Role of the PEST Domains in Proteasomal Degradation of Rett Protein: MeCP2

by

Ladan Kalani

B.Sc., University of Victoria, 2020

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

In the Department of Biochemistry and Microbiology

Supervisory Committee

Role of the PEST Domains in Proteasomal Degradation of Rett Protein: MeCP2

by

Ladan Kalani

B.Sc., University of Victoria, 2020

Dr. Juan Ausió, Department of Biochemistry and Microbiology
Supervisor

Dr. Caren C. Helbing, Department of Biochemistry and Microbiology
Departmental Member

Dr. Raad Nashmi, Department of Biology
Outside Member

Abstract

Located on the X-chromosome is the gene encoding the nuclear protein Methyl CpG binding protein 2 (MeCP2). The instability of this protein causes pleiotropic neurological abnormalities, including the debilitating neurodevelopmental disease Rett syndrome (RTT). MeCP2, an epigenetic regulator abundant in neurons, is involved in pleiotropic molecular interaction. Many deleterious mutations of MeCP2 impact its mRNA or protein levels. Neuron maturation and dendritic arborization are compromised when MeCP2 levels are out of the homeostatic range. The mechanisms the cell uses to maintain MeCP2 levels within a tight range have yet to be fully understood. Several hypotheses addressed the homeostatic mechanisms of MeCP2, which involve miRNAs, N-terminal degradation signals or N-degrons, and the PEST domains that act as degradation switches upon post-translational modifications (PTMs). Our lab hypothesized the involvement of MeCP2 PEST-mediated degradation as a mechanism of its homeostatic regulation; however, this hypothesis has yet to be experimentally proven. I experimentally tested the PEST-mediated degradation of MeCP2 with Rett-causing mutations by integrating MeCP2 constructs that have an altered or deleted PEST domain and used microscopy, FRAP analysis and western blotting to characterize *in vitro* how these constructs behave relative to WT and mutated MeCP2. MeCP2 has Rett-causing mutations that cause lower protein levels, such as T158M; the PEST motif expedites its degradation as deleting it results in higher protein levels. Moreover, mutations that result in higher levels of MeCP2, such as R294X, show stronger DNA binding relative to WT, as assessed by NaCl fractionation. For the first time, we report that the Ct-PEST domain of MeCP2 plays a role in its degradation.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AP	alkaline phosphatase
ASD	autism spectrum disorder
Asx/ST	asparagine/aspartate and serine/threonine turn created by hydrogen bonding
BME	beta-mercaptoethanol
CDKL5	cyclin dependant kinase-like 5
CREB1	cyclic AMP response element binding protein 1
CTD	C-terminal domain
DAPI	4', 6-diamidino-2-phenylindole
DMEM	Delbecco's modified eagle medium
DS	down syndrome
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ESC	embryonic stem cells
FOXG1	forkhead box protein 1
FRAP	fluorescence recovery after photobleaching
GPE	glycine-proline-glutamate
HAT	histone acetyltransferase
HMGN1	high-mobility group N1
HDAC	histone deacetylase
H4	Histone 4
ID	interveining domain
IDP	intrinsically disordered protein
IRSA	international Rett syndrome association
IRSF	international Rett syndrome foundation
LLPS	liquid-liquid phase separation
MBD	methyl binding domain

MBP	methyl binding protein
MeCP2	methyl CpG binding protein 2
MEF2C	myocyte enhancer factor 2C
miRNA	micro RNA
MRE	microRNA recognition element
MS	mass spectrometry
Myb	myeloblastosis, Myb family of genes are transcription factors
NCoR	nuclear receptor coreceptor
NID	NCoR interaction domain
ncRNA	noncoding RNA
NME	N-methionine excision
NRL	nucleosome repeat length
NTD	N-terminal domain
OD	optical density
ODC	ornithine decarboxylase
PAGE	polyacrylamide gel electrophoresis
PBS(T)	phosphate buffer saline (tween)
PCC	Pearson's correlation coefficient
PEST	proline (P), glutamic acid (E), serine (S) and threonine (T)
PTM	post-translational modifications
RE	regulatory elements
RNF4	Ring-finger protein 4
ROI	region of interest
RTT	Rett syndrome
SDS	sodium dodecyl sulphate
SMRT	silencing mediator of retinoic acid and thyroid hormone receptor coreceptor
Ste3p	a-factor receptor in yeast for sexual conjugation
TRD	transcription repression domain

T158M	threonine at position158 mutated to methionine
TBLR1	transduction β-ike protein 1-related protein
UTR	untranslated region
WHD	winged helix domain
WT	wildtype
XCI	X-chromosome inactivation

Acknowledgements

I express my deepest gratitude to my supervisor, Dr. Juan Ausió, for their invaluable mentoring and teachings in developing a critical perspective. I am immensely thankful to my colleagues Katrina Vanessa Good, Cindy Bo Hyun Kim and Dr. Carla Liria Sánchez-Lafuente for their incredible empathy, support and guidance. I am grateful to our collaborators, Dr. John B.Vincent for the generous donation of MeCP2 constructs, Dr. Michael J. Hendzel and Anastasia Roemer for the FRAP analysis, and Dr. Bob Chow and Alberto Ruiz De Chavez for the beautiful confocal fluorescence microscopy imaging. I also wish to thank my committee members, Dr. Caren C. Helbing and Dr. Raad Nashmi, for leading me to become a true scientist. My heartfelt appreciation goes to my parents, who made incredible sacrifices in helping me settle in Canada, and to my friends, especially Teya Sharir, for her genuine kindness and care and Arlo Levi Poole for being the loving catalyst to my accomplishments. Lastly, I thank God, the Spirit, who has bestowed on me this life and awareness, which is the greatest gift and responsibility.

Chapter 1: Introduction

1.1 Chromatin and MeCP2

The eukaryotic genome is highly organized in a dynamic structure called chromatin. 85% of DNA in chromatin participate in DNA-protein complexes called nucleosomes that can be arranged in uniform arrays inside the nucleus, classically referred to as "beads on a string" [1-3]. The nucleosome core particle is the fundamental building block of chromatin consisting of two copies of histone proteins—H2A, H2B, H3, and H4— that are put together into an octamer. 145–147 base pairs (bp) of DNA are wrapped 1.8 turn in a left-handed super-helix around the histone octamer [4,5]. A small stretch of linker DNA attaches the core particles. Incorporation of the linker histone H1 into chromatin stabilizes the nucleosome and allows the formation of higher-order structures that are more compact [6,7]. Linker histone H1 binds to the nucleosome's DNA entry and exit site. It is widely accepted that H1 quickly exchanges in living cells and binds DNA less tightly than the core histones. FRAP studies in Michael Hendzel's lab confirm the dynamic nature of linker histone [8-11].

In humans, each cell holds ~2m of genetic information that is meticulously organized within chromatin inside the small nuclear volume, and chromatin is continuously changing in response to developmental and environmental cues to regulate the accessibility of the transcriptional machinery. Numerous post-translational modifications (PTMs) are added to histones, such as acetylation, ubiquitination, and SUMOylation. The PTMs are found both on the globular part of histones and the histone tails. Like DNA modifications, such as the methylation of CpG dinucleotides, histone PTMs serve as signals where each mark conveys a specific message [12,13].

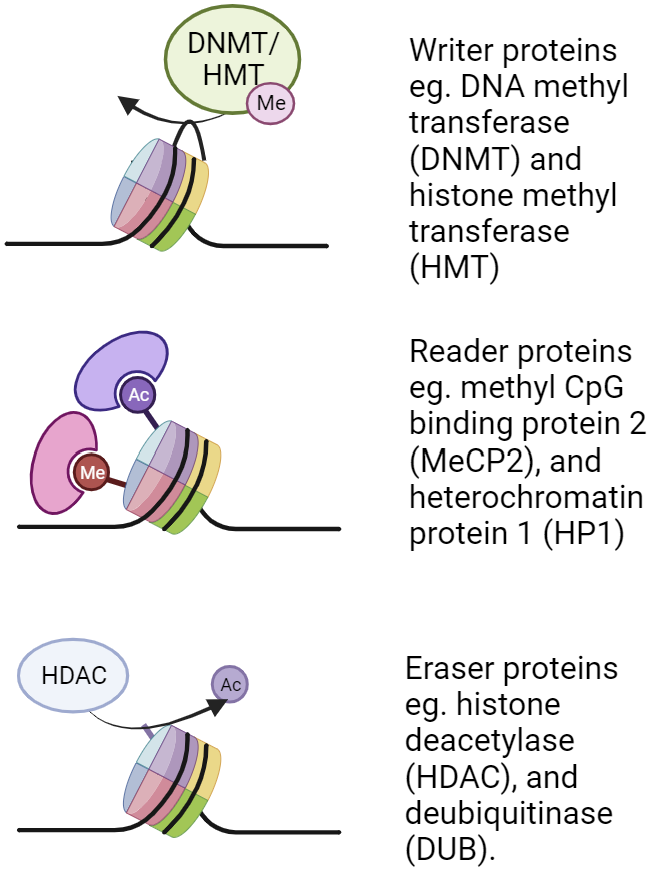


Figure 1.1. Reader, writer, and eraser proteins that regulate the epigenome on DNA and histones. Proteins mediating the epigenomic mark on DNA and histone tails change the accessibility of genes to the transcription machinery [12,13].

