAGCCCTCGAGGAATTCAGATAGACTTAATTTTT	pTYB1-MBD4	This study
NdeI-MBD2-153-414-F	TATGAAGAGTACATCATATGATGGACTGCCCGGCCCTC	pTYB1-MBD2b	This study
EcoRI-MBD2-153-414-R	GAAGAGCCCTCGAGGAATTCCGCCTCATCTCCATCGTC	pTYB1-MBD2b	This study
MBD2-236_414-F	AACAAGGGTAAACCAGAC	pTYB1-MBD2a-C	This study
pTYB1-BB PCR-R-1	CATCATATGTATATCTCCTTCT	pTYB1-MBD2a-C	This study
MBD2-1_152-R	CCTCTTCCCGCTCTCCG	pTYB1-MBD2a-N	This study
MBD2-1_235-R-2	CTGATTGAGGGGGTCATTCCG	pTYB1-MBDTRD; pTYB1-MBD2a-ΔC	This study
HindIII-MBD2-F	ACATATTTATAAGCTTATGCGCGCGCACCCG	pEG- MBD2a-ΔC	This study
SalI-MBD2a-1_235-R	GCACGCATTATCGTCGACCTGATTGAGGGGGTCATTCC	pEG- MBD2a-ΔC	This study
MBD2c-C-terminus-F	ATCAGAACAAGTTCCGGCTCATTAAGAAGCAGACACTAATTGGTCTCG	pEG- MBD2c	This study
MBD2c-C-terminus-R	TCGACGAGACCAATTAGTGTCTGCTTCTTAATGAGCCGGAACTTGTTCTGA	pEG- MBD2c	This study
MG-up	CTCAACAACTAACTACCATCXGGACCAGAAGAGTCATCATGG	dsDNA synthesis	Rothbauer, 2008; Zhang, et al. 2017
CG-up	CTCAACAACTAACTACCATCCGGACCAGAAGAGTCATCA	dsDNA synthesis	Rothbauer, 2008;

	TGG		Zhang, et al. 2017
Fill-In-550	ATTO550- CCA TGA TGA CTC TTC TGG TC	dsDNA synthesis	Rothbauer, 2008; Zhang, et al. 2017
Fill-In-647N	ATTO647N- CCA TGA TGA CTC TTC TGG TC	dsDNA synthesis	Rothbauer, 2008; Zhang, et al. 2017
MGup(20bp)	Cy5-ACTACCATCXGGACCAGAAG	dsDNA synthesis	Zhang, et al. 2017
MGdown(20bp)	CTTCTGGTCXGGATGGTAGT	dsDNA synthesis	Zhang, et al. 2017
CGup(20bp)	Cy5-ACTACCATCCGGACCAGAAG	dsDNA synthesis	Zhang, et al. 2017
CGdown(20bp)	CTTCTGGTCCGGATGGTAGT	dsDNA synthesis	Zhang, et al. 2017
Fw	CGGTACCTAATACGACTCACTATA	dsDNA synthesis (PCR)	Zhang, et al. 2017
Rev 380	GTGCCAAGCTTGCATGC	dsDNA synthesis (PCR)	Zhang, et al. 2017
Rev 800	ATAGGCGTATCACGAGGC	dsDNA synthesis (PCR)	Zhang, et al. 2017
Rev 1600	TGGTCCTGCAACTTTATCCG	dsDNA synthesis (PCR)	Zhang, et al. 2017
Rev 3000	ATTCGTAATCATGGTCATAGCTG	dsDNA synthesis (PCR)	Zhang, et al. 2017

X: mC

Table 4: Bacterial cell line characteristics.

Name	Species	Purpose	Genotype	Reference
Top 10	*Escherichia coli*	plasmid production	F^– mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139	Invitrogen
BL21(DE3)	*Escherichia coli*	IPTG induced protein expression	F – ompT hsdSB (rB – , m~B–) gal dcm (DE3)	Studier and Moffatt, 1986
BL21(DE3) pLysS	*Escherichia coli*	IPTG induced protein expression	F- ompT hsdSB(rB- mB-) gal dcm (DE3) pLysS (CamR)	Studier and Moffatt, 1986
BL21(DE3) Star	*Escherichia coli*	IPTG induced protein expression	F-ompT hsdSB (rB-, mB-) galdcmrne131 (DE3)	New England Biolabs GmbH (NEB)

Table 5: Eukaryotic cell line characteristics.

Name	Species	Type	Genotype	Reference
C2C12	*Mus musculus*	myoblast	wildtype	Yaffe et al, 1977
MEF P	*Mus musculus*	embryonic fibroblast	p53 -/-	Lande-Diner et al, 2007
MEF PM	*Mus musculus*	embryonic fibroblast	p53-/- Dmnt1 -/-	Lande-Diner et al, 2007

Table 6: Imaging systems characteristics.

Microscope/ Company	Lasers/lamps	Filters (ex. & em. [nm])*	Objectives/ lenses	Detection system	Incubation system	Application
Amersham AI600 imager	white transillumination: 470 – 635 nm; UV transillumination: 312 nm	-	-	16-bit Pelltier cooled Fujifilm Super CCD	-	Western blot, coomassie gels
Nikon Eclipse TiE2	SPECTRA X LED 470/24 nm (196 mW) 640/30 nm (231 mW)	em.: Quadbandpass (432/25 nm; 515/25 nm; 595/25 nm; 730/70 nm)	40X air Plan Apo λ DIC (0.95 NA, 230 µm WD)	Nikon Qi2 751600 16.25 MPx	-	Phase separation
Nikon Ti-E	Wide-field epifluorescence; Transmitted light source 88 nm (50 mW)	F36-525 (472/30 & 520/35)	Plan Fluar Multiimersion 20x (0.75 NA, 0.38 µm WD)	Hamamatsu® Electron Multiplying Charge-Coupled Devices (EMCCDs) C9100-50 EMCCD	-	Phase separation
Nikon Eclipse Ti	AOTF ** shutter 488: Sapphire 488-75 561: Sapphire 561-75 CW 640: OBIS 640 (100 mW)	em.: Quadbandpass (432/25 nm; 515/25 nm; 595/25 nm; 730/70 nm)	oil immersion Nikon 100X Apo TIRF (1.49 NA)	Andor iXon DU897BV EMCCD (EM: 300, camera zoom factor 1.5X, 107 nm/px)	-	Single-Molecule Tracking
Confocal microscope Leica TCS SP5-II	405 nm Diode Laser 50 mW; 488 nm Argon ion laser: 458 nm ~5 mW 476 nm ~5 mW 488 nm ~20 mW 496 nm ~5 mW 514 nm ~20 mW	DAPI: ex. 420/30 em. 465/20 FITC: ex. 495/15 em. 530/30 Rhod: ex. 570/20 em. 640/40	HCX PL APO 100x / 1.44 oil Corr CS	HyD Hybrid Detectors	ACU live cell chamber (Olympus)	fixed and live cell imaging (FRAP)

*ex.: excitation; em.: emission; ** AOTF: Acousto-optic tunable filters EtBr: ethidium bromide.

Table 7: Primary and secondary antibody characteristics.

Reactivity	Host	Dilution	Application	Catalog/clone	Company/reference
Anti-GFP (primary)	rabbit	1:100	WB	CCHL Lab	Rothbauer et al, 2008
Anti-rabbit IgG Cy3 (secondary)	donkey	1:1000	WB	711-165-152	Jackson ImmunoResearch

WB: western blot

5 Results

5.1 The MBD family proteins drive heterochromatin organization via phase separation

5.1.1 The phase separation prediction of the MBD family proteins

Previous studies show that the HP1-driven phase separation promotes heterochromatin formation (Strom, Emelyanov et al. 2017, Erdel, Rademacher et al. 2020). Yet, the heterochromatin compartment is a complex context, in which many factors are involved and may play overlapping roles, such as SUV39H1, Dnmt, and members of the MBD protein family. It is interesting to explore if phase separation is a common mechanism underlying heterochromatin organization. The MBD family proteins (MBDs), as DNA methylation readers, are of particular importance.

Proteins that could undergo liquid-liquid phase separation (LLPS) always contain the intrinsically disordered regions (IDRs) or low complexity regions (LCs). First of all, we analyzed the disorder properties of the five members of the MBD protein family using the online PONDR VLXT algorithm for disorder prediction (http://www.pondr.com/), the EMBOSS charge for residue charge analysis (https://www.bioinformatics.nl/cgi-bin/emboss/charge), and the Peptide property calculator (https://pepcalc.com/) for the iso-electric point (pI) prediction (Fig. 1.1). We found that all MBDs are more than 50% disordered as predicted except for MBD4 which is around 48.74%. The disordered regions are distributed across the amino acid sequence flanking the MBD domain (Fig. 1.1A-E). This means that the MBDs could probably undergo LLPS *in vivo* and *in vitro*. Besides, the isoelectric points (pI) of all proteins except for MBD3 are more than 7, predicting probably unspecific binding with negative charged DNA (Fig. 1.1F).

Besides, as the best-characterized member of mCpG binding protein, previous studies support the hypothesis that LLPS is the underlying mechanism of MeCP2 function in the heterochromatin compartments. Firstly, MeCP2 interacts with DNA, methyl cytosines, and nucleosomes via separate domains, and interacts with several chromatin proteins. Like most proteins that could form liquid-like droplets, MeCP2 intrinsically disordered regions consist of mainly positively charged residues (arginines, histidines, and lysines). These residues form electrostatic interactions with the negatively charged amino acids in other proteins and phosphates in DNA or RNA, thus, building multivalent protein-protein/DNA/RNA interactions.

Secondly, MeCP2 foci exhibit liquid-like properties *in vivo*. Brero et al. showed that MeCP2 forms round-shaped foci within the cell nucleus and that foci in close proximity tend to fuse over time. Furthermore, during mitosis, these chromatin clusters undergo fission and reformation. MeCP2 was also shown to promote heterochromatin compartment clustering in

a dose-dependent manner (Brero, Easwaran et al. 2005). Thirdly, purified MeCP2 protein alone showed oblate ellipsoid appearance in electron microscopy analysis (Nikitina, Shi et al. 2007).

Taken together, we propose that the functions of MBDs at the heterochromatin compartments are possibly mediated by (liquid-liquid) phase separation.

5.1.2 Phase separation properties of the MBD family proteins *in vitro*

First of all, we tested if the purified MBD family proteins (MBDs) are capable of phase separation. MBDs were purified from bacteria using chitin beads, but the eluted proteins contained large amounts of DNA. DNA has been recognized as an important regulator for the phase separation of many proteins in test tubes (Shakya and King 2018). Thus, DNaseI and benzonase were applied to remove the nucleic acids during purification (Fig. 1.2A). Unfortunately, we failed to purify the MBD1 protein due to its fast degradation but succeeded in purifying the MBD1ΔCxxC3, the short isoform lacking the CxxC3 domain.

LLPS is mediated by electrostatic interactions and/or hydrophobic interactions. The electrostatic interactions are enhanced at lower salt conditions, and the hydrophobic interactions can be disrupted by adding chemicals like 1,6-hexanediol. Firstly, we checked the phase separation properties of all MBDs using the *in vitro* phase separation assay at different salt concentrations. Both MBD2 and MBD3 formed spherical droplets at all conditions tested (Fig. 1.2C-D). While MeCP2 only formed liquid-like droplets at low salt conditions but did not form droplets at physiological salt concentrations (Fig. 1.2F). MBD1ΔCxxC3 and MBD4 formed irregular protein aggregates rather than liquid-like droplets (Fig. 1.2 B, E).

Then, we analyzed the phase separation properties of MBDs by disrupting the hydrophobic interactions. 1,6-hexanediol is an aliphatic alcohol that specifically disrupts weak hydrophobic interactions and is thus widely applied in the phase separation assays (Alberti, Gladfelter et al. 2019)(Fig. 1.3A). Irregular aggregates formed by MBD1ΔCxxC3 dissolve in a 1,6-hexanediol concentration-dependent manner (Fig. 1.3B). While the irregular aggregates formed by MBD4 became bigger at low 1,6-hexanediol concentrations and dissolved at higher 1,6-hexanediol concentrations (Fig. 1.3B). MBD2 and MBD3 droplets became irregular at lower 1,6-hexanediol concentrations and dissolved at higher 1,6-hexanediol concentrations (Fig. 1.3B). MeCP2 droplets kept spherical at lower 1,6-hexanediol concentrations but became less spherical at higher 1,6-hexanediol concentrations (Fig. 1.3B). Taken together, MeCP2 LLPS is likely mainly driven by electrostatic interactions, while MBD2/3 LLPS is mainly driven by hydrophobic interactions. Besides, disrupting the hydrophobic interactions (+ 1,6-hexanediol) induced a transition of MBD2/3 from spherical droplets to irregular aggregates.

Macromolecules were reported to promote the LLPS of proteins by increasing the local target concentration via the “excluded volume” effect. The artificial PEG 8000 (polyethylene glycol 8000) (Fig. 1.3C) is commonly applied for the *in vitro* phase separation assay to mimic a crowded environment (André and Spruijt 2020). We tested the influence of the crowding environments on the phase separation properties of MBDs in the physiological salt condition (150 mM NaCl). Upon adding PEG 8000, irregular aggregates from MBD1ΔCxxC3 and MBD4 became bigger (Fig. 1.3D). The spherical droplets formed by MBD2 and MBD3 became irregular aggregates, which grew at higher PEG 8000 concentrations (Fig. 1.3D). At physiological salt conditions, MeCP2 did not form any condensate even at higher protein concentrations but formed liquid-like droplets in the presence of PEG 8000, which became bigger at higher concentrations of PEG 8000 (Fig. 1.3D).

In conclusion, MBD2/3 and MeCP2 alone could undergo LLPS in a salt dependent manner. Increasing the viscosity (+ PEG 8000) and disrupting the hydrophobic interactions (+ 1,6-hexanediol) could induce a transition of MBD2/3 from spherical droplets to irregular aggregates. MeCP2 droplets showed a resistance to 1,6-hexanediol and a coexistence of spherical droplets and irregular aggregates at high 1,6-hexanediol concentration. Increasing the viscosity (+ PEG 8000) could promote LLPS of MeCP2. MBD1ΔCxxC3 and MBD4 formed irregular aggregates in a protein, salt, and crowding agent dependent manner, which are driven by both electrostatic interactions and hydrophobic interactions.

5.1.3 The influence of DNA on the phase separation properties of the MBD family proteins *in vitro*

Nucleic acids (DNA or RNA) were reported to promote LLPS of multiple disordered proteins (Shakya and King 2018, Zhou, Song et al. 2019, Chen, Cui et al. 2020). All MBD proteins except for the MBD3 were reported to be enriched at the highly methylated pericentromeric heterochromatin, which forms heterochromatin compartments in mouse cells. Heterochromatin compartments are mainly composed of AT-rich (64% AT) major satellite repeats. The basic unit of major satellite repeats (MajSat) is 234 bp with 8 CpG dinucleotides (Fig. 1.4 A).

Here, we checked whether and how DNA influences the phase separation properties of MBDs using the 234 bp-long MajSat sequence. MBD1ΔCxxC3 formed bigger but still irregular aggregates with increased protein and MajSat DNA concentrations (Fig. 1.4 B). MBD2 and MBD3 still formed liquid-like spherical droplets but the size was unchanged with increased DNA concentrations (Fig. 1.4 C). Besides, DNA was weakly enriched into the MBD3 droplets, which could only be observed at low DNA concentration (Fig. 1.4 D). MBD4, which could form only irregular aggregates without DNA, formed both irregular aggregates and liquid-like spherical droplets at certain protein/DNA concentrations (Fig. 1.4 E). Upon

adding MajSat DNA, MeCP2 formed liquid-like droplets at the physiological salt concentration with 5% dextran T150, and the droplet size increased in a protein and DNA concentration-dependent manner (Fig. 1.4 F). Conclusively, MajSat DNA could promote the LLPS of both MBD4 and MeCP2 and irregular aggregate formation of MBD1ΔCxxC3, but has no significant influence on the phase separation properties of MBD2 and MBD3.

Taken together, by *in vitro* phase separation assay, we found that MBDs adopt distinct phase separation properties with distinct biophysical properties under various conditions. The MeCP2 showed the strongest LLPS properties, followed by MBD2, MBD3, MBD4 and finally MBD1ΔCxxC3 (MeCP2>>MBD2>=MBD3>MBD4>>MBD1ΔCxxC3). The MBD1ΔCxxC3 did not form liquid-like spherical droplets in all conditions tested. Of note, MBD3 showed similar phase properties to MBD2 but required higher protein concentration. This is probably due to their high sequence similarity (76% and 75.2% identity in human and mouse, respectively). Considering the similarities of MBDs in genome-wide distribution, MBDs might probably play differential but somehow redundant roles in heterochromatin organization except for MBD3.

5.1.4 Functions of the MBD family proteins in heterochromatin compartment size

In our previous work, we showed that MeCP2 promotes heterochromatin compartment fusion in mouse cells in a dose-dependent manner (Brero, Easwaran et al. 2005, Agarwal, Becker et al. 2011), but it is not clear if the other MBDs could also modulate heterochromatin compartment dynamics. We firstly checked the influence of MBDs in heterochromatin organization (Fig. 1.5) and morphology (Fig. 1.6) in mouse myoblast C2C12 cells with transient overexpression of GFP tagged MBDs. 36 h after transfection, the C2C12 cells were fixed with 3.7% formaldehyde, stained with DAPI, and analyzed by confocal microscopy. The nuclei and heterochromatin compartments were segmented using the DAPI channel (Fig. 1.5A), and the nuclei and heterochromatin compartment parameters (intensity, size, etc.) were measured. The heterogeneous C2C12 populations after transfection were further subdivided into three groups with low (mean intensity [a.u.]: 0.01-0.5), middle (mean intensity [a.u.]: 0.5-5) and high (mean intensity [a.u.]: ≥5) MBD level based on the average GFP intensities in the nuclei.

Consistent with previous work (Agarwal, Becker et al. 2011), MeCP2 was enriched into the heterochromatin compartments and its overexpression led to increased mean heterochromatin compartment area in a dose-dependent manner (Fig. 1.5D). MBD1, MBD2, and MBD4 were also enriched into heterochromatin compartments and were shown to promote the formation of bigger heterochromatin compartments in a concentration-dependent manner although to different extent (Fig. 1.5A-D). MBD1ΔCxxC3 was localized at and enriched into heterochromatin compartments (Fig. 1.5A-B), but the levels of enrichment

did not change with increased expression levels (Fig. 1.5C). Besides, MBD1ΔCxxC3 overexpression did not affect the heterochromatin compartment size regardless of expression levels (Fig. 1.5D). Thus, the function of MBD1 in heterochromatin compartment fusion must rely on the CxxC3 domain. CxxC3 binds to unmethylated CpG dinucleotides, and the CpG dinucleotides in the major satellites were reported to be ~70% methylated (Casas-Delucchi, Becker et al. 2012, Wen, Li et al. 2014). Thus, the specific CxxC3-CpG interaction largely determines the function of MBD1 in heterochromatin compartment fusion. MBD3, the only member of MBDs with no methyl-CpG binding properties, was evenly distributed in the nucleoplasm outside the nucleolus (DAPI depleted region) (Fig. 1.5A-C) and did not influence the heterochromatin compartment size (Fig. 1.5D).
In summary, MBD1, MBD2, MBD4 and MeCP2 could promote heterochromatin compartment fusion in a dose-dependent manner.

5.1.5 Functions of the MBD family proteins in heterochromatin compartment morphology

As LLPS produces spherical condensates *in vivo* and *in vitro*, we then evaluated the heterochromatin compartment morphology in the presence and absence of MBDs by measuring the aspect ratio (AR) (Fig. 1.6A, right). The aspect ratio of a circle is exactly 1.
We first checked the aspect ratio (AR) of heterochromatin compartments in C2C12 cells expressing GFP (0.703 ± 0.002), which was applied as the control (Fig. 1.6B, green line). We found that all MBDs regulated the heterochromatin compartment morphology to distinct extents. MeCP2 promoted the formation of more spherical heterochromatin compartments (0.727 ± 0.005, $P < 0.0001$) due to the strong LLPS. Surprisingly, the heterochromatin compartments became more spherical upon MBD3 expression (0.735 ± 0.002, $P < 0.0001$). As MBD3 does not accumulate at heterochromatin compartments, the side effects of MBD3 overexpression should be considered such as the crowding effects of MBD3 molecules. Overexpression of MBD2 did not influence the heterochromatin compartment morphology (0.700 ± 0.001, $P = 0.281$), which is consistent with the *in vitro* studies that major satellite DNA did not influence the phase separation properties of MBD2 (Fig. 1.4C). However, heterochromatin compartments showed significantly decreased aspect ratio in cells expressing MBD1 (0.676 ± 0.002), MBD1ΔCxxC3 (0.672 ± 0.002), and MBD4 (0.684 ± 0.003) compared to those expressing GFP ($P < 0.0001$ for all three). This is consistent with the finding that purified MBD1ΔCxxC3 and MBD4 preferably formed irregular aggregates *in vitro*. Thus, there is a strong relationship between droplet morphology *in vitro* and heterochromatin compartment morphology *in vivo*.

5.1.6 Molecular dynamics of the MBD family proteins in heterochromatin compartments

Beside morphology, liquid-like compartments are also characterized by the dynamic exchange of proteins between the droplet and the surrounding soluble proteins. Thus, the fluorescence recovery after photobleaching (FRAP) assay was performed in living C2C12 cells expressing GFP tagged MBDs 20-30 h after transfection. To exclude the possible influence of protein levels, cells with similar protein levels were selected and photobleached (Fig. 1.7). And the intensities were double normalized (Fig. 1.7, right) (Phair, Gorski et al. 2003). For MBD3, which did not form any droplets or aggregate in C2C12 cells, the fluorescence at the photobleached region recovered immediately and was not further analyzed.

To keep all data consistent (compared with fixed cells) and reproducible, we firstly analyzed the fold enrichment of MBDs in the photobleached region of interest (ROI) prebleach relative to the mean fluorescence intensities in the nucleoplasm. MBD2a showed the highest enrichment at ROI, followed by MeCP2, MBD1, MBD1ΔCxxC3 and finally MBD4, which is consistent with the protein fold enrichment in fixed C2C12 populations with high MBDs levels (Fig. 1.7B, D).

Then the molecular diffusion rates and the extent of recovery were evaluated by recovery halftime and immobile fractions separately (Fig. 1.7E-F). All MBDs recovered more than 60% (< 40% immobile fractions) within one minute after bleaching (Fig. 1.7C, E), indicating liquid-like properties. MBD1ΔCxxC3 and MBD4, which caused the formation of less spherical heterochromatin compartments, recovered faster than the others together with lower immobile fractions. While MBD1, which also promoted irregular heterochromatin compartments, showed the slowest recovery and the largest immobile fraction. MBD2 and MeCP2, which did not affect heterochromatin compartment roundness showed intermediate levels of an immobile fraction and recovery rate. Besides, comparison of recovery half time and immobile fractions between MBD1 (39.7 ± 2.9 s and 32.7 ± 2.2%, respectively) and MBD1ΔCxxC3 (18.0 ± 1.3 s and 13.0 ± 1.5%, respectively) showed > 2 fold change, further confirming the importance of CxxC3-CpG interaction for MBD1 function in heterochromatin. Thus, we could summarize that all MBDs exhibit dynamic exchange between heterochromatin compartments and the surrounding environment, but there is no consistent relationship between LLPS properties (or condensates morphology) and mobility.

MBD1 function is highly dependent on the CxxC3, as the MBD1ΔCxxC3 showed significantly decreased heterochromatin compartment size and roundness and is more mobile. This indicates that the protein dynamics is highly influenced by the protein-DNA binding. The binding affinities measured by EMSA assay were reported as: MBD3(MBD) < MBD1(MBD) <

MBD4(MBD) < MBD2(MBD) < MeCP2(MBD) (Buchmuller, Kosel et al. 2020), and are, thus, negatively correlated with the protein dynamics: MBD4 > MBD1ΔCxxC3 > MBD2 > MeCP2. These results indicate direct relationships between methyl-CpG binding and protein dynamics. Yet, other MBDs-protein and MBDs-DNA interaction might also contribute to molecular mobility if strong enough, such as CxxC3-CpG interaction.

5.1.7 Conclusion

In this chapter, we analyzed the phase separation properties of the core MBD members (MBD1-4, MeCP2) *in vivo* and *in vitro* and confirmed that MBDs drive the heterochromatin organization via (liquid-liquid) phase separation with distinct properties.

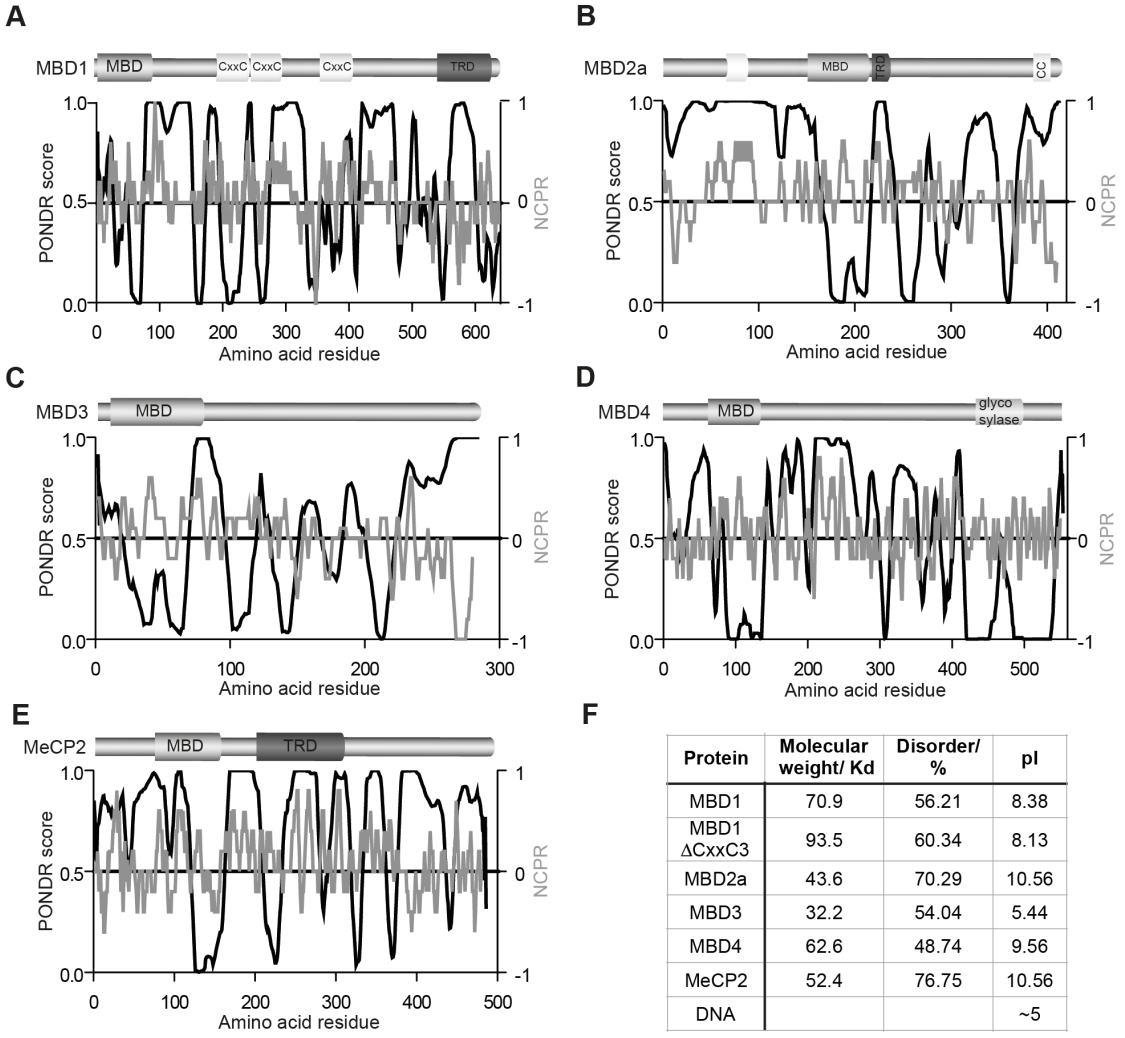
In vitro, all MBDs tested except for MBD1ΔCxxC3 could form liquid-like droplets. The overall LLPS properties are: MeCP2 > MBD2 > MBD3 > MBD4 > MBD1ΔCxxC3 (Fig. 1.8D). MBD3 shows a similar phase diagram with MBD2 likely due to a high sequence similarity of the MBD and C-terminus but requires higher protein concentration. Besides, MBD3 lacks the N-terminus of MBD2, one can imagine that the N-terminus of MBD2 is not essential for MBD2 LLPS but enhances the MBD2 LLPS by decreasing the concentration threshold. Although MBD1ΔCxxC3 formed only irregular aggregates in all conditions tested, it still keeps the possibility of LLPS *in vivo*, as the *in vitro* conditions could not 100% mimic the *in vivo* environments.

In cells, we evaluated the liquid-like behaviors of MBDs in heterochromatin compartments by the fusion properties, heterochromatin compartment morphology, and dynamics exchange between heterochromatin compartments and surroundings, which are displayed by the *in vitro* droplets (McSwiggen, Mir et al. 2019) (Fig. 1.8A-B).

MeCP2 promotes the formation of bigger and more spherical heterochromatin compartments. MBD2 promotes heterochromatin compartment fusion but does not influence the heterochromatin compartment morphology. Both MeCP2 and MBD2 show slower dynamic exchange between heterochromatin compartments and surroundings. MBD3 is not enriched into heterochromatin compartments, and does not influence the heterochromatin compartment size, but makes more spherical heterochromatin compartments probably due to side effects, such as crowding. MBD1 and MBD4 promote heterochromatin compartment fusion, generating bigger but less spherical heterochromatin compartments with different dynamics. The most irregular heterochromatin compartments were observed in cells expressing MBD1ΔCxxC3. Besides, MBD1ΔCxxC3 showed > 2 fold decreased heterochromatin compartment size, roundness, t1/2, and immobile fraction compared to the full-length MBD1. CxxC is an unmethylated CpG dinucleotides binding domain and collaborates with MBD to bind unmethylated and methylated CpG separately. MBD1 has three CxxC motifs but only the third CxxC (CxxC3) shows strong interactions with

unmethylated CpG dinucleotides. CpG dinucleotides in the heterochromatin compartments are around 70% methylated (Casas-Delucchi, Becker et al. 2012, Wen, Li et al. 2014), resulting in a molar ratio of 7:3 (mCpG:CpG) and highlighting the importance of unmethylated CpG binding in heterochromatin compartments. Thus, the MBD1 dynamics is highly dependent on both MBD-mCpG and CxxC3-CpG interactions.

Yet, not all MBDs show strong interaction with unmethylated DNA and the MBDs-DNA interactions are mainly provided by the specific interactions between the MBD and the methylated cytosine (mC). The protein dynamics *in vivo* were: MeCP2 < MBD2 < MBD1ΔCxxC3 < MBD4 (Fig. 1.8B), which is negatively correlated with the MBD-mC binding affinities: MeCP2(MBD) > MBD2(MBD) > MBD4(MBD) > MBD1(MBD)(ΔCxxC3) (> MBD3(MBD)) (Buchmuller, Kosel et al. 2020) (Fig. 1.8C), and LLPS properties: MeCP2 > MBD2 > (MBD3 >) MBD4 ≥ MBD1ΔCxxC3 (Fig. 1.8D). These results indicate a possible relationship among methyl-CpG binding, dynamics and LLPS.



F

Protein	Molecular weight/ Kd	Disorder/ %	pI
MBD1	70.9	56.21	8.38
MBD1 ΔCxxC3	93.5	60.34	8.13
MBD2a	43.6	70.29	10.56
MBD3	32.2	54.04	5.44
MBD4	62.6	48.74	9.56
MeCP2	52.4	76.75	10.56
DNA			~5

Figure 1.1 Analysis of MBD protein sequences.

(A-E) Protein structure (top) and amino acid sequence analysis (bottom) of MBD1 (A), MBD2a (B), MBD3 (C), MBD4 (D), and MeCP2 (e2) (E). NTD: N-terminal domain; MBD: methyl binding domain; ID: intervening domain; NID: N-CoR interacting domain; CTD: C-terminal domain; TRD: transcriptional repression domain. Blackline: PONDR prediction (http://www.pondr.com/) of MeCP2 ordered/disordered regions, >0.5 is considered disordered. Grayline: Protein charge, >0 means positively charged (https://www.bioinformatics.nl/cgi-bin/emboss/charge).
(F) Summary table of the molecular weight, overall disorder, and iso-electric point (pI) prediction. The Isoelectric points (PI) of the MBD protein family are predicted using INNOVAGEN (https://pepcalc.com/).

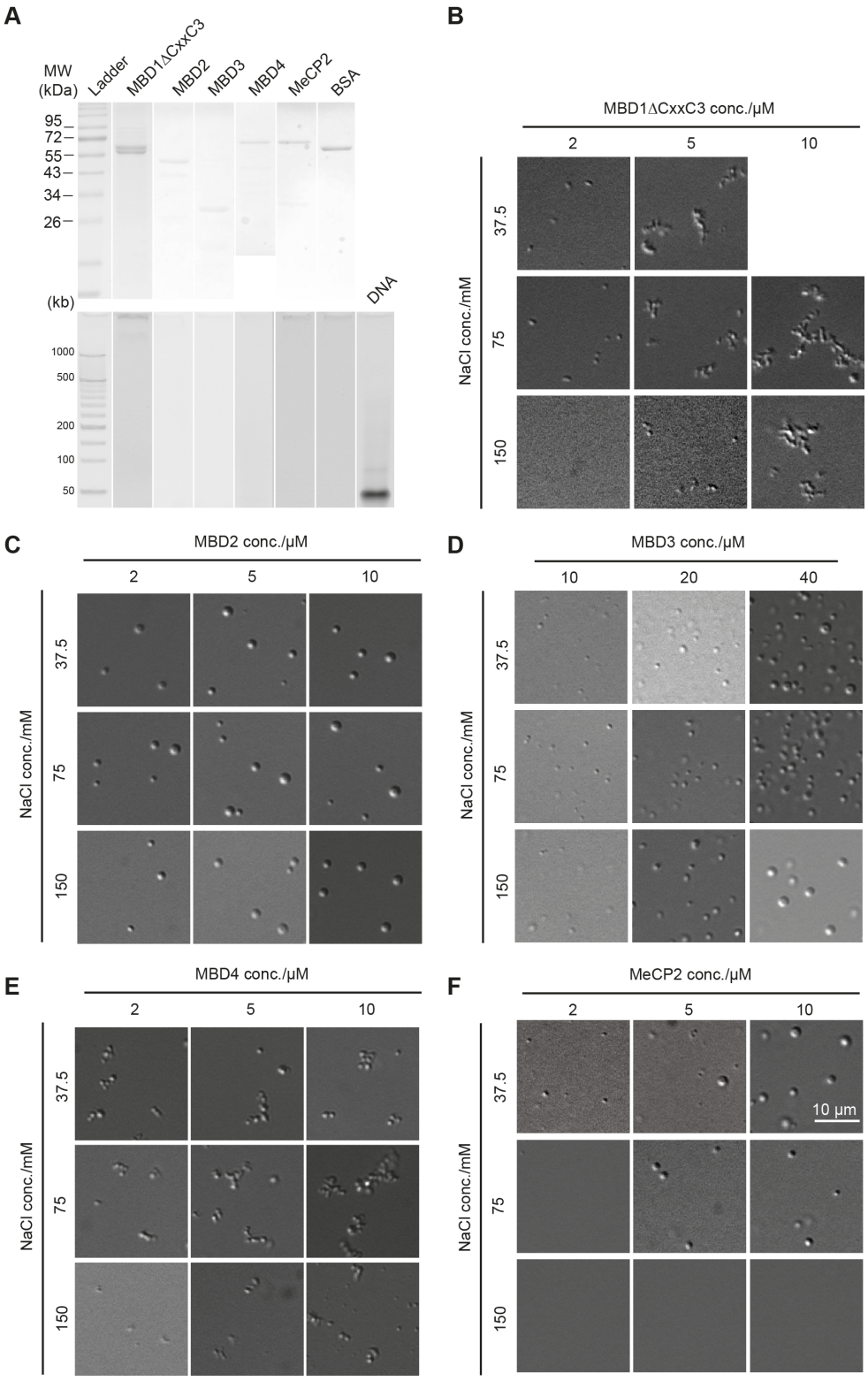


Figure 1.2 Phase separation properties of the MBD family proteins.

(A) Detection of MBDs purity. The MBDs were expressed in bacteria by IPTG induction, purified using chitin beads and eluted by DTT induced self-cleavage. The final protein concentrations were measured by Pierce™ 660nm Protein Assay Reagent. Top: SDS polyacrylamide gel electrophoresis of purified MBDs followed by Coomassie staining. 2 μg each lane. Bottom: TBE polyacrylamide gel electrophoresis of purified MBDs followed by ethidium bromide staining. 10 μg each lane. Negative control: BSA with the same amount. Positive control: synthesized 42 bp DNA.
(B-F) Phase separation of MBDs droplet formation at the indicated salt and protein concentrations. Images were taken under the Nikon Eclipse TiE2 microscope equipped with differential interference contrast (DIC). Scale bar = 10 μm.

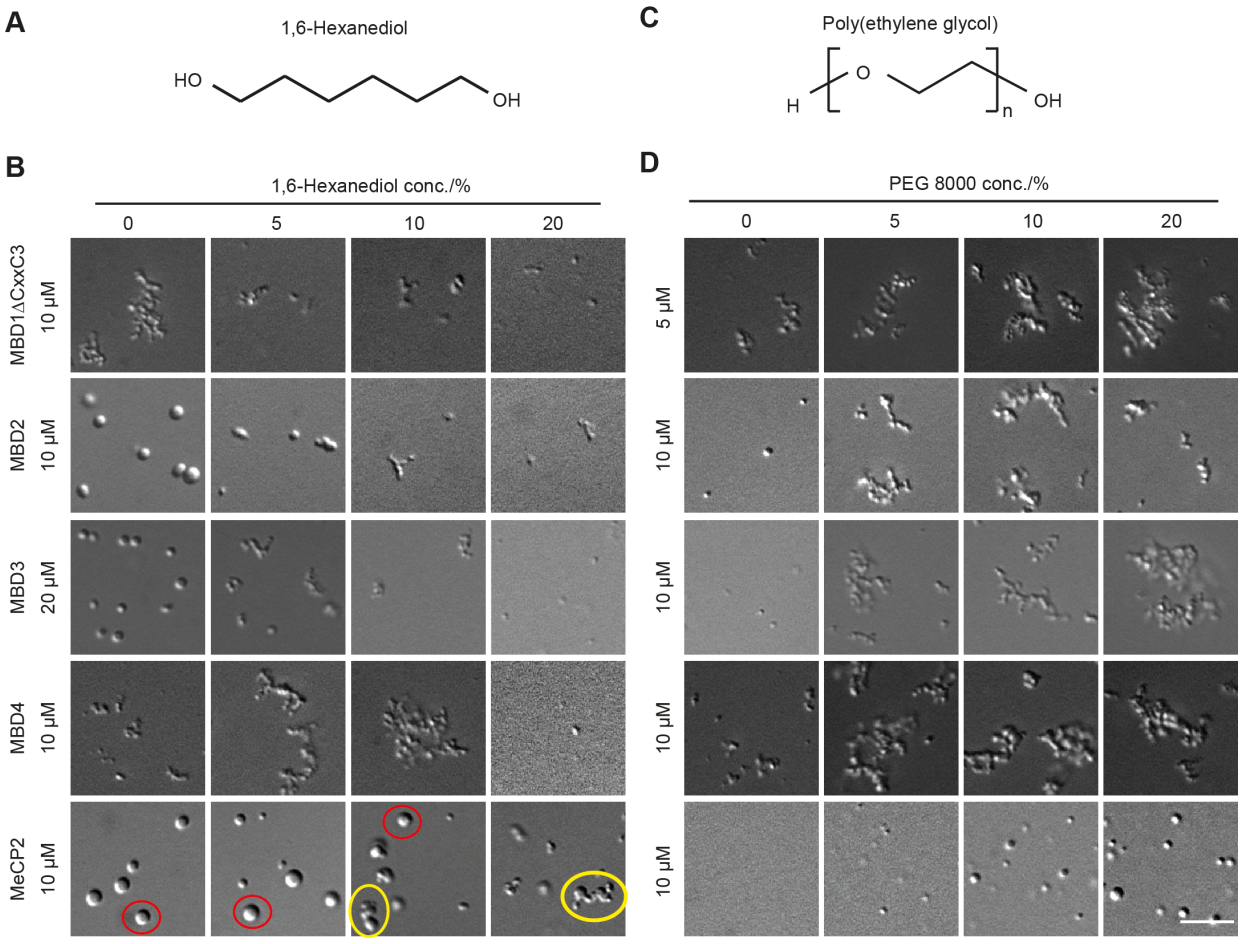


Figure 1.3 The phase separation of the MBD family proteins in the presence of molecular crowder or 1,6-hexanediol.

The *in vitro* phase separation assay in the presence of PEG 8000 or 1,6-hexanediol was done by incubating at room temperature for 45 min. Images were taken under the Nikon Eclipse TiE2 microscope equipped with differential interference contrast (DIC).
(A) Chemical structure of 1,6-hexanediol.
(B) DIC images of MBDs phase separation at different concentrations of 1,6-hexanediol.
(C) Chemical structure of polyethylene glycol (PEG).
(D) DIC images of MBDs phase separation at different concentrations of PEG 8000. Scale bar = 10 μm.