These chemical modifications are placed on DNA and histones by proteins called "writers," which deposit the epigenetic signature. Proteins that bind to the epigenetic marks and initiate a cascade of events are "readers." The proteins that remove the marks are "erasers," allowing vast flexibility and reversibility for many cellular processes or malfunctions **(Fig. 1.1)**. MicroRNAs (miRNA)s, which are a subset of noncoding RNAs (ncRNA) approximately 22 nucleotides long, have been recognized as another regulator of gene expression by pairing partially or entirely with

the target mRNA's 3'untranslated region (3'UTR) and preventing its translation and inducing its degradation accordingly [14-17]. Examples of epigenetic regulators are histone acetyltransferases (HATs), a writer protein that deposits acetyl groups on the epsilon amino group of lysine residues in histone tails [18], histone deacetylases (HDACs), an eraser protein removing acetyl marks from histone tails, and methyl binding proteins (MBPs), reader proteins that bind to methylated DNA or histones and recruit other proteins to regulate gene expression. Often, these regulators interact with each other and form protein complexes that induce changes in nucleosome compaction. Methyl CpG binding protein 2 (MeCP2) is a transcription factor and a methyl-DNA reader protein ubiquitously expressed in vertebrate organisms, yet it is most abundant in neurons at one per every two nucleosomes [19]**.** Like histone H1, MeCP2 binds to the nucleosome's DNA entry/exit site, allowing chromatin compaction. Although controversial, several researchers believe that MeCP2 and a linker histone variant $H1^0$ compete for binding to nucleosomes [20]; however, it is accepted that MeCP2 binds a plethora of binding partners and it is involved in both repression and activation of gene expression [21-23]. MeCP2, in conjugation with other epigenetic regulators, mediates the on or off state of genes during development and differentiation. For a healthy cellular function, the levels of this protein must be maintained within a tight range. Deleterious mutations of this protein or destabilization of its dose results in numerous neuropsychiatric disorders, the most well-known of which is Rett syndrome (RTT), the leading cause of intellectual disability in girls with a frequency of ~ 7 per 100,000 live births [24]. The following chapters discuss in depth the history of RTT, the fundamental mechanisms that MeCP2 is involved in and what is known about the regulatory systems the cell uses to keep this protein within a tight range.

1.1 History of Rett syndrome and MeCP2

Dr. Andreas Rett, a pediatrician from Vienna, was the first to identify Rett syndrome; he reported his observations of 22 patients, all young girls, who shared characteristic phenotypes of hand wringing, mental retardation, speech impairment, and hyperammonemia at L.Boltzman Institute on Brain-Damaged Children in 1966. Dr. Rett continued publishing his observations in German over the next several years. Despite his efforts, the publications did not gain attention because it was written in German; moreover, the misleading inclusion of hyperammonemia in the title of his publications deterred many physicians, as ammonia in blood was not observed by other pediatricians who had patients with symptoms overlapping with what Dr. Rett had first described [25,26]. During the same time in Sweden, a child neurologist, Bengt Hagberg, had visited patients with resemblant pathologies. In 1970, Bengt presented his observation at the European child neurologists' meeting and learned that Dr. Rett had made a strikingly similar report. In 1981, Andreas Rett and Bengt Hagberg met in Toronto and concluded their young patients had been suffering from the same disorder. They agreed to name this condition Rett syndrome. In 1983, Bengt Hagberg and colleagues published their observations in the *Annals of Neurology* [27] in English. The publication enabled an international understanding of this disorder among pediatricians worldwide [28]. During this time in Europe, several families whose children were diagnosed with Rett formed the first not-for-profit International Rett Syndrome Association (IRSA), which exists to this day as the International Rett Syndrome Foundation (IRSF). The foundation funds Rett's research to create treatments worldwide [29].

After the international recognition of Rett syndrome in 1983, it took another three years for a North American institute to be dedicated to researching Rett. Dr. Vanja Holm, a developmental pediatrician who served as Boyer's Medical Director in Seattle, met Dr. Bengt and learned of the

newly identified neurodevelopmental disease [30]. Soon after, Alan Percy, a child neurologist at Baylor College of Medicine in Texas, received a copy of the publication by Bengt and colleagues. Alan confirmed a case of an unidentified neurodevelopmental condition in a young girl as RTT and admitted the patient to Texas Children's Hospital. Several other cases of RTT were confirmed throughout Texas, and multiple pediatricians were informed of this condition. In 1986, Dr. Alan Percy established the Rett Syndrome Clinic within Baylor College of Medicine with colleagues Daniel Glaze and Rebecca Schulz [28,31]. Rett remained a condition that exclusively affected females until 1990.

The first case of Rett-like symptoms appearing in males was published by a child neurologist, Mary Coleman, from Washington, DC, who was informed of Rett at the international medical conference in Paris soon after Bengt's publication [28]. The patient was a 6-year-old boy and had a healthy early development ten months after birth. His parents soon noticed, however, that their child's eyesight was poor, and his motor skills were regressing. He lost his purposeful hand movements and failed to feed himself. His family agreed to keep a helmet on their child as he was described as "clumsy" and unable to coordinate muscle movements, all in agreement with Rett phenotype. Since the cause of Rett remained unknown at the time, no standard testing was designed for diagnosis [32]. When the Rett syndrome clinic was established at Baylor College of Medicine, a post-doctoral researcher, Huda Zoghbi, encountered young girls with Rett syndrome. With support from Alan, Huda Zoghbi dedicated her research to finding the cause of Rett syndrome.

In 1999, Amir E. Ruthie *et al.* from Huda Zoghbi's lab found the cause of RTT to be mutations in the epigenetic regulator Methyl CpG Binding Protein 2 or MeCP2 [33]. MeCP2 was identified in 1992 by Adrian Bird *et al*. It was purified from rat nuclear extracts of various tissues. Using

immunofluorescence in rat fibroblasts, they found that MeCP2 was concentrated in pericentromeric regions while euchromatic regions were stained faintly [34]. At the time, it was unknown that there were two isoforms of this protein. They identified the shorter E2 isoform first; hence, the nomenclature of amino acid position refers to the E2 isoform as the accepted standard in the literature. The more relevant E1 isoform, which is 10 – 15x more abundant in the brain, was described in 2004 by Gevork N Mnatzakanian *et al.,* who inspected the 5' UTR of MeCP2 and found several open reading frames within exon 1, which had gone unnoticed when MeCP2 was identified [35]. The isoforms are discussed in detail in section 1.2 *MeCP2 domains and isoforms.* Rett research has grown steadily since its cause was identified; today, there remains no promising cure for this debilitating disease. There are, however, many reasons to believe Rett will be a reversible condition in humans. Animal studies have demonstrated a degree of reversibility of this disease [36-38]; such studies enable drug testing and therapeutic design, many of which are currently in clinical stages. The most recent FDA-approved pharmaceutical for Rett patients as of March 2023 is Trofinetide or (Glycyl-L-2-methylproyl-L-glutamic acid). The drug is an analogue of glycine-proline-glutamate (GPE), the N-terminal tripeptide of insulin-like growth factor 1. At a dose of 200 mg/kg, there is a significant improvement in mood and disruptive behaviour, ambulation, and verbal communication. Common side effects are mild gastrointestinal disorders, seizures and irritability [23,39]. Due to the ubiquitous expression of MeCP2, pharmacological approaches only partially improve the disease. It remains critical to understand the basic mechanisms of MeCP2, especially those concerning its degradation pathways, and strive to find a cure for RTT.

1.2 MeCP2 domains and isoforms

MeCP2 is an intrinsically disordered protein (IDP) for ~60% of its primary structure [40]. Intrinsically disordered proteins (natively unfolded) have at least one region that naturally fails to fold into α-helices or β-sheets. MeCP2, as its name indicates, preferentially binds methylated DNA and is involved in transcriptional repression and maintenance of silenced heterochromatin regions [41-43]. Recent findings on MeCP2's association with the transcriptional activator CREB1 at the promoter of active genes [23] changed the dogma on its exclusivity to repressed genes and heterochromatin. MeCP2 has four exons, transcribed from the telomere to centromere [44]. Alternative splicing of its mRNA yields two isoforms, E1 and E2. The E1 isoform starts with exon one but entirely skips exon two, producing a 498 amino acid long protein in humans and 484 in mice. The E2 isoform skips most of exon 1, and the start codon begins at exon 2, giving rise to a 486-residue-long protein in humans and 501 residues in mice [45] **(Fig.1.2 A and C)**. The two isoforms only differ within their NTD, which lacks tertiary structure. E1 has a stretch of poly alanine and poly glycine, making it 21 residues longer than E2, as shown in the Clustal omega alignment of the human MeCP2 isoforms **(Fig. 1.2 B)**. The yellow highlighted region of the Alphafold image of MeCP2 shows the unstructured NTD; the black arrow indicates the N-terminal methionine **(Fig. 1.2 D).** Due to the similarity between the two isoforms, it has been suggested that they are functionally redundant, although significant differences still exist between them [46,47]. As mentioned, the longer E1 isoform is more abundant in the brain, except for the dorsal thalamus (DT) [44,46]. Moreover, sex-dependant differential expression of the two isoforms has also been reported. In the neurons and astrocytes of mouse brains, differential accumulation of MeCP2 isoforms has been observed between sexes. Female astrocytes and neurons accumulate similar levels of e1 transcripts, while e2 transcripts are higher in neurons

than in astrocytes. In males, however, both isoforms are expressed equally in neurons and astrocytes. Rastegar *et al.* reported that DNA methylation of MeCP2 regulatory elements (RE) is associated with the differential expression in male and female mice [47], suggesting that DNA methylation at REs of MeCP2 controls the levels of MeCP2 isoforms. In the context of Rett, mutations occurring within exon one are rare. A single missense mutation Ala2Val in exon one has been reported where the missense mutation results in faster proteasomal degradation of MeCP2. This observation transcends this specific mutation as many MBD mutations, such as the most common T158M missense mutation, also result in rapid proteasomal degradation of MeCP2 when compared to WT. See section 1.3.1. *MBD mutations* for a synopsis of what is known about MeCP2 MBD mutations, especially the most common, T158M.