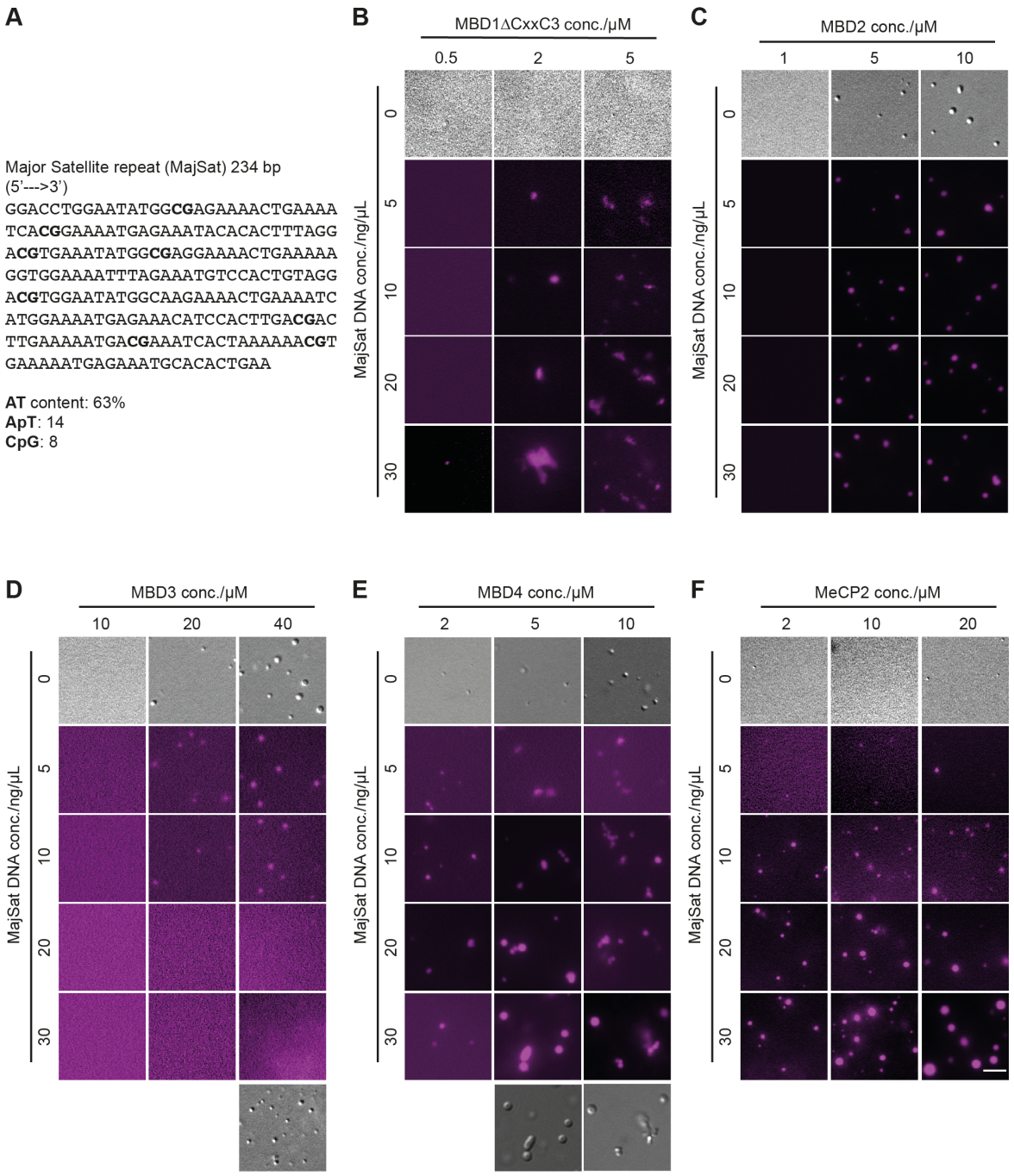


Figure 1.4 Phase separation of the MBD family proteins in the presence of DNA.

(A) Nucleotide sequence of a single major satellite unit (MajSat) from the mouse genome. MajSat is highly AT riched (63%) with 8 CpG dinucleotides.
(B-F) DIC images of MBD1ΔCxxC3 (B), MBD2 (C), MBD3 (D), MBD4 (E), and MeCP2 (F) phase separation with MajSat. The MajSat were synthesized by PCR, purified from agarose gel, and labeled by DRAQ5 for DNA visualization. The DRAQ5 labeled DNA was mixed with purified MBD family proteins at different conditions for *in vitro* droplet formation. Both fluorescent and DIC images were taken under the Nikon Eclipse TiE2 microscope equipped with a 40x objective. Buffer conditions for MBD1ΔCxxC3 and MBD3: 150 mM NaCl; for MBD2, MBD4, and MeCP2: 150 mM NaCl and 5% Dextran T150. Scale bar = 10 μm.

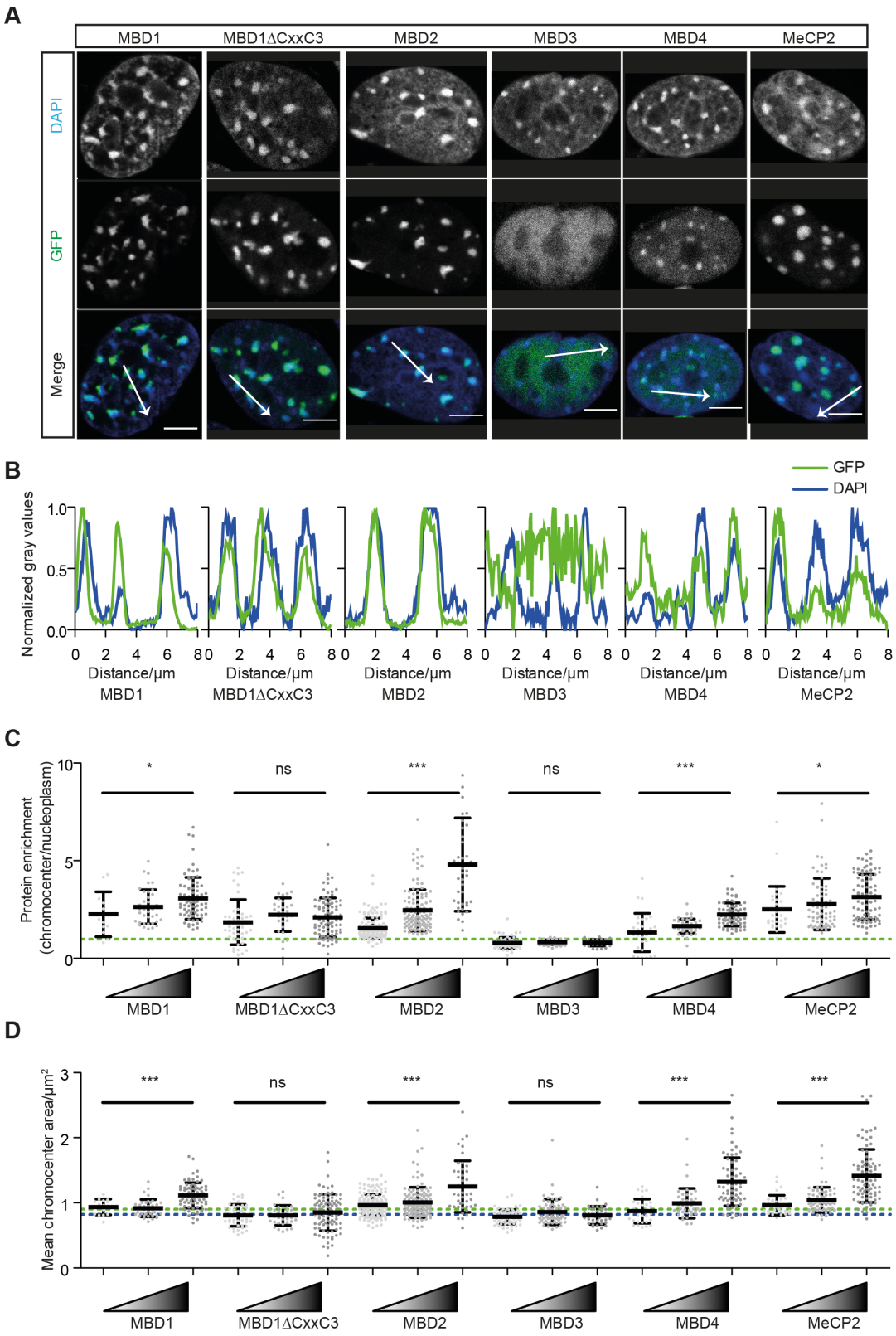


Figure 1.5 Influence of the MBD family proteins on heterochromatin compartment dynamics.

C2C12 cells were transfected with plasmids expressing GFP tagged MBDs, fixed with 3.7% formaldehyde, stained with DAPI (4′,6-diamidino-2-phenylindole) and analyzed using a confocal microscope Leica TCS SP5 II equipped with 100 X objective.
(A) Representative images showing the localization of MBDs *in vivo*. The heterochromatin compartments were marked by dense DAPI signals. Scale bar = 5 µm.

(B) Plot profiles showing the normalized distribution of GFP and DAPI fluorescence intensities measured in transiently transfected C2C12 cells and indicated by lines. Arrows indicate the direction of measurement.
(C, D) Quantification of protein enrichment in heterochromatin compartments (C) and heterochromatin compartment size (D). The cells are analyzed using imageJ. The cell nuclei and heterochromatin compartments were identified by Auto Threshold based on the DAPI intensities. Cells were subgrouped by GFP intensities. The scatter plot represents the distribution of protein fold enrichment (C) and mean heterochromatin compartment area (D) per nucleus. Green dashed lines represent the fold enrichment of GFP fluorescence signals (C) and the mean heterochromatin compartment area (D) in C2C12 cells 36 h after transfection with the plasmid coding for GFP only. The blue dashed line represents the mean heterochromatin compartment area (D) in untransfected C2C12 cells after fixation. Error bars on the histogram correspond to standard deviation (SD). Significances were calculated by unpaired t-test comparing the high expressing samples with low expressing samples as indicated by lines. ns P > 0.05; *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001.

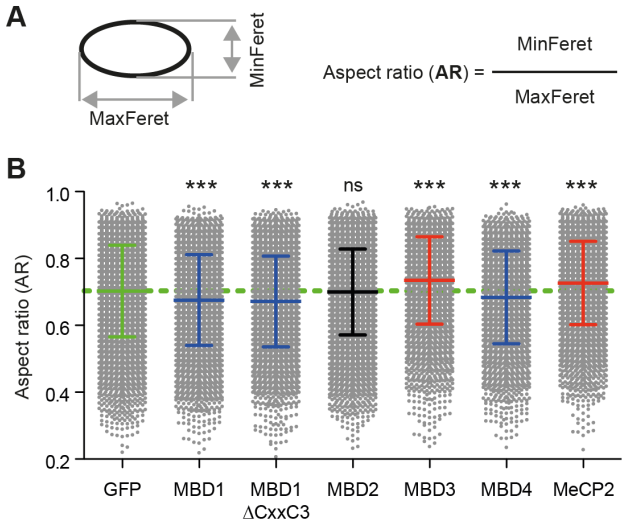


Figure 1.6 Influence of the MBD family proteins on heterochromatin compartment morphology.

(A, left) Schematic graph showing how to measure the minimal and maximal Feret diameters (MinFeret and MaxFeret separately). The Feret diameter (or caliper diameter) is defined as the distance between the two parallel planes restricting the object perpendicular to that direction. Thus, the MinFeret and MaxFeret of heterochromatin compartments are defined as the minimal and maximal distance between the two parallel planes restricting the heterochromatin compartment perpendicular to that direction separately. (A, right) The equation used to calculate the aspect ratio (AR) to evaluate the heterochromatin compartment roundness. Higher AR indicates a more spherical shape.
(B) The scatter plot with mean ± SD represents the distribution of aspect ratio (AR) of heterochromatin compartments in fixed C2C12 cells after transfection. Heterochromatin compartments with mean GFP fluorescence intensity > 0.01 pixels were chosen. Green lines represent the mean AR in C2C12 cells transiently expressing GFP. Red, black, and blue lines represent the increased, unchanged, and increased AR values in C2C12 cells after transfection compared to samples with GFP only, respectively. Significances were calculated by unpaired t-test comparing the MBD expressing samples with GFP expressing samples.ns P = 0.2812; *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001.

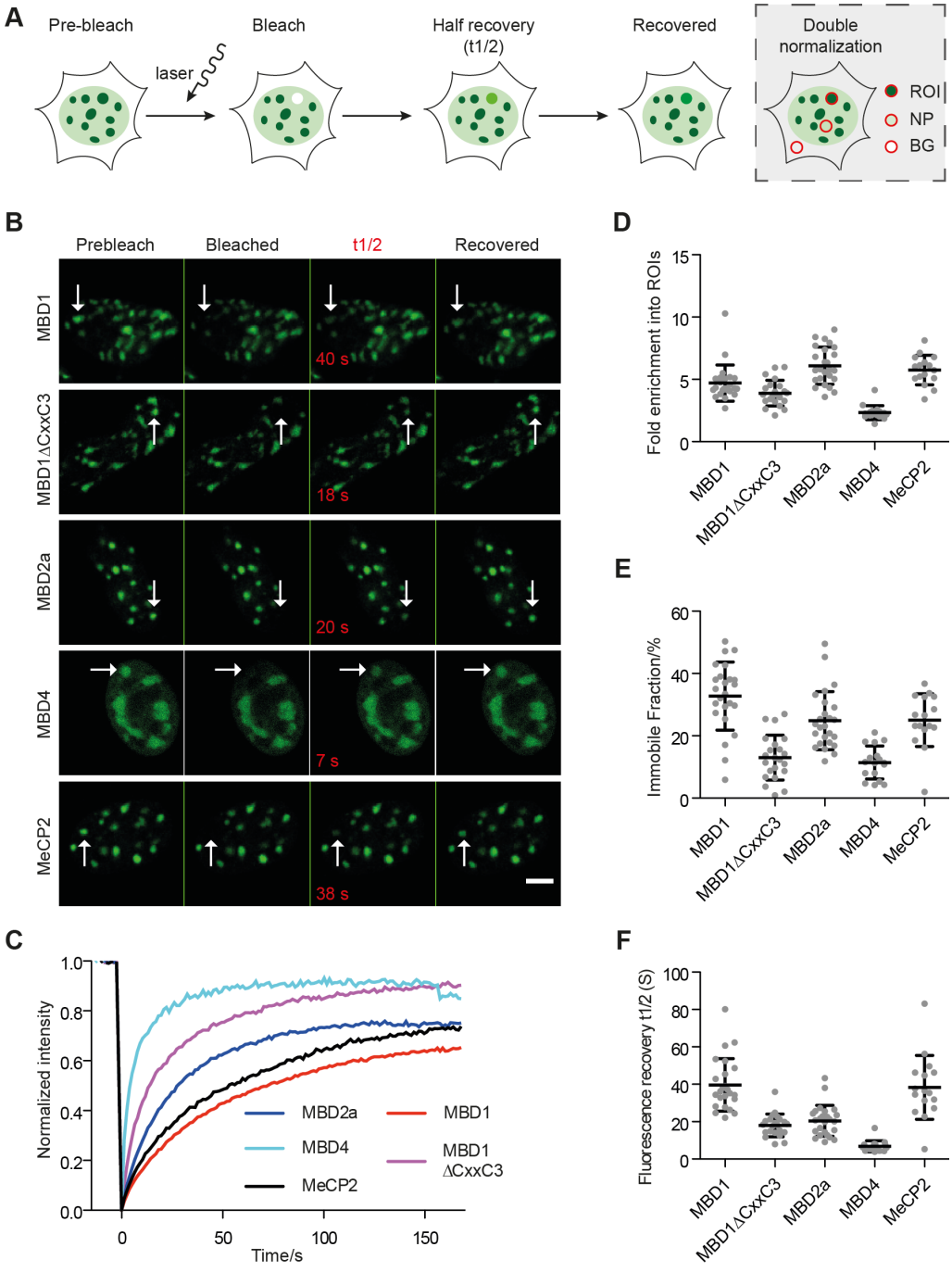


Figure 1.7 Dynamics of the MBD family proteins at heterochromatin compartments measured by fluorescence recovery after photobleaching (FRAP).

(A) Schematic graph showing the experimental procedure. C2C12 cells were transfected with plasmids expressing GFP tagged MBDs by electroporation and incubated for 24 h before FRAP. In the FRAP experiment, a circular heterochromatin compartment area in a C2C12 cell expressing GFP tagged MBDs was rapidly and irreversibly bleached using a 488 laser pulse with 100% laser power, leaving a circular region devoid of GFP signal. Due to molecular exchange between bleached and unbleached regions, the GFP signal increased with time till plateau in the bleached region. The fluorescence changes were imaged. The extent of final recovery depends on the fraction of mobile molecules. The recovery halftime (t1/2) is the time from the bleach point to the timepoint where the fluorescence intensity reaches half of the final recovered intensity. The GFP signal intensity in the nucleoplasm and outside nuclei were applied for double normalization (box with dashed line) (Phair, Gorski et al. 2003). All FRAPs were done at 37 °C, 5% CO2, 90% humidity in an incubation chamber.
(B) Time-lapse confocal images of C2C12 cells expressing GFP-MBDs before, immediately after, half recovered (t1/2) and 150 s after photobleaching. Arrows marked the bleached regions (ROIs). t1/2 on representative series was indicated (red). The scale bar is 5 µm.

(C) Mean FRAP curves in C2C12 cells expressing GFP-MBD1 (n = 24), GFP-MBD1ΔCxxC3 (n = 22), GFP-MBD2a (n= 25), GFP-MBD4 (n = 17) and GFP-MeP2 (n = 16).
(D) Fold enrichment of MBDs in bleached regions relative to the nucleoplasm. The fold enrichment is calculated as (IROI - IBG)/(INP - IBG). ROI: bleached region; BG: background; NP: nucleoplasm. Data were given as mean ± SD.
(E-F) Immobile fraction and fluorescence recovery t1/2 of FRAP curve in (C). Data are plotted as mean ± SD.

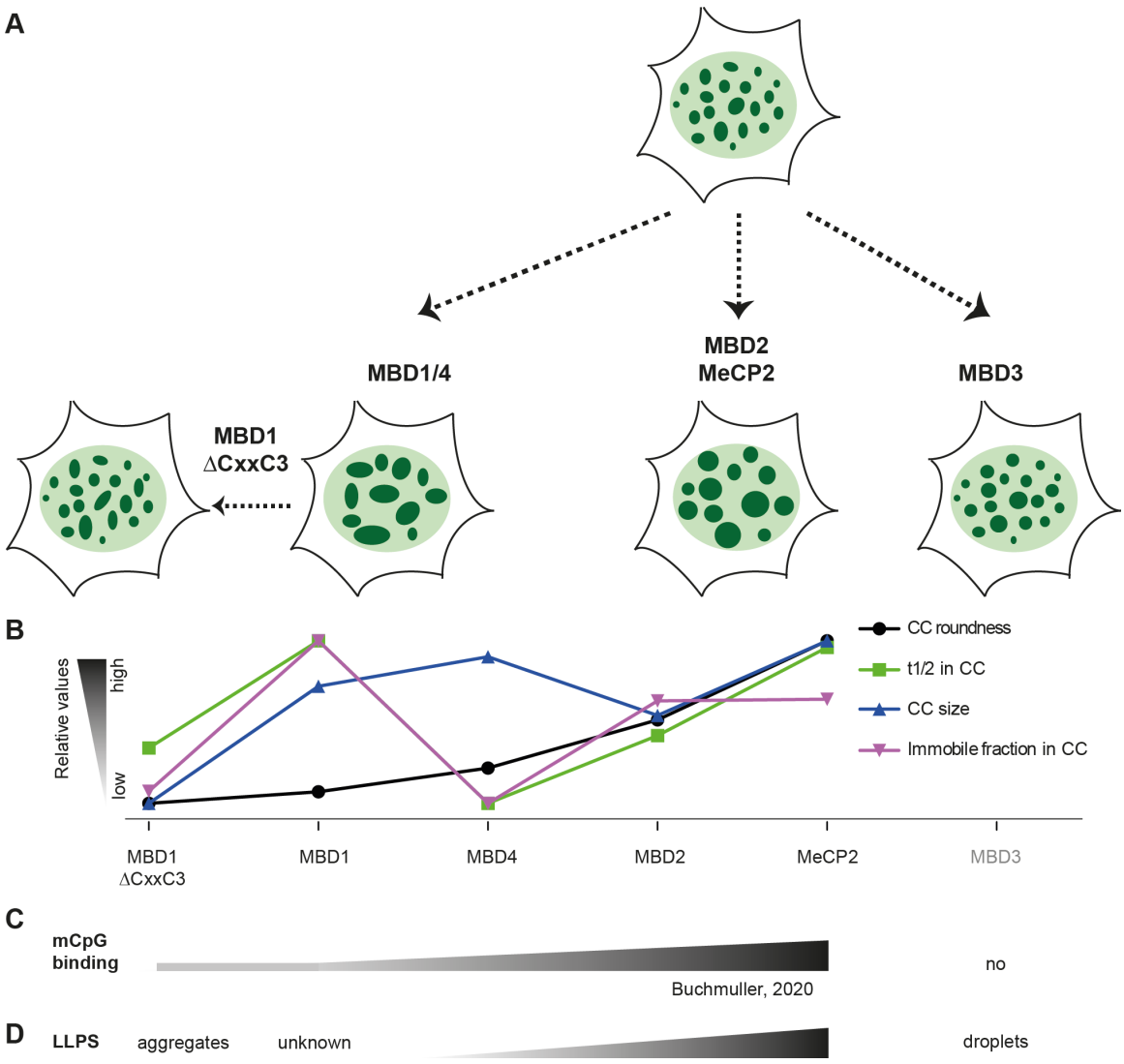


Figure 1.8 Summary of the functions of the MBD family proteins *in vivo* and *in vitro*.

(A) Schematic graph showing the structure of heterochromatin compartments upon MBDs expression.
(B) Comparison of MBDs influence on heterochromatin compartment structure and boundary effects.
(C) Binding affinities of the MBD domain from different MBDs with methylated DNA (Buchmuller, Kosel et al. 2020).
(D) Phase separation properties of MBDs.

5.2 Isoform-specific influence of MBD2 in heterochromatin organization via phase separation

5.2.1 Truncated isoforms of MBD2 have different impacts on heterochromatin organization

MBD2 (methyl-CpG binding domain protein 2) has three isoforms due to alternative splicing and translation start sites (Fig. 2.1A). All isoforms contain the highly conserved MBD-TRD, but retain either the N-terminus (MBD2c) or the C-terminus (MBD2b) or both (MBD2a). These distinctions of the three isoforms correspond to different binding partners and functions (Lu, Loh et al. 2014, Zhang, Hastert et al. 2017). Thus, it becomes relevant to understand the functional variants of MBD2 isoforms.

MBD2a was shown to promote the differentiation of mouse myoblast C2C12 cells to myotubes via interaction with focal adhesion kinase (FAK) accompanied by heterochromatin reorganization (Luo, Zhang et al. 2009). Thus, the C2C12 cells were chosen as the *in vivo* cellular model to check whether/how MBD2 isoforms contribute to the heterochromatin organization.

To exclude the possible influence of GFP, cells with only GFP expression were used as a negative control. All images were taken at the confocal microscopy Leica TCS SP5 II and further processed with the imageJ. The localization of MBD2 isoforms was firstly analyzed. Compared to the GFP only, which showed homogeneous distribution across the cells, all three isoforms were localized in the nuclei and highly enriched locally (Fig. 2.1B). Further plot profiles of both GFP and DAPI fluorescence intensity confirmed the colocalization of GFP dense foci and DAPI dense foci (Fig. 2.1C).

Further, we checked if MBD2 isoforms could influence the heterochromatin compartment fusion, which was found for MBD2a and MeCP2 (Brero, Easwaran et al. 2005, Karaca, Brito et al. 2018, Mok, Zhang et al. 2020). The nuclei and the heterochromatin compartments were segmented based on the DAPI intensities with the same parameters for reproducible cross-comparisons. Considering that the expression levels of MBD2a also plays a role in heterochromatin compartment dynamics, the heterogeneous C2C12 populations after transfection were further subdivided into three groups with low (mean intensity [a.u.]: 0.01-0.5), middle (mean intensity [a.u.]: 0.5-5) and high (mean intensity [a.u.]: ≥5) MBD2 expression based on the average GFP intensities in the nuclei.

The three isoforms had distinct influences on heterochromatin compartment size compared to the control GFP-expressing cells (Fig. 2.1D, green line). The full-length MBD2a promoted the formation of bigger heterochromatin compartments in a dose-dependent manner. A similar tendency was also found for MBD2b whereas to a much less extent. MBD2c showed the least influence on heterochromatin compartment size regardless of MBD2c levels.

Further, we measured the fold enrichment of MBD2 isoforms into heterochromatin compartments via the ratio of the mean GFP fluorescence in heterochromatin compartments to the mean GFP fluorescence in the nuclei outside heterochromatin compartments (nucleoplasm plus euchromatin) (Fig. 2.1E). Compared to the fold enrichment of GFP, which showed no enrichment at heterochromatin compartments (ratio = 0.98 ± 0.01), both MBD2a and MBD2c were enriched at the heterochromatin compartments to a similar extent and in a concentration-dependent manner. MBD2b was also enriched in heterochromatin compartments in a concentration-dependent manner, yet to a far less extent. A possible explanation underlying the differential function and localization was that the C-terminus of MBD2a is essential for heterochromatin compartment fusion while the N-terminus of MBD2a promotes the enrichment of MBD2a at heterochromatin compartments. MBD-TRD is responsible for the MBD2 binding to mC in the highly methylated heterochromatin compartments.

5.2.2 The MBD-TRD of MBD2 is essential for heterochromatin localization and clustering

Due to the distinct functions and distributions of the three MBD2 isoforms, we further checked in detail if/how different regions of MBD2a influence the heterochromatin compartment structure using the same strategy as mentioned above. Here, the C2C12 cells were transfected with the constructs expressing GFP tagged N-terminus (1_152) (MBD2a-N), MBD-TRD (153_235) (MBD2a-MBD-TRD), C-terminus (236-414) (MBD2a-C) and the truncation lacking C-terminus from MBD2a, MBD2a (1_235) (MBD2a-ΔC) (Fig. 2.2A). The samples were analyzed by confocal microscopy after fixation.

MBD2a-N was localized in the nuclei and formed condensates locally, but did not colocalize with heterochromatin compartments (Fig. 2.2B-D). Indeed, the MBD2 nuclear localization signal (NLS) is predicted within the N-terminus and includes the glycine/arginine repeats (G/R). G/R is positively charged and thus probably interacts with negatively charged DNA via weak electrostatic interactions. As arginine-rich motifs are responsible for RNA binding, the MBD2a-N is locally enriched in the RNA-rich nucleolus (Fig. 2.2B) (Bayer, Booth et al. 2005). Both MBD2a-MBD-TRD and MBD2a-ΔC were localized in the nuclei and enriched locally at the heterochromatin compartment regions (Fig. 2.2B-D) due to the specific methyl-CpG binding property of the MBD domain. While the enrichment of MBD2a-MBD-TRD was not concentration-dependent, the enrichment of MBD2a-ΔC was concentration-dependent. Thus, the MBD is responsible for the heterochromatin compartment localization while the N-terminus promotes the enrichment of MBD2 into heterochromatin compartments (Fig. 2.2D). The MBD2a-C was evenly distributed throughout the cells due to a lack of nuclear localization signal (Fig. 2.2B-D).

Further, we analyzed the influence of MBD2a constructs in heterochromatin dynamics (Fig. 2.2E). We found that both MBD2a-N and MBD2a-C alone did not influence the heterochromatin compartment size. The MBD-TRD promoted the heterochromatin compartment growth in a concentration-dependent manner, which was surprisingly suppressed by the N-terminus. The comparison of heterochromatin compartment size in cells expressing high levels of MBD2a-MBD-TRD (1.25 ± 0.04 μm^2) and MBD2a (1.25 ± 0.06 μm^2) (Fig. 2.1D) indicated that the MBD-TRD is essential to promote the heterochromatin compartment fusion (P value = 0.935). As MBD2a-ΔC lacks the C-terminus from MBD2c which was composed of 14 amino acids, we compared the influence of the short C-terminus from MBD2c on heterochromatin compartment fusion but found no significant difference at both low and high expression levels (low levels: 0.91 ± 0.01 μm^2 for MBD2a-ΔC and 0.90 ± 0.02 μm^2 for MBD2c, P value = 0.894; high levels: 1.00 ± 0.02 μm^2 for MBD2a-ΔC and 0.98 ± 0.03 μm^2 for MBD2c, P value = 0.533). Thus, the C-terminus of MBD2c is dispensable for the localization and function of MBD2c in heterochromatin organization.

5.2.3 The N-terminus of MBD2a restricts the MBD2 mobility

Next, we checked the dynamic exchange of MBD2 isoforms and domains between condensates and surrounding environments in live cells by the fluorescence recovery after photobleaching assay (FRAP) (Fig. 2.3A). As MBD2a-N formed aggregates outside the heterochromatin compartments (incorporated into the nucleolus), the signals at the condensates rather than heterochromatin compartments were photobleached. MBD2a-C showed no local enrichment and thus was not included in the FRAP assay.

For all constructs tested, the fluorescence was half recovered within 30 seconds and more than 60% signals were recovered (mobile fractions) after photobleaching (Fig. 2.3B-E), confirming that all constructs have fast molecule exchange between condensates and surrounding environments. In detail, these constructs could be divided into two groups according to the fluorescence recovery rates and immobile fractions. Group A (marked in the yellow box): MBD2a, MBD2c, and MBD-ΔC with slower recovery rates and more immobile fractions; Group B (marked in the red box): MBD2a-N, MBD2a-MBD-TRD, and MBD2b with faster recovery rates and smaller immobile fractions. The constructs within each group showed similar recovery rates and immobile fractions. Structurally, all constructs in group A contain both N-terminus and MBD-TRD while constructs in group B contain either N-terminus or MBD-TRD domain. Thus, the MBD and the N-terminus cooperate to regulate the MBD2 exchange dynamics.

Meanwhile, we measured the fold enrichment of GFP signals in the bleached regions relative to the nucleoplasm. The result showed that the MBD2a, MBD2c, and MBD2a-ΔC (Group A) that contains the N-terminus, had higher fold enrichment at the bleached region

than MBD2b and MBD2a-MBD-TRD (Group B), which lack the N-terminus. These are consistent with MBD2 fold enrichment observed in fixed cells.

To summarize, the MBD-TRD is the core element of MBD2 localization, function, and dynamics *in vivo*. The N-terminus promotes the enrichment and stability of MBD-TRD at heterochromatin compartments, while the C-terminus regulates the MBD-TRD mediated heterochromatin compartment clustering.

5.2.4 The C-terminus drives the MBD2 LLPS

The MBD2 isoforms are highly dynamic inside heterochromatin compartments with distinct mobility and functions, indicating that MBD2 might probably mediate the heterochromatin compartment organization via phase separation. To further dissect the functional difference among MBD2 isoforms *in vitro*, we explored the phase separation properties of MBD2 isoforms using purified proteins from bacteria (Fig. 2.4A). As MBD2a-ΔC and MBD2c showed a similar influence on heterochromatin compartment dynamics, we assumed that MBD2a-ΔC and MBD2c are functionally consistent. The MBD2-ΔC protein was purified and applied to the phase separation assay.

First, we checked the phase separation abilities of the MBD2 isoforms using the *in vitro* phase separation assay at different conditions. Both MBD2a (Fig. 2.4B and E) and MBD2b (Fig. 2.4C and E) formed liquid-like spherical droplets, which decreased at higher salt concentration and lower protein concentration. Moreover, at physiological salt condition (150 mM NaCl), MBD2a formed spherical droplets at both low and high protein concentrations, while MBD2b formed less spherical droplets at low protein concentrations (≤10 μM) (Fig. 2.4C). Besides, MBD2b still formed liquid-like droplets at the higher salt concentration (300 mM) although required higher protein concentration (Fig. 2E). MBD2c (MBD2-ΔC) (Fig. 2.4D-E) formed irregular aggregates in all conditions tested. Besides, the aggregates increased with increased protein concentrations.

Next, we checked the influence of hydrophobic interactions on phase separation properties of MBD2 isoforms by adding 1,6-hexanediol. We found that the droplets or aggregates decreased at high 1,6-hexanediol concentrations (Fig. 2.4F). MBD2a droplets became irregular aggregates at low 1,6-hexanediol concentration (5%). The size of the aggregates decreased at higher 1,6-hexanediol concentrations (10-20%). MBD2b and MBD2c aggregates fall apart with increased 1,6-hexanediol concentration. Yet the MBD2a and the MBD2c exhibited higher resistance to 1,6-hexanediol, as the aggregates still existed at high 1,6-hexanediol concentrations (20%) (Fig. 2.4F).

Besides, the phase separation properties were explored in crowding conditions by adding PEG 8000. We found that in the presence of PEG 8000, the MBD2a spherical droplets turned into irregular aggregates, while MBD2b irregular aggregates turned into liquid-like

spherical droplets, which could grow at higher PEG 8000 concentration. MBD2c (MBD2-ΔC) still formed irregular droplets regardless of crowding environment.

Thus, the condensates formed by the three isoforms of MBD2 had distinct physical properties. The LLPS of MBD2a and MBD2b were driven by both electrostatic interactions and hydrophobic interactions. The phase separation of MBD2c is mainly driven by hydrophobic interactions. For MBD2a and MBD2b, the morphology of condensates (spherical droplets or irregular aggregates) could be oppositely modulated by the strength of weak interactions. One possible explanation is that the N-terminus enhances the MBD-TRD-C mediated self-interaction or provides additional self-interaction sites (Becker, Allmann et al. 2013). Higher weak interaction promotes the formation of more spherical MBD2b condensates.

Then, we checked how different regions of MBD2 contributed to the phase separation properties of MBD2 using the MBD2a truncations purified from bacteria (Fig. 2.4A). We found that MBD2a-N (Fig. 2.5A) and MBD2a-C (Fig. 2.5B) proteins formed irregular aggregates in both protein and salt concentration-dependent manner. MBD2a-MBD-TRD was dissolved in solvents, with no detectable condensates formed (Fig. 2.5C). Besides, the aggregates formed by MBD2a-N were shown to be more resistant to 1,6-hexanediol (Fig. 2.5D) and PEG 8000 (Fig. 2.5E). Consistent with MBD2b, the MBD2a-C aggregates disappeared at high 1,6-hexanediol concentrations (Fig. 2.5D) and became liquid-like spherical droplets in the presence of PEG 8000 (Fig. 2.5E). MBD2a-MBD-TRD did not demix in solutions containing PEG 8000.

Taken together, the C-terminus is responsible for the LLPS of MBD2, which could be enhanced by the MBD-TRD and the N-terminus with increased multivalent weak interactions. Weak interactions modulate a phase transition of MBD2a and MBD2b. With increased weak interactions, the MBD2a and MBD2b could probably undergo a phase transition from aggregates (MBD2b at low concentration) to droplets (MBD2a and MBD2b at higher concentration and MBD2b with PEG 8000) and further to aggregates (MBD2a with PEG 8000). At the same condition, MBD2a and MBD2b show different condensate morphology due to distinct overall weak interactions. The N-terminus, a 100% disordered region, is responsible for irregular aggregate formation rather than LLPS. MBD2c does not have the ability of LLPS due to the absence of the C-terminus.

5.2.5 DNA-protein interactions promote the phase transition of MBD2a

Further, the possible influence of DNA on the phase separation properties of MBD2 was explored, as MBD2 could bind to unmethylated and methylated (Hendrich and Bird 1998, Baubec, Ivanek et al. 2013, Buchmuller, Kosel et al. 2020). Considering that both higher DNA concentration and longer DNA could provide more binding sites (valency), the phase

separation was done by adding DNA with various lengths and concentrations. The DNA was labeled with DRAQ5 for visualization. The condensates were observed under the Nikon Eclipse TiE2 microscope.

The DNA had distinct influences on different MBD2 constructs (Fig. 2.6). MBD2a condensates became more irregular in the presence of longer or more DNA (Fig. 2.6A-B). MBD2b condensates decreased while the irregular MBD2c and MBD2a-N aggregates increased in both DNA length and concentration-dependent manner (Fig. 2.6C-E, G). Besides, the morphology of MBD2b, MBD2c, and MBD2a-N condensates did not change (Fig. 2.6C). The DNA could only be weakly incorporated into the MBD2a-C droplets and did not influence the phase separation properties of MBD2a-C (Fig. 2.6F).

Besides, the protein-DNA binding affinity was quantitatively analyzed via the DNA fold enrichment (Fig. 2.6H). The DNA was highly enriched into the condensates of MBD2a, MBD2c, and MBD2a-N but much weaker enriched into the MBD2b aggregates (≤ 2 fold). This is consistent with the *in vivo* studies showing that the MBD2a and MBD2c showed a higher protein enrichment at heterochromatin compartments than MBD2b. Meanwhile, longer DNA showed less enrichment inside the MBD2b and MBD2a-N aggregates.

5.2.6 Cytosine methylation promotes the phase transition of MBD2a/b

As MBD could specifically bind to methylated DNA (Hendrich and Bird 1998, Buchmuller, Kosel et al. 2020), we further checked the influence of DNA methylation on the phase separation properties of MBD2 constructs (Fig. 2.7A). The 800 bp DNA with various methylation levels was generated by PCR using Q5 polymerase by replacing partially dCTP with dmCTP (Rausch, Zhang et al. 2021). We found that DNA methylation promoted the transition from liquid-like droplets to irregular aggregates for MBD2a and MBD2b in a methylation level-dependent manner but had no influence on the aggregate morphology of MBD2c, MBD2a-N, and MBD2a-MBD-TRD (Fig. 2.7A-B). Higher DNA methylation promotes the formation of bigger condensates of MBD2a and MBD2b together with decreased aggregate numbers (Fig. 2.7C-D). Interestingly, the number of MBD2b condensates decreased sharply at low DNA methylation levels while the size of MBD2b condensates increased only in the presence of highly methylated DNA (Fig. 2.7C-D). Therefore, low DNA methylation inhibits the LLPS of MBD2b, while high DNA methylation promotes the formation of irregular aggregates. The DNA methylation level showed no influence on the MBD2c and MBD2a-N aggregate morphology, size, and number (Fig. 2.7A-D).

Surprisingly, MBD2a-MBD-TRD formed irregular aggregates in the presence of DNA (Fig. 2.7A). Higher DNA methylation levels promoted bigger condensate formation with slightly increased condensate number (Fig. 2.7A-D, F). These data suggest that the DNA methylation could switch on the MBD2a-MBD-TRD mediated phase separation and that the

DNA induced MBD2-MBD-TRD phase separation promotes the heterochromatin compartment clustering in vivo (Fig. 2.2E).

The MBD2-DNA binding affinities were further analyzed by the DNA fold enrichment (Fig. 2.7E). The DNA was highly enriched into the MBD2a, MBD2c, and MBD2a-N aggregates regardless of DNA methylation levels. The DNA was weakly enriched into the MBD2b condensates at low methylation levels but highly enriched at higher methylation levels. The DNA was weakly enriched into the MBD2a-MBD-TRD aggregates regardless of DNA methylation levels.

In summary, DNA methylation promotes the phase transition of MBD2a/b from liquid-like spherical droplets to irregular aggregates with distinct properties. DNA methylation promotes the condensate growth of MBD2b in a methylation level-dependent manner. In the presence of DNA, the MBD2a-MBD-TRD formed irregular aggregates, which grew with methylated DNA and fell apart upon (methyl-)DNA removal. This indicates a possible polymer-polymer phase separation (PPPS) property of MBD2a-MBD-TRD.

As for the protein-DNA interaction, the unspecific interactions between N-terminus and DNA play a major role in DNA enrichment for MBD2a and MBD2c as MBD2a/c and MBD2a-N recruited DNA to a similar level. In the absence of the N-terminus (MBD2b), the specific interactions between the MBD and methylated cytosine (mC) play a cumulative role in DNA enrichment, which could be enhanced by the C-terminus.

5.2.7 Polymer-polymer phase separation properties of MBD2-MBD-TRD in vitro and in vivo

The high dependency of MBD2-MBD-TRD phase separation on methylated DNA points to a hypothesis that polymer-polymer phase separation (PPPS) also underlies the function of MBD2 in heterochromatin organization. To distinguish, the total amount of DNA in heterochromatin compartments was measured in cells expressing MBD2c, MBD2-MBD-TRD, and MBD2b. The percent of heterochromatin did not change in cells with increased MBD2c and MBD2a level, while it increased in cells with increased MBD2b levels (Fig. 2.8A, B). Thus, the MBD2c and MBD2-MBD-TRD might probably adopt PPPS while MBD2b adopts LLPS for heterochromatin organization. In organisms, PPPS and LLPS might collaborate to mediate heterochromatin structure.

5.2.8 Conclusion

In this chapter, we studied the functional differences of the MBD2 isoforms and found that the N-terminus, the MBD-TRD, and the C-terminus played distinct roles in MBD2 localization and function *in vivo* together with altered phase separation behavior *in vitro* (Fig. 2.9A).

The N-terminus is predicted to be 100% disordered and shows a biased charge distribution. The very N-terminal region is enriched with negatively charged amino acid residues, which is followed by an enrichment of positively charged amino acid residues. These lead to a negatively charged region (-) and positively charged region (+), suggesting possible MBD2a-N-MBD2a-N and MBD2a-N-DNA interactions. Indeed, the N-terminus of MBD2a is responsible for nonspecific protein-DNA interaction and shows the ability to form irregular aggregates *in vitro* by itself or in the presence of DNA. *In vivo*, the MBD2a-N was mislocated into the nucleolus rather than heterochromatin compartments due to strong electrostatic RNA binding properties.