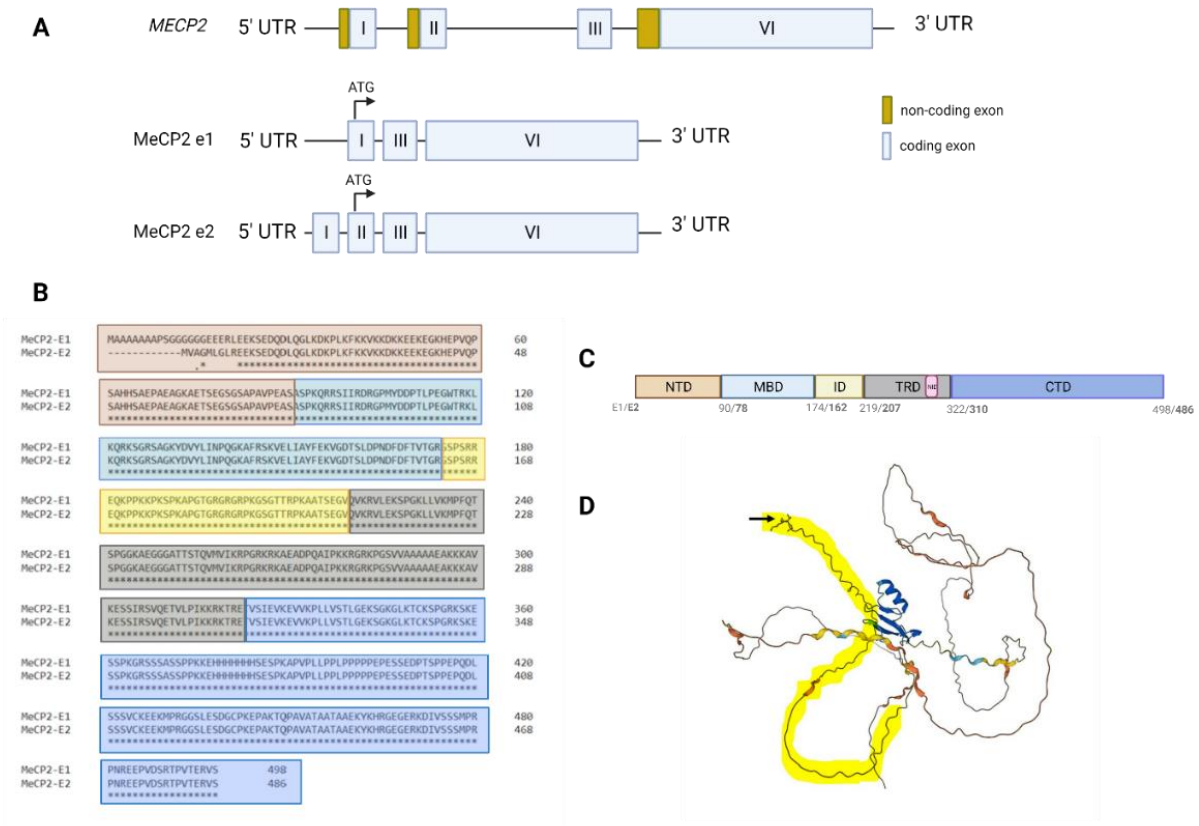


Figure 1.2. Alternative splicing of MeCP2 mRNA yields the E1 and E2 isoform. A. MeCP2 mRNA has four exons, I – VI. The e1 isoform begins with exon one, and the e2 isoform begins with exon two [48,49]. The 3'UTR is not shown to scale as it is unusually long, 8.5 Kbp [44] **B.** Clustal omega alignment of human MeCP2 E1 and E2; the highlighted sequence corresponds to the domains represented in the cartoon diagram. **C.** Domains of MeCP2 with the residue number of E1 and E2: NTD, N-terminal domain; MBD, methyl binding domain; ID, intervening domain; TRD, transcription repression domain; NID., NCoR (nuclear receptor coreceptor) interaction domain; CTD, C-terminal domain. **D.** Alphafold structure prediction of human MeCP2-E2. The arrow points to the N-terminal methionine; the yellow highlighted region is the disordered NTD ending at serine 78 [50,51].

Albeit the NTD is unstructured, it impacts its neighbouring MBD and can stabilize it in conditions of low ionic strength [52]. Our lab had shown the differential stability of the NTDs when tethered to MBD. Cycloheximide chase assay of the two isoforms with a GFP tag in transiently transfected neuroblastoma showed that after 48 hours of the treatment, the E1 levels

dropped to ~10% while the E2 isoform dropped to ~ 40% [53], suggesting the longer NTD of E1 may contribute to the shorter half-life of this isoform **(Fig. 1.3[53]).**

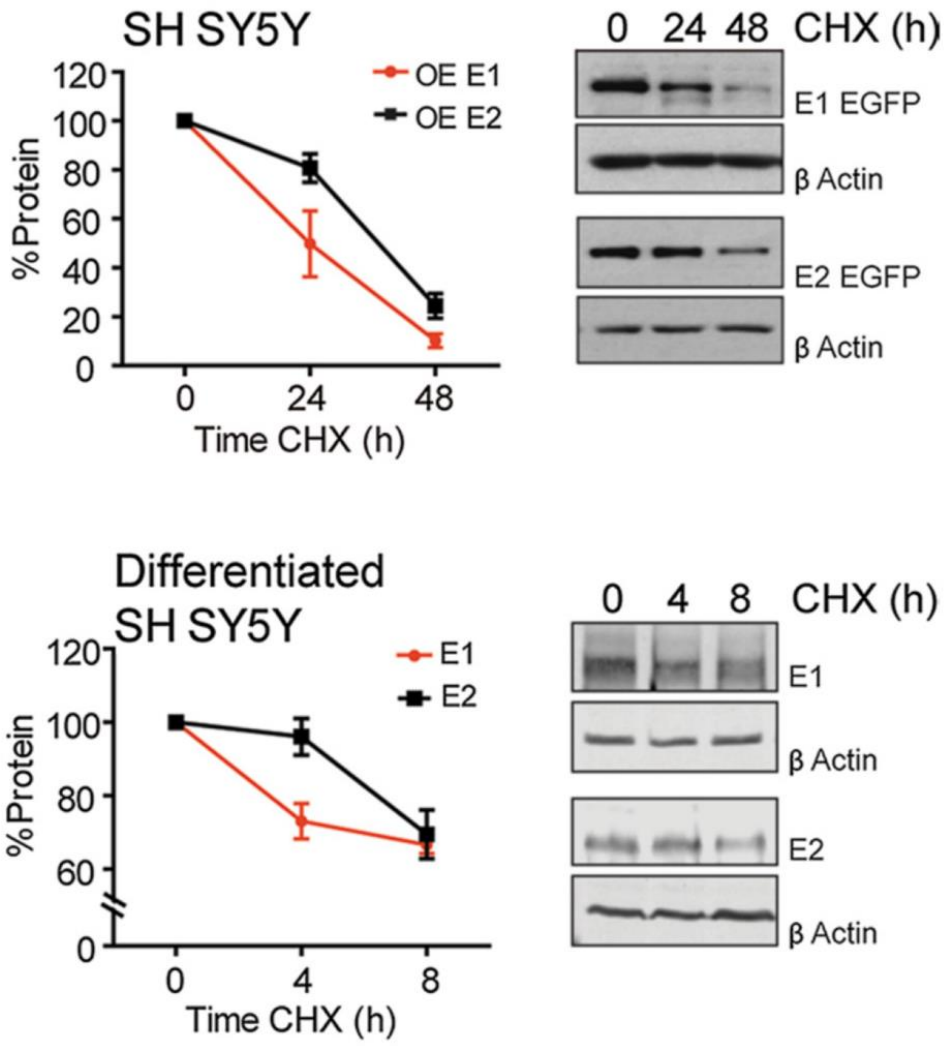


Figure 1.3. MeCP2 isoforms have different half-lives. Cycloheximide chase assay of transiently transfected neuroblastoma SH SY5Y with recombinant MeCP2 E1 and E2 after 48 hours [OE: overexpressed] (top). Endogenous expression of MeCP2 isoforms in differentiated SH SY5Y after 8 hours (bottom) [53], used with permission from Creative Commons Attribution 4.0 International License. Epigenetics & Chromatin, vol 12, MeCP2-E1 isoform is a dynamically expressed, weakly DNA-bound protein with different protein and DNA interactions compared to MeCP2-E2. Khajavi and Sheikh et al.