The MBD-TRD contains the MBD and TRD domain, two conserved domains that are responsible for methylated CpG binding and protein-protein interaction, separately. The MBD-TRD is less disordered (~42% disordered). *In vivo*, the MBD2a-MBD-TRD was enriched into the heterochromatin compartments and showed a similar influence on heterochromatin compartment size to MBD2a. Thus, the MBD-TRD retains the majority of MBD2a functions and is the minimal requirement for MBD2 function in heterochromatin organization. *In vitro*, the MBD-TRD alone did not phase separate in all conditions tested, while adding methylated DNA promotes the condensate formation. This is a typical characteristic of polymer-polymer phase separation (PPPS) (Erdel and Rippe 2018), indicating that the PPPS might also play a role in MBD2 function in heterochromatin compartments.

The C-terminus is highly disordered (~60% disordered) and is responsible for the NuRD complex assembly. *In vitro*, the purified MBD2a-C displayed no (or weak) DNA binding affinity but showed the ability of LLPS, indicating that the C-terminus drives the LLPS of MBD2a and MBD2b.

The three isoforms of MBD2 found *in vivo* adopt different combinations of the three regions, the N-terminus, the MBD-TRD, and the C-terminus, and thus exhibit distinct functions *in vivo* and different phase separation behaviors *in vitro*.

MBD2a was enriched into heterochromatin compartments in a protein concentration-dependent manner and promoted heterochromatin compartment fusion together with a slow exchange between heterochromatin compartments and the surrounding context *in vivo*. *In vitro*, the MBD2a formed irregular aggregates under low weak interactions (at low protein concentrations, or destroying the hydrophobic interactions), which changed into liquid-like spherical droplets with increased weak interactions (lower salt and higher protein concentrations), and further transformed into irregular aggregates with higher weak interactions (adding PEG or methyl-DNA) (Fig. 2.9B). These studies suggest that the phase separation behavior of MBD2a is strictly regulated by the strength of weak interactions.

MBD2b exhibits an overall similar but weaker function than the full-length MBD2a *in vivo* and *in vitro*. Compared to MBD2a, MBD2b was less enriched into the heterochromatin compartments and promoted the heterochromatin compartment fusion to a less extent yet with faster mobility *in vivo*. Meanwhile, the unmethylated DNA was less enriched into the MBD2b condensates than into the MBD2a condensates *in vitro*. Besides, MBD2b showed a delayed phase transition behavior compared to MBD2a, as MBD2b requires a higher protein concentration and DNA methylation to induce a phase transition (Fig. 2.9B). These observations are consistent with the N-terminal properties of DNA binding and irregular aggregate formation. Besides, the (G/R), a glycine-arginine repeat within the positively charged N-terminus, was reported to weaken the NuRD complex assembly and MBD-mCpG interaction when the arginine residues were methylated (Tan and Nakielny 2006). Taken together, the N-terminus enhances the phase behavior and functions of MBD2a.

Indeed, MBD2a and MBD2b performed distinct functions in ES cell differentiation although both were increasingly expressed during differentiation (Du, Luu et al. 2015). MBD2b could fully rescue the differentiation block in MBD3 KO ES cells, while the full-length MBD2a could only partially rescue the differentiation block (Schmolka, Bhaskaran et al. 2021). Besides, previous work showed that the MBD2a could be located at unmethylated sites *in vivo* mediated by the NuRD complex (Baubec, Ivanek et al. 2013). Here, we confirmed the unmethylated DNA binding affinity of the N-terminus and found that the MBD2b is less enriched at the heterochromatin compartments than MBD2a. These data indicate that the genome-wide distribution of MBD2a is probably different from that of MBD2b, especially at unmethylated sites.

MBD2c retains the N-terminus and MBD-TRD, showed similar molecular dynamics and enrichment into heterochromatin compartments but did not influence the heterochromatin compartment size as compared to MBD2a. *In vitro*, the MBD2c formed only irregular aggregates (Fig. 2.9B). These indicate that the C-terminal induced LLPS is responsible for heterochromatin compartment fusion. The MBD2c plays distinct roles in heterochromatin compartmentalization via phase separation compared to MBD2a and MBD2b.

MBD2c is highly expressed in pluripotent stem cells and the expression decreased during differentiation, indicating a role of MBD2c in pluripotent state maintenance (Lu, Loh et al. 2014). Indeed, the MBD2c promotes the ES cell self-renew, while the MBD2a/b promotes the ES cell differentiation (Lu, Loh et al. 2014, Schmolka, Bhaskaran et al. 2021). Molecularly, the MBD2c competes with MBD2a/b for promoters of OCT4 and NANOG, inhibits the MBD2a/b induced assembly and localization of NuRD complex locally, and thus inhibited the MBD2-NuRD induced OCT4 and NANOG gene repression (Lu, Loh et al. 2014). Here, we provided an alternative mechanism underlying the isoform-specific functions

of MBD2 isoforms. The MBD2a and MBD2b promote the ES cell differentiation via LLPS induced heterochromatin reorganization.

The MBD-TRD and the C-terminus are capable of PPPS and LLPS, respectively. The PPPS is featured with invariant chromatin bodies with increased protein concentration, as the chromatin region that contains binding sites for bridging factors (proteins) is not changed (Erdel and Rippe 2018). The LLPS is characterized by an increased size of chromatin bodies with increased protein concentration due to increased multivalent interactions (Erdel and Rippe 2018). We measured the occupancy of heterochromatin compartments in cells expressing MBD2c, MBD-TRD, MBD2b, and full-length MBD2a, respectively. The percent of heterochromatin did not increase with increased MBD-TRD and MBD2c levels but increased with increased MBD2b levels. Thus, the MBD2b might adopt the LLPS process in heterochromatin compartments, while the MBD2c adopts the PPPS process. The MBD2a adopts alternative or both processes depending on the context conditions.

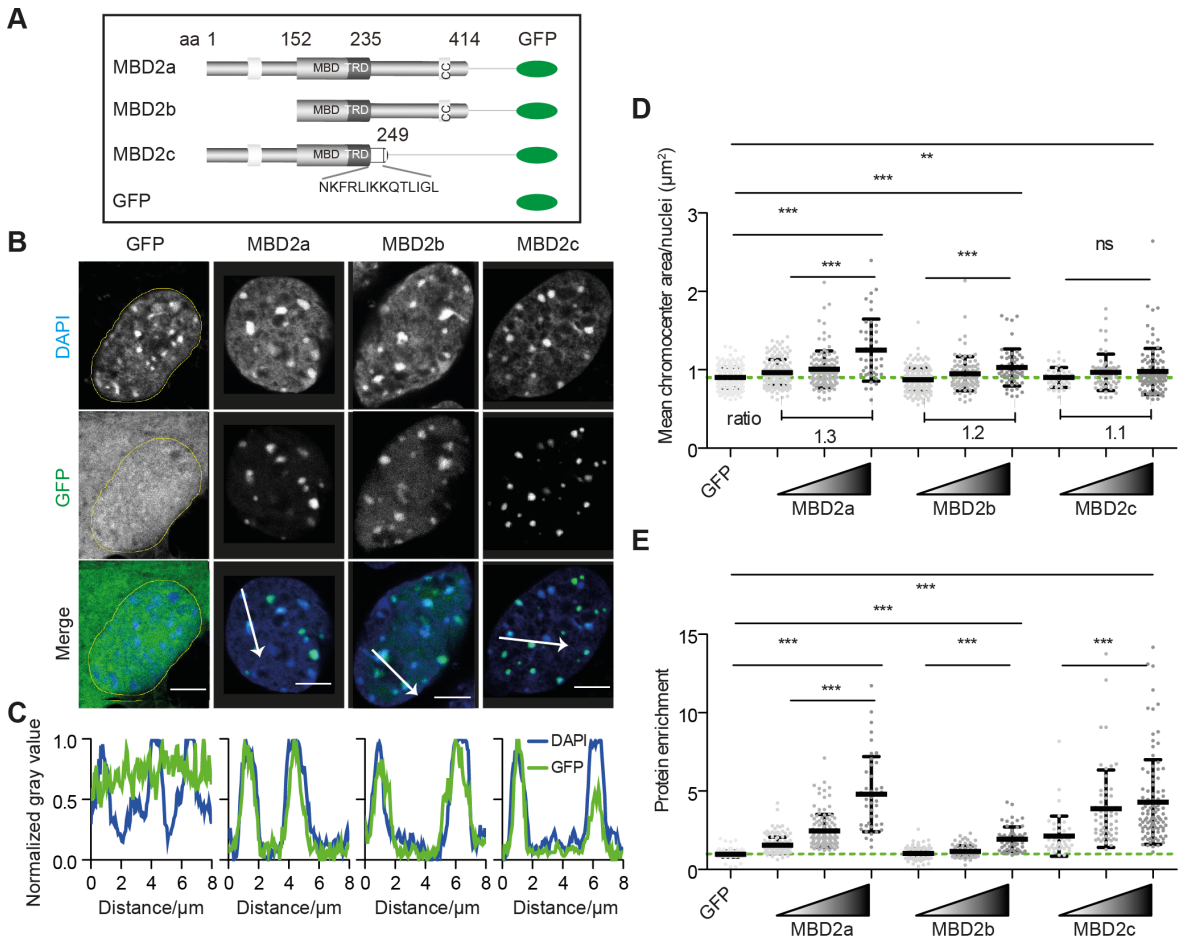


Figure 2.1 Truncated isoforms of MBD2 have impaired functions in heterochromatin compartment fusion.

(A) Schematic representation of MBD2 constructs from Mus musculus (house mouse). Fusion proteins with GFP tag at the C-terminus were expressed in C2C12 cells. MBD2a is the full-length MBD2 containing both N- and C-terminal disordered regions. MBD2b lacks the disordered N-terminus due to alternative translational start sites (TSS). MBD2c utilizes an alternative shorter C-terminus due to alternative splicing. aa: amino acids. MBD:

methyl-DNA binding domain. TRD: transcriptional repression domain. Green ellipse: GFP. The line between MBD2 isoforms and GFP does not represent the length of the linker region in between.
(B) Representative images showing the distribution of MBD2 isoforms *in vivo* with individual optical sections. C2C12 myoblast cells were transfected with plasmids expressing the three isoforms of MBD2, which were fused with a GFP coding sequence at the C-terminus. Cells were fixed 36 h after transfection with formaldehyde and the genome was stained with DAPI before microscope observation. The confocal microscopy Leica TCS SP5 II equipped with 100 X objective was applied in this study. Cells with GFP expression, which has homogeneous distribution across cells, were applied as a control to exclude the possible influence of GFP on heterochromatin compartment size. The Yellow dashed line tracked the nuclei boundary. All images were taken under the same microscope settings. Scale bar = 5 µm.
(C) Fluorescence intensity profile and co-localization analyses of MBD2 isoforms. Graphs represent relative GFP (green line) and DAPI (blue line) fluorescence intensity profiles, which were calculated on images in (B; white arrows) using ImageJ.
(D-E) Quantification of the heterochromatin compartment size (D) and protein enrichment in heterochromatin compartments (E) relative to MBD2 expression levels. Heterochromatin compartments were segmented by DAPI staining and cells were subgrouped by GFP fluorescence intensities using the images obtained from confocal microscope as described above. Scatter plots represent the distribution of the mean heterochromatin compartment areas (D) per nucleus and protein fold enrichment in the heterochromatin compartments (E) in C2C12 cells, related to (B). Green dashed lines represent the mean heterochromatin compartment area (D) and the fold enrichment of GFP fluorescence signals (E) in C2C12 cells fixed 36 h after transfection with the plasmid expressing GFP only. The blue dashed line represents the mean heterochromatin compartment area (D) in untransfected C2C12 cells after fixation. Error bars on the histogram correspond to SD (standard deviation). Significances were calculated by unpaired t-test comparing the high expressing samples with low expressing samples as indicated by lines. ns P > 0.05; *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001.

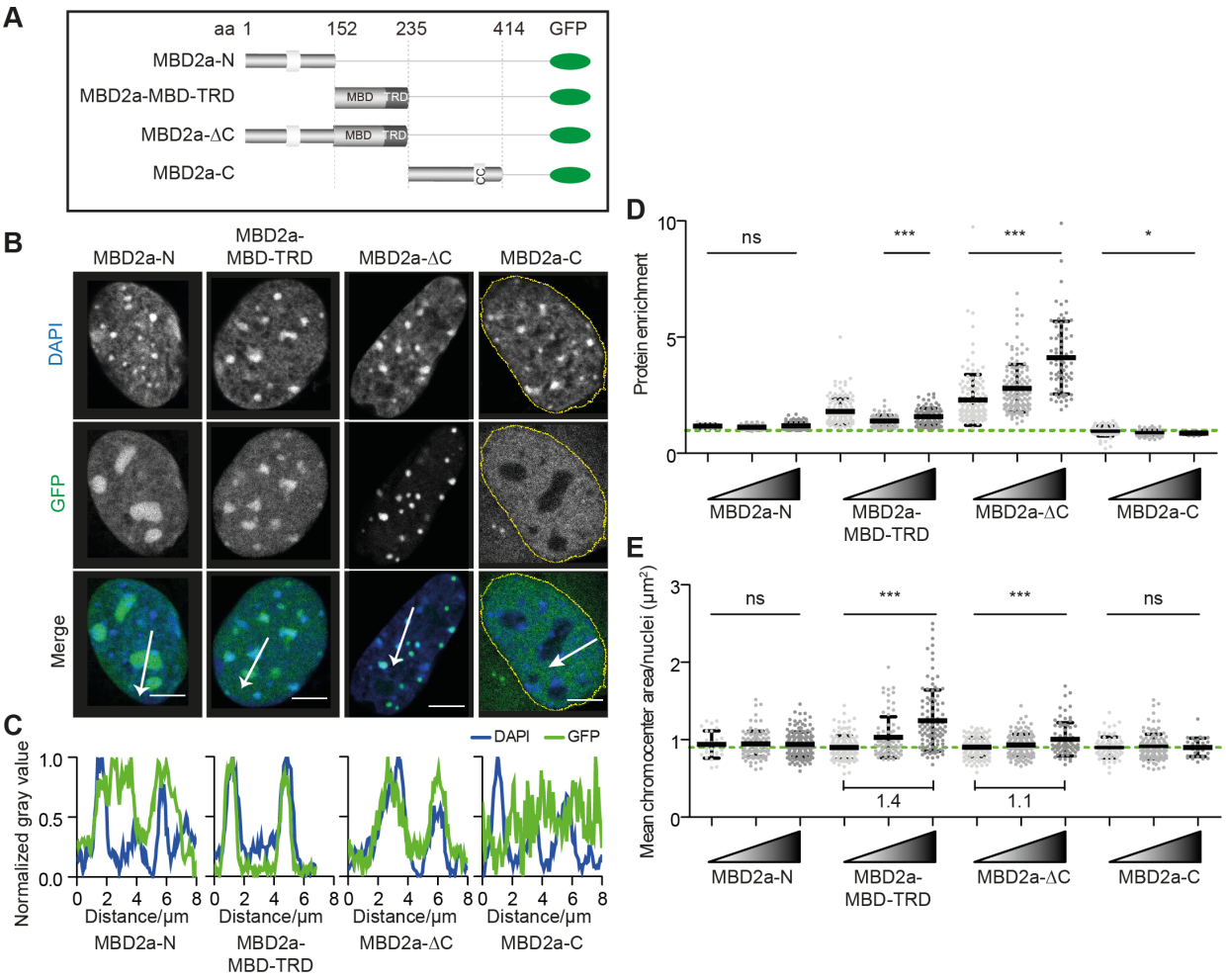


Figure 2.2 Contributions of different MBD2a regions to heterochromatin compartments.

(A) Schematic representation of MBD2a-GFP fusion constructs. Green ellipse: GFP.
(B) Representative images showing the distribution of MBD2a constructs *in vivo*. Cells were transfected, fixed, DAPI stained, and visualized by confocal microscopy (Leica TCS SP5 II) equipped with a 100 X objective. The

yellow dashed line marked the nuclei boundary. All images were taken with the same microscope settings. Scale bar = 5 µm.
(C) Fluorescence intensity profile and co-localization analyses of MBD2a constructs. Graphs represent relative GFP (green line) and DAPI (blue line) fluorescence intensity profiles, which were calculated on the images in (B; white arrows).
(D-E) Quantification of protein enrichment in heterochromatin compartments (D) and heterochromatin compartment size (E) relative to the expression levels of MBD2a constructs. Heterochromatin compartments were segmented based on the DAPI intensities and cells were subgrouped by GFP fluorescence intensities using the images obtained from confocal microscope. Scatter plots represent protein fold enrichment in the heterochromatin compartments (D) per nucleus and the distribution of the mean heterochromatin compartment areas (E) in C2C12 cells, related to (B). The green dashed lines represent the fold enrichment of GFP fluorescence signals (D) and the mean heterochromatin compartment area (E) in C2C12 cells fixed 36 h after transfection with the plasmid expressing GFP only. Error bars on the histogram correspond to SD values (standard deviation). Significances were calculated by unpaired t-test comparing the high expressing samples with low expressing samples as indicated by lines. ns P > 0.05; *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001.

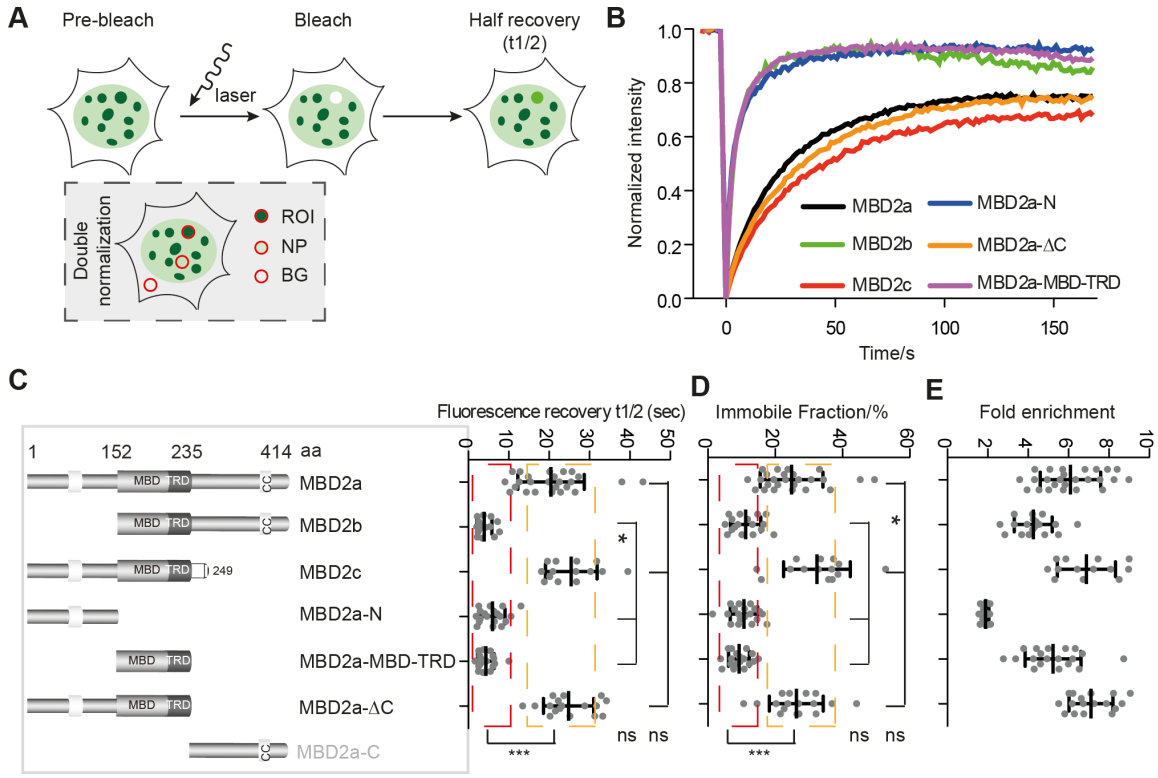


Figure 2.3 Dynamics of MBD2 constructs at nuclear condensates.

(A) Schematic graph showing the process of FRAP in heterochromatin compartments. In the FRAP experiment, the GFP fluorescence in the circular heterochromatin compartment area in a cell was rapidly and irreversibly bleached using a 488 laser pulse with 100% laser power, leaving a circular region devoid of GFP signal. Due to molecular exchange between bleached and unbleached regions, the GFP signal increased with time till reaching a plateau in the bleached region. The images were taken using a time-lapse confocal microscope Leica TCS SP5 II equipped with 100 X objective. The extent of final recovery depends on the fraction of mobile molecules. The recovery halftime (t1/2) is the time from the bleach point to the time point where the fluorescence intensity reaches half of the final recovered intensity. The GFP signal intensity in the nucleoplasm and outside nuclei were used for double normalization (box with dashed line).
(B) Mean FRAP curves in C2C12 cells expressing GFP-MBD2a (n = 25), GFP-MBD2b (n = 18), GFP-MBD2c (n= 14), GFP-MBD2a-N (n = 19) , GFP-MBD2-ΔC (n = 17) and GFP-MBD2a-MBD-TRD (n = 19). MBD2a-C did not form any condensates in cells, so no FRAP was done with this construct.
(C-E) Fluorescence recovery half time (1/2) (C), the immobile fraction (D), and the fold enrichment (E) of MBD2 constructs in bleached regions relative to the nucleoplasm. The fold enrichment was calculated as (IROI - IBG)/(INP - IBG). I: mean GFP fluorescence intensity; ROI: bleached region; BG: background; NP: nucleoplasm. Data were given as mean ± SD. Significances were calculated by unpaired t-test comparing the high expressing samples with low expressing samples as indicated by lines. ns P > 0.05; *P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001.

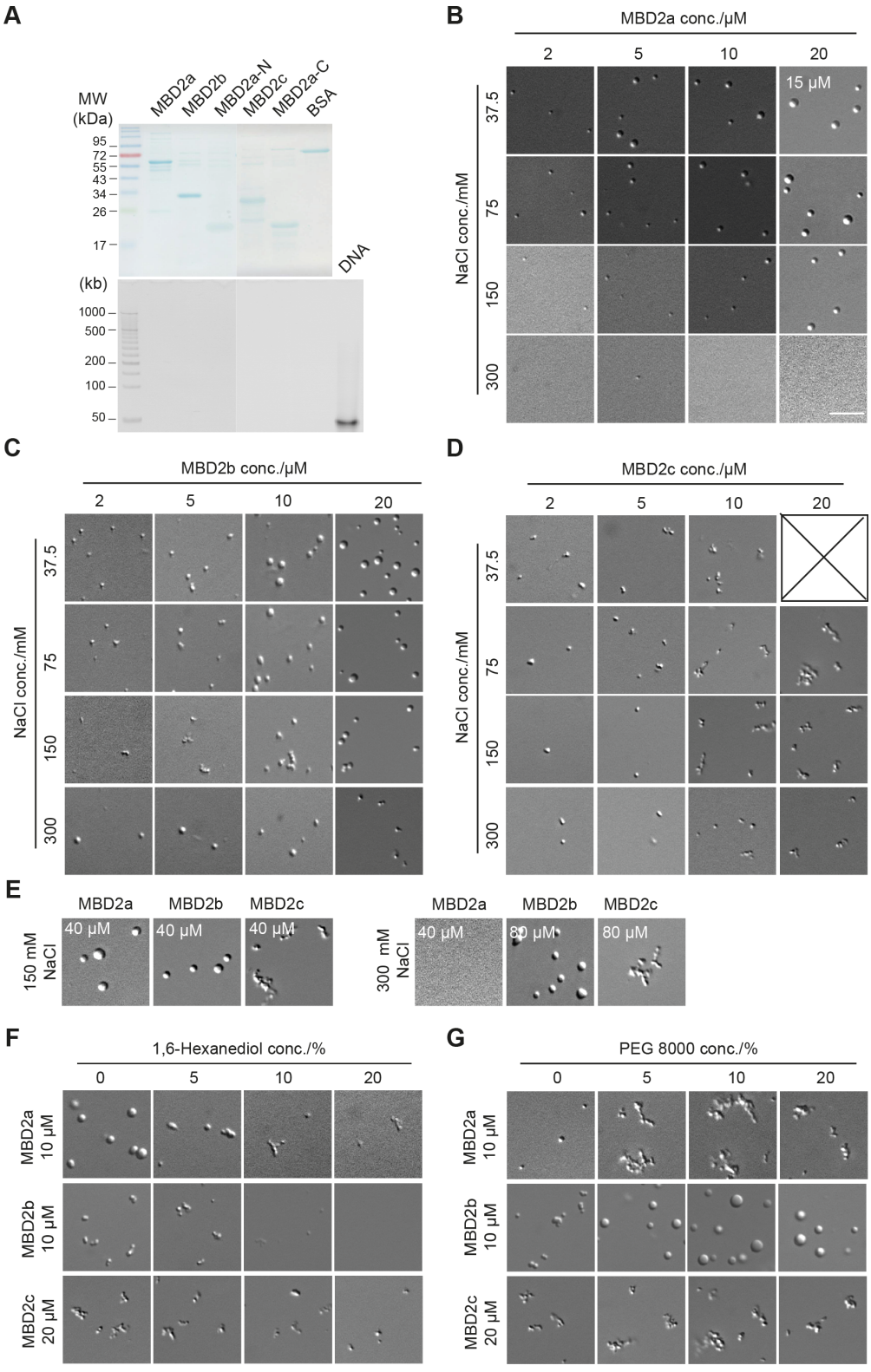


Figure 2.4 Phase separation properties of MBD2 isoforms.

(A) Detection of protein purity. Different MBD2 constructs were expressed in bacteria by IPTG induction, purified using chitin beads, and finally stored in buffer containing 300 mM (for MBD2a) and 500 mM NaCl (all other proteins). Top: SDS polyacrylamide gel electrophoresis of purified MBDs followed by Coomassie staining. 2 μg each lane. Bottom: TBE polyacrylamide gel electrophoresis of purified MBDs followed by ethidium bromide staining. 10 μg each lane. Negative control 1: BSA with the same amount. Positive control 2: synthesized ~140 ng 42 bp dsDNA.

(B-E) Phase diagram of MBD2 isoforms droplet formation at indicated salt and protein concentrations. The *in vitro* phase separation assay was done at different protein and salt concentrations. The mixtures were transferred to chambers made of double-sided tapes and sealed with coverslips 45 min after incubation at room temperature. The droplets were observed using a Nikon Eclipse TiE2 microscope equipped with differential interference contrast (DIC) microscopy. Scale bars = 10 μm. conc.: concentration.
(F-G) Phase diagram of MBD2 isoforms phase separation. 10 μM proteins were incubated in buffer containing 150 mM NaCl and various concentrations of 1,6-hexanediol (F) or PEG 8000 (G) at room temperature for 45 min before imaging. conc.: concentration.

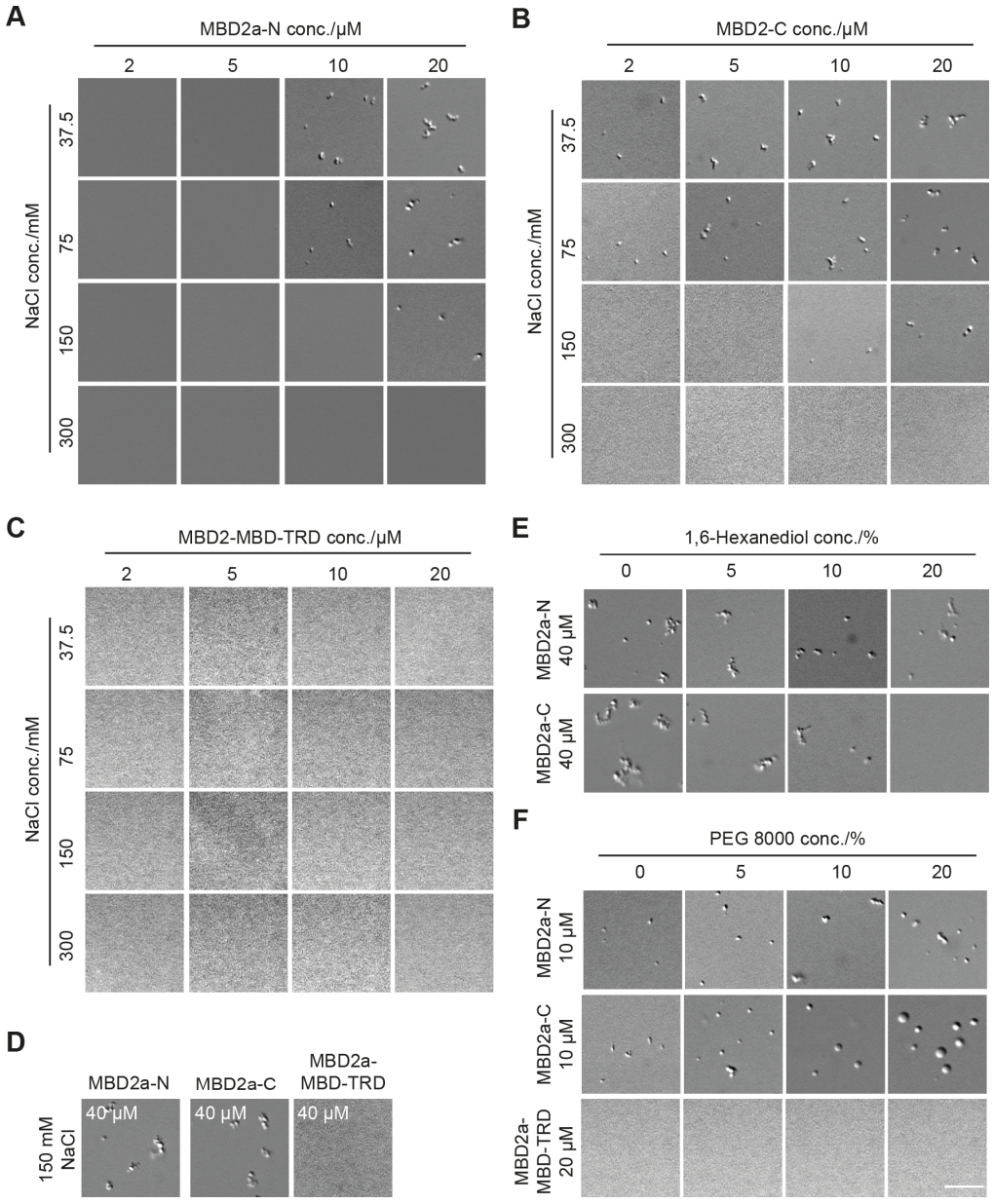


Figure 2.5 Phase separation properties of MBD2a truncations.

(A-D) DIC images showing the LLPS properties of MBD2a-N (A), MBD2a-C (B), and MBD2a-MBD-TRD (C) truncations at indicated salt and protein concentrations. conc.: concentration.
(E-F) DIC images showing the LLPS properties of MBD2a-N, MBD2a-C, and MBD2a-MBD-TRD truncations in the presence of 1,6-Hexanediol (E) or crowding reagents (F). 10 μM MBD2a-N and MBD2a-C and 20 μM MBD2a-MBD-TRD were incubated with various concentrations of 1,6-hexanediol and PEG 8000 in buffer containing 150 mM NaCl for *in vitro* phase separation and imaged. conc.: concentration. Scale bar = 10 μm.

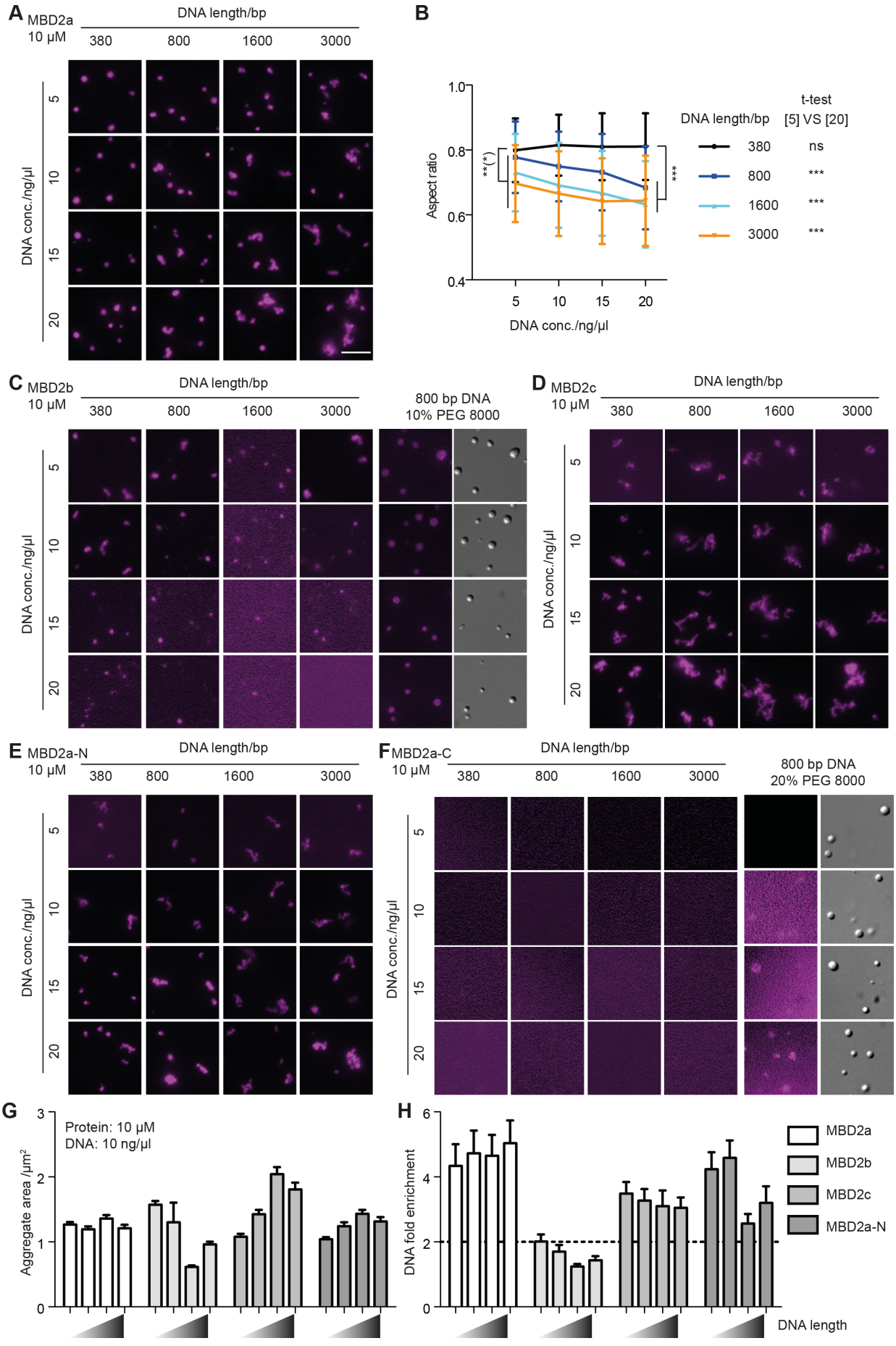

A
MBD2a
10 µM
DNA length/bp
380
800
1600
3000
DNA conc./ng/µl
5
10
15
20
B
Aspect ratio
1.0
0.8
0.6
0.4
**(*)
***
5
10
15
20
DNA conc./ng/µl
t-test
DNA length/bp
[5] VS [20]
380
ns
800
***
1600
***
3000
***
C
MBD2b
10 µM
DNA length/bp
380
800
1600
3000
800 bp DNA
10% PEG 8000
D
MBD2c
10 µM
DNA length/bp
380
800
1600
3000
E
MBD2a-N
10 µM
DNA length/bp
380
800
1600
3000
F
MBD2a-C
10 µM
DNA length/bp
380
800
1600
3000
800 bp DNA
20% PEG 8000
G
Protein: 10 µM
DNA: 10 ng/µl
Aggregate area /µm²
3
2
1
0
H
DNA fold enrichment
6
4
2
0
MBD2a
MBD2b
MBD2c
MBD2a-N
DNA length

Figure 2.6 Influence of DNA on the phase separation properties of MBD2 constructs.

The phase separation assay was done by incubating 10 μM protein with DNA of different lengths and concentrations in buffer containing 150 mM NaCl for 45 min at room temperature. The DNA was labeled by DRAQ5. The condensates were visualized by the Nikon Eclipse TiE2 microscope. The condensates were recognized based on the DRAQ5 intensities inside/outside condensates. Droplets with size > 0.1 μm2 were identified and droplet parameters were measured. The aspect ratio (AR) was calculated as the ratio of maximal Feret diameter to minimal Feret diameter. The fold enrichment of DRAQ5 represented the fold enrichment of DNA. The DNA (DRAQ5) fold enrichment was calculated as the ratio of DRAQ5 intensity within droplets to DRAQ5 overall mean intensity per image. Three images for each condition.
(A) Fluorescence images showing the phase separation behavior of MBD2a in the presence of DNA. conc.: concentration. Scale bar = 10 μm.
(B) Graph showing the condensate roundness changes formed by MBD2a in the presence of DNA with different lengths and concentrations. Data were given as mean ± SD. Significances were calculated by unpaired t-test comparing the high expressing samples with low expressing samples as indicated by lines. ns $P > 0.05$; *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$.
(C-F) Fluorescence and DIC images showing the phase separation behavior of MBD2b (C), MBD2c (D), MBD2a-N (E), and MBD2a-C (F) in the presence of DNA. conc.: concentration.
(G-H) Separated bar graph showing the average condensate area (G) and average DNA fold enrichment (H) within condensates. The condensates formed in mixtures containing 10 μM protein, 10 ng/μl DNA with different lengths, and 150 mM NaCl were applied for quantitative analysis. Error bars correspond to the standard error of the mean (SEM) for condensate area and standard deviation (SD) for fold enrichment.

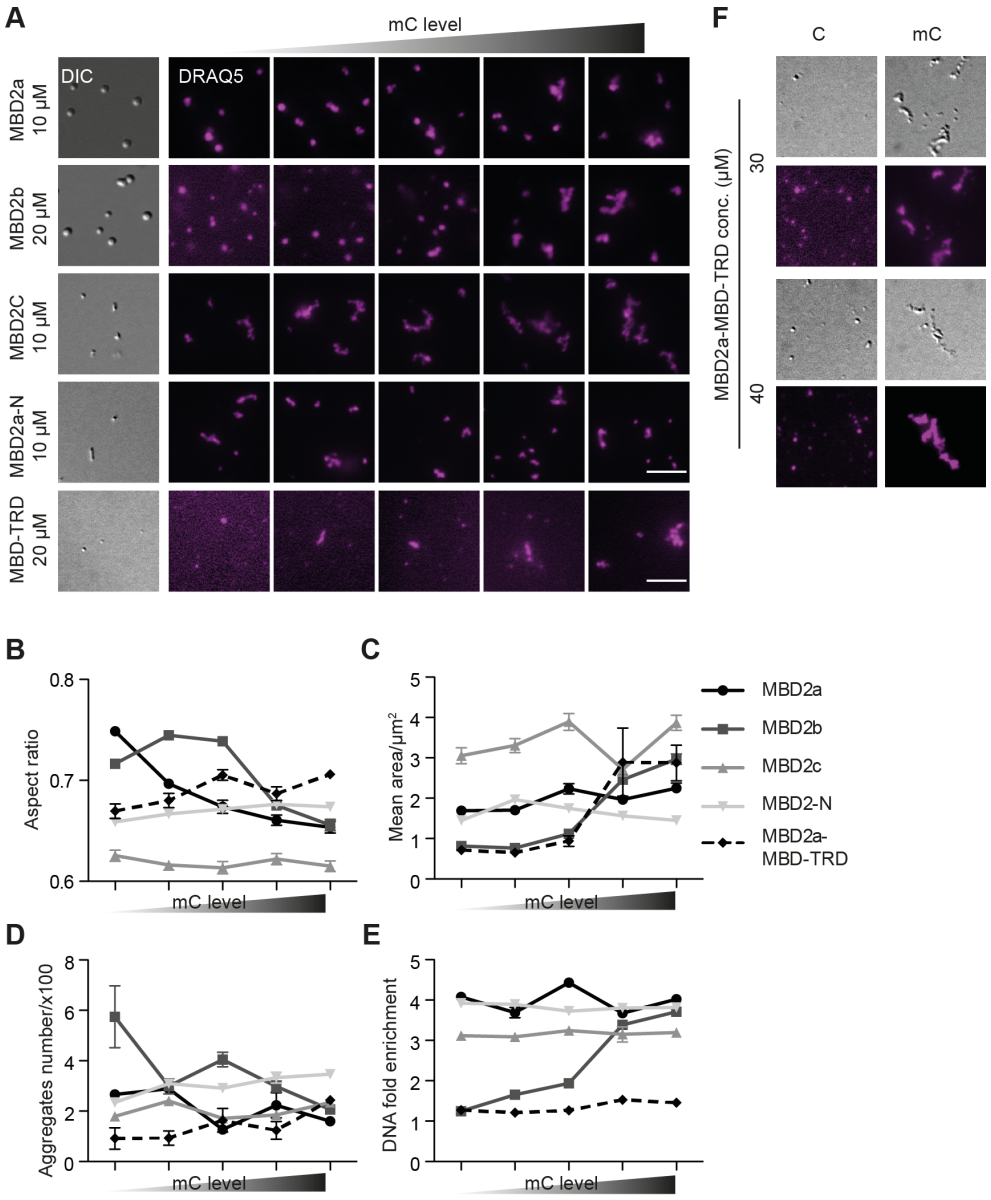


Figure 2.7 DNA methylation promotes phase transition of MBD2a and MBD2b.

The 800 bp DNA with various methylation levels was generated by PCR using Q5 polymerase by replacing partially dCTP with dmCTP. The DNA was labeled by DRAQ5 to visualize the distribution of DNA. The *in vitro* phase separation assay was done by incubating proteins in buffer containing DNA of different methylation levels at room temperature for 45 min. NaCl: 150 mM. The condensates were visualized by the Nikon Eclipse TiE2 microscope. The condensates were recognized based on the DRAQ5 intensities inside/outside condensates. Droplets with size > 0.1 µm2 were identified and droplet parameters were measured. The aspect ratio (AR) was calculated as the ratio of maximal Feret diameter to minimal Feret diameter. The fold enrichment of DRAQ5 represented the fold enrichment of DNA. The DNA fold enrichment was calculated as the ratio of DRAQ5 intensity within droplets to DRAQ5 overall mean intensity per image. Three images for each condition.
(A) Fluorescence and DIC images showing the phase separation behavior of MBD2a, MBD2b, MBD2c, MBD2a-N, and MBD2a-MBD-TRD in the presence of DNA with different methylation levels. DNA for MBD2a, MBD2b, MBD2c, and MBD2a-N: 10 ng/µl; DNA for MBD2a-MBD-TRD: 20 ng/µl. Scale bar = 10 µm.
(B-E) Graphs showing the influence of DNA methylation on condensate size (B), number (C), roundness (D), and DNA enrichment into protein condensates (E). Error bars correspond to the standard error (SEM).
(F) Fluorescence and DIC images showing the phase separation behavior of MBD2a-MBD-TRD with unmethylated (C) and methylated (mC) DNA. DNA: 20 ng/µl.

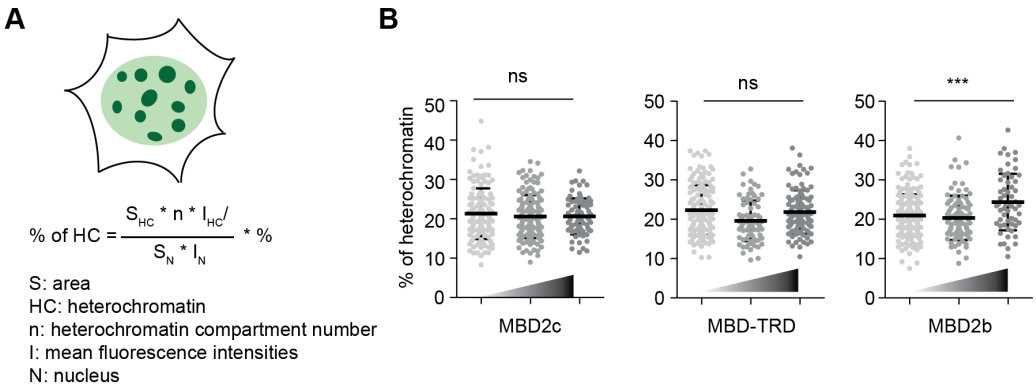


Figure 2.8 MBD2b overexpression increased the total amount of heterochromatin

(A) Equation applied to quantify the percent of heterochromatin. the percent of heterochromatin was calculated as the percent ratio of total DAPI intensity in heterochromatin compartments (S_{HC} * n * I_{HC}) to total DAPI intensity in the whole nuclei (S_N * I_N).
(B) Scatter plot showing the heterochromatin compartment % of the total nuclei volume in C2C12 cell expressing MBD2c, MBD-TRD, and MBD2b. Data were given as mean ± SD. Significances were calculated by unpaired t-test comparing the high expressing samples with low expressing samples as indicated by lines. ns $P > 0.05$; *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$.

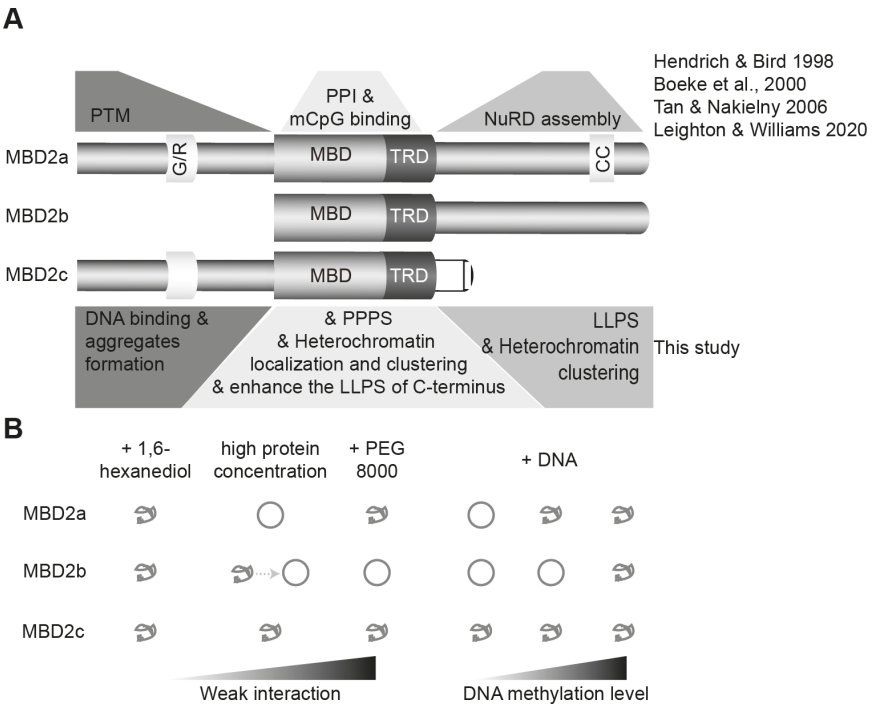


Figure 2.9 Different regions of MBD2a show distinct functions in vivo and in vitro.

(A) Schematic graph showing region-dependent functions of MBD2 from publications (up) (Hendrich and Bird 1998, Boeke, Ammerpohl et al. 2000, Tan and Nakielny 2006, Leighton and Williams Jr 2020) and this study (down). The G/R containing N-terminus is post-translationally modified and is responsible for nonspecific DNA binding and aggregate formation. The MBD-TRD shows the ability of protein-protein (PPI) and protein-mCpG interactions, polymer-polymer phase separation (PPPS), and heterochromatin localization and clustering. The C-terminus participates in the NuRD complex assembly, liquid-liquid phase separation (LLPS), and heterochromatin clustering in the presence of the MBD-TRD, which provides the heterochromatin localization ability. The LLPS properties of the C-terminus could be enhanced by the upstream sequence (MBD-TRD and N-terminus).
(B) Graphic representation elucidating the phase behaviors of MBD2 isoforms under various homo and hetero interaction strengths.

5.3 MeCP2-induced heterochromatin organization is driven by oligomerization-based liquid-liquid phase separation and restricted by DNA methylation

5.3.1 MeCP2 forms condensates with characteristics of liquid-like droplets in physiological concentrations

In section 5.1, we already showed that purified MeCP2 could form liquid-like droplets in a NaCl and protein concentration-dependent manner. To further characterize the phase separation properties of MeCP2 *in vitro*, both untagged MeCP2 and GEP tagged MeCP2 (GFP-MeCP2) were purified which are free of DNA (Fig. 3.1A). Firstly, based on the MeCP2 levels in the mouse brain (6 x 10^6 molecules per nucleus) (Skene, Illingworth et al. 2010) and the size of the nuclei (200-800 μm^3, 500 μm^3 was taken as average) (Wegiel, Flory et al. 2015), we calculated the MeCP2 concentration (~ 10-50 μM in mouse brain cells).

Secondly, the phase separation properties of MeCP2 were checked. The turbidity assay was applied to check the MeCP2 absorbance at λ = 340 nm in buffer containing different concentrations of NaCl and MeCP2. The result showed that the turbidity increased at higher protein concentration and lower salt concentration (Fig. 3.1B). The droplets were observed under the microscope after incubating at room temperature for 45 min to induce droplet formation. We observed that MeCP2 alone in the range of physiological concentrations (~10 μM) formed distinct spherical condensates at low salt concentration independent of the cations used in the buffer (Na^+, K^+, NH^{4+}). The number and size of the condensates increased with increasing protein concentration but decreased with increasing salt concentration till no droplet formed at physiological salt conditions (150 mM salt) (Fig. 3.1C, E-F). Time-lapse imaging showed that these condensates could fuse into bigger ones (Fig. 3.1D), indicating that the MeCP2 condensates are highly dynamic and liquid-like in NaCl buffer. This is reminiscent of the fusion of MeCP2 labeled heterochromatin domains in cells as we have previously shown (Brero, Easwaran et al. 2005). Together with previously shown results in 5.1 that MeCP2 showed resistance to 1,6-hexanediol, the MeCP2 LLPS is probably mainly driven by electrostatic interactions rather than hydrophobic interactions. Bio-macromolecules are commonly used to mimic nuclear crowding environments and were shown to promote LLPS. We found that in physiological salt conditions dextran T150 could promote LLPS of MeCP2 in a concentration-dependent manner (Fig. 3.1G). Yet, the crowders themselves were not enriched inside the protein droplets (Fig. 3.1H).

Thirdly, the MeCP2 distribution was analyzed based on sedimentation and fluorescent intensity. The MeCP2 levels in the buffer phase (supernatant (S)) and the droplets were quantified by droplet sedimentation assay (Fig. 3.2A-B). 10 μM MeCP2 were incubated in different conditions for *in vitro* phase separation assay, then the droplets were pelleted by

centrifuge at 14000 rpm for 15 min at room temperature. The supernatants were applied for the SDS-PAGE gel electrophoresis and followed by coomassie staining (Fig. 3.2A). The result showed that more MeCP2 molecules were concentrated into the droplets with higher concentrations of PEG 8000 (Fig. 3.2B), indicating a decreased MeCP2 concentration threshold for LLPS when in crowding environments.

The droplet properties were further quantitatively analyzed based on the fluorescent intensity inside and outside droplets. The fluorescent intensities of GFP-MeCP2 in buffer with no droplet formation (150 mM NaCl) were measured under a microscope and plotted versus GFP-MeCP2 concentration to generate the standard curves (Fig. 3.2C-E). 10 μM untagged and GFP-tagged MeCP2 mixture (molar ratio: 99:1) was applied for *in vitro* phase separation assay (Fig. 3.2C-J). For in-solution protein concentration, the droplets were pelleted by centrifuge at 14000 rpm for 15 min. The supernatants were applied for imaging, and the mean GFP intensities in supernatants were measured (Fig. 3.2C). For in-droplet protein concentration, the MeCP2 mixtures after *in vitro* phase separation were applied for microscope detection, and the fluorescence intensities in droplets were measured (Fig. 3.2C, F). The MeCP2 concentrations in solution and droplets were quantified based on the GFP intensity in solution and inside the droplets interpolated from a standard curve of GFP-MeCP2 in solution (Fig. 3.2D-E). In the low salt condition without PEG 8000, MeCP2 concentration in droplets (121.7 ± 20.5 μM) and the solution (8.6 ± 0.5 μM) revealed a 14 fold enrichment of MeCP2 in droplets following LLPS (Fig. 3.2G). Yet, the MeCP2 concentration in solution did not decrease significantly (8.6 ± 0.3 μM compared to the total initial concentration 10 μM) probably because only a small fraction (~25%) of MeCP2 underwent LLPS in this condition (Fig. 3.2B). Then, we analyzed the MeCP2 concentration in the droplets at physiological salt (150 mM) and protein (10 μM) conditions with various crowder concentrations (Fig. 3.2F, H). With increasing crowder concentration, the MeCP2 concentration in the droplets increased from 45.63 ± 0.11 μM at 5% PEG 8000 to 104.01 ± 1.27 μM at 20% PEG 8000 (Fig. 3.2H, J). This means that the (crowding) environment influences the MeCP2 distribution. *In vivo*, MeCP2 was reported to be 9 fold enriched at heterochromatin compartments in NIH3T3 fibroblast cells and 5 fold enriched in Pmi28 myoblast cells (Agarwal, Hardt et al. 2007, Müller‐Ott, Erdel et al. 2014), which is of a similar magnitude to the MeCP2 enrichment in the droplets. Besides, all droplets were highly spherical (Fig. 3.2 I-J), further suggesting that they are liquid-like.

In conclusion, purified MeCP2 showed the ability to form liquid-like droplets by itself in physiological protein concentrations, which we defined as the minimal MeCP2-LLPS system.

5.3.2 DNA promotes the liquid-liquid phase separation of MeCP2

We next investigated the influence of DNA on *in vitro* MeCP2 phase separation properties. Firstly, template DNA was synthesized with different lengths (Fig. 3.3A-B) and labeled with the DNA dye DRAQ5 to detect the DNA distribution during MeCP2 LLPS. By mixing template DNA with MeCP2 in the physiological salt condition (150 mM NaCl) in which MeCP2 alone does not form droplets (150 mM NaCl), we found that MeCP2 could form liquid-like droplets with longer DNA (800, 1600, or 3000 bp) but not with short (380 bp) DNA (Fig. 3.3D) together with enrichment of DNA (DRAQ5) in the droplets. This indicates a multi-valency induced phase separation property of MeCP2 with DNA. The DRAQ5 signals were further applied for droplet segmentation and quantitative analysis. Longer and higher concentrations of DNA were shown to promote the formation of more and bigger droplets (Fig. 3.3E). Thus, DNA promoted (switched on) *de novo* phase separation of MeCP2 in physiological conditions in a DNA length and concentration-dependent manner, highlighting the importance of multiple ionic interactions among MeCP2 and DNA molecules for LLPS.

5.3.3 DNA methylation restricts the size of MeCP2 droplets

MeCP2 was originally recognized as a methyl-cytosine binding protein (Lewis, Meehan et al. 1992). Recently, MeCP2 was found to bind methylated cytosine at non-CpG sites (mCpH, H=A, G, C). mCpH was shown to increase during early brain development, in parallel with MeCP2 postnatal accumulation. MeCP2 interactions with noncanonical cytosine methylation were shown to be important for normal brain function (Chen, Chen et al. 2015, Jang, Shin et al. 2017, Tillotson, Cholewa-Waclaw et al. 2020). Thus, we checked if/how CpG methylation (mCpG) and non-canonical cytosine methylation (mC) play a role in MeCP2 LLPS. As the 800 bp template DNA could already promote the LLPS of MeCP2 under physiological conditions, we further generated 800 bp DNA templates with methylated cytosines. CpG methylated DNA (mCpG) was generated by the M.SssI methyltransferase, validated with the methylation-sensitive enzyme HpaII (Fig. 3.3C). DNA with all cytosine methylated was generated by replacing all dCTP with dmCTP (Fig. 3.3C). Then unmethylated DNA (CpG), mCpG DNA, and mC DNA were labeled with DRAQ5 and introduced into the *in vitro* phase separation assay of MeCP2 (mixed with 1% GFP-MeCP2).

We found that all three kinds of DNA could be incorporated into the MeCP2 droplets (Fig. 3.4A). The GFP signals were used for droplet segmentation. Quantification of droplet size showed that both protein and DNA promoted the LLPS of MeCP2 in a concentration-dependent manner (Fig. 3.4B-F). Surprisingly, the addition of mCpG DNA and mC DNA to the mixture restricted the increase in droplet size in most conditions tested (Fig. 3.4B-F) compared to the unmethylated DNA. Moreover, MeCP2 formed far smaller droplets with mC DNA than with mCpG DNA (Fig. 3.4G), indicating a non-CpG methylation-dependent

influence on MeCP2 phase separation. Alternatively, cytosine methylation cumulatively restricts the MeCP2 droplet growth regardless of methylation context.

In summary, the cumulative weak DNA-MeCP2 interaction enhances the MeCP2 LLPS, while the stronger specific 5mC-MeCP2 restricts the growth of MeCP2 droplets.

To measure the binding affinities of MeCP2 to (methyl)-DNA, we applied the microscale thermophoresis (MST) technique. Using 20 and 42 bp template DNA with single mCpG methylation or no methylation (Fig. 3.4H), we found that MeCP2 had higher binding affinities for methylated and longer DNA. This is due to, on the one hand, specific 5mC-MeCP2 interaction and, on the other hand, cumulative multivalent nonspecific DNA-MeCP2 interaction. With the longer DNA, adding methylation did not affect the binding affinity. With the shorter DNA, adding methylation increased the binding affinity (Fig. 3.4I). We interpret this as, in the longer template, the multiple MeCP2-DNA interactions dominate the population K_D, whereas, in the shorter template with less possible MeCP2-DNA interactions, the 5mC-MeCP2 interaction becomes more prominent. Accordingly, MeCP2-R168X, an RTT-related nonsense mutation lacking all regions downstream of the MBD that are responsible for DNA interaction, showed an overall decreased binding affinity with DNA (Fig. 3.4I).

5.3.4 MeCP2 liquid-like droplet formation requires self-oligomerization

RTT-related MeCP2 nonsense mutations frequently occur in the intervening domain (ID) and transcriptional repression domain (TRD). Previously, we reported that the ID-TRD is responsible for the homo-dimerization of MeCP2 (Becker, Allmann et al. 2013). As LLPS can be induced by weak homo/hetero-interactions, we examined if the ID-TRD driven self-interaction is required for MeCP2 LLPS.

Firstly, as the MeCP2 RTT mutation R168X lacks the ID-TRD while the R255X retains partial ID-TRD and R294X retains most ID-TRD (Fig. 3.5A), we examined whether these nonsense mutations retain the ability of self-interaction by *in vitro* pull-down assay and MST assay. The *in vitro* pull-down experiment was done by incubating the immobilized MeCP2 truncations with purified GFP tagged MeCP2 truncations, followed by western blot analysis (Fig. 3.5B-C). The GFP-MeCP2-R255X, GFP-MeCP2-R294X, and the GFP-MeCP2 were pulled down by MeCP2-R255X, MeCP2-R294X, and MeCP2 separately, while the GFP-MeCP2-R168X was not pulled down by MeCP2-R168X (Fig. 3.5C).

We, then, quantified the self-interaction strength of full-length MeCP2, R255X, and R168X by MST assay. The GFP tagged MeCP2 and the two truncations were incubated with corresponding untagged proteins and the dissociation constants (Kd) were calculated (Fig. 3.5D). We found that MeCP2-R255X had a higher dissociation constant (Kd) (320.36 ± 14.35 nM) than the MeCP2 (121.66 ± 11.00 nM), while the shortest truncation MeCP2-

R168X showed a micromolar scale dissociation constant (1970 ± 19.91 nM). These data indicate that the shortest nonsense mutation MeCP2-R168X is not capable of self-interaction, the MeCP2-R255X is capable of self-interaction but with a lower binding affinity than the MeCP2. In summary, the amino acid sequence from position 168 to 254 is responsible and essential for MeCP2 self-interaction, which could be enhanced by the downstream C-terminal region.

Secondly, we checked the phase separation properties of MeCP2 R168X, R255X, R294X and MeCP2 full length (FL) compared to the MeCP2 using the turbidity assay and the phase separation droplet assay. We registered the absorbance changes at 340 nm of MeCP2 and truncations at low salt conditions (37.5 mM NaCl). The turbidity of MeCP2 increased with higher protein concentration, while there was no turbidity change for MeCP2 R168X, R255X, and R294X (Fig. 3.5E). Then, we analyzed the phase separation properties of the three MeCP2 truncations and found that R168X failed to form liquid-like droplets in all tested conditions, while the MeCP2 R255X and R294X could form liquid-like droplets at higher protein concentration (≥ 20 μM) or in the presence of crowding agents (Fig. 3.5F).

In conclusion, self-interaction is essential for MeCP2 LLPS. Both self-interaction and LLPS are enhanced by the MeCP2 C-terminus.

5.3.5 RTT-related MeCP2 nonsense mutations impact the phase separation properties of MeCP2

Previously, we already showed that MeCP2 RTT-related missense mutations disrupted its ability of chromatin organization (Agarwal, Becker et al. 2011, Bertulat, De Bonis et al. 2012, Zhang, Rausch et al. 2017). But little is known about the functions of nonsense mutations on chromatin organization. The above work confirmed that MeCP2 nonsense mutations impaired the function of LLPS. Thus, we examined whether MeCP2 nonsense mutations influence the phase separation properties of MeCP2 with methylated and unmethylated DNA. As the MeCP2 R168X showed no phase separation properties, the MeCP2 R255X and R294X were chosen, and the purified proteins were incubated with unmethylated DNA of different lengths (380bp, 800 bp, 1600 bp, 3000 bp) for *in vitro* phase separation. R255X showed weak phase separation properties (Fig. 3.6A-B). Similar to full length (FL) MeCP2, R294X showed overall DNA length- and concentration-dependent phase separation properties (Fig. 3.6A-B). Yet the ability of R294X phase separation increased with increased DNA length till 1600 bp and followed by a slight decrease with the longest 3000 bp DNA (Fig. 3.6A-B). Besides, the R294X showed weaker phase separation properties compared to the full-length MeCP2 (FL) (Fig. 3.6A-B). Thus, the nonsense mutations closer to the N-terminus showed much weaker phase separation properties. Considering the observation that the nonsense mutations closer to the N-terminus showed more severe symptoms, the

severity of RTT nonsense mutations was inversely correlated with phase separation properties.

As MeCP2 tracks the DNA methylation density *in vivo*. We checked the phase separation properties of MeCP2 RTT nonsense mutations with DNA of different methylation levels. The 3000 bp DNA with different methylation levels was generated by replacing partially dCTP with dmCTP for polymerase chain reaction (PCR). Purified proteins were incubated with DNA of different concentrations and methylation levels for *in vitro* phase separation assay. Still, the R255X showed rather weak phase separation properties regardless of DNA concentrations and methylation levels (Fig. 3.6C-D). Consistent with the phase separation of full-length MeCP2 with methylated 800 bp DNA, methylated 3000 bp DNA restricted the droplet size in a methylation level-dependent manner (Fig. 3.6C-D). Surprisingly, DNA methylation promoted the phase separation of R294X (Fig. 3.6C-D), indicating that the ability of MeCP2 full length and truncations are distinctly regulated by DNA methylation.

5.3.6 RTT-related MeCP2 nonsense mutations impact the heterochromatin compartment clustering with distinct dynamics

In vivo, the functions of MeCP2 nonsense mutations in genome-wide heterochromatin organization were examined. The C2C12 myoblast cells were transfected with plasmids expressing GFP tagged MeCP2 nonsense mutations. 36 hours after incubation, the cells were fixed and DAPI stained for microscope imaging. The heterochromatin compartments (DAPI dense regions) were segmented based on the DAPI intensity. The heterochromatin compartment size and protein fold enrichment were further measured (Fig. 3.7 A). Previous publications observed that the MeCP2 promoted the heterochromatin compartment fusion in a concentration-dependent manner. Thus, the cell populations were subdivided into three groups based on the protein levels (Fig. 3.7 B-C). The fold enrichment of R168X is similar to the DAPI enrichment, indicating that the enrichment of R1168X is mediated by the DNA compaction levels rather than DNA methylation levels (Fig. 3.7 B). Functionally, the R168X did not influence the heterochromatin compartment size significantly (Fig. 3.7 C). MeCP2 R255X, R294X, and full length showed increased enrichment at heterochromatin compartments in a concentration-dependent manner. Consistently, full-length MeCP2 promoted the formation of bigger heterochromatin compartments in a concentration-dependent manner ($P < 0.001$). Yet for R255X and R293, high expression levels promoted the formation of bigger droplets than full MeCP2 together with higher protein enrichment, while middle expression levels showed no influence on heterochromatin compartment size ($P > 0.05$) together with lower protein enrichment. Considering that protein enrichment or protein levels inside heterochromatin compartments might influence the heterochromatin compartment size. The mean heterochromatin compartment size was plotted versus mean

GFP intensity in heterochromatin compartments per nuclei and the linear trendlines were generated (Fig. 3.7 D). MeCP2 full length showed a concentration-dependent influence on heterochromatin compartment growth, which was impaired by the RTT-related nonsense mutations. The more the MeCP2 C-terminus was deleted, the weaker it promoted heterochromatin compartment growth. R294X showed a similar phase property and influence in heterochromatin compartments. This is consistent with the observation that CTD deletion does not affect MeCP2 function or recruitment in cells and that MeCP2 CTD truncation has no overt phenotypic consequences in this mouse model (Guy, Alexander-Howden et al. 2018).

Besides, we also checked the dynamics of MeCP2 nonsense mutations in heterochromatin compartments by FRAP assay in C2C12 cells with ectopic overexpression. A whole heterochromatin compartment was photobleached by a 100% 488 nm laser and the time-lapse images were taken using the Leica SP5 confocal microscope (Fig. 3.7 A). More than 50% of molecules were mobile and the fluorescence was half recovered within one minute, showing that they are highly dynamic (Fig. 3.7 E-G). MeCP2 R168X signals recovered rather fast with > 80% mobile fraction, suggesting that R168X has a weak interaction with chromatin although retains the MBD domain. For the truncations with longer C-terminus, the fluorescence recovery required a longer duration. Consistently, the ID and TRD of MeCP2 were shown to enhance the MBD-DNA binding (Ghosh, Nikitina et al. 2010). Surprisingly, similar to R255X, the full-length MeCP2 had intermediate mobility and immobile fraction. These observations demonstrate a functional connection between CTD and MeCP2 molecular dynamics. In summary, the ID and TRD inhibit MeCP2 dynamics by enhancing the MBD-DNA interaction, while the CTD promotes MeCP2 mobility.

The severities of MeCP2 nonsense mutations were: R168X > R255X > R294X > FL; the phase separation properties were: R168X (no) < R255X < R294X ≤ FL; the phase separation properties with methylated DNA were: R168X (no) < R255X < FL (and R294X); the function in heterochromatin compartment growth were: R168X (little) < R255X < R294X < FL; the protein dynamics were: R168X > R255X ≈ FL > R294X. These observations strongly correlated the severities of MeCP2 RTT mutations with LLPS-mediated heterochromatin organization and protein dynamics. Moreover, self-interaction mediated MeCP2 LLPS drives the heterochromatin compartment clustering.

5.3.7 AT-hook2 regulates the phase separation properties of MeCP2

MeCP2 contains three highly conserved AT-hook-like motifs (AT-hook1: aa 184–195; AT-hook 2: aa 264–273; AT-hook 3: aa 341–364) (Fig. 3.8A). All three AT-hooks bear the homology to the high-mobility group AT-hook family of DNA bending proteins (HMGA) and are responsible for binding with AT-rich DNA (Baker, Chen et al. 2013). AT-hook1 is located

within the ID region; AT-hook2 is located at the boundary of the ID and NID domain; AT-hook3 is located at the CTD. AT-hook2 is disrupted by the RTT-related nonsense mutation R270X (Fig. 3.8A), and its disruption caused more severe phenotypes in RTT patients and mouse models (compared to G273X, RTT mutation retaining the AT-hook 2). Here we examined if disrupting AT-hook 2 impacts MeCP2 phase separation and heterochromatin organization.

Firstly, the MeCP2 R270X, G273X, and MeCP2ΔAT-hook 2 (full-length MeCP2 with a deletion of AT-hook 2 sequence) were purified and applied for *in vitro* phase separation assay (Fig. 3.8B). In low salt concentration (37.5 mM NaCl), MeCP2 R270X and R273X only formed liquid-like droplets at high protein concentration. MeCP2-ΔAT-hook2 showed similar phase separation properties to full-length MeCP2, forming droplets in both low and high protein concentrations. At physiological salt concentration (150 mM NaCl), all three constructs formed liquid-like droplets in the presence of PEG 80008.

Next, the phase separation properties of all three constructs were examined in the presence of 3000 bp unmethylated DNA by *in vitro* phase separation assay and subsequent quantitative analysis (Fig. 3.8C-D). All three constructs showed a weaker phase separation ability compared to the full-length MeCP2, demonstrating the observation that the MeCP2 RTT-related nonsense mutations comprise the phase separation properties of MeCP2. Yet the R270X and G273X showed a similar phase separation property.

Given that the AT-hook binds with the AT-rich DNA sequence, the AT-rich major satellite DNA repeat (234 bp) (MajSat) was applied. The *in vitro* phase separation assay was done with majSat DNA in buffer containing 5% Dextran and 150 mM NaCl (Fig. 3.8E-F). R273X showed a stronger phase separation property than R270X. Consistently, MeCP2 full length formed bigger droplets than MeCP2ΔAT-hook2. These results demonstrated a role of AT-hook 2 in MeCP2 phase separation. This observation is also in agreement with the finding that MeCP2-R270X failed to facilitate NA oligomerization, while MeCP2-G273X facilitated NA oligomerization (Baker, Chen et al. 2013).

Then we checked the influence of AT-hook2 in heterochromatin organization based on the heterochromatin compartment size. C2C12 cells were transfected with plasmids expressing MeCP2 full length, MeCP2 R270X, G273X, and MeCP2ΔAT-hook 2 and incubated for 36 hours. Then the cells were fixed, DAPI stained, and observed under a confocal microscope Leica SP5. The heterochromatin compartments were segmented based on the DAPI intensity and heterochromatin compartment parameters were measured. Similar to the nonsense mutation R255X, MeCP2 R270X, G273X, and MeCP2ΔAT-hook 2 showed impaired functions on heterochromatin compartment size (Fig. 3.9A). Surprisingly, R270X and G273X showed a similar influence on heterochromatin compartment size. Thus, the

functions of MeCP2 R270X and G273X in heterochromatin compartment size are the same or not distinguishable by 2D imaging, higher resolution 3D images are required.

We then checked the molecular dynamics of the three constructs by FRAP in C2C12 cells 24 hours after transfection (Fig. 3.9B). MeCP2 R270X, G273X, and MeCP2ΔAT-hook 2 were dynamic, as the fluorescence recovered within 80 seconds for all these constructs (Fig. 3.9C). G273X exhibited a slower recovery compared to the R270X, together with a higher immobile fraction (Fig. 3.9C-D). Consistently, full-length MeCP2 required more time to recover the signals and had more immobile fractions compared to MeCP2ΔAT-hook 2 (Fig. 3.9C-D). This result is consistent with publications showing that MeCP2 AT-hook domains constrain the diffusion of unbound MeCP2 molecules by single-molecule tracking (Piccolo, Liu et al. 2019).

Taken together, the AT-hook2 regulates the MeCP2 LLPS via interactions with AT-rich DNA/chromatin with more confined mobility.

5.3.8 Molecule dynamics of MeCP2 in droplets and in cells

The molecule dynamics of various disordered proteins inside condensates *in vivo* and *in vitro* were commonly examined by FRAP analysis. Recently, single-molecule tracking was applied to dissect the protein dynamics under the single-molecule level *in cellulo*, yet the single-molecule protein mobility in droplets is still blank. Here, we further investigated whether and how LLPS impacts MeCP2 dynamics *in cellulo* and in the minimal LLPS system *in vitro* using a fluorescence-based high-resolution single-molecule tracking (SMT) approach.

In cellulo, the C2C12， MEF-P, and MEF-PM (DNA methylation deficient) cells were transfected with plasmids expressing GFP-MeCP2 and MaSat-mRFP. In vitro, 1% GFP-MeCP2 and 99% untagged MeCP2 were mixed for *in vitro* phase separation assay at different conditions. The mixtures were then transferred to chambers made of double-sided tapes and sealed with coverslips. Cells and droplets were observed under the Nikon Eclipse Ti microscope equipped with oil immersion Nikon 100x Apo TIRF (1.49 NA objective and highly inclined laminar optical sheet illumination (HILO) (Tokunaga, Imamoto et al. 2008) (Fig. 3.10A, C). Fluorescence was photobleached by a 488 nm laser. For single-molecule tracking, fluorescence bleaching was stopped when about 10-20 spots per time point were detected in TrackMate (Fig. 3.10A, C). Then, the track coordinates were imported to SMTracker to determine the number of populations, diffusion coefficients, and population weights (Fig. 3.10B, D).

In cellulo, as the MeCP2 bound fractions and diffusion of the unbound fractions are similar in euchromatin and heterochromatin regions, the MeCP2 movement within the whole nucleus was measured. Heterogeneous mobility with three dynamic populations with distinct diffusion

was detected (Fig. 3.10B). Overall, increasing MeCP2 concentration led to a general slow down of MeCP2 movement (Fig. 3.10B). Furthermore, MeCP2 R111G, an RTT-related mutation with no methylcytosine binding properties exhibited only the fast and slow population. Further, the influence of DNA methylation levels on MeCP2 dynamics was explored using MEF-P (high DNA methylation level) and MEF-PM (low DNA methylation level) with ectopic GFP-MeCP2 expression. Overall, MeCP2 mobility is decreased in cells with higher DNA methylation levels. Thus, the stable MeCP2-mC interactions promote the static population in cells.

In vitro, MeCP2 alone showed an isotropic diffusion fitted to a one-population model with a D of 0.530 $\mu m^2 s^{-1}$ (Fig. 3.10D) in droplets (fast population only). The addition of crowders led to an anisotropic MeCP2 movement fitted to a two-population model (Fig. 3.10D). From the 2 populations, one (fast population) diffuses similarly as in the absence of crowders (0.530 $\mu m^2 s^{-1}$). The second (slow) population was almost 10-fold more constrained with a D of 0.059 $\mu m^2 s^{-1}$ and occupied around 16% of MeCP2 molecules. Considering that crowding agents affect molecule mobility in a size-dependent manner and that MeCP2 molecules could cluster via multivalent weak self-interactions, the slow population occurs due to weak self-oligomerization. Further adding DNA increased the second slow population due to enhanced weak interactions (Fig. 3.10D). Surprisingly, DNA methylation promoted the emergence of the third static population due to strong MeCP2-mC interactions, confirming that the MeCP2-mC interactions contribute to the static population found *in cellulo*.

In conclusion, in the crowding environments *in vivo* and *in vitro*, stable MeCP2-mC interactions contribute to the static population, multivalent weak interactions (MeCP2-MeCP2/DNA) contribute to the properly confined slow population, and free MeCP2 molecules make up the fast population.

5.3.9 Discussion

We demonstrated that MeCP2 alone could form liquid-like spherical droplets via self-interaction-induced LLPS (Fig. 3.11A). The self-interaction is mediated by the amino acid sequence from position 168 to 254 and enhanced by the downstream sequence. The downstream sequence after the self-interaction region increased the strength and enhanced the phase separation properties. The strength of self-interaction positively correlates with LLPS properties (Fig. 3.11B). The fold enrichment of MeCP2 in the droplets is similar to that in the heterochromatin compartments (Agarwal, Hardt et al. 2007, Müller‐Ott, Erdel et al. 2014). MeCP2 LLPS could be promoted by various factors that somehow mimic the *in vivo* physiological conditions (Fig. 3.11B). DNA promotes the LLPS of MeCP2 in both DNA length and concentration-dependent manner due to increased sites for multivalent weak interactions. The kind of weak interactions was characterized by the emergence of a

properly confined diffusion population in the droplets and heterochromatin compartments. Cytosine methylation restricts the size of MeCP2 condensates *in vivo* and *in vitro* with more confined overall mobility (Fig. 3.11B). The specific MeCP2-mC interaction contributes to the emergency of the static population with low to static diffusion in MeCP2 condensates *in vivo* and *in vitro*.

Self-interaction-induced MeCP2 phase separation is mainly driven by electrostatic interaction.

LLPS is mainly driven by electrostatic and/or hydrophobic interactions. To declare how the MeCP2 droplets are driven, we examined the phase separation properties of MeCP2 at different salt conditions and in the absence/presence of 1,6-hexanediol. We found that increasing the electrostatic interactions at lower salt concentration enhances the phase separation properties of MeCP2 while disrupting the hydrophobic interactions by 1,6-hexanediol does not dissolve the MeCP2 droplets easily. These indicate that the MeCP2 phase separation is probably mainly driven by electrostatic interactions.

We further explored which region of MeCP2 is responsible for the self-interaction and found that the MeCP2 R168X, which shows no self-interaction, failed to form condensates. MeCP2 R255X, which shows intermediate self-interaction, succeeded in forming liquid-like spherical droplets. This is consistent with *in vivo* experiments, which found that the R168X (or MBD) is less enriched in the heterochromatin with higher mobility (Kumar, Kamboj et al. 2008, Schmiedeberg, Skene et al. 2009) and causes the most severe RTT phenotypes in patients (Schanen, Houwink et al. 2004, Neul, Fang et al. 2008, Schmidt, Zhang et al. 2020). Thus, self-interaction is a prerequisite for MeCP2 phase separation (Fig. 3.11A). Stronger self-interaction promotes the formation of bigger droplets with a more restricted but homogeneous diffusion population (Fig. 3.9A-C).

Crowdedness and DNA promote MeCP2 LLPS with distinct effects on MeCP2 mobility

In the real biological system, high densities of macromolecules (protein, nucleic acids, etc.) not only perform function-related activities but also show an influence of crowding via nonspecific “excluded volume” effect (Richter, Nessling et al. 2008). The macromolecular crowding could increase the local concentration of targets via decreasing the available space and thus promotes protein stabilization (Homouz, Perham et al. 2008), influences biochemical reaction (binding, enzymatic activities) (Zhou, Rivas et al. 2008) and restricts molecular diffusion (Dix and Verkman 2008). The eukaryotic cell is a heterogeneous system with various crowding states. 50-400mg/ml macromolecules in the cell make it 5-40% crowded (Biswas, Kundu et al. 2018), 100 mg/ml protein and 50 mg/ml chromatin in the nuclei makes it around 15% crowded (Hancock , Handwerger, Cordero et al. 2005, Richter,

Nessling et al. 2008), 110mg/ml macromolecules in the nucleoplasm make the nucleoplasm around 11% crowded (Rippe 2007). Depending on condensation levels, chromatin varies from euchromatin to facultative heterochromatin and finally to constitutive heterochromatin with higher compaction levels, lower mobility, and increased crowding environments.

Dextran T150 (~150 kDa) and PEG 8000 (~8 kDa) were introduced to the *in vitro* phase separation system to declare how the crowding environment influences the MeCP2 activities. We found that both Dextran T150 and PEG 8000 could promote the MeCP2 LLPS in a dose-dependent manner by increasing the MeCP2 local concentration and/or (2) enhancing the self-interaction (Fig. 3.11B). The first is confirmed by the finding that in-droplet MeCP2 concentration increased at higher crowding environments. The second is confirmed by the overall slower diffusion and the segmentation of the second confined population, which was not detectable in droplets formed in low salt conditions with no crowders (Fig. 3.11B). *In vivo*, MeCP2 showed an overall faster diffusion in the cells with lower MeCP2 levels (similar level in mature neurons) (Fig. 3.11B). Meanwhile, it has been reported that lower temperature promotes the LLPS of multiple proteins, such as LAF-1, a key component of P granules (Schuster, Dignon et al. 2020). These indicate that enhancing multivalent weak interactions could promote the LLPS, which could be achieved by increasing protein concentration or buffer crowdedness. Besides, Dextran T150 and PEG 8000 showed different potentials on MeCP2 LLPS due to the crossing effect, which is affected by physical parameters such as size and shape (Richter, Nessling et al. 2008).

DNA is commonly considered an important factor for LLPS of multiple proteins, especially nucleic proteins involved in DNA replication, damage repair, and chromatin architecture (Laflamme and Mekhail 2020). Both MeCP2-MeCP2 and MeCP2-DNA interactions occurred in droplets containing both protein and DNA. Quantification of MeCP2 binding affinity with itself (121.66 ± 11.00 nM) and with DNA (138.73 ± 22.10 nM) exhibited a similar binding affinity. However, DNA increased the binding frequency by providing more binding sites, leading to overall more restricted diffusions and increased second population (Fig. 3.11B). Besides, The influence of crowding environments on molecules is highly molecule size-dependent (Minton 2001), the bigger molecules are much more strongly affected than smaller molecules. Thus, the free molecules and the molecules undergoing oligomerization are distinguished as two separate populations. As DNA is a much larger molecule than MeCP2, crowding agents may directly and more pronouncedly affect its ability to move, observed as slowed MeCP2 diffusion. Alternatively, the DNA may also contribute to the crowding effect.

Thus, DNA promotes MeCP2 LLPS via providing additional sites for multivalent weak interaction between MeCP2 and DNA. Crowding agents enhance the frequency of multivalent weak interactions by increasing local protein concentrations. Besides, crowding

agents promote population segmentation due to a size-dependent influence of crowding agents on molecule dynamics.

Cytosine methylation negatively regulates droplet size and reduce MeCP2 mobility

Originally, MeCP2 was found to specifically recognize and bind methyl-CpG (mCpG) via the MBD domain (Lewis, Meehan et al. 1992), which could be enhanced by both flanking regions. More recently, researchers found that non-CpG methylation contributes around 25% of all cytosine methylation (5mC) in neurons, which is recognized as an essential target of MeCP2 (Chen, Gu et al. 2011, Guo, Ma et al. 2011, Chen, Chen et al. 2015, Jang, Shin et al. 2017, Tillotson, Cholewa-Waclaw et al. 2020). Thus, we checked how DNA methylation influences MeCP2 phase separation properties under different contexts (mCpG and mC). Both mCpG and mC restrict the droplet size in most (> 50%) conditions tested. Contrary to publications (Fan, Zhang et al. 2020, Wang, Hu et al. 2020), who found that DNA with mCpG promotes the growth of MeCP2 droplets in most conditions tested. We thus examined the influence of cytosine methylation in MeCP2 mediated heterochromatin compartment clustering using Dnmt1 knockout MEF fibroblast cells. Recovered cytosine methylation via Dnmt1 overexpression decreased the heterochromatin compartment size, while MeCP2 overexpression increased the heterochromatin compartment size, which is in parallel with the *in vitro* phase separation assay (data not shown). This is also in agreement with the fact that highly condensed heterochromatin is hallmarked with high cytosine methylation and inactivity (Bhaumik, Smith et al. 2007, Wen, Li et al. 2014, Saksouk, Simboeck et al. 2015, Penagos-Puig and Furlan-Magaril 2020). The apparent inconsistency to our results can be explained by the fact that while we base this conclusion on measuring individual droplet compartment size to directly relate to the size of heterochromatin compartments within the cell nucleus, the other authors measured the sum of all the droplets. The latter measures whether DNA methylation promotes LLPS altogether and not whether the individual compartments can grow in size to the same extent. By comparing to the equivalent *in vivo* situation, e.g., using cells with lower DNA methylation and RTT mutants of MeCP2 deficient on 5mC binding, we found our measurement and effect to be a good representation of heterochromatin compartmentalization dynamics in cells.