The methyl binding domain (MBD) and the transcriptional repression domain (TRD) are well characterized, as the deleterious mutations of MeCP2 are often clustered within these two domains [38,54,55]. The trypsin digest of MeCP2 yields seven trypsin-resistant fragments. Four of these overlap with the MBD and TRD. **(Fig. 1.4 A)**, suggesting that these domains contain a degree of tertiary structure, restricting trypsin access [40]. These findings were corroborated after the MBD structure was identified to be composed of four antiparallel β-sheets on the N-terminal and an α-helical region at the C-terminal, as predicted by the MOLSCRIPT algorithm **(Fig. 1.4 B, C)** [56-58]. The X-ray analysis of the co-crystalized MBD-DNA revealed contact between methylated DNA and hydrophilic surfaces, including tightly bound water molecules. Their finding suggests that MeCP2 recognizes DNA by hydration of the major groove of methylated DNA**.**

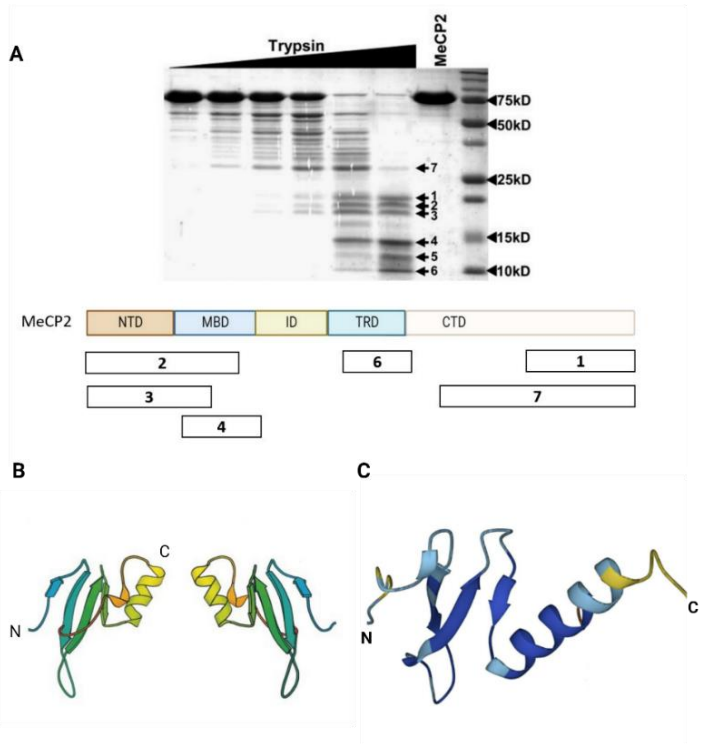


Figure 1.4. MBD of MeCP2 has a defined secondary structure. A. Trypsin digest of MeCP2 exhibits seven trypsin-resistant fragments (top). Cartoon presentation of MeCP2 domains aligned with seven trypsin-resistant fragments [40](reprinted from the J. Biol. Chem., open access; Intrinsic Disorder and Autonomous Domain Function in the Multifunctional Nuclear Protein, MeCP2; 2007; Adams & Hansen et al.**B.** MOLSCRIPT [58] diagram of MeCP2 MBD [56] (Reprinted from J. Mol. Biol., Vol 29, Solution structure of the domain of MeCP2 that binds methylated DNA, Robert and Bird et al. 1999, with permission from Elsevier). The β-sheets are by the N-terminal and the α-helices by the C-terminal of MBD domain **C.** AlphaFold structure of MeCP2 MBD. Dark and light blue correspond to high and moderate confidence in structural prediction [50,51].

The Asx-ST (Asp/Asn-S/T) motif within MBD lies near the C-terminus of the MBD, 13 residues after the defined α-helix of MBD. The motif contacts DNA via hydrogen bonding between Val159 of Asx-ST and oxygen of the phosphate backbone of DNA. Thr158 stabilizes the tandem Asx-ST, forming hydrogen bonding between the Asx turn and ST motif **(Fig. 1.5[57]).**

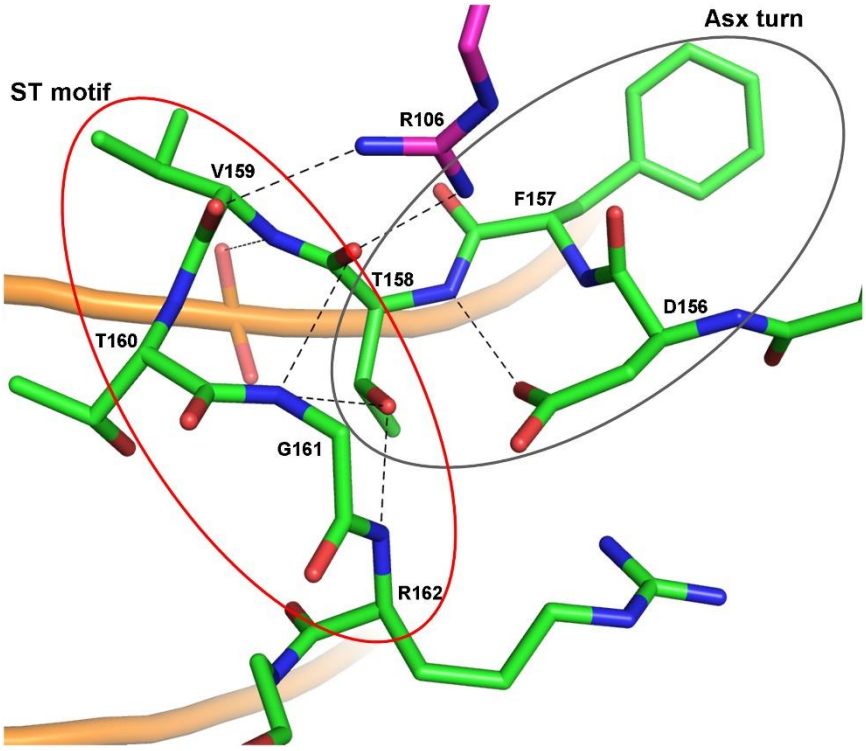


Figure 1.5. Asx-ST motif of MeCP2. T158 is essential for the stability of the Asx turn and ST motif. D156, F157, and T158 make up the Asx-turn. The critical hydrogen bonding of this motif is between the COOH of Asp156 (red) and the nitrogen of Thr158 amino group (blue). ST motif comprises T158, V159, T169, G161, R162. Critical hydrogen bonding is between the COOH group of Thr158 and the nitrogen of amino groups of Gly161 and Arg162 [57]. Reprinted from *Molecular Cell*, vol 29 (4), Ho and Walkinshaw et al. MeCP2 Binding to DNA Depends upon Hydration at Methyl-CpG. (2008), with permission from Elsevier.PBD domain identifier of MeCP2 MBD with methylated DNA: 6C1Y.

The transcriptional repression domain (TRD) of MeCP2 is critical for maintaining a repressed transcriptional environment. Histone deacetylase (HDAC) containing complexes such as Sin3A and NCoR-SMRT have been identified to interact with MeCP2 by its TRD in co-fractionation

experiments [59,60]. Common Rett-causing mutations within TRD are clustered in residues 302-306, with R306C ranking as the most common Rett-causing missense mutation. The mutation hot spot identified the specific region within TRD, residues 302 – 306 that binds the co-repressor NCoR-SMRT through the direct binding subunit of the NCoR complex, TBLR1.

Rett mouse models with null MeCP2, MBD or TRD mutations recapitulate the Rett phenotype. MBD and TRD of MeCP2 account for most of MeCP2's critical functions, DNA binding and recruitment of repressor complexes. Intriguingly, when MeCP2-null mice are introduced with a truncated MeCP2 that only contains the MBD and segments of TRD that are responsible for nuclear localization and binding to co-repressor complex, NCoR-SMRT, many of the phenotypes were reversed to near WT levels. The improvements include increased survival from 16 weeks to 50 weeks and improved locomotion, measured by latency to fall, to near WT levels [61]. However, this should not be interpreted that the NTD, ID, and CTD domains of MeCP2 are unimportant. As previously discussed, the NTD is the domain that varies between the two isoforms, and this variation is enough to cause significant differences in the isoform half-life [53]. The intervening domain (ID) stabilizes the MBD as its deletion significantly reduces protein levels; however, the mRNA levels remained unaltered [52,61]. Moreover, two PEST domains, discussed in detail in the next chapter, have been suggested to mediate MeCP2's homeostasis [62].

1.2.1 PEST domains

PEST domains were first described by Rogers *et al.* in 1986 as protein regions enriched in proline (P), glutamic acid (E), serine (S) and threonine (T), flanked by positively charged residues such as lysine (K) or arginine (R), that target the protein for degradation [63]. PEST domains are commonly found in short-lived proteins [63] and IDPs enriched in disorder-promoting residues [64-66]. PEST domains have marginal stability due to the secondary structures, such as

polyproline type II (PPII) and alpha helices. Upon being phosphorylated at a serine or threonine adjacent to a proline, the conformational stability of the PEST domain is impaired. Proteolysis occurs following the structural impairment [63,67].

Rogers et al. developed an algorithm, PEST-find, which identifies putative PEST motifs and gives each candidate a score between -50 to +50. Scores greater than +5 are considered strong PEST motifs [68]. Soon after PEST-mediated degradation was postulated, other researchers published their findings on PEST domains acting as degradation signals, resulting in rapid proteolysis [69-71]. Seldomly, PEST domains are found on long-lived proteins; in such instances, they serve as intermolecular signals for proteolysis [63]. Proteolytic pathways associated with PEST-containing proteins are the ubiquitin-proteosome system (UPS) and calpain cleavage. Calpain proteolysis involves the binding of calpain to an acidic proline in the PEST motif, which drives proteolysis in a calcium-dependent manner [72]. A bona fide example of PEST-mediated proteolysis is mouse ornithine decarboxylase (ODC), an enzyme for polyamine biosynthesis. The half-life of ODC is ~ 1hr; however, deletion of its PEST motif, located at its C-terminus, increased its half-life to at least 4 hours [73]. As an extension to this finding, Ghoda *et al.* reported that ODC in a protozoan parasite, *Trypanosoma brucei,* has a half-life of over 6 hours. By transfecting ODC from mouse and *T. brucei* to ODC deficient cells, they observed in cells transfected with mouse ODC, its expression was less stable than the cells transfected with *T. brucei* ODC*,* and the stability of protozoan ODC was linked to lack of the C-terminus containing PEST motif which was exclusive to mammals [74].

There are contrasting observations regarding the deletion of PEST domains and increased stability. Contrary to the observation in mouse ODC stability after deletion of its PEST domains,

deletion of this motif in the oncogene, c-Myb, did not affect protein turnover [75]. This seeming contradiction suggests the PEST motifs function as regulatory regions that either directly result in proteolysis or indirectly result in signalling pathways that will eventually lead to degradation, such as PEST-like containing protein Ste3p in yeast that acts as an endocytosis signal upon ubiquitination, followed by vacuolar degradation [64,76]. It is important to note that the PEST domains are always present. Yet, they do not result in immediate destabilization and rapid proteolysis because the PEST-containing proteins must be marked for degradation. Phosphorylation is one of the many mechanisms that can activate the PEST domain. Serine or threonine phosphorylation precedes the attachment of the polyubiquitin chain to a flanking lysine of the PEST domain, recruiting the proteasome [68,77].

MeCP2 has two strong PEST domains near the N- and C-terminus. The C-terminal PEST is the stronger candidate with a score of 17.29, generated by the PESTfind algorithm[78] **(Fig. 1.6).** Our lab has postulated that MeCP2 proteolysis may be controlled by the PTMs of this motif [62,79]. One of the well-characterized MeCP2's phosphorylation is at Ser80, critical for chromatin association [80], and it acts in opposition to another essential phosphorylation at Ser421[81], which has been incorrectly associated with the Ct-PEST [62]. Ser421 in mice corresponds to S423 in humans, and it lies outside the Ct-PEST **(Fig. 1.6; red box)**; however, phosphorylation at Ser401 and ubiquitination at Lys82 has been identified as a putative PTM in mass-spec analysis [79,82]. It remains unclear if MeCP2 PEST domains play a role in its degradation, and if they do, to what extent are they driving proteolysis. As discussed in section 1.4.2.2, *PEST-mediated degradation*, this project focuses on finding an answer for the role of PEST domains in MeCP2 degradation.

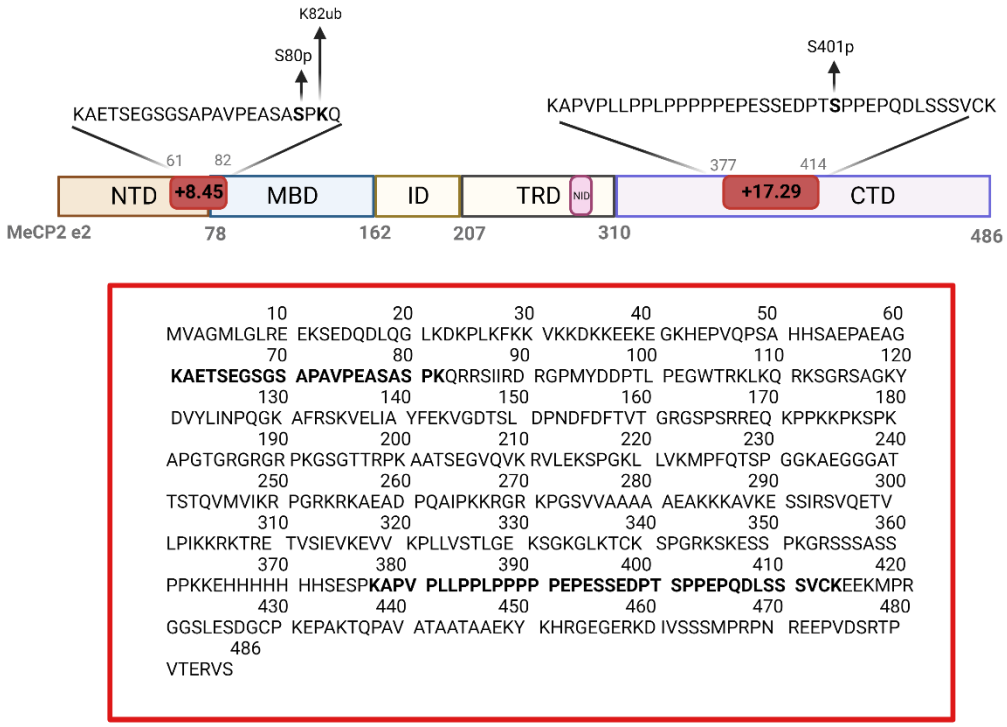


Figure 1.6. MeCP2 has two PEST motifs. Cartoon presentation of MeCP2 domains shows the N and C-terminal PEST in red and their PEST score generated by the PESTfind algorithm[78] (top). The Human MeCP2-E2 protein sequence depicts the exact position of the two PEST domains in bold (bottom). Phosphorylation at S80 and S401 and ubiquitination at K82 have been identified as MeCP2 PTMs [80,82].

1.3 Rett causing mutations

Although MeCP2 has been associated with numerous neuropsychiatric diseases, Rett remains the most severe and debilitating MeCP2-related disorder. 90 to 95% of Rett mutations are within MeCP2, yet atypical Rett phenotype may not be caused by *MECP2* mutations; instead, the mutations are in cyclin-dependant kinase-like (CDKL) 5 protein[83] and forkhead box protein (FOXG) 1[84]. Recent findings have demonstrated that MeCP2 drives liquid-liquid phase separation (LLPS) when in complex with DNA. LLPS has been studied in the context of MeCP2. Evidence supports the formation of a membrane-less compartment inside the cell following

LLPS [85], in agreement with MeCP2 function inside the nucleus as it facilitates tight chromatin packing. MeCP2-DNA interaction is a multivalent interaction. MeCP2 binds DNA through its intervening domain, MBD, and the TRD [86]. Common Rett-causing mutations such as T158M, R133C, and R306H have been shown to impair the *in vitro* phase separation of MeCP2 in conjugation with methylated DNA. Intriguingly, the two isoforms exhibit different capacities for driving LLPS [87]. Impaired LLPS of mutated MeCP2 suggests how these mutations disable MeCP2 from forming dense chromocenters inside nuclei. Confocal fluorescence microscopy images of various MBD mutations correlated with the observation that MeCP2's fundamental capacity in developing dense heterochromatin regions is abolished or severely impaired, depending on the position of the MBD mutation [88]. As depicted in **Table 1**, the Rett-causing mutations, their frequency and how they impact protein levels are tabulated. The following topics discuss common Rett-causing mutations, such as T158M and R306C, that are within the MBD and the TRD, respectively.

Table 1. Classification of Rett-causing mutations based on their position in MeCP2 domains.

Domain of MeCP2	Mutation	Molecular phenotype	Protein levels *in vivo* relative to WT	% of RTT cases	Median life span in mice	Reference
N-terminal Domain (NTD)	MeCP2-E1- A2V	Methionine at the P1 position and the penultimate valine (P'1) were acetylated *in vitro*. N-methionine excision (NME) was absent when compared with E1. MeCP2 colocalization to chromocenters was unaffected.	~1/4	0.2%	Unknown	88-90
Methyl Binding Domain (MBD)	T158M	Impaired stability of Asx-ST motif. Impaired binding to chromocenters due to reduced affinity for mCpG.	~1/3	8.79%	12 weeks	38,88,90-92
	P152R/A	Impaired clustering of MeCP2 to chromocenters.	Unknown	1.5%	Unknown	88,93
	R133C	Impaired stability of 4th β-sheet and α-helix. Reduced binding to hmC (retained binding to mC). Impaired interaction with TCF20 complex. Reduced MeCP2-driven LLPS.	~1/2 in the hippocampus; indistinguishable in cortex and cerebellum	4.24%	42 weeks	54,88,92,94-96
	R111G	Abolished binding to mC; disrupted nuclear localization; disrupted binding to TCF20 complex. Reduced MeCP2-driven LLPS.	unknown	>1%	6 weeks	88,95-98
	R106W	Impaired stability of 2nd beta-sheet and reduced affinity to mCpG. Decreased MeCP2-driven LLPS.	~1/3	2.76%	10 weeks	88,92,96,99-101
Transcription Repression Domain (TRD)	R294X	Increased binding to DNA. Clustering to chromocenters was unaffected.	Indistinguishable	6.1%	35 weeks	90,102-104
NCoR/SMRT Interaction Domain (NID)	R306C	Abolished interaction between NCoR/SMRT by disrupting TBLR1 binding. Disrupted binding to ATRX complex. Reduced HDAC activity in the forebrain of mice.	Indistinguishable	6.8%	18 weeks	55,105,106