In vivo, MeCP2 exhibited three diffusion populations with heterogeneous distribution across heterochromatic regions. In Dnmt1 deficient cells, the second confined and the third static population decreased, leading to an increased fast population. R111G, an RTT-related mutation within MBD and devoid of specific MeCP2-mC interaction, showed overall faster diffusion, and only two diffusion populations were identified, losing the static population. These indicate that the specific and strong MeCP2-mC interaction is responsible for the fractionation of the third static population. In droplets, MeCP2 molecules exhibit two diffusion

populations with unmethylated DNA and three populations with methylated DNA in the presence of crowders, confirming that the third population is induced by the specific and strong interaction between MBD and methylated cytosine.

RTT-related nonsense mutations comprised the MeCP2-induced heterochromatin organization

MeCP2 was shown to be associated with the neurological disorder Rett syndrome (RTT), as mutations in this gene were found in about 80% of Rett patients (Amir, Van den Veyver et al. 1999). Mutations causing RTT and related neurological disorders have been identified along the entire MeCP2 locus, but effects vary depending on the mutation type and location. Previously, we already elucidated that RTT-related MeCP2 missense mutations within the MBD domain showed impaired function in the large-scale chromatin organization (Agarwal, Becker et al. 2011). Yet, little is known about the influence of nonsense mutations on MeCP2 function in the cell nucleus. Here, we checked the influence of MeCP2 nonsense mutations on heterochromatin organization and protein LLPS.

The severities of MeCP2 nonsense mutations in RTT patients and mice models were: R168X > R255X > R270X > G273X > R294X; the phase separation properties were: R168X < R255X < R270X < R273X < R294X < FL; the function in heterochromatin compartment growth were: R168X < R270X ≈ G273X < **R255X** < R294X < FL. These observations strongly correlated the severities of MeCP2 RTT mutations with LLPS-mediated heterochromatin organization. The protein dynamics were: R168X > FL ≈ R255X ≈ R270X > R273X ≈ R294X. Considering that the amino acid sequence from 168 to 254 contains the AT-hook 1 and is responsible for self-interaction, the R255X showed slower mobility due to multivalent weak interactions. Besides, the AT-hook 2, which is retained by G273X but absent from R270X. The AT-hook2 mediated MeCP2-DNA interaction regulates the MeCP2 mobility. Rather, full-length MeCP2 moved faster than the R294X, a nonsense mutation retaining the majority TRD (NID) domain. These Indicate that the CTD could regulate the dynamics of MeCP2.

In summary, we establish *in vitro* and *in vivo* the parameters determining the LLPS properties of MeCP2 and how MeCP2 diffusion and binding properties change when inside the droplets as well as in cells and how these contribute to heterochromatin compartment organization and kinetics.

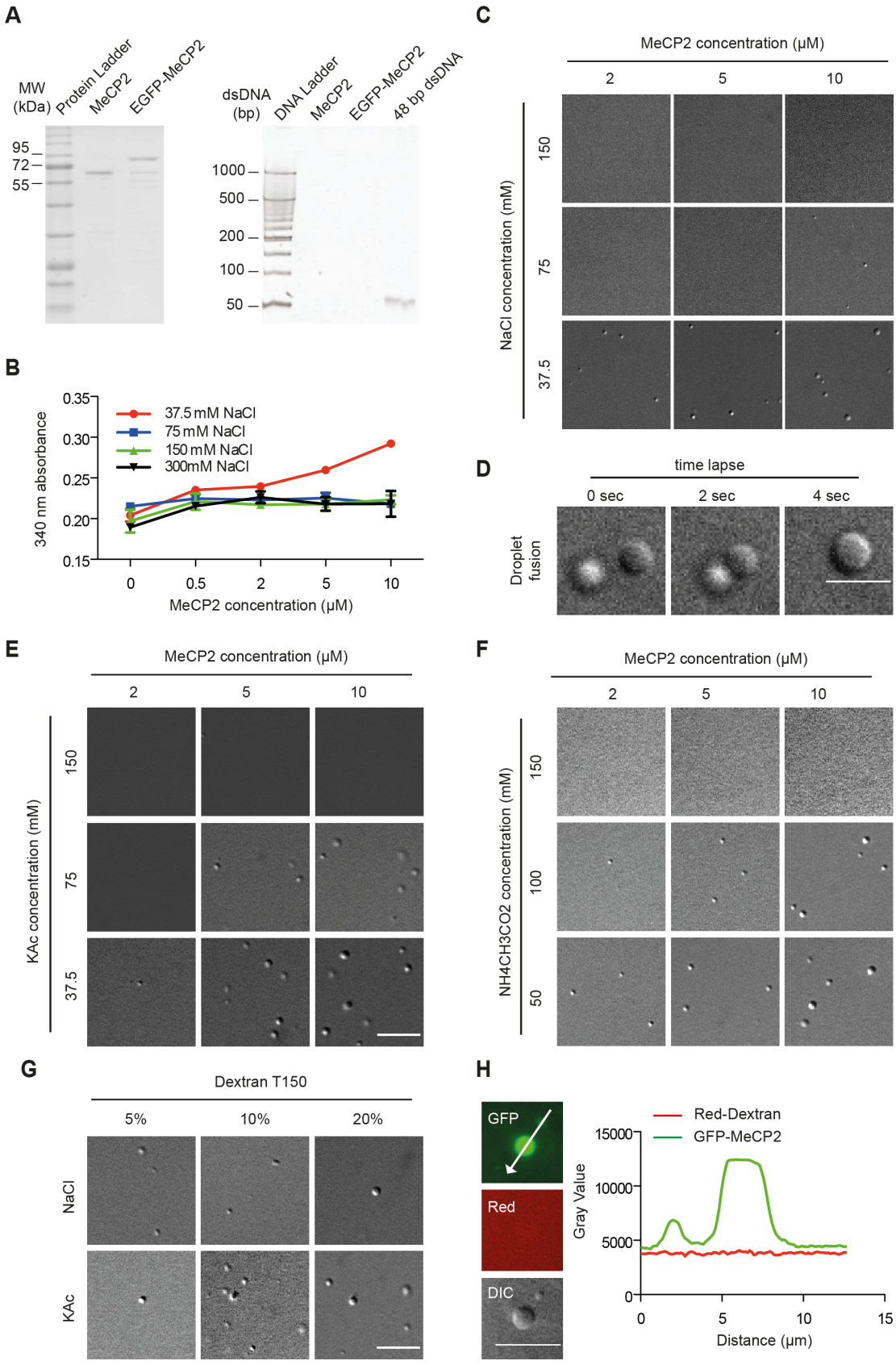
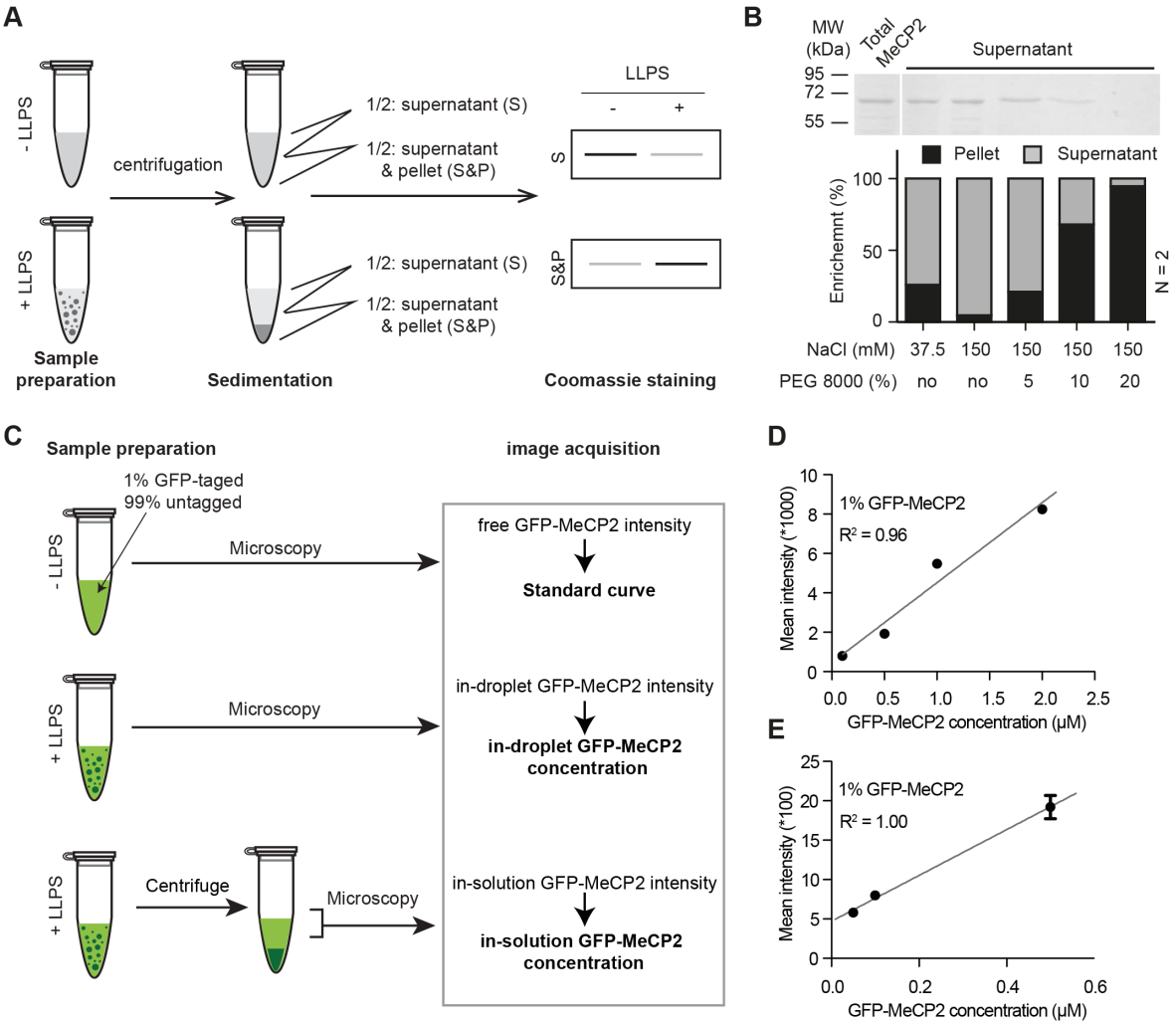


Figure 3.1 Purified MeCP2 forms liquid-like droplets in physiologically crowding environments.

The *in vitro* phase separation assay was done at different conditions by incubating at room temperature for 45 min. The mixtures after phase separation were then transferred to chambers made of double-sided tapes and sealed with coverslips. The droplets were observed using a Nikon Eclipse TiE2 microscope equipped with

differential interference contrast microscopy. Alternatively, the mixtures were transferred to a 384-well plate and incubated for 45 min at room temperature. The absorbance was measured at λ = 340 nm at room temperature to check the turbidity of MeCP2 solutions.
(A) Validation of MeCP2 purity. The MeCP2 and GFP-MeCP2 proteins were expressed in bacteria by IPTG induction, purified using chitin beads, and eluted by DTT. The final protein concentrations were measured by Pierce™ 660nm Protein Assay Reagent. 2 μg and 10 μg purified proteins were then used for SDS polyacrylamide gel electrophoresis and tris borate EDTA polyacrylamide gel electrophoresis, respectively. Left: SDS polyacrylamide gel electrophoresis of purified human MeCP2 and GFP-MeCP2 followed by Coomassie staining. 2μg each lane. Right: tris borate EDTA polyacrylamide gel electrophoresis of purified human MeCP2 and GFP-MeCP2 followed by ethidium bromide (EtBr) staining. 10 μg each lane. ~140 ng 42 bp DNA was used as a positive control.
(B) Turbidity assay showing a protein and salt concentration-dependent phase behavior of MeCP2. Absorbance was plotted as mean ± SD.
(C-G) DIC images of MeCP2 droplets under various protein and salt conditions. Scale bars = 10 μm. Time-lapse DIC images showing a fusion event of MeCP2 in a buffer containing 37.5 mM NaCl and 10 μM MeCP2 (D) (Movie S1). Scale bar = 10 μm.
(H) Fluorescence images of droplets formed by human GFP-MeCP2 in the presence of Texas Red labeled dextran. The phase separation was conducted in a buffer containing 10 μM MeCP2, 150 mM NaCl, 18% Dextran T150 and 0.1% 10 kDa Texas Red labeled dextran. Fluorescent images (GFP and Red) and DIC images were taken using the Nikon Eclipse TiE2 microscope. The plot profile of GFP and Red intensity was conducted to analyze the distribution of dextran. Scale bar = 10 μm.



J

NaCl (mM)	37.5	150	150	150
PEG 8000 (%)	0	5	10	20
Total MeCP2 concentration/ µM	10	10	10	10
In-droplet MeCP2 concentration/ µM	122	43	114	87

Figure 3.2 Distribution of MeCP2 in droplets and solution

The *in vitro* phase separation assay was done at different conditions by incubating at room temperature for 45 min in the presence or absence of GFP-MeCP2 and followed by droplet sedimentation or microscope detection.
(A) Schematic graph showing the droplet sedimentation assay and quantification. In brief, MeCP2 solutions with or without droplet formation were centrifuged at 14000 rpm for 15 min to pellet (if any) the droplets. The clear supernatants were transferred to new tubes with 1X loading buffer (to avoid destroying the pellet, the top half was collected) followed by 95 °C for 5 min. The supernatants were then loaded to SDS-PAGE gel and the gel was stained by coomassie for quantification.
(B) Quantification of MeCP2 distribution in solution (supernatant) and droplets (pellets). Top: Coomassie staining result of the supernatants after SDS-PAGE electrophoresis. Bottom: Quantitative analysis for SDS-PAGE gel above. Replicates (N) = 2.
(C) Schematic drawing showing how to quantify the protein concentration in the droplets and solution. Top: the fluorescent intensities of the gradient concentrations of GFP-MeCP2 were measured under the microscope, and were plotted versus the corresponding protein concentrations to generate the standard curve. Middle and bottom: in-droplet and in-solution protein concentration measurement. In brief, 1% GFP-MeCP2 was mixed with 99% untagged MeCP2 and the mixtures were incubated at RT for 45 min for LLPS. The droplet-free solutions were collected by centrifugation as mentioned in A and the in-solution fluorescent intensities were measured by microscopy. The in-solution protein concentration was measured by the in-solution GFP-MeCP2 fluorescent intensity relative to the standard curve. To measure the protein concentration in the droplets, the droplet-containing solutions (no centrifugation) were directly used for microscopy. Then, the droplets were segmented and the in-droplet fluorescent intensities were measured using ImageJ. The in-droplet protein concentrations were measured by the in-droplet GFP-MeCP2 fluorescence intensity relative to the standard curve.
(D-E) Standard curve of fluorescence intensity versus low (E) and high (F) protein concentration.
(F) Fluorescence images and DIC images of MeCP2 phase-separated droplets. Both fluorescent images (GFP) and DIC images were taken under the Nikon Eclipse TiE2 microscope equipped with differential interference contrast (DIC) microscopy. Total MeCP2 concentration: 10 μM. Scale bar = 10 μm.
(G) Schematic graph showing the MeCP2 concentration in solution and droplets formed by 10 μM MeCP2 with 37.5 mM NaCl.
(H-I) In-droplet protein concentration (H) and droplet aspect ratio (I) for experiments in (F).
The GFP channel was applied for droplet segmentation by a bandpass filter and threshold based on the mean intensity in/out droplets. Droplets with size > 0.1 μm2 were identified and droplet parameters were measured. The aspect ratio (G) was calculated as the ratio of maximal Feret diameter to minimal Feret diameter. The in-droplet MeCP2 concentration (H) was measured by the GFP intensity inside the droplets. Replicates (N) = 3.
(J) Mean values of aspect ratio and in-droplet MeCP2 concentration from (G-H). n: number of droplets.

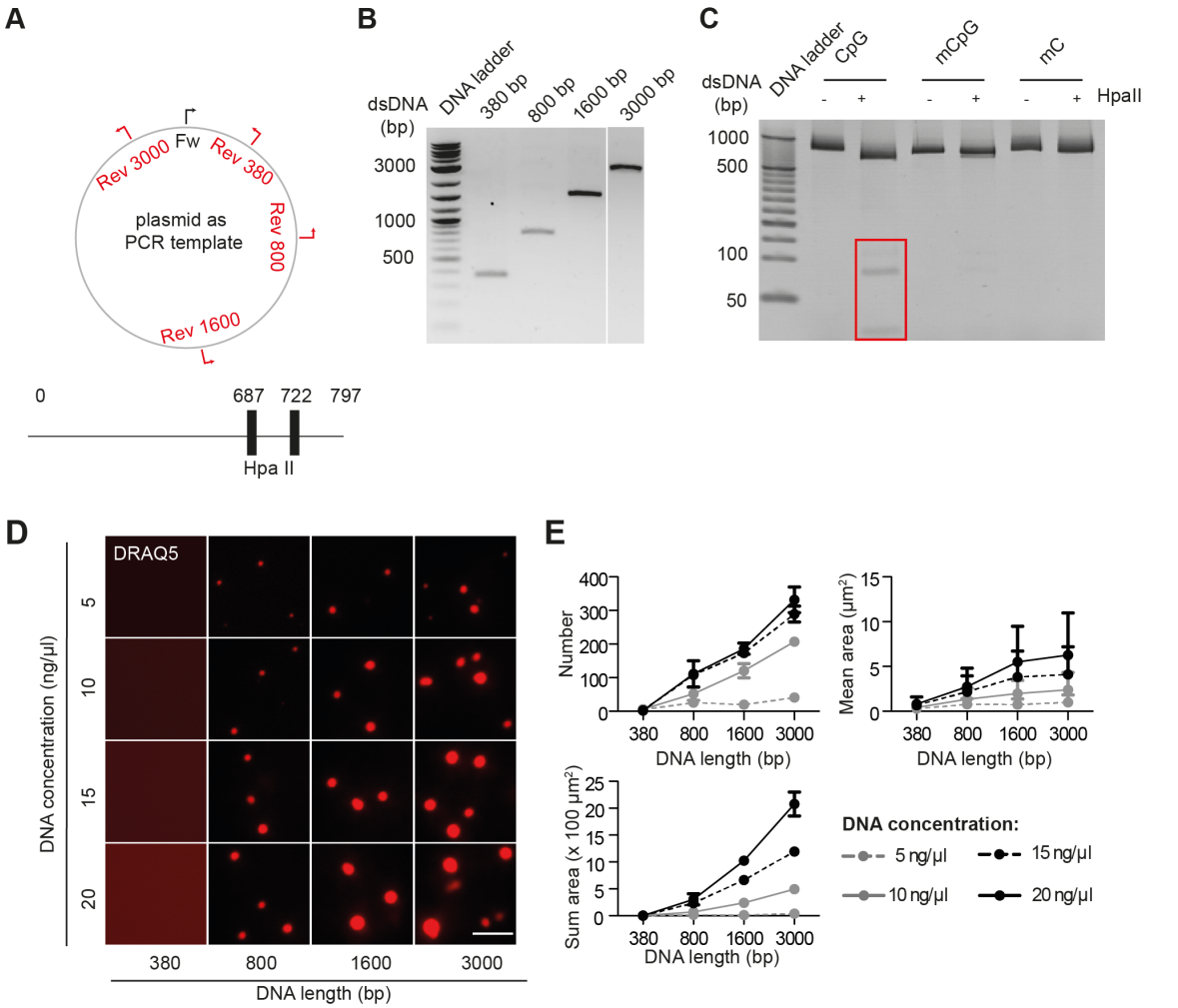


Figure 3.3 DNA promotes the liquid-liquid phase separation of MeCP2.

(A) Top: Schematic graph showing that different reverse primers were applied to generate PCR products with different lengths using plasmid DNA as PCR template. Fw: forward primer. Rev: reverse primer. Bottom: Schematic graph showing the positions of two HpaII restriction sites on the 800 bp DNA template.
(B) Agarose gel electrophoresis after PCR. PCR products with various lengths as described in (A) were analyzed by agarose gel electrophoresis followed by ethidium bromide staining and imaging.
(C) Unmethylated (CpG), CpG methylated (mCpG), and all cytosine methylated (mC) DNA templates were generated by PCR. The CpG methylated (mCpG) was generated by incubating the purified 800 bp DNA with the CpG methyltransferase (M.SssI). All cytosine methylated DNA was synthesized by replacing the dCTP with an equal amount of dmCTP. Different DNA templates were incubated with HpaII at 37 °C for 2 h, followed by 15% native tris borate EDTA polyacrylamide gel electrophoresis, ethidium bromide staining, and imaging. In the non-methylated templates, the HpaII enzyme can cleave the DNA generating additional 35 bp and 75 bp DNA fragments (red box).
(D) Fluorescent images of MeCP2 droplets in the presence of DRAQ5 labeled DNA with different concentrations and lengths. The synthesized DNA was labeled with DRAQ5. The *in vitro* phase separation assay at different conditions was done by incubation at room temperature for 45 min. The fluorescent and DIC images were taken using the Nikon Eclipse TiE2 microscope. MeCP2: 3 µM, NaCl: 150 mM, no PEG. Scale bar = 10 µm.
(E) Quantification of size, area, and the number of droplets from (A). The red channel was applied for droplet segmentation by a bandpass filter and threshold based on the mean intensity in/out droplets. Droplets with size > 0.1 µm2 were considered and droplet parameters were measured. ≥ 3 images were taken for each condition. The droplets number and sum droplet area per image and mean droplet area were plotted with mean ± SD (standard deviation).

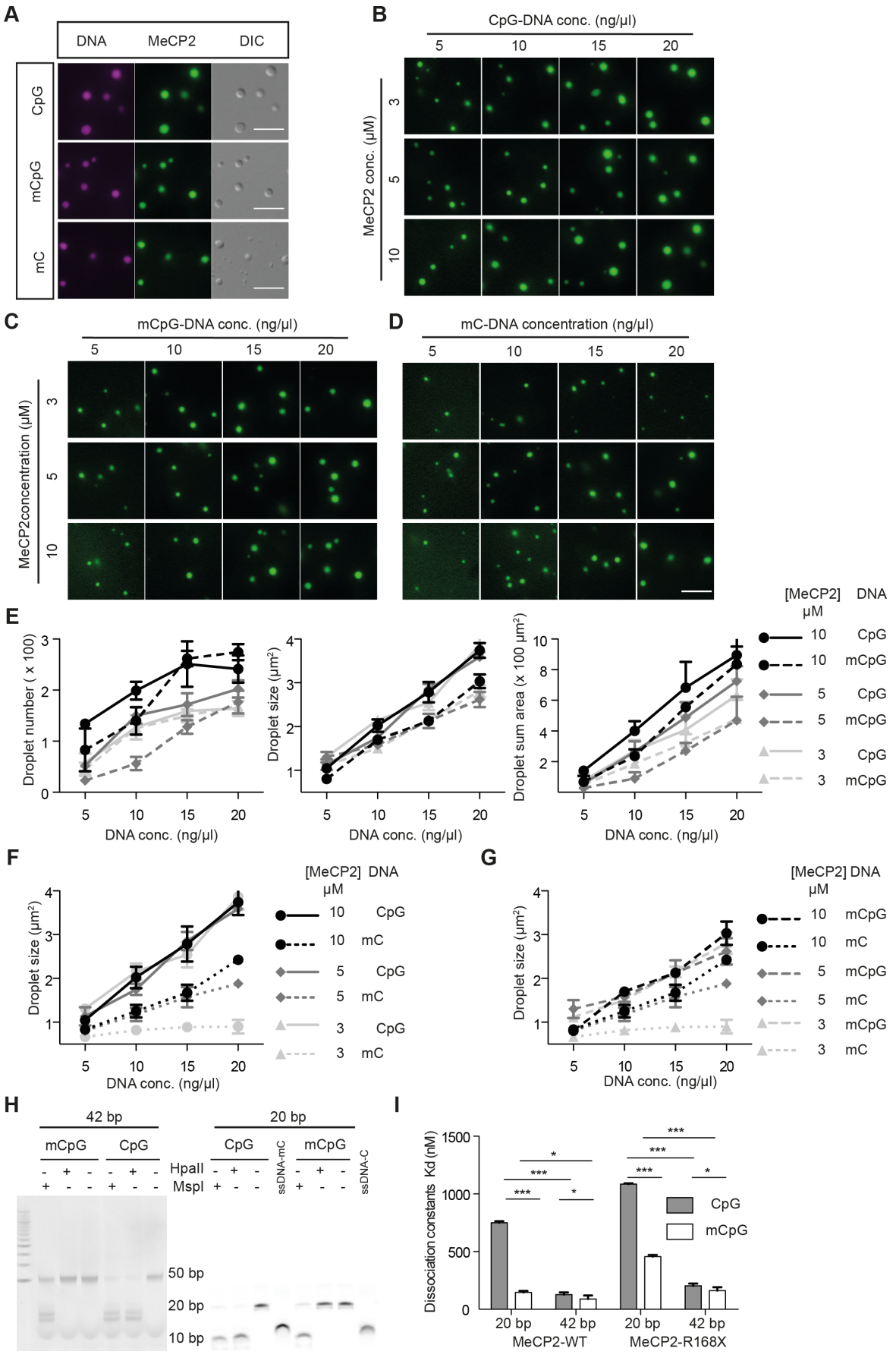

A
DNA
MeCP2
DIC
CpG
mCpG
mC
B
CpG-DNA conc. (ng/µl)
5
10
15
20
MeCP2 conc. (µM)
3
5
10
C
mCpG-DNA conc. (ng/µl)
MeCP2concentration (µM)
D
mC-DNA concentration (ng/µl)
E
Droplet number (x 100)
Droplet size (µm²)
Droplet sum area (x 100 µm²)
DNA conc. (ng/µl)
[MeCP2] µM
DNA
10 CpG
10 mCpG
5 CpG
5 mCpG
3 CpG
3 mCpG
F
10 CpG
10 mC
5 CpG
5 mC
3 CpG
3 mC
G
10 mCpG
10 mC
5 mCpG
5 mC
3 mCpG
3 mC
H
42 bp
20 bp
mCpG
CpG
HpaII
MspI
ssDNA-mC
ssDNA-C
50 bp
20 bp
10 bp
I
Dissociation constants Kd (nM)
CpG
mCpG
20 bp
42 bp
MeCP2-WT
MeCP2-R168X

Figure 3.4 DNA methylation restricts the droplet formation of MeCP2.

(A) Fluorescent images of MeCP2 droplets formed in the presence of 800 bp DNA with different methylation patterns. The synthesized 800 bp DNA was labeled with DRAQ5. 1% GFP-MeCP2 was mixed with 99% untagged MeCP2. The *in vitro* phase separation assay was done at different conditions by incubating at room temperature for 45 min. Then, the mixtures were transferred to chambers made of double-sided tapes and sealed with coverslips. The droplets were observed using a Nikon Eclipse TiE2 microscope based on the fluorescent and DIC images. MeCP2: 10 μM, NaCl: 150 mM, DNA: 20 ng/μl, no PEG. Scale bars = 10 μm.
(B-D) Fluorescent images of droplets by MeCP2 in the presence of DNA with no methylation (CpG) (B), CpG methylation (mCpG) (C), and non-canonical all cytosine methylation (mC) (D). MeCP2 and DNA concentrations are as mentioned. NaCl: 150 mM, no PEG. Representative images of the GFP channel. Scale bar = 10 μm.
(E) Quantitative measurements showing the influence of CpG DNA and mCpG DNA on the LLPS properties of MeCP2 based on droplet number, droplet size (mean droplet area), and sum droplet area from (B-C). The droplets were segmented by a bandpass filter and threshold based on the mean GFP intensities in/out droplets. Droplets with size > 0.1 μm^2 were considered and droplet parameters were measured.
(F) Quantitative measurements showing the influence of CpG DNA and non-canonical mC DNA on the LLPS properties of MeCP2 based on droplet size from (B) and (D).
(G) Quantitative measurements showing the influence of mCpG DNA and non-canonical mC DNA on the LLPS properties of MeCP2 based on droplet size from (C-D).
(H) DNA methylation detection. 1 μl of 10 μM 42 bp and 20 bp dsDNA with and without CpG methylation was treated with HpaII and MspI respectively for 2 h ar 37 °C and then applied to tris borate EDTA polyacrylamide gel electrophoresis. The gel was stained with ethidium bromide and imaged.
(I) Left: MST analysis of the dissociation constants of MeCP2 and RTT related truncation MeCP2-R168X with methylated or unmethylated DNA synthesized in (H). Student t-test was applied, ns (not significant), $P > 0.05$; *, $P < 0.05$; **, $P < 0.001$; ***, $P < 0.0001$. Right: Table showing the mean values of the dissociation constant (Kd).

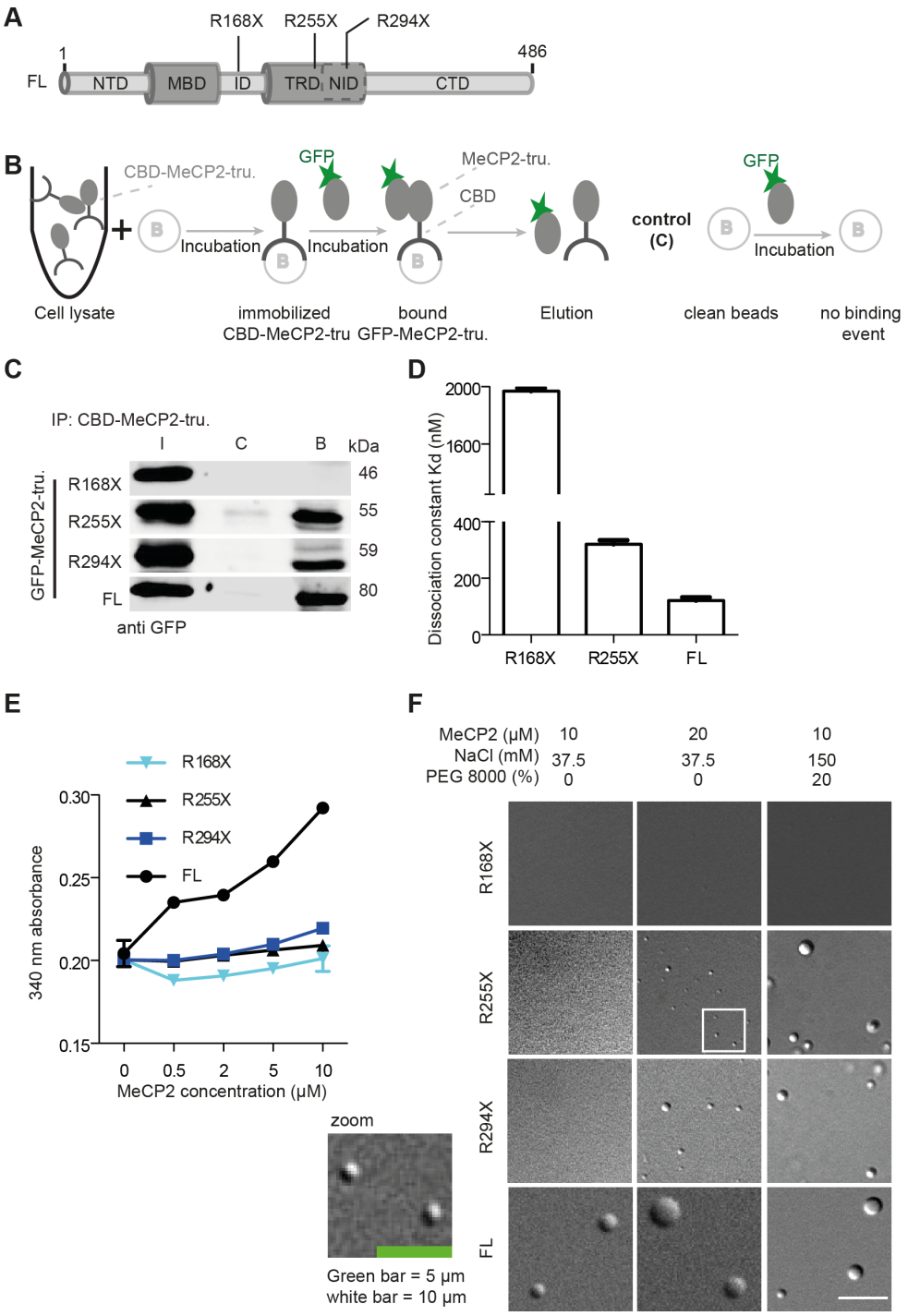


Figure 3.5 Oligomerization of MeCP2 is a prerequisite for liquid-liquid phase separation *in vitro*

(A) Schematic graph showing three representative nonsense mutations in the MeCP2 gene in patients with Rett syndrome.
(B) Schematic graph showing the *in vitro* pull-down experiments. Shortly, the GFP tagged proteins were purified and stored in a buffer containing 0.5 M NaCl using chitin beads (A-D). The untagged MeCP2-CBD and MeCP2-truncations (MeCP2-tru.)-CBD were purified using immobilized chitin beads by incubating the bacterial lysates with the beads for 3 h at 4 °C with rotation. The pull-down assay was conducted by incubating purified GFP-MeCP2 or truncations in solution with immobilized MeCP2 or truncations in PBS plus 150 mM NaCl and 0.05% NP-40. The bound fractions were released by adding 40 µl loading buffer and boiling, followed by western blot for further analysis.
(C) Western blot analysis of the *in vitro* pull-down assays as mentioned in (B).
(D) Dissociation constant (Kd) values measured by microscale thermophoresis assay (MST).

(E) Turbidity changes of the full-length MeCP2 and RTT-related nonsense mutations with increasing protein concentrations. The phase separation was done by incubating proteins of various concentrations in a buffer containing 37.5 mM NaCl at room temperature for 45 min in a 384-well plate. The absorbance was measured at λ = 340 nm at room temperature. Absorbance was plotted as mean ± SD.
(F) DIC images showing the phase separation properties of MeCP2 full length and truncations at different conditions. The *in vitro* phase separation assay at the conditions indicated was done by incubating at room temperature for 45 min. The DIC images were taken using the Nikon Eclipse TiE2 microscope. The boxed region was shown at higher magnification. Scale bar (green) = 5 μm; Scale bar (white) = 10 μm.

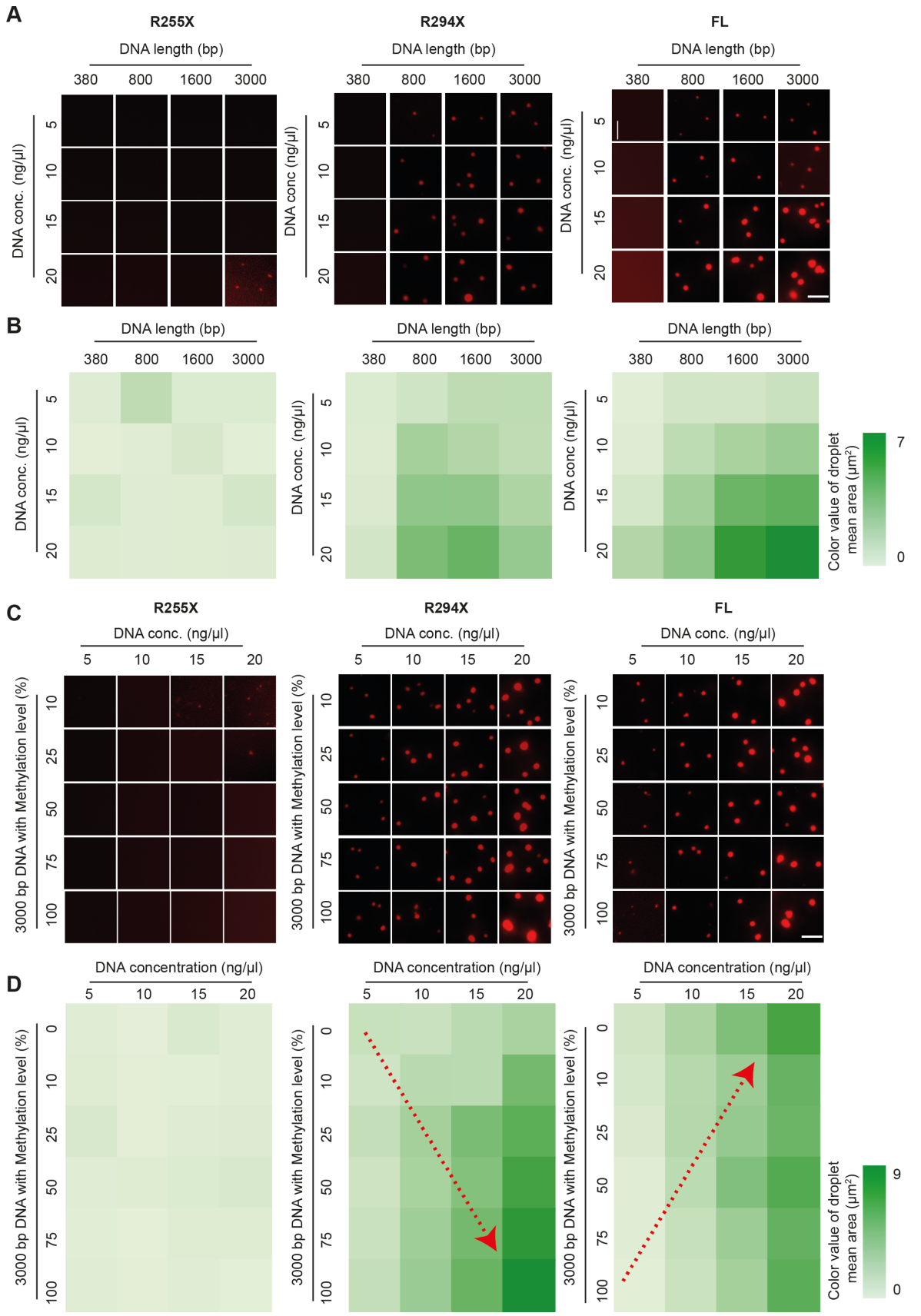


Figure 3.6 RTT-related nonsense mutations affect MeCP2 LLPS with unmethylated and methylated DNA.

The synthesized DNA was labeled with DRAQ5. The *in vitro* phase separation assay at different conditions was done by incubation at room temperature for 45 min. Then, the mixtures were transferred to chambers made of

double-sided tapes and sealed with coverslips. The fluorescent and DIC images were taken using the Nikon Eclipse TiE2 microscope. MeCP2: 3 μM, NaCl: 150 mM, no PEG.
(A) Fluorescent images of MeCP2 droplets in the presence of DRAQ5 labeled DNA with different concentrations and lengths. MeCP2 and DNA concentrations (conc.) are as mentioned. Representative images of the QRAQ5 (Cy5) channel are shown. Scale bar = 10 μm.
(B) Heatmap graphs showing the influence of mCpG on the LLPS properties of MeCP2 based on droplet size from (A). Droplets were segmented by a bandpass filter and thresholded based on the mean fluorescence intensities in/out droplets. Droplets with size > 0.1 μm2 were considered. The droplet parameters were measured and plotted. conc.: concentration.
(C) Fluorescent images showing the phase properties of MeCP2 nonsense mutations in the presence of methylated DNA. 3000 bp DNA with various methylation levels was synthesized by replacing partially the dCTP with dmCTP during polymerase chain reaction (PCR). DNA methylation levels and concentrations (conc.) are as mentioned. Representative images of the QRAQ5 (Cy5) channel are shown. Scale bar = 10 μm.
(D) Heatmap graphs showing the influence of mCpG on the LLPS properties of MeCP2 based on droplet size from (C). Droplets were segmented by a bandpass filter and thresholded based on the mean fluorescence intensities in/out droplets. Droplets with size > 0.1 μm2 were considered and droplet parameters were measured. conc.: concentration. Arrows indicate increased droplet size.