1.3.1 MBD mutations

The most common MBD mutation, T158M, accounts for 9 % of all Rett cases [38,90]. This mutation destabilizes the Asx-ST motif. **Fig. 1. 5** depicts the Asx turn comprising ^{156}DFT^{158} and the immediate ST motif ^{158}TVTG^{161}. T158 is critical for the stability of this motif through H-bonding. Both female and male mouse models with T158M mutations have been created. The mice develop neurological abnormalities such as abnormal gait, reduced brain weight, and reduced survival. A critical observation in this mouse model was that the level of this protein, but not the mRNA, drops to ~ 30% of that of WT, yet the mutation does not have a dominant negative effect. Transgenic mice that over-express the T158M protein exhibit improved phenotype and a degree of reversibility of the Rett phenotype. Brain weight and survival of mice overexpressing the T158M mutation were indistinguishable from the WT [38]. Surprisingly, the overexpression of mutated protein does not lead to abnormal neuronal arborization, as observed when WT MeCP2 is overexpressed [107-110]. Moreover, incorporating the proteasome inhibitor MG132 restored the degradation of T158M MeCP2 while it did not affect the WT **(Fig.1.7).** These observations provided new routes for designing Rett treatments through gene therapy or through pharmacological administration of proteasome inhibitors to alleviate fast degradation of T158M.

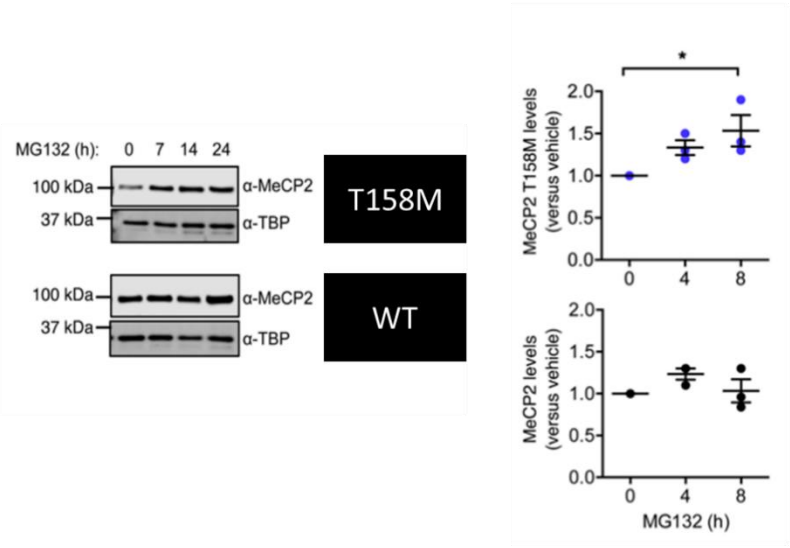


Figure 1.7. Proteosome is responsible for degrading MeCP2 T158M. Incorporating the proteasome inhibitor MG132 restores the declining levels of MeCP2 T158M [38]. Used with permission of the American Society for Clinical Investigation, from Elevating expression of MeCP2 T158M rescues DNA binding and Rett syndrome-like phenotypes by Lamonica and Cui et al. vol 127 (5). 2017. Permission is conveyed through Copyright Clearance Center, Inc.

Binding analysis and stability of the MBD domain with Rett causing mutations enabled the categorization of the MBD mutants into three groups **(Fig. 1.8)**. Group three mutants are mutations that destabilize the protein the most, and group one is comparatively less destabilizing. Group three mutations include the first residue within the Asx-turn, D156E. In addition, P152R, S134C, and L100V are within this cluster[98]. As depicted in **Fig. 1.8**, P152R amongst cluster 3 mutants conserves the most DNA binding (lowest on the y-axis). In the context of Rett, P152R is the 4th most common MBD mutation; however, it is within the disordered region [90]. Rett patients with P152R have a diverse degree of phenotypic severity. The normal phenotype of girls with this mutation is typical Rett phenotype; however, a report of a girl with atypical Rett with P152R has been published where the patient became microcephalic with mild mental retardation and behaviour disturbances [111]. The correlation between genotype and phenotype is only reliable in

males (XY). In females, X-chromosome inactivation (XCI) creates a mosaic expression of genes on the two X-chromosomes, making the genotype-to-phenotype comparisons more complex. It is plausible that P152R, similar to T158M, becomes rapidly degraded and that the proteasome is responsible for their breakdown. It remains to be determined how the 26S ubiquitin-proteosome system (UPS) drives the degradation of MeCP2 with MBD mutants and whether the PEST domains play a part in the proteasome recruitment.

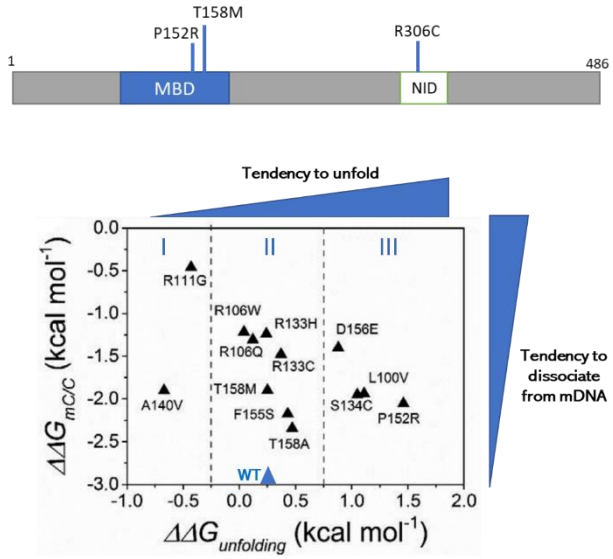


Figure 1.8. Categorization of Rett causing MBD mutations based on biophysical analysis. Top, cartoon presentation of MeCP2 primary structure. Bottom, three groups of MBD mutations are based on how the mutation impairs stability (X-axis) and affinity for DNA (Y-axis). Reprinted with permission from ACS Chemical Biology. Binding analysis of Methyl CpG binding domain of MeCP2 and Rett Syndrome mutations. Yang and Cao, et al. vol 11. 2006. [98].

1.3.2 TRD mutations

Another hot spot of Rett-causing mutations is concentrated within a narrow region of TRD that directly binds the NCoR-SMRT co-repressor complex [55,105]. The most common TRD mutation is R306C, accounting for ~6% of all Rett cases [90]. This mutation prevents the binding of MeCP2 with the TBLR1 component of the NCoR-SMRT complex. In addition to R306C, NID mutations include P302R, K304E, K305R/E, and R306C. Protein levels in mouse models with R306C MeCP2 were indistinguishable from WT [55,105], suggesting that NID mutations do not destabilize the protein. Colocalization of MeCP2 to target DNA in mouse ES cells appeared unaffected in R306C. The colocalization was unchanged since R306 lies outside the MBD [55].

1.4 MeCP2 homeostasis

The levels of MeCP2 must be kept within a tight range for its proper function. Numerous mechanisms regulate MeCP2, many involving miRNAs binding *MECP2* 3'UTR; however, less is known about mechanisms at play for maintaining protein levels of MeCP2. The following sections elaborate on the mechanisms that play a role in maintaining MeCP2 homeostasis before and after translation.

1.4.1 MeCP2 regulation before translation

At the mRNA level, various proteins and microRNAs have been identified to regulate levels of *MECP2.* These regulators include negative regulators such as high mobility group N1 protein (HMGN1) [112], several miRNAs such as miRNA132-212 [113], and positive regulators such as myocyte enhancer factor 2C (MEF2C) [114].

1.4.1.1 Negative regulators of MeCP2

HMGN1 is a ubiquitous nuclear protein, and like MeCP2, it binds to nucleosome core particles. [115,116]. HMGN1 was found to be a negative regulator of MeCP2 by binding to its promoter. The observation that resulted in this finding was that in patients with Down syndrome (DS), MeCP2 expression is reduced [117]; more specifically, there was a ~50% increase in HMGN1 protein and a 30% decrease of *MeCP2* transcripts in the brain of DS patients. Western blot analysis confirmed protein levels of MeCP2 were also reduced [112]. Similar to MeCP2's sensitive dose dependency, too little and too much HMGN1 results in abnormal behaviour. In mice that overexpress HMGN1, hyperactivity and abnormal social interest are observed, similar to DS mouse models [118]. In contrast, null Hmgn1 mice exhibit hypoactivity, aberrant social behaviour, and other traits that mirror models of autism spectrum disorder (ASD) [119]. These data suggest irregularities in the homeostasis of proteins that affect MeCP2 dose and consequently impact the epigenetic machinery, contribute to the etiology of behavioural problems in ASD and DS.

There are miRNA recognition elements (MREs) in MeCP2 3' UTR, and several miRNAs have been predicted to bind this region, such as miRNA194, miRNA-7b, miRNA-212 and miRNA-132 [113,120-122]. CREB-induced miRNAs132/212 are known for regulating synaptic transmission [113,121]. miRNA-132, enriched in the brain, has been shown to bind MeCP2 3'-UTR and negatively regulate MeCP2 expression [123]. *MeCP2* has multiple polyadenylation sites, which results in several 3'UTRs ranging between ~ 2 to 10 Kbp. Longer transcripts found in the brain are highly conserved across mammals [124]; the miRNA-132 binding site is only available in the longer 3'UTR transcripts in the brain. A shorter 3' UTR of *MeCP2* is present in visceral tissues such as muscles. Induction of miRNA-132 expression by treatment with forskolin or KCl

decreased MeCP2 levels [113,123]. Moreover, the use of oligonucleotides that are complimentary to the miR132 recognition element (MRE) on *MeCP2* 3' UTR, designed to block the interaction between miRNA132 and MeCP2, resulted in higher MeCP2 levels in cortical neurons relative to oligonucleotides non-specific to the MRE, corroborating the observation that miRNA-132 negatively regulates this protein [113].

There is bi-directional feedback autoregulation between MeCP2 and miRNA7-b. miRNA-7b can become hypermethylated at the 5'-flanking regions. MeCP2 recruitment to methylated CpG islands has been shown to diminish the expression of miRNA-7b. Conversely, the binding of miRNA-7b to 3'UTR of MeCP2 inhibits the levels of this protein. This autoregulation between MeCP2 and miRNA-7b occurred during postnatal murine neuronal maturation when MeCP2 levels increased and regulated healthy neuronal development [122].

1.4.1.2 Positive regulators of MeCP2

Another bi-directional regulation of MeCP2 is with the myocyte enhancer factor 2C (MEF2C), one of MeCP2's binding partners. MeCP2 binds to the murine *Mef2c* promoter and represses transcription [125]. MEF2C was found to be mutated or truncated in patients with severe mental retardation [114,126]. Mutations or truncation of this transcription factor diminished expression of MeCP2 and CDKL5; as mentioned, atypical Rett patients have mutations in CDKL5 [83]. Luciferase gene assay analysis of MeCP2 and MEF2C demonstrated that MEF2C activates MeCP2 while deleterious MEF2C mutations diminished transcriptional activation of the MeCP2 promoter [127]. The mechanistic details of how MEF2C positively regulates MeCP2 remain elusive.

1.4.2 MeCP2 regulation after translation

1.4.2.1 N-methionine excision

Nearly all polypeptides translated by the ribosome begin with a methionine (Met), which marks the translation initiation. The N-terminal methionine is co-translationally excised by Met-aminopeptidases bound to the ribosome if the penultimate residue is small enough, such as Ala, Gly, Val, Ser, Thr, Cys, and Pro. N-methionine excision (NME) is evolutionarily conserved among prokaryotes and eukaryotes, and it is applied to more than 50% of nascent proteins [128,129]. In ~ 80% of nascent human proteins, the Nt-Met or the penultimate residue after NME undergoes Nαt-acetylation [130]. Nt-acetylation acts as a degradation signal (degrons) targeted by ubiquitin ligases for proteolysis; however, there are records of numerous long-lived stable proteins that become co-translationally acetylated at the second position after NME; a plausible explanation for this contrast has been the inaccessibility of the Nt-domain to the proteasomal machinery created by higher order structures [131]. It has been postulated that the destabilizing Nt residue reduces the half-life of the protein, and this phenomenon is referred to as the N-end rule [132-134]. In *Saccharomyces cerevisiae*, the E3 ubiquitin ligase, Doa10, adds the polyubiquitin chain to Nt-acetylated Met, Ala, Gly, Ser, Thr and Cys [135]. Our lab has previously analyzed the Nt-co and post-translational modification of MeCP2 by expressing the NTD of the E1 isoform in HEK293 cells. MS analysis of the MeCP2-E1 isoform showed no Nt-Met, indicating complete NME; in addition, the Ala that became exposed following NME was acetylated. Several MS reads exhibited the cleavage of the first two, three, five and six Nt-Ala following NME and the acetylation of the following exposed Nt residue (**Fig1.9, top**) [89]. For the shorter MeCP2 E2 isoform, MS analysis revealed the retention and acetylation of Nt-Met and acetylation of Val at

position P'1. At the same time, a few peptide reads showed NME and acetylation of the penultimate Val [53] (**Fig.1.9, bottom**).

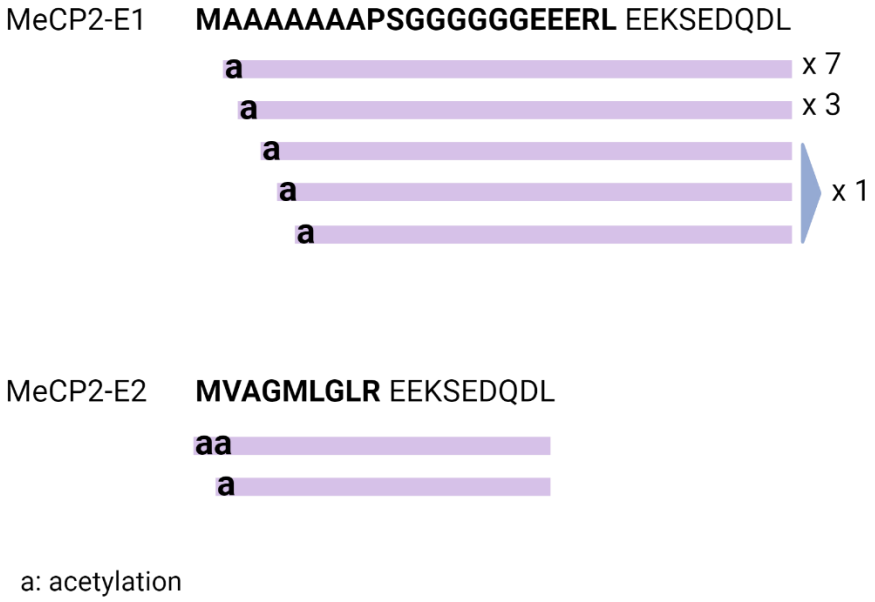


Figure 1.9. N-methionine excision and acetylation of N-terminal residues of MeCP2 identified by mass-spectrometry sequencing. Most mass spec reads found NME and acetylation of the penultimate alanine for MeCP2 E1, top. The acetylation and retention of Nt-methionine or NME and acetylation of the penultimate valine were observed for the E2 isoform, bottom. The figure was adapted and modified from Sheikh et al. [53].

The NME and the acetylation of the penultimate residue are found in higher peptide reads in the E1 relative to E2. According to the N-end rule, the E1 isoform would be more destabilized by its acetylated Nt-residue than its shorter E2 isoform that retains Nt-Met [134]. Cycloheximide chase assays of MeCP2 isoforms in the transient transfected SH-SY5Y neuroblastoma cells showed that after 4 hours, the E1 isoform was ~30% less than the E2 [53]. Moreover, the lower folding stability of the first two domains of MeCP2-E1, NTD-MBD, and its lower affinity for DNA suggests a higher presence of the E1 isoform in solution, expediting its degradation as it is

exposed to the proteolytic machinery [53,136]. The higher degradation rate of E1 isoform is essential for its involvement in dynamic processes in neurons throughout development.

1.4.2.2 PEST-mediated degradation

The PEST-mediated degradation in the context of Rett-causing mutations re-sparked our interest after Dr. Zhaolan Zhou and colleagues published their findings in 2017 with the T158M mouse model. They found that protein levels *in vivo* were reduced to 1/3rd of WT and that by inhibiting the 26S UPS by MG132, the mutated MeCP2 levels were restored [38]. Our lab has previously suggested the PEST-mediated degradation as a potential proteolytic pathway for MeCP2; however, it had never been experimentally tested. As mentioned in section 1.2.1, PEST domains are insufficient to induce proteolysis and must be 'activated' to recruit the proteolytic machinery. A pre-requisite of this activation is phosphorylation on a serine or threonine within the PEST that induces structural instability, especially for the stability imposed by the polyproline type II secondary structure [62,63,67]. MeCP2's polyproline region is located within the Ct-PEST motif. However, as of today, the only phosphorylation event identified by MS in this region is at S401, situated within the Ct-PEST polyproline stretch of MeCP2 **(Fig. 1.6)**. It remains unclear whether the PEST domains play a role in the recruitment of the proteasome when MeCP2 has deleterious mutations. To conclude about the role of PEST domains in the degradation of MeCP2 with MBD mutations, I have dedicated the focus of this thesis to addressing the PEST hypothesis in the degradation of MeCP2 with MBD mutations.

1.5 C2C12 myoblasts, an excellent model for neuronal cells

The C2C12 cells were created by Blau and Webster in 1983 [137] from a subclone of C2 cells generated by Yaffe and Saxel in 1977 [138]. C2 cells were generated from primary cultures prepared from the thigh muscle of a two-month-old healthy female C3H mouse 17 h after thigh muscle crush injury. This cell line was used as a control for studying muscular dystrophy [138,139]. One disadvantage of the C2 cell line was that other cell types, such as adipocytes and fibroblasts, were present, and the proportion of these cell types varied in diseased versus healthy cells. Blaue et al. generated the C2C12 subclone by creating a culture condition that enabled the isolation of muscle cells to circumvent this variability due to a heterogeneous population of cells. C2C12 cells undergo differentiation at high confluency from myoblasts to polynucleated myotubes, and the rearrangement of centromeric domains occurs in terminal differentiation in both C2C12 and mouse neurons [140-142]. Moreover, as the myoblasts differentiate, the pericentric DNA methylation and MeCP2 expression increases [142,143]. MeCP2 cells similarly increase in neuronal cells as neurons mature and undergo differentiation [144]. The clustering of chromocenters, addressed in the co-localization analysis in Chapter 4.1, increases in myogenesis [142]. The visual analysis of MeCP2 for FRAP and co-localization is facilitated by the defined chromocenters that house densely methylated DNA where MeCP2 binds. These striking similarities between C2C12 and neuronal cells and the defined pericentromeric structures make C2C12 an excellent model for studying MeCP2.