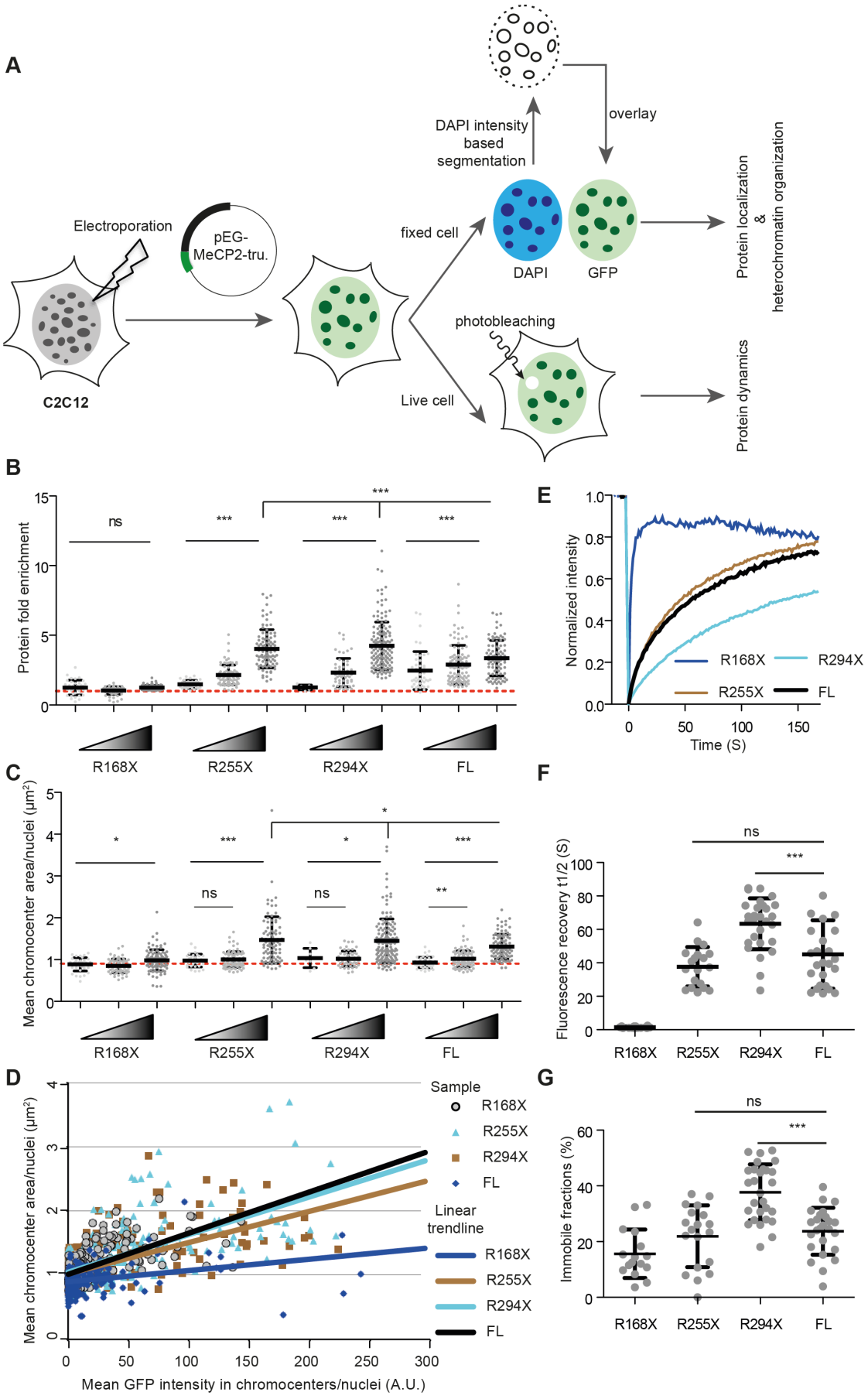

A
Electroporation
pEG-MeCP2-tru.
C2C12
fixed cell
Live cell
DAPI intensity based segmentation
overlay
DAPI
GFP
photobleaching
Protein localization & heterochromatin organization
Protein dynamics
B
Protein fold enrichment
ns
***
R168X
R255X
R294X
FL
C
Mean chromocenter area/nuclei (µm²)
*
**
D
Sample
R168X
R255X
R294X
FL
Linear trendline
Mean GFP intensity in chromocenters/nuclei (A.U.)
E
Normalized intensity
Time (S)
F
Fluorescence recovery t1/2 (S)
G
Immobile fractions (%)

Figure 3.7 The influence of RTT-related nonsense mutations on MeCP2 driven chromatin organization and dynamics *in vivo*.

(A) Schematic overview of the cell transfection, heterochromatin compartment segmentation, and FRAP assay. The C2C12 cells were transfected with plasmids expressing GFP-MeCP2 nonsense mutations (truncations (tru.)). 24 - 36 hours after transfection, the cells were either fixed to analyze the function of MeCP2 truncations on heterochromatin compartment size or applied for live-cell imaging and FRAP analysis.
(B-C) Quantification of protein enrichment in heterochromatin compartments (B) and heterochromatin compartment size (C). The cells are analyzed using imageJ. The cell nuclei and heterochromatin compartments were identified by Auto Threshold based on the DAPI intensities (A). Cells were subgrouped by GFP intensities. The scatter plot represents the distribution of protein fold enrichment (B) and mean heterochromatin compartment area (C) per nucleus. Red dashed lines represent the fold enrichment of GFP fluorescence signals (B) and the mean heterochromatin compartment area (C) in C2C12 cells 36h after transfection with the plasmid expressing GFP only. Error bars correspond to standard deviation (SD). Significances were calculated by unpaired t-test comparing the middle and high expressing samples with low expressing samples as indicated by lines. ns $P > 0.05$; *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$.
(D) Scatter graph with fitted linear regression illustrating the influence of actual protein concentrations inside heterochromatin compartments in heterochromatin compartment size. The mean heterochromatin compartment size per nuclei was plotted versus the mean fluorescent intensity of GFP-MeCP2 truncations within heterochromatin compartments per nuclei.
(E) Mean FRAP curves in C2C12 cells expressing GFP-MeCP2 FL (full length) (n = 25), GFP-MeCP2 R168X (n = 16), GFP-MeCP2 R255X (n = 19), GFP-MeCP2 R294X (n = 29).
(F-G) Fluorescence recovery t1/2 (F) and immobile fractions (G) of FRAP curve in (E). Data were given as mean ± SD. Significances were calculated by unpaired t-test comparing the high expressing samples with low expressing samples as indicated by lines. ns $P > 0.05$; *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$.

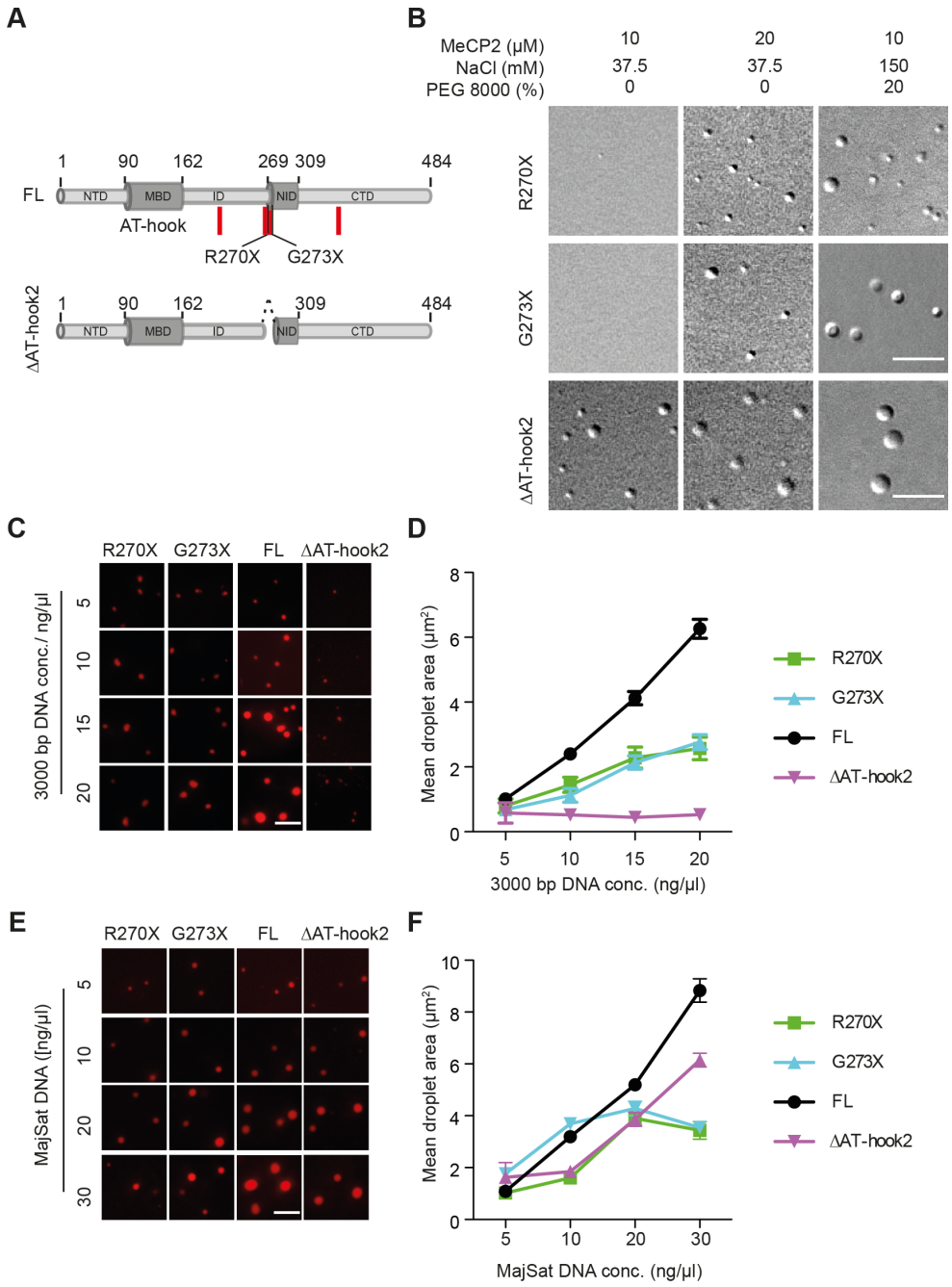


Figure 3.8 AT-hook2 regulates the MeCP2 phase properties.

The *in vitro* phase separation assay at the conditions indicated was done by incubating at room temperature for 45 min. Then, the mixtures were transferred to chambers made of double-sided tapes and sealed with coverslips. The DIC and fluorescence images were taken using the Nikon Eclipse TiE2 microscope.
(A) Top: schematic graph showing the distribution of three AT-hook-like motifs (red boxes) across the MeCP2 coding sequence and two AT-hook 2 related nonsense mutations found in patients with Rett syndrome. Bottom: schematic graph showing the MeCP2 coding sequence lacking the AT-hook 2 region.
(B) DIC images showing the phase separation properties of MeCP2 R270X, G273X and MeCP2ΔAT-hook 2 at different conditions. Scale bar (green) = 5 μm; Scale bar (white) = 10 μm.
(C) Fluorescent images showing the phase separation properties of MeCP2 R270X, G273X, MeCP2ΔAT-hook 2, and full-length MeCP2 in the presence of DRAQ5 labeled 3000 bp DNA with different concentrations (conc.). Representative images of the QRAQ5 (Cy5) channel are shown. MeCP2: 3 μM, NaCl: 150 mM, no crowding agents. Scale bar = 10 μm.
(D) Line graph showing the droplet size formed by MeCP2 R270X, G273X, MeCP2ΔAT-hook 2, and full-length MeCP2 in the presence of DRAQ5 labeled 3000 bp DNA with different concentrations from (C). The droplets were segmented by a bandpass filter and threshold based on the mean fluorescence intensity in/out droplets.

Droplets with size > 0.1 µm2 were considered and droplet parameters were measured. ≥ 3 images were taken for each condition. The mean droplet area was plotted with mean ± SEM (standard error of the mean).
(E) Fluorescent images showing the phase separation properties of MeCP2 R270X, G273X, MeCP2ΔAT-hook 2, and full-length MeCP2 in the presence of DRAQ5 labeled 234 bp major satellite repeat (MajSat) with different concentrations (conc.). Representative images of the QRAQ5 (Cy5) channel are shown. MeCP2: 10 µM, NaCl: 150 mM, 5% dextran T150. Scale bar = 10 µm.
(F) Line graph showing the droplet size formed by MeCP2 R270X, G273X, MeCP2ΔAT-hook 2, and full-length MeCP2 in the presence of DRAQ5 labeled 234 bp major satellite repeat (MajSat) with different concentrations (conc.) from (F). The droplets were segmented by a bandpass filter and threshold based on the mean fluorescence intensity in/out droplets. Droplets with size > 0.1 µm2 were considered and droplet parameters were measured. ≥ 3 images were taken for each condition. The mean droplet area was plotted with mean ± SEM (standard error of the mean).

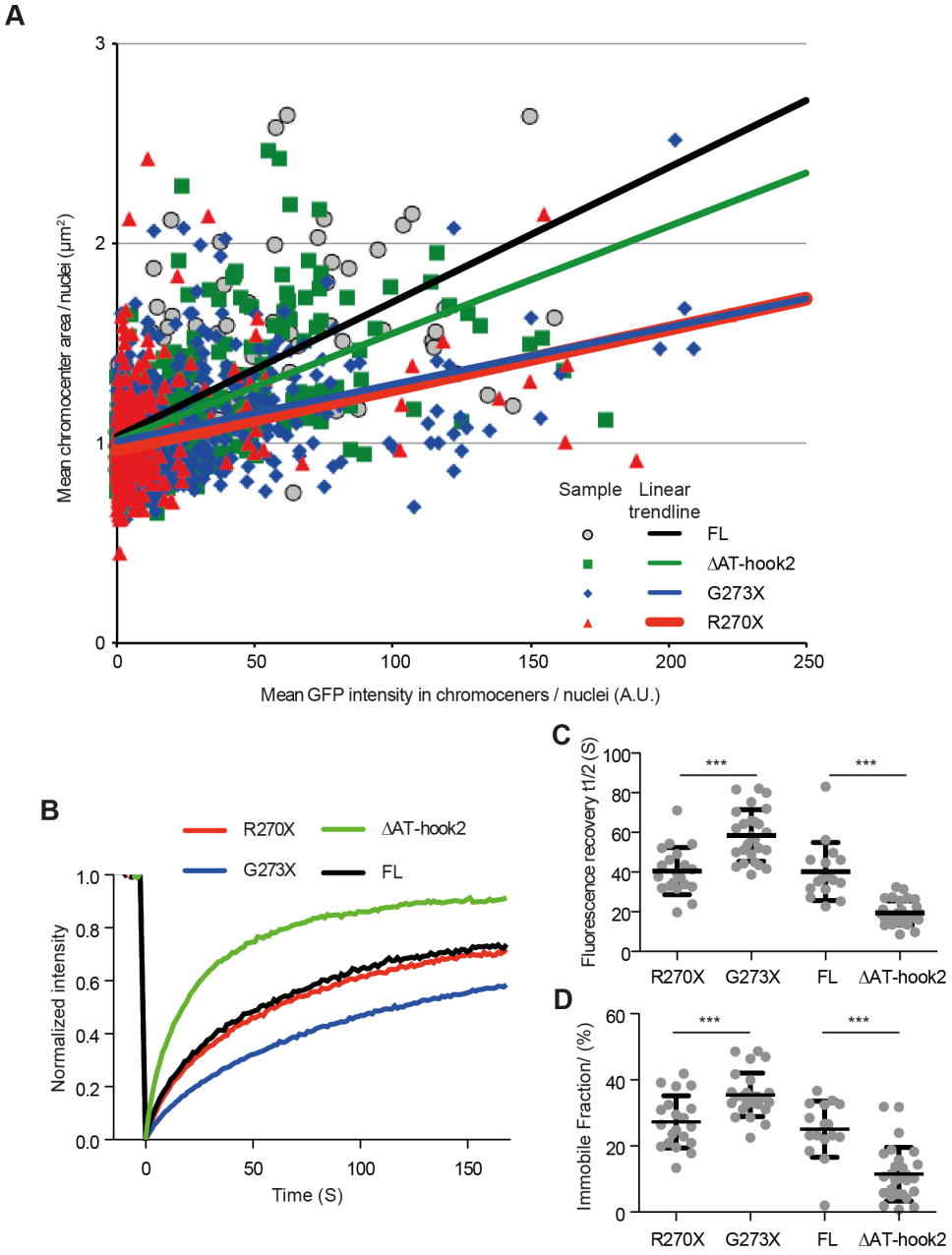


Figure 3.9 AT-hook 2 function in MeCP2 mediated heterochromatin organization and MeCP2 dynamics.

The C2C12 cells were transfected with plasmids expressing MeCP2 R270X, G273X, MeCP2ΔAT-hook 2, and full-length MeCP2. 24 - 36 hours after transfection, the cells were either fixed to analyze the function of MeCP2 truncations on heterochromatin compartment size or applied for live-cell imaging and FRAP analysis.
(A) Scatter graph with fitted linear regression illustrating the influence of actual protein concentrations inside heterochromatin compartments in heterochromatin compartment size in fixed cells. The cells are analyzed using

imageJ. The cell nuclei and heterochromatin compartment were identified by Auto Threshold based on the DAPI intensities and the heterochromatin compartment parameters were measured for quantitative analysis. The mean heterochromatin compartment size per nuclei was plotted versus the mean fluorescent intensity of GFP-MeCP2 truncations within heterochromatin compartments per nuclei.
(B) Mean FRAP curves in C2C12 cells expressing GFP-MeCP2 FL (full length) (n = 25), GFP-MeCP2 R168X (n = 16), GFP-MeCP2 R255X (n = 19), GFP-MeCP2 R294X (n = 29).
(C-D) Fluorescence recovery t1/2 (C) and immobile fractions (D) of FRAP curve in (B). Data were given as mean ± SD. Significances were calculated by unpaired t-test comparing the high expressing samples with low expressing samples as indicated by lines. ns, P > 0.05; *, P ≤ 0.05; **, P ≤ 0.01; ***, P ≤ 0.001.

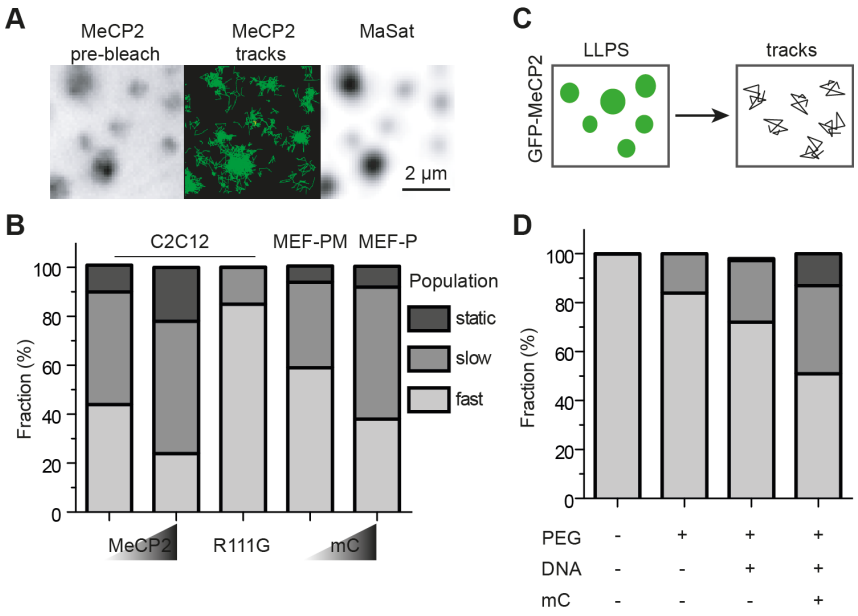


Figure 3.10 Single-molecule dynamics of MeCP2 in different conditions in cellulo and in vitro. (in collaboration with Hector Romero, AG Cardoso)

(A) Exemplary images of myoblast C2C12 cells with GFP-MeCP2 expression and the MeCP2 tracks. The pre-bleached image corresponds to an image taken before photobleaching using a 488 nm laser. MajSat overview image corresponds to an image taken before photobleaching using a 561 nm laser to elucidate the heterochromatin compartments. MeCP2 tracks were obtained from movies after eliminating the bleaching frames. Scale bar = 2 µm.
(B) Gaussian-mixture model comparison of MeCP2 populations and population fractions in different conditions *in cellulo*.
(C) Scheme graph of LLPS and single-molecule tracking. 10 µM MeCP2 with 1% GFP-MeCP2 was induced for droplet formation at different conditions. Then the mixtures were then transferred to chambers made of double-sided tapes and sealed with coverslips. Droplets were observed under the Nikon Eclipse Ti microscope equipped with oil immersion Nikon 100x Apo TIRF (1.49 NA objective and highly inclined laminar optical sheet illumination (HILO) (Tokunaga, Imamoto et al. 2008). Tracks and mobility populations were analyzed by TrackMate and SMTracker, separately.
(D) Gaussian-mixture model comparison of MeCP2 populations and population fractions in different conditions *in vitro*.

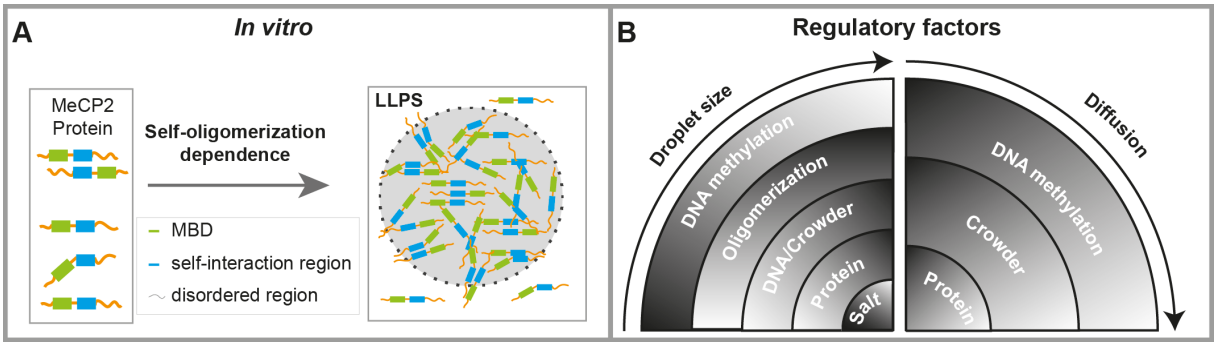


Figure 3.11 Summary of factors and conditions affecting MeCP2 LLPS and diffusion properties leading to a model comparing minimal and cell systems.

(A) Schematic chart depicting the conditions required for MeCP2 LLPS *in vitro*. MeCP2 is composed of the structured MBD domain, which is surrounded by the highly disordered amino acid sequences including the self-interaction domain (amino acids 168 to 254). Purified MeCP2 molecules locally condense to form a liquid-like spherical phase (LLPS). Self-interaction drives the formation of liquid-like droplets.
(B) Regulatory factors for MeCP2 LLPS and dynamics. The droplet size is positively regulated by protein, crowders, and unmethylated DNA concentration and negatively regulated by DNA methylation and salt concentration. MeCP2 mobility is reduced by increased concentration of protein, crowders, and methylated DNA.

6 References

Agarwal, N., A. Becker, K. L. Jost, S. Haase, B. K. Thakur, A. Brero, T. Hardt, S. Kudo, H. Leonhardt and M. C. Cardoso (2011). "MeCP2 Rett mutations affect large scale chromatin organization." Human molecular genetics 20(21): 4187-4195.

Agarwal, N., T. Hardt, A. Brero, D. Nowak, U. Rothbauer, A. Becker, H. Leonhardt and M. C. Cardoso (2007). "MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation." Nucleic Acids Research 35(16): 5402-5408.

Alberti, S., A. Gladfelter and T. Mittag (2019). "Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates." Cell 176(3): 419-434.

Allshire, R. C. and H. D. Madhani (2018). "Ten principles of heterochromatin formation and function." Nat Rev Mol Cell Biol 19(4): 229-244.

Amir, R. E., I. B. Van den Veyver, M. Wan, C. Q. Tran, U. Francke and H. Y. Zoghbi (1999). "Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2." Nature genetics 23(2): 185-188.

André, A. A. M. and E. Spruijt (2020). "Liquid–liquid phase separation in crowded environments." Int J Mol Sci 21(16): 5908-5908.

Angeloni, A. and O. Bogdanovic (2019). "Enhancer DNA methylation: implications for gene regulation." Essays in biochemistry 63(6): 707-715.

Angrisano, T., F. Lembo, R. Pero, F. Natale, A. Fusco, V. E. Avvedimento, C. B. Bruni and L. Chiariotti (2006). "TACC3 mediates the association of MBD2 with histone acetyltransferases and relieves transcriptional repression of methylated promoters." Nucleic Acids Res 34(1): 364-372.

Anosova, I., S. Melnik, K. Tripsianes, F. Kateb, I. Grummt and M. Sattler (2015). "A novel RNA binding surface of the TAM domain of TIP5/BAZ2A mediates epigenetic regulation of rRNA genes." Nucleic Acids Res 43(10): 5208-5220.

Aran, D., G. Toperoff, M. Rosenberg and A. Hellman (2011). "Replication timing-related and gene body-specific methylation of active human genes." Human molecular genetics 20(4): 670-680.

Arumugam, K., W. Shin, V. Schiavone, L. Vlahos, X. Tu, D. Carnevali, J. Kesner, E. O. Paull, N. Romo, P. Subramaniam, J. Worley, X. Tan, A. Califano and M. P. Cosma (2020). "The Master Regulator Protein BAZ2B Can Reprogram Human Hematopoietic Lineage-Committed Progenitors into a Multipotent State." Cell Rep 33(10): 108474-108474.

Babinchak, W. M. and W. K. Surewicz (2020). "Liquid-Liquid Phase Separation and Its Mechanistic Role in Pathological Protein Aggregation." J Mol Biol 432(7): 1910-1925.

Bachman, M., S. Uribe-Lewis, X. Yang, M. Williams, A. Murrell and S. Balasubramanian (2014). "5-Hydroxymethylcytosine is a predominantly stable DNA modification." Nat Chem 6(12): 1049-1055.

Bader, S., M. Walker, H. A. McQueen, R. Sellar, E. Oei, S. Wopereis, Y. Zhu, A. Peter, A. P. Bird and D. J. Harrison (2003). "MBD1, MBD2 and CGBP genes at chromosome 18q21 are infrequently mutated in human colon and lung cancers." Oncogene 22(22): 3506-3510.

Baker, S. A., L. Chen, A. D. Wilkins, P. Yu, O. Lichtarge and H. Y. Zoghbi (2013). "An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders." Cell 152(5): 984-996.

Bannister, A. J., P. Zegerman, J. F. Partridge, E. A. Miska, J. O. Thomas, R. C. Allshire and T. Kouzarides (2001). "Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain." Nature 410(6824): 120-124.

Batham, Lim and Rao (2019). "SETDB-1: A Potential Epigenetic Regulator in Breast Cancer Metastasis." Cancers 11(8): 1143-1143.

Baubec, T., R. Ivanek, F. Lienert and D. Schubeler (2013). "Methylation-dependent and -independent genomic targeting principles of the MBD protein family." Cell 153(2): 480-492.

Bayer, T. S., L. N. Booth, S. M. Knudsen and A. D. Ellington (2005). "Arginine-rich motifs present multiple interfaces for specific binding by RNA." RNA 11(12): 1848-1857.

Baymaz, H. I., A. Fournier, S. Laget, Z. Ji, P. W. Jansen, A. H. Smits, L. Ferry, A. Mensinga, I. Poser, A. Sharrocks, P. A. Defossez and M. Vermeulen (2014). "MBD5 and MBD6 interact with the human PR-DUB complex through their methyl-CpG-binding domain." Proteomics 14(19): 2179-2189.

Beaulaurier, J., E. E. Schadt and G. Fang (2019). "Deciphering bacterial epigenomes using modern sequencing technologies." Nature Reviews Genetics 20(3): 157-172.

Becker, A., L. Allmann, M. Hofstätter, V. Casà, P. Weber, A. Lehmkuhl, H. D. Herce and M. C. Cardoso (2013). "Direct homo-and hetero-interactions of MeCP2 and MBD2." PLoS One 8(1): e53730-e53730.

Bellacosa, A. and A. C. Drohat (2015). "Role of base excision repair in maintaining the genetic and epigenetic integrity of CpG sites." DNA repair 32: 33-42.

Berger, J. and A. Bird (2005). "Role of MBD2 in gene regulation and tumorigenesis." Biochemical Society Transactions 33(6): 1537-1540.

Berger, S. L., T. Kouzarides, R. Shiekhattar and A. Shilatifard (2009). "An operational definition of epigenetics." Genes Dev 23(7): 781-783.

Bertulat, B., M. L. De Bonis, F. Della Ragione, A. Lehmkuhl and M. Milden (2012). "MeCP2 Dependent Heterochromatin Reorganization during Neural."

Bertulat, B., M. L. De Bonis, F. Della Ragione, A. Lehmkuhl, M. Milden, C. Storm, K. L. Jost, S. Scala, B. Hendrich and M. D'Esposito (2012). "MeCP2 dependent heterochromatin reorganization during neural differentiation of a novel Mecp2-deficient embryonic stem cell reporter line." PLoS One 7(10): e47848-e47848.

Bewick, A. J., L. Ji, C. E. Niederhuth, E.-M. Willing, B. T. Hofmeister, X. Shi, L. Wang, Z. Lu, N. A. Rohr and B. Hartwig (2016). "On the origin and evolutionary consequences of gene body DNA methylation." Proceedings of the National Academy of Sciences 113(32): 9111-9116.

Bhaumik, S. R., E. Smith and A. Shilatifard (2007). "Covalent modifications of histones during development and disease pathogenesis." Nature structural & molecular biology 14(11): 1008-1016.

Bird, A., M. Taggart, M. Frommer, O. J. Miller and D. Macleod (1985). "A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA." Cell 40(1): 91-99.

Biswas, S., J. Kundu, S. K. Mukherjee and P. K. Chowdhury (2018). "Mixed macromolecular crowding: a protein and solvent perspective." ACS omega 3(4): 4316-4330.

Boeke, J., O. Ammerpohl, S. Kegel, U. Moehren and R. Renkawitz (2000). "The minimal repression domain of MBD2b overlaps with the methyl-CpG-binding domain and binds directly to Sin3A." J Biol Chem 275(45): 34963-34967.

Boija, A., I. A. Klein, B. R. Sabari, A. Dall'Agnese, E. L. Coffey, A. V. Zamudio, C. H. Li, K. Shrinivas, J. C. Manteiga, N. M. Hannett, B. J. Abraham, L. K. Afeyan, Y. E. Guo, J. K. Rimel, C. B. Fant, J. Schuijers, T. I. Lee, D. J. Taatjes and R. A. Young (2018). "Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains." Cell 175(7): 1842-1855.e1816.

Bosshard, M., E. Markkanen and B. van Loon (2012). "Base excision repair in physiology and pathology of the central nervous system." Int J Mol Sci 13(12): 16172-16222.

Botchkarev, V. A., M. R. Gdula, A. N. Mardaryev, A. A. Sharov and M. Y. Fessing (2012). "Epigenetic regulation of gene expression in keratinocytes." Journal of Investigative Dermatology 132(11): 2505-2521.

Brandeis, M., M. Ariel and H. Cedar (1993). "Dynamics of DNA methylation during development." Bioessays 15(11): 709-713.

Brangwynne, C. P., C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher and A. A. Hyman (2009). "Germline P granules are liquid droplets that localize by controlled dissolution/condensation." Science 324(5935): 1729-1732.

Brangwynne, C. P., P. Tompa and R. V. Pappu (2015). "Polymer physics of intracellular phase transitions." Nature Physics 11(11): 899-899.

Brero, A., H. P. Easwaran, D. Nowak, I. Grunewald, T. Cremer, H. Leonhardt and M. C. Cardoso (2005). "Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation." The Journal of cell biology 169(5): 733-743.

Brown, S. E., M. J. Suderman, M. Hallett and M. Szyf (2008). "DNA demethylation induced by the methyl-CpG-binding domain protein MBD3." Gene 420(2): 99-106.

Buchmuller, B. C., B. Kosel and D. Summerer (2020). "Complete profiling of Methyl-CpG-binding domains for combinations of cytosine modifications at CpG dinucleotides reveals differential read-out in normal and Rett-associated states." Scientific Reports 10(1): 1-9.

Camarena, V., L. Cao, C. Abad, A. Abrams, Y. Toledo, K. Araki, M. Araki, K. Walz and J. I. Young (2014). "Disruption of Mbd5 in mice causes neuronal functional deficits and neurobehavioral abnormalities consistent with 2q23.1 microdeletion syndrome." EMBO Mol Med 6(8): 1003-1015.

Casas-Delucchi, C. S., A. Becker, J. J. Bolius and M. C. Cardoso (2012). "Targeted manipulation of heterochromatin rescues MeCP2 Rett mutants and re-establishes higher order chromatin organization." Nucleic acids research 40(22): e176-e176.

Casas-Delucchi, C. S., J. G. van Bemmel, S. Haase, H. D. Herce, D. Nowak, D. Meilinger, J. H. Stear, H. Leonhardt and M. C. Cardoso (2012). "Histone hypoacetylation is required to maintain late replication timing of constitutive heterochromatin." Nucleic acids research 40(1): 159-169.

Cassina, V., M. Manghi, D. Salerno, A. Tempestini, V. Iadarola, L. Nardo, S. Brioschi and F. Mantegazza (2016). "Effects of cytosine methylation on DNA morphology: An atomic force microscopy study." Biochimica et Biophysica Acta (BBA)-General Subjects 1860(1): 1-7.

Cervoni, N. and M. Szyf (2001). "Demethylase activity is directed by histone acetylation." Journal of Biological Chemistry 276(44): 40778-40787.

Chahrour, M., S. Y. Jung, C. Shaw, X. Zhou, S. T. C. Wong, J. Qin and H. Y. Zoghbi (2008). "MeCP2, a key contributor to neurological disease, activates and represses transcription." Science 320(5880): 1224-1229.

Charlop-Powers, Z., L. Zeng, Q. Zhang and M. M. Zhou (2010). "Structural insights into selective histone H3 recognition by the human Polybromo bromodomain 2." Cell Res 20(5): 529-538.

Chatagnon, A., S. Bougel, L. Perriaud, J. Lachuer, J. Benhattar and R. Dante (2009). "Specific association between the methyl-CpG-binding domain protein 2 and the hypermethylated region of the human telomerase reverse transcriptase promoter in cancer cells." Carcinogenesis 30(1): 28-34.

Chatagnon, A., L. Perriaud, N. Nazaret, S. Croze, J. Benhattar, J. Lachuer and R. Dante (2011). "Preferential binding of the methyl-CpG binding domain protein 2 at methylated transcriptional start site regions." Epigenetics 6(11): 1295-1307.

Chen, H., Y. Cui, X. Han, W. Hu, M. Sun, Y. Zhang, P.-H. Wang, G. Song, W. Chen and J. Lou (2020). "Liquid–liquid phase separation by SARS-CoV-2 nucleocapsid protein and RNA." Cell Research 30(12): 1143-1145.

Chen, H., X. Zheng and Y. Zheng (2014). "Age-associated loss of lamin-B leads to systemic inflammation and gut hyperplasia." Cell 159(4): 829-843.

Chen, L., K. Chen, L. A. Lavery, S. A. Baker, C. A. Shaw, W. Li and H. Y. Zoghbi (2015). "MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome." Proceedings of the National Academy of Sciences 112(17): 5509-5514.

Chen, R. Z., S. Akbarian, M. Tudor and R. Jaenisch (2001). "Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice." Nat Genet 27(3): 327-331.

Chen, Z., F. Miao, B. H. Braffett, J. M. Lachin, L. Zhang, X. Wu, D. Roshandel, M. Carless, X. A. Li and J. D. Tompkins (2020). "DNA methylation mediates development of HbA1c-associated complications in type 1 diabetes." Nature metabolism 2(8): 744-762.

Chen, Z. P., D. S. Gu, Z. P. Zhou, X. L. Chen, Z. X. Guo, W. T. Du, J. Ge, Q. Ren and R. C. Yang (2011). "Decreased expression of MBD2 and MBD4 gene and genomic-wide hypomethylation in patients with primary immune thrombocytopenia." Hum Immunol 72(6): 486-491.

Cheutin, T., A. J. McNairn, T. Jenuwein, D. M. Gilbert, P. B. Singh and T. Misteli (2003). "Maintenance of stable heterochromatin domains by dynamic HP1 binding." Science 299(5607): 721-725.

Choi, Y. J., N. J. Yoo and S. H. Lee (2015). "Mutation and expression of a methyl-binding protein 6 (MBD6) in gastric and colorectal cancers." Pathol Oncol Res 21(3): 857-858.

Collins, A. L., J. M. Levenson, A. P. Vilaythong, R. Richman, D. L. Armstrong, J. L. Noebels, J. David Sweatt and H. Y. Zoghbi (2004). "Mild overexpression of MeCP2 causes a progressive neurological disorder in mice." Hum Mol Genet 13(21): 2679-2689.

Cook, P. C., H. Owen, A. M. Deaton, J. G. Borger, S. L. Brown, T. Clouaire, G. R. Jones, L. H. Jones, R. J. Lundie, A. K. Marley, V. L. Morrison, A. T. Phythian-Adams, E. Wachter, L. M. Webb, T. E. Sutherland, G. D. Thomas, J. R. Grainger, J. Selfridge, A. N. McKenzie, J. E. Allen, S. C. Fagerholm, R. M. Maizels, A. C. Ivens, A. Bird and

A. S. MacDonald (2015). "A dominant role for the methyl-CpG-binding protein Mbd2 in controlling Th2 induction by dendritic cells." Nat Commun 6: 6920-6920.

Cooper, D. N. and H. Youssoufian (1988). "The CpG dinucleotide and human genetic disease." Hum Genet 78(2): 151-155.

Cross, S. H., R. R. Meehan, X. Nan and A. Bird (1997). "A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins." Nat Genet 16(3): 256-259.

Desai, M. A., H. D. Webb, L. M. Sinanan, J. N. Scarsdale, N. M. Walavalkar, G. D. Ginder and D. C. Williams Jr (2015). "An intrinsically disordered region of methyl-CpG binding domain protein 2 (MBD2) recruits the histone deacetylase core of the NuRD complex." Nucleic Acids Res 43(6): 3100-3113.

Devailly, G., M. Grandin, L. Perriaud, P. Mathot, J.-G. Delcros, Y. Bidet, A.-P. Morel, J.-Y. Bignon, A. Puisieux and P. Mehlen (2015). "Dynamics of MBD2 deposition across methylated DNA regions during malignant transformation of human mammary epithelial cells." Nucleic Acids Research 43(12): 5838-5854.

Dix, J. A. and A. S. Verkman (2008). "Crowding effects on diffusion in solutions and cells." Annu. Rev. Biophys. 37: 247-263.

Dixon, J. R., D. U. Gorkin and B. Ren (2016). "Chromatin domains: the unit of chromosome organization." Molecular Cell 62(5): 668-680.

Dolinoy, D. C., J. R. Weidman and R. L. Jirtle (2007). "Epigenetic gene regulation: linking early developmental environment to adult disease." Reprod Toxicol 23(3): 297-307.

dos Santos, R. L., L. Tosti, A. Radzisheuskaya, I. M. Caballero, K. Kaji, B. Hendrich and J. C. Silva (2014). "MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner." Cell Stem Cell 15(1): 102-110.

Du, Q., P.-L. Luu, C. Stirzaker and S. J. Clark (2015). "Methyl-CpG-binding domain proteins: readers of the epigenome." Epigenomics 7(6): 1051-1073.

Duncan, B. K. and J. H. Miller (1980). "Mutagenic deamination of cytosine residues in DNA." Nature 287(5782): 560-561.

Dunn, D. B. and J. D. Smith (1955). "Occurrence of a new base in the deoxyribonucleic acid of a strain of Bacterium coli." Nature 175(4451): 336-337.

Ehrlich, M., M. A. Gama-Sosa, L.-H. Huang, R. M. Midgett, K. C. Kuo, R. A. McCune and C. Gehrke (1982). "Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells." Nucleic Acids Research 10(8): 2709-2721.

Entrambasaguas, L., M. Ruocco, K. J. F. Verhoeven, G. Procaccini and L. Marín-Guirao (2021). "Gene body DNA methylation in seagrasses: inter-and intraspecific differences and interaction with transcriptome plasticity under heat stress." Scientific Reports 11(1): 1-15.

Erdel, F., A. Rademacher, R. Vlijm, J. Tünnermann, L. Frank, R. Weinmann, E. Schweigert, K. Yserentant, J. Hummert and C. Bauer (2020). "Mouse heterochromatin adopts digital compaction states without showing hallmarks of HP1-driven liquid-liquid phase separation." Molecular cell 78(2): 236-249. e237.

Erdel, F. and K. Rippe (2018). "Formation of chromatin subcompartments by phase separation." Biophysical journal 114(10): 2262-2270.

Fan, C., H. Zhang, L. Fu, Y. Li, Y. Du, Z. Qiu and F. Lu (2020). "Rett mutations attenuate phase separation of MeCP2." Cell discovery 6(1): 1-4.

Feng, S., S. E. Jacobsen and W. Reik (2010). "Epigenetic reprogramming in plant and animal development." Science 330(6004): 622-627.

Fernandez, P., P. Scaffidi, E. Markert, J.-H. Lee, S. Rane and T. Misteli (2014). "Transformation resistance in a premature aging disorder identifies a tumor-protective function of BRD4." Cell Reports 9(1): 248-260.

Fichou, Y., J. Nectoux, N. Bahi-Buisson, H. Rosas-Vargas, B. Girard, J. Chelly and T. Bienvenu (2009). "The first missense mutation causing Rett syndrome specifically affecting the MeCP2_e1 isoform." Neurogenetics 10(2): 127-127.

Forlani, G., E. Giarda, U. Ala, F. Di Cunto, M. Salani, R. Tupler, C. Kilstrup-Nielsen and N. Landsberger (2010). "The MeCP2/YY1 interaction regulates ANT1 expression at 4q35: novel hints for Rett syndrome pathogenesis." Hum Mol Genet 19(16): 3114-3123.

Fudenberg, G., M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur and L. A. Mirny (2016). "Formation of chromosomal domains by loop extrusion." Cell Reports 15(9): 2038-2049.

Fujita, H., R. Fujii, S. Aratani, T. Amano, A. Fukamizu and T. Nakajima (2003). "Antithetic effects of MBD2a on gene regulation." Molecular and Cellular Biology 23(8): 2645-2657.

Fujita, N., S. Takebayashi, K. Okumura, S. Kudo, T. Chiba, H. Saya and M. Nakao (1999). "Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms." Mol Cell Biol 19(9): 6415-6426.

Gabel, H. W., B. Kinde, H. Stroud, C. S. Gilbert, D. A. Harmin, N. R. Kastan, M. Hemberg, D. H. Ebert and M. E. Greenberg (2015). "Disruption of DNA-methylation-dependent long gene repression in Rett syndrome." Nature 522(7554): 89-93.

Galvao, T. C. and J. O. Thomas (2005). "Structure-specific binding of MeCP2 to four-way junction DNA through its methyl CpG-binding domain." Nucleic Acids Research 33(20): 6603-6609.

Georgel, P. T., R. A. Horowitz-Scherer, N. Adkins, C. L. Woodcock, P. A. Wade and J. C. Hansen (2003). "Chromatin compaction by human MeCP2 assembly of novel

secondary chromatin structures in the absence of DNA methylation." Journal of Biological Chemistry 278(34): 32181-32188.

Ghosh, R. P., R. A. Horowitz-Scherer, T. Nikitina, L. S. Shlyakhtenko and C. L. Woodcock (2010). "MeCP2 binds cooperatively to its substrate and competes with histone H1 for chromatin binding sites." Molecular and cellular biology 30(19): 4656-4670.

Ghosh, R. P., T. Nikitina, R. A. Horowitz-Scherer, L. M. Gierasch, V. N. Uversky, K. Hite, J. C. Hansen and C. L. Woodcock (2010). "Unique physical properties and interactions of the domains of methylated DNA binding protein 2." Biochemistry 49(20): 4395-4410.

Gibson, B. A., L. K. Doolittle, M. W. G. Schneider, L. E. Jensen, N. Gamarra, L. Henry, D. W. Gerlich, S. Redding and M. K. Rosen (2019). "Organization of chromatin by intrinsic and regulated phase separation." Cell 179(2): 470-484. e421.

Gnanapragasam, M. N., J. N. Scarsdale, M. L. Amaya, H. D. Webb, M. A. Desai, N. M. Walavalkar, S. Z. Wang, S. Zu Zhu, G. D. Ginder and D. C. Williams Jr (2011). "p66Alpha-MBD2 coiled-coil interaction and recruitment of Mi-2 are critical for globin gene silencing by the MBD2-NuRD complex." Proc Natl Acad Sci U S A 108(18): 7487-7492.

Gonzales, M. L., S. Adams, K. W. Dunaway and J. M. LaSalle (2012). "Phosphorylation of distinct sites in MeCP2 modifies cofactor associations and the dynamics of transcriptional regulation." Mol Cell Biol 32(14): 2894-2903.

Gopalakrishnan, S., B. O. Van Emburgh and K. D. Robertson (2008). "DNA methylation in development and human disease." Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 647(1-2): 30-38.

Grafstrom, R. H., R. Yuan and D. L. Hamilton (1985). "The characteristics of DNA methytation in an in vitro DNA synthesizing system from mouse fibroblasts." Nucleic Acids Research 13(8): 2827-2842.

Greenstein, R. A. and B. Al-Sady (2019). "Epigenetic fates of gene silencing established by heterochromatin spreading in cell identity and genome stability." Current genetics 65(2): 423-428.

Grewal, S. I. and D. Moazed (2003). "Heterochromatin and epigenetic control of gene expression." Science 301(5634): 798-802.

Guillemette, B., P. Drogaris, H. H. Lin, H. Armstrong, K. Hiragami-Hamada, A. Imhof, E. Bonneil, P. Thibault, A. Verreault and R. J. Festenstein (2011). "H3 lysine 4 is acetylated at active gene promoters and is regulated by H3 lysine 4 methylation." PLoS Genet 7(3): e1001354-e1001354.

Gunther, K., M. Rust, J. Leers, T. Boettger, M. Scharfe, M. Jarek, M. Bartkuhn and R. Renkawitz (2013). "Differential roles for MBD2 and MBD3 at methylated CpG islands,

active promoters and binding to exon sequences." Nucleic Acids Res 41(5): 3010-3021.

Guo, H., E. Bettella, P. C. Marcogliese, R. Zhao, J. C. Andrews, T. J. Nowakowski, M. A. Gillentine, K. Hoekzema, T. Wang, H. Wu, S. Jangam, C. Liu, H. Ni, M. H. Willemsen, B. W. van Bon, T. Rinne, S. J. C. Stevens, T. Kleefstra, H. G. Brunner, H. G. Yntema, M. Long, W. Zhao, Z. Hu, C. Colson, N. Richard, C. E. Schwartz, C. Romano, L. Castiglia, M. Bottitta, S. U. Dhar, D. J. Erwin, L. Emrick, B. Keren, A. Afenjar, B. Zhu, B. Bai, P. Stankiewicz, K. Herman, S. Mercimek-Andrews, J. Juusola, A. B. Wilfert, R. Abou Jamra, B. Büttner, H. C. Mefford, A. M. Muir, I. E. Scheffer, B. M. Regan, S. Malone, J. Gecz, J. Cobben, M. M. Weiss, Q. Waisfisz, E. K. Bijlsma, M. J. V. Hoffer, C. A. L. Ruivenkamp, S. Sartori, F. Xia, J. A. Rosenfeld, R. A. Bernier, M. F. Wangler, S. Yamamoto, K. Xia, A. P. A. Stegmann, H. J. Bellen, A. Murgia and E. E. Eichler (2019). "Disruptive mutations in TANC2 define a neurodevelopmental syndrome associated with psychiatric disorders." Nat Commun 10(1): 4679-4679.

Guo, J. U., D. K. Ma, H. Mo, M. P. Ball, M.-H. Jang, M. A. Bonaguidi, J. A. Balazer, H. L. Eaves, B. Xie and E. Ford (2011). "Neuronal activity modifies the DNA methylation landscape in the adult brain." Nature neuroscience 14(10): 1345-1351.

Guo, J. U., Y. Su, J. H. Shin, J. Shin, H. Li, B. Xie, C. Zhong, S. Hu, T. Le and G. Fan (2014). "Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain." Nature neuroscience 17(2): 215-222.

Guo, J. U., Y. Su, J. H. Shin, J. Shin, H. Li, B. Xie, C. Zhong, S. Hu, T. Le, G. Fan, H. Zhu, Q. Chang, Y. Gao, G. L. Ming and H. Song (2014). "Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain." Nat Neurosci 17(2): 215-222.

Guy, J., B. Alexander-Howden, L. FitzPatrick, D. DeSousa, M. V. Koerner, J. Selfridge and A. Bird (2018). "A mutation-led search for novel functional domains in MeCP2." Human molecular genetics 27(14): 2531-2545.

Guy, J., B. Hendrich, M. Holmes, J. E. Martin and A. Bird (2001). "A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome." Nat Genet 27(3): 322-326.

Hagberg, B., J. Aicardi, K. Dias and O. Ramos (1983). "A Progressive Syndrome of Autism, Dementia, Ataxia, and Loss of Purposeful Hand Use in Girls - Retts Syndrome - Report of 35 Cases." Annals of Neurology 14(4): 471-479.

Hancock, R. Packing of the polynucleosome chain in interphase chromosomes: evidence for a contribution of crowding and entropic forces, Elsevier.

Handwerger, K. E., J. A. Cordero and J. G. Gall (2005). "Cajal bodies, nucleoli, and speckles in the Xenopus oocyte nucleus have a low-density, sponge-like structure." Molecular biology of the cell 16(1): 202-211.

Hashimoto, H., Y. Liu, A. K. Upadhyay, Y. Chang, S. B. Howerton, P. M. Vertino, X. Zhang and X. Cheng (2012). "Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation." Nucleic Acids Res 40(11): 4841-4849.

Hendrich, B., C. Abbott, H. McQueen, D. Chambers, S. Cross and A. Bird (1999). "Genomic structure and chromosomal mapping of the murine and human Mbd1, Mbd2, Mbd3, and Mbd4 genes." Mamm Genome 10(9): 906-912.

Hendrich, B. and A. Bird (1998). "Identification and characterization of a family of mammalian methyl-CpG binding proteins." Mol Cell Biol 18(11): 6538-6547.

Hendrich, B., J. Guy, B. Ramsahoye, V. A. Wilson and A. Bird (2001). "Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development." Genes Dev 15(6): 710-723.

Hendrich, B. and S. Tweedie (2003). "The methyl-CpG binding domain and the evolving role of DNA methylation in animals." Trends Genet 19(5): 269-277.

Hennig, W. (1999). "Heterochromatin." Chromosoma 108(1): 1-9.

Ho, K. L., I. W. McNae, L. Schmiedeberg, R. J. Klose, A. P. Bird and M. D. Walkinshaw (2008). "MeCP2 binding to DNA depends upon hydration at methyl-CpG." Molecular cell 29(4): 525-531.

Homouz, D., M. Perham, A. Samiotakis, M. S. Cheung and P. Wittung-Stafshede (2008). "Crowded, cell-like environment induces shape changes in aspherical protein." Proceedings of the National Academy of Sciences 105(33): 11754-11759.

Horike, S.-i., S. Cai, M. Miyano, J.-F. Cheng and T. Kohwi-Shigematsu (2005). "Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome." Nature genetics 37(1): 31-40.

Hotchkiss, R. D. (1948). "The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography." J. Biol. Chem. 175: 315-332.

Huisinga, K. L., B. Brower-Toland and S. C. Elgin (2006). "The contradictory definitions of heterochromatin: transcription and silencing." Chromosoma 115(2): 110-122.

Ichimura, T., S. Watanabe, Y. Sakamoto, T. Aoto, N. Fujita and M. Nakao (2005). "Transcriptional repression and heterochromatin formation by MBD1 and MCAF/AM family proteins." J Biol Chem 280(14): 13928-13935.

Ichino, L., B. A. Boone, L. Strauskulage, C. J. Harris, G. Kaur, M. A. Gladstone, M. Tan, S. Feng, Y. Jami-Alahmadi, S. H. Duttke, J. A. Wohlschlegel, X. Cheng, S. Redding and S. E. Jacobsen (2021). "MBD5 and MBD6 couple DNA methylation to gene silencing through the J-domain protein SILENZIO." Science.

Inoue, A., L. Jiang, F. Lu, T. Suzuki and Y. Zhang (2017). "Maternal H3K27me3 controls DNA methylation-independent imprinting." Nature 547(7664): 419-424.

Isaac, R. S., S. Sanulli, R. Tibble, M. Hornsby, M. Ravalin, C. S. Craik, J. D. Gross and G. J. Narlikar (2017). "Biochemical basis for distinct roles of the heterochromatin proteins Swi6 and Chp2." Journal of molecular biology 429(23): 3666-3677.

Ito-Ishida, A., S. A. Baker, R. V. Sillitoe, Y. Sun, J. Zhou, Y. Ono, J. Iwakiri, M. Yuzaki and H. Y. Zoghbi (2020). "MeCP2 levels regulate the 3D structure of heterochromatic foci in mouse neurons." Journal of Neuroscience 40(45): 8746-8766.

Iurlaro, M., G. Ficz, D. Oxley, E. A. Raiber, M. Bachman, M. J. Booth, S. Andrews, S. Balasubramanian and W. Reik (2013). "A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation." Genome Biol 14(10): R119-R119.

Iyer, L. M., D. Zhang and L. Aravind (2016). "Adenine methylation in eukaryotes: Apprehending the complex evolutionary history and functional potential of an epigenetic modification." Bioessays 38(1): 27-40.

Jaffer, S., P. Goh, M. Abbasian and A. C. Nathwani (2018). "Mbd3 Promotes Reprogramming of Primary Human Fibroblasts." Int J Stem Cells 11(2): 235-241.

Jang, H. S., W. J. Shin, J. E. Lee and J. T. Do (2017). "CpG and non-CpG methylation in epigenetic gene regulation and brain function." Genes 8(6): 148-148.

Janssen, A., S. U. Colmenares and G. H. Karpen (2018). "Heterochromatin: guardian of the genome." Annual review of cell and developmental biology 34: 265-288.

Jin, Q., L. R. Yu, L. Wang, Z. Zhang, L. H. Kasper, J. E. Lee, C. Wang, P. K. Brindle, S. Y. Dent and K. Ge (2011). "Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation." EMBO J 30(2): 249-262.

Johnson, W. L. and A. F. Straight (2017). "RNA-mediated regulation of heterochromatin." Current opinion in cell biology 46: 102-109.

Johnson, W. L., W. T. Yewdell, J. C. Bell, S. M. McNulty, Z. Duda, R. J. O'Neill, B. A. Sullivan and A. F. Straight (2017). "RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin." Elife 6: e25299-e25299.

Jones, P. L., G. J. Veenstra, P. A. Wade, D. Vermaak, S. U. Kass, N. Landsberger, J. Strouboulis and A. P. Wolffe (1998). "Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription." Nat Genet 19(2): 187-191.

Jongeneel, C. V. (2000). "Searching the expressed sequence tag (EST) databases: panning for genes." Brief Bioinform 1(1): 76-92.

Jørgensen, H. F., I. Ben-Porath and A. P. Bird (2004). "Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains." Mol Cell Biol 24(8): 3387-3395.

Kaji, K., I. M. Caballero, R. MacLeod, J. Nichols, V. A. Wilson and B. Hendrich (2006). "The NuRD component Mbd3 is required for pluripotency of embryonic stem cells." Nat Cell Biol 8(3): 285-292.

Kaji, K., J. Nichols and B. Hendrich (2007). "Mbd3, a component of the NuRD co-repressor complex, is required for development of pluripotent cells." Development 134(6): 1123-1132.

Kalish, J. M., C. Jiang and M. S. Bartolomei (2014). "Epigenetics and imprinting in human disease." Int J Dev Biol 58(2-4): 291-298.

Kanaan, N. M., C. Hamel, T. Grabinski and B. Combs (2020). "Liquid-liquid phase separation induces pathogenic tau conformations in vitro." Nature Communications 11(1): 1-16.

Karaca, K. G., D. V. C. Brito, B. Zeuch and A. M. M. Oliveira (2018). "Adult hippocampal MeCP2 preserves the genomic responsiveness to learning required for long-term memory formation." Neurobiology of learning and memory 149: 84-97.

Karymov, M. A., M. Tomschik, S. H. Leuba, P. Caiafa and J. Zlatanova (2001). "DNA methylation ‐dependent chromatin fiber compaction in vivo and in vitro: requirement for linker histone." The FASEB Journal 15(14): 2631-2641.

Kernohan, K. D., Y. Jiang, D. C. Tremblay, A. C. Bonvissuto, J. H. Eubanks, M. R. Mann and N. G. Berube (2010). "ATRX partners with cohesin and MeCP2 and contributes to developmental silencing of imprinted genes in the brain." Dev Cell 18(2): 191-202.

Kernohan, K. D., D. Vernimmen, G. B. Gloor and N. G. Berube (2014). "Analysis of neonatal brain lacking ATRX or MeCP2 reveals changes in nucleosome density, CTCF binding and chromatin looping." Nucleic Acids Res 42(13): 8356-8368.

Kersh, E. N. (2006). "Impaired memory CD8 T cell development in the absence of methyl-CpG-binding domain protein 2." J Immunol 177(6): 3821-3826.

Klemm, S. L., Z. Shipony and W. J. Greenleaf (2019). "Chromatin accessibility and the regulatory epigenome." Nat Rev Genet 20(4): 207-220.

Ko, M., J. An, W. A. Pastor, S. B. Koralov, K. Rajewsky and A. Rao (2015). "TET proteins and 5 ‐methylcytosine oxidation in hematological cancers." Immunological reviews 263(1): 6-21.

Koenhen, D. M., M. H. V. Mulder and C. A. Smolders (1977). "Phase separation phenomena during the formation of asymmetric membranes." Journal of applied polymer science 21(1): 199-215.

Koh, C. W. Q., Y. T. Goh, J. D. W. Toh, S. P. Neo, S. B. Ng, J. Gunaratne, Y.-G. Gao, S. R. Quake, W. F. Burkholder and W. S. S. Goh (2018). "Single-nucleotide-resolution sequencing of human N 6-methyldeoxyadenosine reveals strand-asymmetric clusters associated with SSBP1 on the mitochondrial genome." Nucleic Acids Research 46(22): 11659-11670.

Kokura, K., S. C. Kaul, R. Wadhwa, T. Nomura, M. M. Khan, T. Shinagawa, T. Yasukawa, C. Colmenares and S. Ishii (2001). "The Ski protein family is required for MeCP2-mediated transcriptional repression." J Biol Chem 276(36): 34115-34121.

Kondo, E., Z. Gu, A. Horii and S. Fukushige (2005). "The thymine DNA glycosylase MBD4 represses transcription and is associated with methylated p16(INK4a) and hMLH1 genes." Mol Cell Biol 25(11): 4388-4396.

Krainer, G., T. J. Welsh, J. A. Joseph, J. R. Espinosa, S. Wittmann, E. de Csilléry, A. Sridhar, Z. Toprakcioglu, G. Gudiškytė, M. A. Czekalska, W. E. Arter, J. Guillén-Boixet, T. M. Franzmann, S. Qamar, P. S. George-Hyslop, A. A. Hyman, R. Collepardo-Guevara, S. Alberti and T. P. J. Knowles (2021). "Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions." Nat Commun 12(1): 1085-1085.

Kriaucionis, S. and N. Heintz (2009). "The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain." Science 324(5929): 929-930.

Kudo, S., Y. Nomura, M. Segawa, N. Fujita, M. Nakao, C. Schanen and M. Tamura (2003). "Heterogeneity in residual function of MeCP2 carrying missense mutations in the methyl CpG binding domain." Journal of Medical Genetics 40(7): 487-493.

Kumar, A., S. Kamboj, B. M. Malone, S. Kudo, J. L. Twiss, K. J. Czymmek, J. M. LaSalle and N. C. Schanen (2008). "Analysis of protein domains and Rett syndrome mutations indicate that multiple regions influence chromatin-binding dynamics of the chromatin-associated protein MECP2 in vivo." Journal of cell science 121(7): 1128-1137.

Laflamme, G. and K. Mekhail (2020). "Biomolecular condensates as arbiters of biochemical reactions inside the nucleus." Communications Biology 3(1): 1-8.

Laget, S., M. Joulie, F. Le Masson, N. Sasai, E. Christians, S. Pradhan, R. J. Roberts and P. A. Defossez (2010). "The human proteins MBD5 and MBD6 associate with heterochromatin but they do not bind methylated DNA." PLoS One 5(8): e11982-e11982.

Laget, S., B. Miotto, H. G. Chin, P. O. Estève, R. J. Roberts, S. Pradhan and P. A. Defossez (2014). "MBD4 cooperates with DNMT1 to mediate methyl-DNA repression and protects mammalian cells from oxidative stress." Epigenetics 9(4): 546-556.

Lagger, S., J. C. Connelly, G. Schweikert, S. Webb, J. Selfridge, B. H. Ramsahoye, M. Yu, C. He, G. Sanguinetti and L. C. Sowers (2017). "MeCP2 recognizes cytosine

methylated tri-nucleotide and di-nucleotide sequences to tune transcription in the mammalian brain." PLoS genetics 13(5): e1006793-e1006793.

Larson, A. G., D. Elnatan, M. M. Keenen, M. J. Trnka, J. B. Johnston, A. L. Burlingame, D. A. Agard, S. Redding and G. J. Narlikar (2017). "Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin." Nature 547(7662): 236-240.

Le Guezennec, X., M. Vermeulen, A. B. Brinkman, W. A. Hoeijmakers, A. Cohen, E. Lasonder and H. G. Stunnenberg (2006). "MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties." Mol Cell Biol 26(3): 843-851.

Lee, D.-H., P. Singh, S. Y. Tsai, N. Oates, A. Spalla, C. Spalla, L. Brown, G. Rivas, G. Larson and T. A. Rauch (2010). "CTCF-dependent chromatin bias constitutes transient epigenetic memory of the mother at the H19-Igf2 imprinting control region in prospermatogonia." PLoS genetics 6(11): e1001224-e1001224.

Lee, J. H., S. J. Park and K. Nakai (2017). "Differential landscape of non-CpG methylation in embryonic stem cells and neurons caused by DNMT3s." Sci Rep 7(1): 11295-11295.

Lee, M. R., N. Prasain, H. D. Chae, Y. J. Kim, C. Mantel, M. C. Yoder and H. E. Broxmeyer (2013). "Epigenetic regulation of Nanog by MiR‐302 cluster‐MBD2 completes induced pluripotent stem cell reprogramming." Stem Cells 31(4): 666-681.

Leighton, G. and D. C. Williams Jr (2020). "The methyl-CpG–Binding domain 2 and 3 proteins and formation of the nucleosome remodeling and deacetylase complex." Journal of molecular biology 432(6): 1624-1639.

Leoh, L. S., B. van Heertum, J. De Rijck, M. Filippova, L. Rios-Colon, A. Basu, S. R. Martinez, S. S. Tungteakkhun, V. Filippov, F. Christ, M. De Leon, Z. Debyser and C. A. Casiano (2012). "The stress oncoprotein LEDGF/p75 interacts with the methyl CpG binding protein MeCP2 and influences its transcriptional activity." Mol Cancer Res 10(3): 378-391.

Levone, B. R., S. C. Lenzken, M. Antonaci, A. Maiser, A. Rapp, F. Conte, S. Reber, J. Mechtersheimer, A. E. Ronchi, O. Mühlemann, H. Leonhardt, M. C. Cardoso, M. D. Ruepp and S. M. L. Barabino (2021). "FUS-dependent liquid-liquid phase separation is important for DNA repair initiation." J Cell Biol 220(5).

Lewis, A. and A. Murrell (2004). "Genomic imprinting: CTCF protects the boundaries." Current Biology 14(7): R284-R286.

Lewis, J. D., R. R. Meehan, W. J. Henzel, I. Maurer-Fogy, P. Jeppesen, F. Klein and A. Bird (1992). "Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA." Cell 69(6): 905-914.

Li, C. H., E. L. Coffey, A. Dall'Agnese, N. M. Hannett, X. Tang, J. E. Henninger, J. M. Platt, O. Oksuz, A. V. Zamudio and L. K. Afeyan (2020). "MeCP2 links heterochromatin condensates and neurodevelopmental disease." Nature 586(7829): 440-444.

Li, E., T. H. Bestor and R. Jaenisch (1992). "Targeted mutation of the DNA methyltransferase gene results in embryonic lethality." Cell 69(6): 915-926.

Li, L., N. Li, N. Liu, F. Huo and J. Zheng (2020). "MBD2 Correlates with a Poor Prognosis and Tumor Progression in Renal Cell Carcinoma." OncoTargets and therapy 13: 10001-10001.

Li, P., S. Banjade, H.-C. Cheng, S. Kim, B. Chen, L. Guo, M. Llaguno, J. V. Hollingsworth, D. S. King and S. F. Banani (2012). "Phase transitions in the assembly of multivalent signalling proteins." Nature 483(7389): 336-336.

Li, S., B. Wu, Y. Ling, M. Guo, B. Qin, X. Ren, C. Wang, H. Yang, L. Chen, Y. Liao, Y. Liu, X. Peng, C. Xu, Z. Wang, Y. Shen, J. Chen, L. Liu, B. Niu, M. Zhu, L. Liu, F. Li, T. Zhu, Z. Zhu, X. Zhou and H. Lu (2021). "Epigenetic Landscapes of Single-Cell Chromatin Accessibility and Transcriptomic Immune Profiles of T Cells in COVID-19 Patients." Front Immunol 12: 625881-625881.

Li, X., B. Z. Barkho, Y. Luo, R. D. Smrt, N. J. Santistevan, C. Liu, T. Kuwabara, F. H. Gage and X. Zhao (2008). "Epigenetic regulation of the stem cell mitogen Fgf-2 by Mbd1 in adult neural stem/progenitor cells." J Biol Chem 283(41): 27644-27652.

Li, Y., H. Wang, J. Muffat, A. W. Cheng, D. A. Orlando, J. Lovén, S.-m. Kwok, D. A. Feldman, H. S. Bateup and Q. Gao (2013). "Global transcriptional and translational repression in human-embryonic-stem-cell-derived Rett syndrome neurons." Cell Stem Cell 13(4): 446-458.

Li, Z., J. Xia, M. Fang and Y. Xu (2019). "Epigenetic regulation of lung cancer cell proliferation and migration by the chromatin remodeling protein BRG1." Oncogenesis 8(11): 66-66.

Liang, M. R., Y. Zeng, S. Y. Zeng, J. W. Zhang, B. C. Yang, Z. Y. Zhang, F. Y. Liu, Y. Luo, Y. Zou, F. Wang and O. P. Huang (2019). "The Expression of MBD6 Is Associated with Tumor Size in Uterine Leiomyomas." Genet Test Mol Biomarkers 23(8): 523-532.

Liao, W., M. Li, H. Wu, S. Jia, N. Zhang, Y. Dai, M. Zhao and Q. Lu (2017). "Down-regulation of MBD4 contributes to hypomethylation and overexpression of CD70 in CD4(+) T cells in systemic lupus erythematosus." Clin Epigenetics 9: 104-104.

Lienert, F., C. Wirbelauer, I. Som, A. Dean, F. Mohn and D. Schübeler (2011). "Identification of genetic elements that autonomously determine DNA methylation states." Nature genetics 43(11): 1091-1097.

Lister, R. and J. R. Ecker (2009). "Finding the fifth base: genome-wide sequencing of cytosine methylation." Genome research 19(6): 959-966.

Lister, R., E. A. Mukamel, J. R. Nery, M. Urich, C. A. Puddifoot, N. D. Johnson, J. Lucero, Y. Huang, A. J. Dwork, M. D. Schultz, M. Yu, J. Tonti-Filippini, H. Heyn, S. Hu, J. C. Wu, A. Rao, M. Esteller, C. He, F. G. Haghighi, T. J. Sejnowski, M. M. Behrens and J. R. Ecker (2013). "Global epigenomic reconfiguration during mammalian brain development." Science 341(6146): 1237905-1237905.

Lister, R., M. Pelizzola, R. H. Dowen, R. D. Hawkins, G. Hon, J. Tonti-Filippini, J. R. Nery, L. Lee, Z. Ye, Q. M. Ngo, L. Edsall, J. Antosiewicz-Bourget, R. Stewart, V. Ruotti, A. H. Millar, J. A. Thomson, B. Ren and J. R. Ecker (2009). "Human DNA methylomes at base resolution show widespread epigenomic differences." Nature 462(7271): 315-322.

Liu, C., Z. Q. Teng, N. J. Santistevan, K. E. Szulwach, W. Guo, P. Jin and X. Zhao (2010). "Epigenetic regulation of miR-184 by MBD1 governs neural stem cell proliferation and differentiation." Cell Stem Cell 6(5): 433-444.

Liu, K., C. Xu, M. Lei, A. Yang, P. Loppnau, T. R. Hughes and J. Min (2018). "Structural basis for the ability of MBD domains to bind methyl-CG and TG sites in DNA." J Biol Chem 293(19): 7344-7354.

Lodygin, D., J. Diebold and H. Hermeking (2004). "Prostate cancer is characterized by epigenetic silencing of 14-3-3σ expression." Oncogene 23(56): 9034-9041.

Loi, M., V. Cenni, S. Duchi, S. Squarzoni, C. Lopez-Otin, R. Foisner, G. Lattanzi and C. Capanni (2016). "Barrier-to-autointegration factor (BAF) involvement in prelamin A-related chromatin organization changes." Oncotarget 7(13): 15662-15662.

Lu, J., J. Qian, Z. Xu, S. Yin, L. Zhou, S. Zheng and W. Zhang (2021). "Emerging Roles of Liquid-Liquid Phase Separation in Cancer: From Protein Aggregation to Immune-Associated Signaling." Front Cell Dev Biol 9: 631486-631486.

Lu, Y., Y.-H. Loh, H. Li, M. Cesana, S. B. Ficarro, J. R. Parikh, N. Salomonis, C.-X. D. Toh, S. T. Andreadis and C. J. Luckey (2014). "Alternative splicing of MBD2 supports self-renewal in human pluripotent stem cells." Cell Stem Cell 15(1): 92-101.

Lücking, C. B. and A. Brice (2000). "Alpha-synuclein and Parkinson's disease." Cellular and Molecular Life Sciences CMLS 57(13): 1894-1908.

Ludwig, A. K., P. Zhang and M. C. Cardoso (2016). "Modifiers and readers of DNA modifications and their impact on genome structure, expression, and stability in disease." Frontiers in genetics 7: 115-115.

Ludwig, A. K., P. Zhang, F. D. Hastert, S. Meyer, C. Rausch, H. D. Herce, U. Müller, A. Lehmkuhl, I. Hellmann and C. Trummer (2017). "Binding of MBD proteins to DNA

blocks Tet1 function thereby modulating transcriptional noise." Nucleic acids research 45(5): 2438-2457.

Lunyak, V. V., R. Burgess, G. G. Prefontaine, C. Nelson, S.-H. Sze, J. Chenoweth, P. Schwartz, P. A. Pevzner, C. Glass and G. Mandel (2002). "Corepressor-dependent silencing of chromosomal regions encoding neuronal genes." Science 298(5599): 1747-1752.

Luo, S. W., C. Zhang, B. Zhang, C. H. Kim, Y. Z. Qiu, Q. S. Du, L. Mei and W. C. Xiong (2009). "Regulation of heterochromatin remodelling and myogenin expression during muscle differentiation by FAK interaction with MBD2." The EMBO journal 28(17): 2568-2582.

Lyst, M. J., J. Connelly, C. Merusi and A. Bird (2016). "Sequence‐specific DNA binding by AT‐hook motifs in Me CP 2." FEBS letters 590(17): 2927-2933.

Lyst, M. J., R. Ekiert, D. H. Ebert, C. Merusi, J. Nowak, J. Selfridge, J. Guy, N. R. Kastan, N. D. Robinson and F. de Lima Alves (2013). "Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor." Nature neuroscience 16(7): 898-898.

Macdonald, J. L., A. Verster, A. Berndt and A. J. Roskams (2010). "MBD2 and MeCP2 regulate distinct transitions in the stage-specific differentiation of olfactory receptor neurons." Mol Cell Neurosci 44(1): 55-67.

Magdinier, F. and A. P. Wolffe (2001). "Selective association of the methyl-CpG binding protein MBD2 with the silent p14/p16 locus in human neoplasia." Proc Natl Acad Sci U S A 98(9): 4990-4995.

Markouli, M., D. Strepkos and C. Piperi (2021). "Structure, Activity and Function of the SETDB1 Protein Methyltransferase." Life (Basel) 11(8).

Martin, V., H. F. Jorgensen, A. S. Chaubert, J. Berger, H. Barr, P. Shaw, A. Bird and P. Chaubert (2008). "MBD2-mediated transcriptional repression of the p14ARF tumor suppressor gene in human colon cancer cells." Pathobiology 75(5): 281-287.

May, S., H. Owen, T. J. Phesse, K. R. Greenow, G. R. Jones, A. Blackwood, P. C. Cook, C. Towers, A. M. Gallimore, G. T. Williams, M. Stürzl, N. Britzen-Laurent, O. J. Sansom, A. S. MacDonald, A. P. Bird, A. R. Clarke and L. Parry (2018). "Mbd2 enables tumourigenesis within the intestine while preventing tumour-promoting inflammation." J Pathol 245(3): 270-282.

McCarty, M. and O. T. Avery (1946). "Studies on the chemical nature of the substance inducing transformation of pneumococcal types: II. Effect of desoxyribonuclease on the biological activity of the transforming substance." The Journal of experimental medicine 83(2): 89-89.

McCord, R. P., A. Nazario-Toole, H. Zhang, P. S. Chines, Y. Zhan, M. R. Erdos, F. S. Collins, J. Dekker and K. Cao (2013). "Correlated alterations in genome organization, histone methylation, and DNA–lamin A/C interactions in Hutchinson-Gilford progeria syndrome." Genome Research 23(2): 260-269.

McCullough, A. K., M. L. Dodson and R. S. Lloyd (1999). "Initiation of base excision repair: glycosylase mechanisms and structures." Annu Rev Biochem 68: 255-285.

McSwiggen, D. T., M. Mir, X. Darzacq and R. Tjian (2019). "Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences." Genes & Development 33(23-24): 1619-1634.

Meehan, R. R., J. D. Lewis, S. McKay, E. L. Kleiner and A. P. Bird (1989). "Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs." Cell 58(3): 499-507.

Mellén, M., P. Ayata, S. Dewell, S. Kriaucionis and N. Heintz (2012). "MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system." Cell 151(7): 1417-1430.

Menafra, R., A. B. Brinkman, F. Matarese, G. Franci, S. J. Bartels, L. Nguyen, T. Shimbo, P. A. Wade, N. C. Hubner and H. G. Stunnenberg (2014). "Genome-wide binding of MBD2 reveals strong preference for highly methylated loci." PLoS One 9(6): e99603-e99603.

Meng, H., D. J. Harrison and R. R. Meehan (2015). "MBD4 interacts with and recruits USP7 to heterochromatic foci." J Cell Biochem 116(3): 476-485.

Millar, C. B., J. Guy, O. J. Sansom, J. Selfridge, E. MacDougall, B. Hendrich, P. D. Keightley, S. M. Bishop, A. R. Clarke and A. Bird (2002). "Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice." Science 297(5580): 403-405.

Minton, A. P. (2001). "The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media." Journal of Biological Chemistry 276(14): 10577-10580.

Misteli, T., A. Gunjan, R. Hock, M. Bustin and D. T. Brown (2000). "Dynamic binding of histone H1 to chromatin in living cells." Nature 408(6814): 877-877.

Mo, K., S. Wu, X. Gu, M. Xiong, W. Cai, F. E. Atianjoh, E. E. Jobe, X. Zhao, W. F. Tu and Y. X. Tao (2018). "MBD1 Contributes to the Genesis of Acute Pain and Neuropathic Pain by Epigenetic Silencing of Oprm1 and Kcna2 Genes in Primary Sensory Neurons." J Neurosci 38(46): 9883-9899.

Mok, R. S. F., W. Zhang, T. I. Sheikh, I. R. Fernandes, L. C. DeJong, M. R. Hildebrandt, M. Mufteev, D. C. Rodrigues, W. Wei and A. Piekna (2020). "Wide phenotypic spectrum of human stem cell-derived excitatory neurons with Rett syndrome-associated MECP2 mutations." bioRxiv.

Molliex, A., J. Temirov, J. Lee, M. Coughlin, A. P. Kanagaraj, H. J. Kim, T. Mittag and J. P. Taylor (2015). "Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization." Cell 163(1): 123-133.

Moore, M. A. (1977). "Theory of the polymer coil-globule transition." Journal of Physics A: Mathematical and General 10(2): 305-305.

Mor, N., Y. Rais, D. Sheban, S. Peles, A. Aguilera-Castrejon, A. Zviran, D. Elinger, S. Viukov, S. Geula, V. Krupalnik, M. Zerbib, E. Chomsky, L. Lasman, T. Shani, J. Bayerl, O. Gafni, S. Hanna, J. D. Buenrostro, T. Hagai, H. Masika, G. Vainorius, Y. Bergman, W. J. Greenleaf, M. A. Esteban, U. Elling, Y. Levin, R. Massarwa, Y. Merbl, N. Novershtern and J. H. Hanna (2018). "Neutralizing Gatad2a-Chd4-Mbd3/NuRD Complex Facilitates Deterministic Induction of Naive Pluripotency." Cell Stem Cell 23(3): 412-425.e410.

Moréra, S., I. Grin, A. Vigouroux, S. Couvé, V. Henriot, M. Saparbaev and A. A. Ishchenko (2012). "Biochemical and structural characterization of the glycosylase domain of MBD4 bound to thymine and 5-hydroxymethyuracil-containing DNA." Nucleic Acids Res 40(19): 9917-9926.

Mullegama, S. V. and S. H. Elsea (2016). "Clinical and Molecular Aspects of MBD5-Associated Neurodevelopmental Disorder (MAND)." Eur J Hum Genet 24(9): 1235-1243.

Müller‐Ott, K., F. Erdel, A. Matveeva, J. P. Mallm, A. Rademacher, M. Hahn, C. Bauer, Q. Zhang, S. Kaltofen and G. Schotta (2014). "Specificity, propagation, and memory of pericentric heterochromatin." Molecular systems biology 10(8): 746-746.

Muschol, M. and F. Rosenberger (1997). "Liquid–liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization." The Journal of chemical physics 107(6): 1953-1962.

Muyle, A. and B. S. Gaut (2019). "Loss of gene body methylation in Eutrema salsugineum is associated with reduced gene expression." Molecular biology and evolution 36(1): 155-158.

Myers, K. A., C. Marini, G. L. Carvill, A. McTague, J. Panetta, C. Stutterd, T. Stanley, S. Marin, J. Nguyen, C. Barba, A. Rosati, R. H. Scott, H. C. Mefford, R. Guerrini and I. E. Scheffer (2021). "Phenotypic Spectrum of Seizure Disorders in MBD5-Associated Neurodevelopmental Disorder." Neurol Genet 7(2): e579-e579.

Nan, X., F. J. Campoy and A. Bird (1997). "MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin." Cell 88(4): 471-481.

Nan, X., R. R. Meehan and A. Bird (1993). "Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2." Nucleic Acids Research 21(21): 4886-4892.

Nan, X., H. H. Ng, C. A. Johnson, C. D. Laherty, B. M. Turner, R. N. Eisenman and A. Bird (1998). "Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex." Nature 393(6683): 386-389.

Nativio, R., K. S. Wendt, Y. Ito, J. E. Huddleston, S. Uribe-Lewis, K. Woodfine, C. Krueger, W. Reik, J. M. Peters and A. Murrell (2009). "Cohesin Is Required for Higher-Order Chromatin Conformation at the Imprinted IGF2-H19 Locus." PLoS genetics 5(11).

Neul, J. L., P. Fang, J. Barrish, J. Lane, E. B. Caeg, E. O. Smith, H. Zoghbi, A. Percy and D. G. Glaze (2008). "Specific mutations in methyl-CpG-binding protein 2 confer different severity in Rett syndrome." Neurology 70(16): 1313-1321.

Ng, H.-H., Y. Zhang, B. Hendrich, C. A. Johnson, B. M. Turner, H. Erdjument-Bromage, P. Tempst, D. Reinberg and A. Bird (1999). "MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex." Nature genetics 23(1): 58-61.

Ng, H. H., P. Jeppesen and A. Bird (2000). "Active repression of methylated genes by the chromosomal protein MBD1." Mol Cell Biol 20(4): 1394-1406.

Ng, H. H., F. Robert, R. A. Young and K. Struhl (2003). "Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity." Mol Cell 11(3): 709-719.

Nielsen, J. B., K. F. Henriksen, C. Hansen, A. Silahtaroglu, M. Schwartz and N. Tommerup (2001). "MECP2 mutations in Danish patients with Rett syndrome: high frequency of mutations but no consistent correlations with clinical severity or with the X chromosome inactivation pattern." European Journal of Human Genetics 9(3): 178-184.

Nikitina, T., R. P. Ghosh, R. A. Horowitz-Scherer, J. C. Hansen, S. A. Grigoryev and C. L. Woodcock (2007). "MeCP2-chromatin interactions include the formation of chromatosome-like structures and are altered in mutations causing Rett syndrome." Journal of Biological Chemistry 282(38): 28237-28245.

Nikitina, T., X. Shi, R. P. Ghosh, R. A. Horowitz-Scherer, J. C. Hansen and C. L. Woodcock (2007). "Multiple modes of interaction between the methylated DNA binding protein MeCP2 and chromatin." Molecular and cellular biology 27(3): 864-877.

Nishikawaji, T., Y. Akiyama, S. Shimada, K. Kojima, T. Kawano, Y. Eishi, Y. Yuasa and S. Tanaka (2016). "Oncogenic roles of the SETDB2 histone methyltransferase in gastric cancer." Oncotarget 7(41): 67251-67265.

Nissar, B., S. A. Kadla, K. A. Wani, I. A. Shah and B. A. Ganai (2021). "Promoter CpG island hypermethylation and down regulation of XRCC1 gene can augment in the gastric carcinogenesis events." Molecular Biology Reports 48(1): 405-412.

Noma, K., C. D. Allis and S. I. Grewal (2001). "Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries." Science 293(5532): 1150-1155.

Oldfield, C. J. and A. K. Dunker (2014). "Intrinsically disordered proteins and intrinsically disordered protein regions." Annual review of biochemistry 83: 553-584.

Olson, C. O., S. Pejhan, D. Kroft, K. Sheikholeslami, D. Fuss, M. Buist, A. A. Sher, M. R. Del Bigio, Y. Sztainberg and V. M. Siu (2018). "MECP2 Mutation Interrupts Nucleolin–mTOR–P70S6K Signaling in Rett Syndrome Patients." Frontiers in genetics 9.

Olson, C. O., R. M. Zachariah, C. D. Ezeonwuka, V. R. B. Liyanage and M. Rastegar (2014). "Brain region-specific expression of MeCP2 isoforms correlates with DNA methylation within Mecp2 regulatory elements." PLoS One 9(3): e90645-e90645.

Ooi, S. K. T., C. Qiu, E. Bernstein, K. Li, D. Jia, Z. Yang, H. Erdjument-Bromage, P. Tempst, S.-P. Lin and C. D. Allis (2007). "DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA." Nature 448(7154): 714-717.

Pakravan, D., G. Orlando, V. Bercier and L. Van Den Bosch (2021). "Role and therapeutic potential of liquid-liquid phase separation in amyotrophic lateral sclerosis." J Mol Cell Biol 13(1): 15-28.

Pal, S., S. N. Vishwanath, H. Erdjument-Bromage, P. Tempst and S. Sif (2004). "Human SWI/SNF-associated PRMT5 methylates histone H3 arginine 8 and negatively regulates expression of ST7 and NM23 tumor suppressor genes." Molecular and Cellular Biology 24(21): 9630-9645.

Parker, M. W., M. Bell, M. Mir, J. A. Kao, X. Darzacq, M. R. Botchan and J. M. Berger (2019). "A new class of disordered elements controls DNA replication through initiator self-assembly." Elife 8.

Patil, V., C. Cuenin, F. Chung, J. R. R. Aguilera, N. Fernandez-Jimenez, I. Romero-Garmendia, J. R. Bilbao, V. Cahais, J. Rothwell and Z. Herceg (2019). "Human mitochondrial DNA is extensively methylated in a non-CpG context." Nucleic Acids Res 47(19): 10072-10085.

Pearson, E. C., D. L. Bates, T. D. Prospero and J. O. Thomas (1984). "Neuronal nuclei and glial nuclei from mammalian cerebral cortex: nucleosome repeat lengths, DNA contents and H1 contents." European journal of biochemistry 144(2): 353-360.

Peña-Hernández, R., R. Aprigliano, S. Carina Frommel, K. Pietrzak, S. Steiger, M. Roganowicz, L. Lerra, J. Bizzarro and R. Santoro (2021). "BAZ2A-mediated repression via H3K14ac-marked enhancers promotes prostate cancer stem cells." EMBO Rep: e53014-e53014.

Penagos-Puig, A. and M. Furlan-Magaril (2020). "Heterochromatin as an important driver of genome organization." Frontiers in Cell and Developmental Biology 8: 982-982.

Pessina, F., U. Gioia, O. Brandi, S. Farina, M. Ceccon, S. Francia and F. d'Adda di Fagagna (2021). "DNA Damage Triggers a New Phase in Neurodegeneration." Trends Genet 37(4): 337-354.

Peters, A. H. F. M., D. O'Carroll, H. Scherthan, K. Mechtler, S. Sauer, C. Schöfer, K. Weipoltshammer, M. Pagani, M. Lachner and A. Kohlmaier (2001). "Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability." Cell 107(3): 323-337.

Phair, R. D., S. A. Gorski and T. Misteli (2003). "Measurement of dynamic protein binding to chromatin in vivo, using photobleaching microscopy." Methods in enzymology 375: 393-414.

Piccolo, F. M., Z. Liu, P. Dong, C. L. Hsu, E. I. Stoyanova, A. Rao, R. Tjian and N. Heintz (2019). "MeCP2 nuclear dynamics in live neurons results from low and high affinity chromatin interactions." Elife 8.

Pidugu, L. S., H. Bright, W. J. Lin, C. Majumdar, R. P. Van Ostrand, S. S. David, E. Pozharski and A. C. Drohat (2021). "Structural Insights into the Mechanism of Base Excision by MBD4." J Mol Biol 433(15): 167097-167097.

Pitcher, M. R., J. A. Herrera, S. A. Buffington, M. Y. Kochukov, J. K. Merritt, A. R. Fisher, N. C. Schanen, M. Costa-Mattioli and J. L. Neul (2015). "Rett syndrome like phenotypes in the R255X Mecp2 mutant mouse are rescued by MECP2 transgene." Human molecular genetics 24(9): 2662-2672.

Poleshko, A., C. L. Smith, S. C. Nguyen, P. Sivaramakrishnan, K. G. Wong, J. I. Murray, M. Lakadamyali, E. F. Joyce, R. Jain and J. A. Epstein (2019). "H3K9me2 orchestrates inheritance of spatial positioning of peripheral heterochromatin through mitosis." Elife 8.

Pulukuri, S. M. and J. S. Rao (2006). "CpG island promoter methylation and silencing of 14-3-3σ gene expression in LNCaP and Tramp-C1 prostate cancer cell lines is associated with methyl-CpG-binding protein MBD2." Oncogene 25(33): 4559-4572.

Qin, W., A. Stengl, E. Ugur, S. Leidescher, J. Ryan, M. C. Cardoso and H. Leonhardt (2021). "HP1β carries an acidic linker domain and requires H3K9me3 for phase separation." Nucleus 12(1): 44-57.

Rai, K., I. J. Huggins, S. R. James, A. R. Karpf, D. A. Jones and B. R. Cairns (2008). "DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45." Cell 135(7): 1201-1212.

Ramsahoye, B. H., D. Biniszkiewicz, F. Lyko, V. Clark, A. P. Bird and R. Jaenisch (2000). "Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a." Proc Natl Acad Sci U S A 97(10): 5237-5242.

Rausch, C., P. Zhang, C. S. Casas-Delucchi, J. L. Daiß, C. Engel, G. Coster, F. D. Hastert, P. Weber and M. C. Cardoso (2021). "Cytosine base modifications regulate DNA duplex stability and metabolism." Nucleic Acids Res.

Ray, S., N. Singh, R. Kumar, K. Patel, S. Pandey, D. Datta, J. Mahato, R. Panigrahi, A. Navalkar and S. Mehra (2020). "α-Synuclein aggregation nucleates through liquid–liquid phase separation." Nature chemistry 12(8): 705-716.

Razin, A., C. Webb, M. Szyf, J. Yisraeli, A. Rosenthal, T. Naveh-Many, N. Sciaky-Gallili and H. Cedar (1984). "Variations in DNA methylation during mouse cell differentiation in vivo and in vitro." Proceedings of the National Academy of Sciences 81(8): 2275-2279.

Reber, S., D. Jutzi, H. Lindsay, A. Devoy, J. Mechtersheimer, B. R. Levone, M. Domanski, E. Bentmann, D. Dormann, O. Mühlemann, S. M. L. Barabino and M. D. Ruepp (2021). "The phase separation-dependent FUS interactome reveals nuclear and cytoplasmic function of liquid-liquid phase separation." Nucleic Acids Res 49(13): 7713-7731.

Reese, K. J., S. Lin, R. I. Verona, R. M. Schultz and M. S. Bartolomei (2007). "Maintenance of paternal methylation and repression of the imprinted H19 gene requires MBD3." PLoS Genet 3(8): e137-e137.

Reik, W. (2007). "Stability and flexibility of epigenetic gene regulation in mammalian development." Nature 447(7143): 425-432.

Reik, W., W. Dean and J. Walter (2001). "Epigenetic reprogramming in mammalian development." Science 293(5532): 1089-1093.

Riback, J. A., C. D. Katanski, J. L. Kear-Scott, E. V. Pilipenko, A. E. Rojek, T. R. Sosnick and D. A. Drummond (2017). "Stress-triggered phase separation is an adaptive, evolutionarily tuned response." Cell 168(6): 1028-1040. e1019.

Richards, E. J. and S. C. Elgin (2002). "Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects." Cell 108(4): 489-500.

Richter, K., M. Nessling and P. Lichter (2008). "Macromolecular crowding and its potential impact on nuclear function." Biochimica et Biophysica Acta (BBA)-Molecular Cell Research 1783(11): 2100-2107.

Rippe, K. (2007). "Dynamic organization of the cell nucleus." Current opinion in genetics & development 17(5): 373-380.

Robertson, K. D. (2005). "DNA methylation and human disease." Nat Rev Genet 6(8): 597-610.

Robertson, K. D. and A. P. Wolffe (2000). "DNA methylation in health and disease." Nat Rev Genet 1(1): 11-19.

Roloff, T. C., H. H. Ropers and U. A. Nuber (2003). "Comparative study of methyl-CpG-binding domain proteins." BMC Genomics 4(1): 1-1.

Rossetto, D., A. W. Truman, S. J. Kron and J. Côté (2010). "Epigenetic modifications in double-strand break DNA damage signaling and repair." Clin Cancer Res 16(18): 4543-4552.

Rothbauer, U., K. Zolghadr, S. Muyldermans, A. Schepers, M. C. Cardoso and H. Leonhardt (2008). "A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins." Molecular & Cellular Proteomics 7(2): 282-289.

Sabari, B. R., A. Dall'Agnese, A. Boija, I. A. Klein, E. L. Coffey, K. Shrinivas, B. J. Abraham, N. M. Hannett, A. V. Zamudio, J. C. Manteiga, C. H. Li, Y. E. Guo, D. S. Day, J. Schuijers, E. Vasile, S. Malik, D. Hnisz, T. I. Lee, Cisse, II, R. G. Roeder, P. A. Sharp, A. K. Chakraborty and R. A. Young (2018). "Coactivator condensation at super-enhancers links phase separation and gene control." Science 361(6400).

Saksouk, N., E. Simboeck and J. Déjardin (2015). "Constitutive heterochromatin formation and transcription in mammals." Epigenetics & chromatin 8(1): 1-17.

Salomon, R. and A. M. Kaye (1970). "Methylation of mouse DNA in vivo: di-and tripyrimidine sequences containing 5-methylcytosine." Biochimica et Biophysica Acta (BBA)-Nucleic Acids and Protein Synthesis 204(2): 340-351.

Samaco, R. C., J. D. Fryer, J. Ren, S. Fyffe, H. T. Chao, Y. Sun, J. J. Greer, H. Y. Zoghbi and J. L. Neul (2008). "A partial loss of function allele of methyl-CpG-binding protein 2 predicts a human neurodevelopmental syndrome." Hum Mol Genet 17(12): 1718-1727.

Sanborn, A. L., S. S. Rao, S.-C. Huang, N. C. Durand, M. H. Huntley, A. I. Jewett, I. D. Bochkov, D. Chinnappan, A. Cutkosky and J. Li (2015). "Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes." Proceedings of the National Academy of Sciences 112(47): E6456-E6465.

Sanulli, S., M. J. Trnka, V. Dharmarajan, R. W. Tibble, B. D. Pascal, A. L. Burlingame, P. R. Griffin, J. D. Gross and G. J. Narlikar (2019). "HP1 reshapes nucleosome core to promote phase separation of heterochromatin." Nature 575(7782): 390-394.

Saunders, C. J., B. E. Minassian, E. W. C. Chow, W. Zhao and J. B. Vincent (2009). "Novel exon 1 mutations in MECP2 implicate isoform MeCP2_e1 in classical Rett syndrome." American Journal of Medical Genetics Part A 149(5): 1019-1023.

Schanen, C., E. J. F. Houwink, N. Dorrani, J. Lane, R. Everett, A. Feng, R. M. Cantor and A. Percy (2004). "Phenotypic manifestations of MECP2 mutations in classical and atypical Rett syndrome." American Journal of Medical Genetics Part A 126(2): 129-140.

Schmidt, A., H. Zhang and M. C. Cardoso (2020). "MeCP2 and Chromatin Compartmentalization." Cells 9(4): 878-878.

Schmiedeberg, L., P. Skene, A. Deaton and A. Bird (2009). "A temporal threshold for formaldehyde crosslinking and fixation." PLoS One 4(2): e4636-e4636.

Schmolka, N., J. Bhaskaran, I. D. Karemaker and T. Baubec (2021). "Dissecting the roles of MBD2 isoforms in regulating NuRD complex function during cellular differentiation." bioRxiv.

Schubeler, D. (2015). "Function and information content of DNA methylation." Nature 517(7534): 321-326.

Schultz, D. C., K. Ayyanathan, D. Negorev, G. G. Maul and F. J. Rauscher 3rd (2002). "SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins." Genes Dev 16(8): 919-932.

Schulz, W. A., C. Steinhoff and A. R. Florl (2006). "Methylation of endogenous human retroelements in health and disease." DNA Methylation: Development, Genetic Disease and Cancer: 211-250.

Schuster, B. S., G. L. Dignon, W. S. Tang, F. M. Kelley, A. K. Ranganath, C. N. Jahnke, A. G. Simpkins, R. M. Regy, D. A. Hammer and M. C. Good (2020). "Identifying sequence perturbations to an intrinsically disordered protein that determine its phase-separation behavior." Proceedings of the National Academy of Sciences 117(21): 11421-11431.

Seabra, C. M., T. Aneichyk, S. Erdin, D. J. C. Tai, C. E. F. De Esch, P. Razaz, Y. An, P. Manavalan, A. Ragavendran, A. Stortchevoi, C. Abad, J. I. Young, P. Maciel, M. E. Talkowski and J. F. Gusella (2020). "Transcriptional consequences of MBD5 disruption in mouse brain and CRISPR-derived neurons." Mol Autism 11(1): 45-45.

Severin, P. M., X. Zou, H. E. Gaub and K. Schulten (2011). "Cytosine methylation alters DNA mechanical properties." Nucleic Acids Res 39(20): 8740-8751.

Shah, P. P., G. Donahue, G. L. Otte, B. C. Capell, D. M. Nelson, K. Cao, V. Aggarwala, H. A. Cruickshanks, T. S. Rai and T. McBryan (2013). "Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape." Genes & Development 27(16): 1787-1799.

Shakya, A. and J. T. King (2018). "DNA local-flexibility-dependent assembly of phase-separated liquid droplets." Biophysical journal 115(10): 1840-1847.

Sheikh, T. I., A. M. de Paz, S. Akhtar, J. Ausió and J. B. Vincent (2017). "MeCP2_E1 N-terminal modifications affect its degradation rate and are disrupted by the Ala2Val Rett mutation." Human molecular genetics 26(21): 4132-4141.

Shimbo, T., Y. Du, S. A. Grimm, A. Dhasarathy, D. Mav, R. R. Shah, H. Shi and P. A. Wade (2013). "MBD3 localizes at promoters, gene bodies and enhancers of active genes." PLoS Genet 9(12): e1004028-e1004028.

Shimoda, N., K. Hirose, R. Kaneto, T. Izawa, H. Yokoi, N. Hashimoto and Y. Kikuchi (2014). "No evidence for AID/MBD4-coupled DNA demethylation in zebrafish embryos." PLoS One 9(12): e114816-e114816.

Shin, Y. and C. P. Brangwynne (2017). "Liquid phase condensation in cell physiology and disease." Science 357(6357).

Simon, S. A. and B. C. Meyers (2011). "Small RNA-mediated epigenetic modifications in plants." Current opinion in plant biology 14(2): 148-155.

Sjolund, A. B., A. G. Senejani and J. B. Sweasy (2013). "MBD4 and TDG: multifaceted DNA glycosylases with ever expanding biological roles." Mutat Res 743-744: 12-25.

Skene, P. J., R. S. Illingworth, S. Webb, A. R. W. Kerr, K. D. James, D. J. Turner, R. Andrews and A. P. Bird (2010). "Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state." Molecular cell 37(4): 457-468.

Smith, Z. D. and A. Meissner (2013). "DNA methylation: roles in mammalian development." Nat Rev Genet 14(3): 204-220.

Sperlazza, M. J., S. M. Bilinovich, L. M. Sinanan, F. R. Javier and D. C. Williams Jr (2017). "Structural Basis of MeCP2 Distribution on Non-CpG Methylated and Hydroxymethylated DNA." J Mol Biol 429(10): 1581-1594.

Stirzaker, C., J. Z. Song, W. Ng, Q. Du, N. J. Armstrong, W. J. Locke, A. L. Statham, H. French, R. Pidsley and F. Valdes-Mora (2017). "Methyl-CpG-binding protein MBD2 plays a key role in maintenance and spread of DNA methylation at CpG islands and shores in cancer." Oncogene 36(10): 1328-1338.

Strom, A. R., A. V. Emelyanov, M. Mir, D. V. Fyodorov, X. Darzacq and G. H. Karpen (2017). "Phase separation drives heterochromatin domain formation." Nature 547(7662): 241-245.

Su, L. K., K. W. Kinzler, B. Vogelstein, A. C. Preisinger, A. R. Moser, C. Luongo, K. A. Gould and W. F. Dove (1992). "Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene." Science 256(5057): 668-670.

Sun, Q. Y., L. W. Ding, J. F. Xiao, W. Chien, S. L. Lim, N. Hattori, L. Goodglick, D. Chia, V. Mah, M. Alavi, S. R. Kim, N. B. Doan, J. W. Said, X. Y. Loh, L. Xu, L. Z. Liu, H. Yang, T. Hayano, S. Shi, D. Xie, D. C. Lin and H. P. Koeffler (2015). "SETDB1 accelerates tumourigenesis by regulating the WNT signalling pathway." J Pathol 235(4): 559-570.

Suzuki, M., T. Yamada, F. Kihara-Negishi, T. Sakurai and T. Oikawa (2003). "Direct association between PU.1 and MeCP2 that recruits mSin3A-HDAC complex for PU.1-mediated transcriptional repression." Oncogene 22(54): 8688-8698.

Szulwach, K. E., X. Li, Y. Li, C. X. Song, H. Wu, Q. Dai, H. Irier, A. K. Upadhyay, M. Gearing, A. I. Levey, A. Vasanthakumar, L. A. Godley, Q. Chang, X. Cheng, C. He

and P. Jin (2011). "5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging." Nat Neurosci 14(12): 1607-1616.

Szyf, M. (2010). "Epigenetic therapeutics in autoimmune disease." Clin Rev Allergy Immunol 39(1): 62-77.

Tahiliani, M., K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu and L. Aravind (2009). "Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1." Science 324(5929): 930-935.

Talkowski, M. E., S. V. Mullegama, J. A. Rosenfeld, B. W. M. Van Bon, Y. Shen, E. A. Repnikova, J. Gastier-Foster, D. L. Thrush, S. Kathiresan and D. M. Ruderfer (2011). "Assessment of 2q23. 1 microdeletion syndrome implicates MBD5 as a single causal locus of intellectual disability, epilepsy, and autism spectrum disorder." The American Journal of human genetics 89(4): 551-563.

Tan, C. P. and S. Nakielny (2006). "Control of the DNA methylation system component MBD2 by protein arginine methylation." Mol Cell Biol 26(19): 7224-7235.

Tate, P. H. and A. P. Bird (1993). "Effects of DNA methylation on DNA-binding proteins and gene expression." Curr Opin Genet Dev 3(2): 226-231.

Tillotson, R., J. Cholewa-Waclaw, K. Chhatbar, J. Connelly, S. A. Kirschner, S. Webb, M. V. Koerner, J. Selfridge, D. Kelly and D. De Sousa (2020). "Neuronal non-CG methylation is an essential target for MeCP2 function." bioRxiv.

Tokunaga, M., N. Imamoto and K. Sakata-Sogawa (2008). "Highly inclined thin illumination enables clear single-molecule imaging in cells." Nature methods 5(2): 159-161.

Ukai, H., H. Ishii‐Oba, M. Ukai‐Tadenuma, T. Ogiu and H. Tsuji (2003). "Formation of an active form of the interleukin‐2/15 receptor β‐chain by insertion of the intracisternal A particle in a radiation‐induced mouse thymic lymphoma and its role in tumorigenesis." Molecular Carcinogenesis: Published in cooperation with the University of Texas MD Anderson Cancer Center 37(2): 110-119.

Vaillant, I. and J. Paszkowski (2007). "Role of histone and DNA methylation in gene regulation." Current opinion in plant biology 10(5): 528-533.

Van Esch, H., M. Bauters, J. Ignatius, M. Jansen, M. Raynaud, K. Hollanders, D. Lugtenberg, T. Bienvenu, L. R. Jensen, J. Gecz, C. Moraine, P. Marynen, J. P. Fryns and G. Froyen (2005). "Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males." Am J Hum Genet 77(3): 442-453.

Vanyushin, B. F. (2005). "Adenine methylation in eukaryotic DNA." Molecular Biology 39(4): 473-481.

Varley, K. E., J. Gertz, K. M. Bowling, S. L. Parker, T. E. Reddy, F. Pauli-Behn, M. K. Cross, B. A. Williams, J. A. Stamatoyannopoulos and G. E. Crawford (2013). "Dynamic DNA methylation across diverse human cell lines and tissues." Genome Research 23(3): 555-567.

Vashi, N. and M. J. Justice (2019). "Treating Rett syndrome: from mouse models to human therapies." Mammalian Genome 30(5): 90-110.

Vukic, M. and L. Daxinger (2019). "DNA methylation in disease: Immunodeficiency, Centromeric instability, Facial anomalies syndrome." Essays in biochemistry 63(6): 773-783.

Wade, P. A., A. Gegonne, P. L. Jones, E. Ballestar, F. Aubry and A. P. Wolffe (1999). "Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation." Nature genetics 23(1): 62-66.

Wade, P. A., P. L. Jones, D. Vermaak and A. P. Wolffe (1998). "A multiple subunit Mi-2 histone deacetylase from Xenopus laevis cofractionates with an associated Snf2 superfamily ATPase." Current Biology 8(14): 843-848.

Walavalkar, N. M., N. Gordon and D. C. Williams (2013). "Unique features of the anti-parallel, heterodimeric coiled-coil interaction between methyl-cytosine binding domain 2 (MBD2) homologues and GATA zinc finger domain containing 2A (GATAD2A/p66α)." Journal of Biological Chemistry 288(5): 3419-3427.

Walsh, C. P. and T. H. Bestor (1999). "Cytosine methylation and mammalian development." Genes & Development 13(1): 26-34.

Walter, J. and M. Paulsen Imprinting and disease, Elsevier.

Wang, C., F. Wang, Z. Li, Q. Cao, L. Huang and S. Chen (2018). "MeCP2-mediated epigenetic regulation in senescent endothelial progenitor cells." Stem cell research & therapy 9(1): 1-7.

Wang, H., Z.-Q. Huang, L. Xia, Q. Feng, H. Erdjument-Bromage, B. D. Strahl, S. D. Briggs, C. D. Allis, J. Wong and P. Tempst (2001). "Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor." Science 293(5531): 853-857.

Wang, L., Y. Gao, X. Zheng, C. Liu, S. Dong, R. Li, G. Zhang, Y. Wei, H. Qu and Y. Li (2019). "Histone modifications regulate chromatin compartmentalization by contributing to a phase separation mechanism." Molecular cell 76(4): 646-659. e646.

Wang, L., M. Hu, M.-Q. Zuo, J. Zhao, D. Wu, L. Huang, Y. Wen, Y. Li, P. Chen and X. Bao (2020). "Rett syndrome-causing mutations compromise MeCP2-mediated liquid–liquid phase separation of chromatin." Cell Research 30(5): 393-407.

Wang, L., Y. Liu, R. Han, U. H. Beier, R. M. Thomas, A. D. Wells and W. W. Hancock (2013). "Mbd2 promotes foxp3 demethylation and T-regulatory-cell function." Mol Cell Biol 33(20): 4106-4115.

Wang, X., M. J. Cairns and J. Yan (2019). "Super-enhancers in transcriptional regulation and genome organization." Nucleic Acids Res 47(22): 11481-11496.

Wang, X., J. Shi, G. Cai, E. Zheng, D. Liu, Z. Wu and Z. Li (2019). "Overexpression of MBD3 Improves Reprogramming of Cloned Pig Embryos." Cell Reprogram 21(5): 221-228.

Weaver, I. C. G., I. C. Hellstrom, S. E. Brown, S. D. Andrews, S. Dymov, J. Diorio, T.-Y. Zhang, M. Szyf and M. J. Meaney (2014). "The methylated-DNA binding protein MBD2 enhances NGFI-A (egr-1)-mediated transcriptional activation of the glucocorticoid receptor." Philosophical Transactions of the Royal Society B: Biological Sciences 369(1652): 20130513-20130513.

Wegiel, J., M. Flory, I. Kuchna, K. Nowicki, S. Y. Ma, H. Imaki, J. Wegiel, J. Frackowiak, B. M. Kolecka and T. Wierzba-Bobrowicz (2015). "Neuronal nucleus and cytoplasm volume deficit in children with autism and volume increase in adolescents and adults." Acta neuropathologica communications 3(1): 1-17.

Wen, L., X. Li, L. Yan, Y. Tan, R. Li, Y. Zhao, Y. Wang, J. Xie, Y. Zhang and C. Song (2014). "Whole-genome analysis of 5-hydroxymethylcytosine and 5-methylcytosine at base resolution in the human brain." Genome biology 15(3): 1-17.

Williams, S. R., S. V. Mullegama, J. A. Rosenfeld, A. I. Dagli, E. Hatchwell, W. P. Allen, C. A. Williams and S. H. Elsea (2010). "Haploinsufficiency of MBD5 associated with a syndrome involving microcephaly, intellectual disabilities, severe speech impairment, and seizures." Eur J Hum Genet 18(4): 436-441.

Williamson, A. K., Z. Zhu and Z. M. Yuan (2018). "Epigenetic mechanisms behind cellular sensitivity to DNA damage." Cell Stress 2(7): 176-180.

Wong, E., K. Yang, M. Kuraguchi, U. Werling, E. Avdievich, K. Fan, M. Fazzari, B. Jin, A. M. Brown, M. Lipkin and W. Edelmann (2002). "Mbd4 inactivation increases Cright-arrowT transition mutations and promotes gastrointestinal tumor formation." Proc Natl Acad Sci U S A 99(23): 14937-14942.

Wood, K. H., B. S. Johnson, S. A. Welsh, J. Y. Lee, Y. Cui, E. Krizman, E. S. Brodkin, J. A. Blendy, M. B. Robinson, M. S. Bartolomei and Z. Zhou (2016). "Tagging methyl-CpG-binding domain proteins reveals different spatiotemporal expression and supports distinct functions." Epigenomics 8(4): 455-473.

Wood, K. H. and Z. Zhou (2016). "Emerging Molecular and Biological Functions of MBD2, a Reader of DNA Methylation." Front Genet 7: 93-93.

Woodcock, C. B., J. R. Horton, J. Zhou, M. T. Bedford, R. M. Blumenthal, X. Zhang and X. Cheng (2020). "Biochemical and structural basis for YTH domain of human YTHDC1

binding to methylated adenine in DNA." Nucleic Acids Research 48(18): 10329-10341.

Woodcock, D. M., P. J. Crowther and W. P. Diver (1987). "The majority of methylated deoxycytidines in human DNA are not in the CpG dinucleotide." Biochem Biophys Res Commun 145(2): 888-894.

Wu, T. P., T. Wang, M. G. Seetin, Y. Lai, S. Zhu, K. Lin, Y. Liu, S. D. Byrum, S. G. Mackintosh and M. Zhong (2016). "DNA methylation on N 6-adenine in mammalian embryonic stem cells." Nature 532(7599): 329-333.

Xiao, C.-L., S. Zhu, M. He, D. Chen, Q. Zhang, Y. Chen, G. Yu, J. Liu, S.-Q. Xie and F. Luo (2018). "N6-methyladenine DNA modification in the human genome." Molecular Cell 71(2): 306-318. e307.

Xie, W., C. L. Barr, A. Kim, F. Yue, A. Y. Lee, J. Eubanks, E. L. Dempster and B. Ren (2012). "Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome." Cell 148(4): 816-831.

Xu, C., K. Liu, M. Lei, A. Yang, Y. Li, T. R. Hughes and J. Min (2018). "DNA Sequence Recognition of Human CXXC Domains and Their Structural Determinants." Structure 26(1): 85-95 e83.

Yakovlev, D. A., A. A. Kuznetsova, O. S. Fedorova and N. A. Kuznetsov (2017). "Search for Modified DNA Sites with the Human Methyl-CpG-Binding Enzyme MBD4." Acta Naturae 9(1): 88-98.

Yan, S. J., S. J. Lim, S. Shi, P. Dutta and W. X. Li (2011). "Unphosphorylated STAT and heterochromatin protect genome stability." The FASEB Journal 25(1): 232-241.

Yildirim, O., R. Li, J. H. Hung, P. B. Chen, X. Dong, L. S. Ee, Z. Weng, O. J. Rando and T. G. Fazzio (2011). "Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells." Cell 147(7): 1498-1510.

Yoder, J. A., C. P. Walsh and T. H. Bestor (1997). "Cytosine methylation and the ecology of intragenomic parasites." Trends in Genetics 13(8): 335-340.

Yu, X., A. Azzo, S. M. Bilinovich, X. Li, M. Dozmorov, R. Kurita, Y. Nakamura, D. C. Williams and G. D. Ginder (2019). "Disruption of the MBD2-NuRD complex but not MBD3-NuRD induces high level HbF expression in human adult erythroid cells." Haematologica 104(12): 2361-2371.

Yuan, L. and D.-S. Huang (2019). "A network-guided association mapping approach from DNA methylation to disease." Scientific Reports 9(1): 1-16.

Yusufzai, T. M. and A. P. Wolffe (2000). "Functional consequences of Rett syndrome mutations on human MeCP2." Nucleic acids research 28(21): 4172-4179.

Zbinden, A., M. Pérez-Berlanga, P. De Rossi and M. Polymenidou (2020). "Phase Separation and Neurodegenerative Diseases: A Disturbance in the Force." Dev Cell 55(1): 45-68.

Zhang, M., J. N. Kimatu, K. Xu and B. Liu (2010). "DNA cytosine methylation in plant development." Journal of Genetics and Genomics 37(1): 1-12.

Zhang, P., F. D. Hastert, A. K. Ludwig, K. Breitwieser, M. Hofstätter and M. C. Cardoso (2017). "DNA base flipping analytical pipeline." Biology Methods and Protocols 2(1): bpx010-bpx010.

Zhang, P., C. Rausch, F. D. Hastert, B. Boneva, A. Filatova, S. J. Patil, U. A. Nuber, Y. Gao, X. Zhao and M. C. Cardoso (2017). "Methyl-CpG binding domain protein 1 regulates localization and activity of Tet1 in a CXXC3 domain-dependent manner." Nucleic Acids Res 45(12): 7118-7136.

Zhang, Y., R. Iratni, H. Erdjument-Bromage, P. Tempst and D. Reinberg (1997). "Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex." Cell 89(3): 357-364.

Zhang, Y., G. LeRoy, H.-P. Seelig, W. S. Lane and D. Reinberg (1998). "The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities." Cell 95(2): 279-289.

Zhang, Y., H.-H. Ng, H. Erdjument-Bromage, P. Tempst, A. Bird and D. Reinberg (1999). "Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation." Genes & Development 13(15): 1924-1935.

Zhao, X., T. Ueba, B. R. Christie, B. Barkho, M. J. McConnell, K. Nakashima, E. S. Lein, B. D. Eadie, A. R. Willhoite, A. R. Muotri, R. G. Summers, J. Chun, K. F. Lee and F. H. Gage (2003). "Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function." Proc Natl Acad Sci U S A 100(11): 6777-6782.

Zhou, H., Z. Song, S. Zhong, L. Zuo, Z. Qi, L. J. Qu and L. Lai (2019). "Mechanism of DNA‐induced phase separation for transcriptional repressor VRN1." Angewandte Chemie 131(15): 4912-4916.

Zhou, H.-X., G. Rivas and A. P. Minton (2008). "Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences." Annu. Rev. Biophys. 37: 375-397.

Zhou, Y. and I. Grummt (2005). "The PHD finger/bromodomain of NoRC interacts with acetylated histone H4K16 and is sufficient for rDNA silencing." Curr Biol 15(15): 1434-1438.

Zhu, D., J. Fang, Y. Li and J. Zhang (2009). "Mbd3, a component of NuRD/Mi-2 complex, helps maintain pluripotency of mouse embryonic stem cells by repressing trophectoderm differentiation." PLoS One 4(11): e7684-e7684.

Zhu, D., S. B. Hunter, P. M. Vertino and E. G. Van Meir (2011). "Overexpression of MBD2 in glioblastoma maintains epigenetic silencing and inhibits the antiangiogenic function of the tumor suppressor gene BAI1." Cancer Res 71(17): 5859-5870.

Ziller, M. J., F. Muller, J. Liao, Y. Zhang, H. Gu, C. Bock, P. Boyle, C. B. Epstein, B. E. Bernstein, T. Lengauer, A. Gnirke and A. Meissner (2011). "Genomic distribution and inter-sample variation of non-CpG methylation across human cell types." PLoS Genet 7(12): e1002389-e1002389.

7 Appendix

7.1 Abbreviations

A

Apc^{Min}	Adenomatous Polyposis coli Min
AR	Aspect ratio
AT	Adenine-thymine
AT-hook	AT rich DNA binding motif with a predicted "hook" structure

B

BAZ2	Bromodomain adjacent to zinc finger domain protein 2 (BAZ2)
BER	Base excision repair
Bromo	Bromodomain

C

CBD	Chitin-binding domain
CC	Coiled-coil domain
CDK2AP1	Cyclin-Dependent Kinase 2 Associated Protein 1
CHD	Chromodomain Helicase DNA Binding Protein
ChIP-seq	Chromatin immunoprecipitation-sequencing
CRC	Colorectal cancer
CTCF	CCCTC-binding factor
CTD	C-terminal domain
CxxC	Unmethylated-CpG-binding zinc finger

D

DAPI	4,6-Diamidine2-phenylindole dihydrochloride
DC	Dendritic cells
DDT	DNA binding domain
DIC	Differential interference contrast
DMEM	Dulbecco’s modified Eagle medium
Dnmt	DNA methyltransferase
DRG	Dorsal root ganglion

E

ESC	Embryonic stem cell
EST	Expressed sequence tag
EtBr	Ethidium bromide

F

FAK	Focal adhesion kinase
FCS	Fetal calf serum
FRAP	Fluorescence recovery after photobleaching

G

G/R	Glycine/arginine
GC	Gastric cancer
Glycosylase	DNA glycosylase function

H

H3K27me3	Histone 3 lysine 27 trimethylation
H3K4ac	Histone H3 lysine 4 acetylation
H3K36me3	Histone 3 lysine 36 trimethylation
H3K9me3	Histone 3 lysine 9 trimethylation
H4R3me2	Dimethylation of histone H4 at arginine 3
HDAC	Histone deacetylase
HDCC	Histone deacetylase core sub-complex
HILO	Highly inclined laminar optical sheet illumination
HP1	Heterochromatin protein
5hmC	Hydroxymethyl-cytosine

I

ICM	Inner cell mass
ID	Intervening domain
IDR	Intrinsically disordered regions
ITP	Primary immune thrombocytopenia

K

Kd	Dissociation constants

L

LCR	Low complexity region
LLPS	Liquid-liquid phase separation

M

MajSat	Major satellite repeat
MAND	MBD5-associated neurodevelopmental disorder
MBD	Methyl-CpG binding domain
MBD1	Methyl-CpG binding domain protein 1
MBD2	Methyl-CpG binding domain protein 2
MBD3	Methyl-CpG binding domain protein 3
MBD4	Methyl-CpG binding domain protein 4
MBDs	MBD family proteins
MBP	Methyl-CpG binding proteins
5mC	5-methylcytosine
MCAF	MBD1-containing chromatin-associated factor
mCpG	Cytosine methylation at CpG dinucleotides
MeCP2	Methyl-CpG binding protein 2
MinFeret/ MaxFeret	Minimal/ Maximal Feret diameters (and
MST	Microscale thermophoresis
MTA	Metastasis Tumor-Associated
MTase	DNA Methyltransferase

N

NA	Nucleosomal array
NID	N-CoR/SMRT interacting domain
NLS	Nuclear localization signal
NSCLC	Non-small cell lung cancer
NSPC	Neural stem/progenitor cell
NTD	N-terminal domain
NuRD	Nucleosome remodeling and histone deacetylase

O

ORN	Olfactory receptor neuron

P

P-W-W-P	Proline-tryptophan-tryptophan-proline

Pab1	Poly(A)-binding protein
PBS	Phosphate buffered saline
PCa	Prostate cancer
PCR	Polymerase chain reaction
PEG 8000	Polyethylene Glycol 8000
pI	Iso-electric point
PPPS	Polymer-polymer phase separation
P rich	Proline rich domain
PRMT	Protein arginine methyltransferase
PTM	Post-translational modifications

R

RBBP	Retinoblastoma Binding Protein
RHA	RNA helicase A
RT	Room temperature
RTT	RETT syndrome

S

SETDB1	SET domain bifurcated histone lysine methyltransferase 1
Sin3a	SIN3 transcription regulator

T

t1/2	Half recovery time
TAD	Topologically associating domains
Tet	Ten-eleven translocation methylcytosine dioxygenase
Th	T helper
TRD	Transcriptional repression domain
TSS	Translational start site
Tud	Tudor domain

U

U	Uracil
UL	Uterine leiomyoma

S

α-Syn	α-Synuclein
SET	Suvar3-9, Enhancer-of-zeste, Trithorax domain

7.2 Declaration

Ehrenwörtliche Erklärung:

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit entsprechend den Regeln guter wissenschaftlicher Praxis selbstständig und ohne unzulässige Hilfe Dritter angefertigt habe.

Sämtliche aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie sämtliche von Anderen direkt oder indirekt übernommenen Daten, Techniken und Materialien sind als solche kenntlich gemacht. Die Arbeit wurde bisher bei keiner anderen Hochschule zu Prüfungszwecken eingereicht.

Darmstadt, den …27. 01. 2022…

Hui Zhang

.....

7.3 Acknowledgements

Firstly, I would like to express my sincere gratitude to Prof. Dr. M. Cristina Cardoso for giving me the opportunity to perform my PhD thesis. Her constant guidance and patient scientific discussions and suggestions helped me throughout the time of experiment and writing of the thesis. I can not imagine how my thesis would look without her supervision. She is more like a friend to me and also to all lab members. She has been giving me so many opportunities to improve myself in teaching and collaboration. All of these will benefit my scientific career. Last, I would like to thank her for all she has been doing to keep a safe and friendly atmosphere in the lab during the hard time of the coronavirus. She asked everyone to strictly follow the rules in and outside the lab and no one in the lab was infected. Without her optimistic character and proper organization, the lab could not work normally and I could not finish my thesis in time.

I would like to thank Prof. Dr. Bodo Laube, my second supervisor, and Prof. Dr. Heinrich Leonhardt for their scientific suggestions and contributions to my thesis.

Given my huge thanks to Hector Romero and Annika Schmidt for the interaction project about MeCP2 quantification and dynamics. I would like to thank Katalina Gagova for MBD1ΔCxxC3 purification and phase separation and Malte Eck for checking the phase separation properties of MBDs. I want to thank Annika Schmidt for the translation "**Zusammenfassung**".

Thanks to everyone in the Cardoso lab for their help in experiment design, instrument instructions, image analysis, questions, and suggestions. The Cardoso lab is an international lab. Although from different countries with different cultures, people in the lab are really nice and empathetic. It is nice to enjoy the food from different regions of the world. In particular, I want to thank Anne Lehmkuhl for preparing the cells, Manuela Milden for her suggestions in protein purification and instruction for multiple instruments, Diana Imblan for preparing the competent cells, and Alexander Rapp for providing the macros for FRAP analysis.

I would like to thank Peng Zhang and Weihua Qin for their help during the beginning of my study in Germany.

Xinmei Ma, I gave my most appreciation to you for spending so many pleasant times with you. You were always accompanying me during my hard times. Your smile cued me.

Also, I would like to thank the China Scholarship Council (CSC) for providing the opportunity to go abroad and study.

Lastly, I want to thank my family for their support in living and studying. Mom and dad, thanks for your belief and encouragement when I am down. Thanks to my older brother and sister-in-law for taking good care of mom and dad, especially when mom was sick.

8 Curriculum vitae

Personal Information

Name:	Hui Zhang
Gender:	Male
Date of birth:	08/04/1990
Nationality:	China
Contact address:	Technische Universität Darmstadt
	Schnittspahnstr.10
	64287 Darmstadt, Germany
E-mail:	zhanghui20139@gmail.com

Education Background

2017.09 – present PhD student

Phase Separation of the Methyl-CpG-Binding Domain (MBD) Protein Family Underlies Heterochromatin Organization

Technische Universität Darmstadt

Supervisor: Prof. M. Cristina Cardoso

2015.09 - 2017.07 Master in biology

Major in developmental and cellular biology

Harbin Institute of Technology

Supervisor: Prof. Yu Li

2011.09 - 2015.07 Bachelor in biotechnology

School of Life Science and Technology, Harbin Institute of Technology

Positions

2017.09 – 2021.08 PhD fellowship from China scholarship council (CSC)

2021.09 - 2022.02 Research assistant in the group of Cristina Cardoso

Teaching

Supervision and planning of the bachelor internship and bachelor thesis for Robin Bruckhoff and Malte Eck.

Supervision of practical course for master students.

Publications

Chen, Y., **Zhang, H.**, Han, F., Yue, L., Qiao, C., Zhang, Y., Dou, P., Liu, W., Li, Y. (2018) The depletion of MARVELD1 leads to murine placenta accreta via integrin β4-dependent trophoblast cell invasion, Journal of cellular physiology 233, 2257-2269.

Schmidt, A.#, **Zhang, H**.# and Cardoso, M. C. (2020) MeCP2 and chromatin compartmentalization. Cells 9: 878. (# equal contribution)

Zhang H.#, Romero H.#, Schmidt A., Gagova K., Qin W., Bertulat B., Lehmkuhl A., Milden M., Eck M., Meckel T., Leonhardt H., Cardoso, M. C. MeCP2-induced heterochromatin organization is driven by oligomerization-based liquid-liquid phase separation and restricted by DNA methylation. Nucleus (accepted). DOI: 10.1080/19491034.2021.2024691 (# equal contribution)

Manuscripts under preparation

Zhang H., Eck M., Gagova K., Cardoso, M. C. Phase separation is involved in the MBDs-induced heterochromatin organization.

Zhang H., Gagova K., Eck M., Cardoso, M. C. Isoform-specific influence of MBD2 in heterochromatin organization via phase separation.