1.6 Research Questions, Objectives and Significance

I aim to address the role of MeCP2's PEST domains in the degradation of MeCP2 with MBD mutations. The specific questions I seek to address are:

1) Is the reduced protein availability, as observed in T158M, recapitulated in MeCP2 with other MBD mutations that have been predicted to have a destabilizing effect?

2) How do MeCP2's MBD mutations differ from other mutants, such as the NID mutation R306C, in DNA binding and protein levels?

3) How does deleting or modifying the Ct-PEST in MeCP2 T158M affect DNA binding and protein levels?

1.6.1 Significance

The dose-dependency of MeCP2 for healthy neuronal maturation has been recognized soon after MeCP2 mutations were identified to cause Rett syndrome. The most common Rett-causing missense mutation, T158M, resulted in lower protein due to its higher proteasomal degradation [38]. However, some non-canonical mutations within MeCP2 CTD, such as P322L, have also caused reduced protein levels [145]. Another MeCP2-related disorder resulting from higher levels of this protein is MeCP2 duplication syndrome, sharing many phenotypes with Rett due to impaired neuronal maturation. Despite the importance of MeCP2 homeostasis for healthy neuronal maturation, significant knowledge gaps remain in which mechanisms regulate MeCP2 at protein levels. Our lab has hypothesized PEST-mediated degradation of MeCP2 to regulate MeCP2 levels at a protein level; however, it has never been experimentally tested. In this work, the PEST hypothesis for degradation of MeCP2 is experimentally tested for the first time. These

findings enhance our understanding of MeCP2 degradation mechanism at a protein level, and create a model for further analysis of the molecular interactions that regulate MeCP2 dynamics.

Chapter 2: Materials and Methods

Section 2.3 of this chapter is the FRAP analysis Anastasia Roemer conducted from Dr. Michael Hendzel's lab. I completed the remaining experiments. Dr. Bob Chow and Alberto Ruiz assisted in the confocal imaging of the transiently transfected mouse myoblasts from the biology department at the University of Victoria.

Table 2. List of reagents and their distributor/ manufacturer.

SDS and Western blotting	**Company, Country, Catalog #**
Coomassie Brilliant blue G 250	Sigma-Aldrich, St. Louis, MO, USA; 6104-58-1
Nitrocellulose membrane 0.2uM	Bio-Rad Laboratories (Canada) Ltd; 1620112
Clarity ™ Western ECL Substrate	Bio-Rad Laboratories (Canada) Ltd; 1705060
Plasmid prep	
DH5-alpha	In-house
Ampicillin	Sigma-Aldrich, St. Louis, MO, USA; A9518
Kanamycin	Sigma-Aldrich, St. Louis, MO, USA; BP861
Plasmid Prep kit	HiSpeed Plasmid Midi Kit, Qiagen, Venlo, Netherlands; 12643
Reagents for cell-culture	
DMEM, high glucose	Invitrogen, Waltham, MA, USA. 11965092
Fetal Bovine Serum (FBS)	Sigma-Aldrich, St. Louis, MO, USA; F1051
PBS, pH 7.4 Gibco™	Invitrogen, Waltham, MA, USA. 10010023
Lipofectamine™ 3000 Transfection Reagent	Invitrogen, Waltham, MA, USA. L3000015
Trypsin_EDTA (0.5%), no phenol red	Invitrogen, Waltham, MA, USA. 15400054
Reagents for nuclear fractionation	
Protease inhibitor	cOmplete™ Mini, EDTA-free Protease Inhibitor Cocktail, Roche, Laval, QC, Canada; 5056489001
Phosphatase Inhibitor	PhosSTOP ™ Roche; 4906837001
Alkaline phosphatase	Antarctic Phosphatase, New England Biolabs Inc; M0289S
Reagents for confocal fluorescence microscopy	
16% Paraformaldehyde	Electron Microscopy Sciences, Hatfield, PA, USA; 15170
Coverslips	Bellco Glass Inc., Vineland, NJ, USA. 50-194-4702
Immumount	Thermo Shandon Limited, Cheshire, UK. 9990402
DAPI (4',6-Diamidino-2-Phenylindole, Dihydrochloride)	Invitrogen, Waltham, MA, USA. D1306

2.1 Cell culture and transfection

Mouse myoblasts, C2C12, were generously donated by Dr. Perry Howard, University of Victoria. Cells were grown at 5% CO2 at 37ºC in DMEM, high glucose (Invitrogen; 11965092) with 10% FBS (Sigma-Aldrich; F1051). Cells were 50-70% confluent at the time of transfection with lipofectamine 3000™ (Invitrogen; L3000015) following the manufacturer's protocol. Human MeCP2_E1 constructs and the expression vectors WT, T158M, P152R and R306C on pcDNA3.1_Ct-GFP were donated by Dr. John Vincent. The MeCP2 double mutants constructs modPEST+T158M, nullPEST+T158M, and the nonsense mutation R294X were cloned onto expression vector pcDNA3.1_Ct-GFP_AP717 ordered from Invitrogen. Cells were harvested 48-52 hours after transfection. Harvested cells were subjected to confocal microscopy for colocalization analysis of MeCP2 to pericentromeric regions of C2C12 nuclei, or they were subjected to NaCl extractions, HCl isolation and western blotting for quantification of total MeCP2 or MeCP2 phosphorylation at S80. Table 3 contains the list and source of the MeCP2 constructs.

Table 3. MeCP2 constructs for the *in vitro* characterization.

Species-protein-isoform	MeCP2 construct	Expression vector	Provider
Homo sapien - MeCP2-E1	WT	pcDNA3.1_CT-GFP	Dr. John Vincent
	T158M	pcDNA3.1_CT-GFP	Dr. John Vincent
	P152R	pcDNA3.1_CT-GFP	Dr. John Vincent
	R306C	pcDNA3.1_CT-GFP	Invitrogen
	Modified PEST+T158M	pcDNA3.1_CT-GFP_AP717	Invitrogen
	nullPEST+T158M	pcDNA3.1_CT-GFP_AP717	Invitrogen
	R294X	pcDNA3.1_CT-GFP_AP717	Invitrogen

2.2 Confocal fluorescence microscopy

Cells were fixed on coverslips (Bellco Glass Inc; 501944702) by 4% paraformaldehyde (Electron Microscopy Sciences; 15170) in PBS at -20ºC for 10 minutes. Following three washes with PBS, cells were stained with DAPI. (Invitrogen; D1306) to a final 1µg/ml concentration. The cell membrane was lysed with triton-X to minimize background autofluorescence to a final concentration of 0.1% triton in PBS. Cells were incubated with the diluted DAPI and triton mixture for 20 minutes at room temperature in the dark. Cells were washed with PBS three times, and coverslips were mounted face down on a glass slide using Immu-Mount™ (Thermo Shandon Limited; 9990402) and were stored in a dark place at 4 ºC until imaging. Confocal fluorescence images were taken using CFI Plan Flour 100XS Oil (NA 1.3) Nikon objective lens on the Nikon C2 confocal microscope. Channel 1 (DAPI) had a laser wavelength of 405nm and laser power of 10, while channel 2 (GFP) had a laser wavelength of 488nm and a laser power of 5. Images were taken in Z-stacks with a 2µm step size. Pixel width and height were 0.103 µm, and the image size was 2048 pixels. Images were captured with NIS-Elements imaging software (https://www.microscope.healthcare.nikon.com/products/software/nis-elements).

2.2.1 Pearson's correlation coefficient and scatter plots

Colocalization of MeCP2 to DAPI stained chromocenters was analyzed with ImageJ (https://imagej.nih.gov/ij/download.html). The background was subtracted by measuring its mean and removing it from the image. The Colocalization Finder plugin generated the scatter plot and Pearson's correlation coefficient $(PCC)_r$. The equation for PCC (r) is as follows:

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

X and y are the pixel values of red and green. N is the number of observations of both red and green values as they are done in pairs. The final number r will be between -1 and 1, where -1 is interpreted as anticorrelation where the variables tend to avoid each other; where the red pixels (DNA) would be present, the green pixels (MeCP2) are absent. If the r value is 0, it implies no correlation between the red and green variables, and r of 1 implies perfect correlation[146], which is not achieved in practice. Hence, the correlation values for this colocalization analysis are between 0 and 1.

2.3 FRAP analysis

For the FRAP analysis of wildtype and mutant MeCP2 proteins, C2C12 cells were seeded and transfected at approximately 60-70% confluency on 35 mm MatTek dishes. The Qiagen Effectene transfection kit was used with some modifications to the protocol. 100 μl of EC buffer was added along with 3 μl of Enhancer and 1 μg of DNA. This was incubated at room temperature for 20 minutes. 5 μl of Effectene was added, and the solution was incubated for another 20 minutes at room temperature before being added to the cells. Media was changed the following morning, and FRAP experiments were performed using a Zeiss LSM710 laser scanning confocal microscope with a 63x 1.4NA Oil DIC Plan-Apochromat objective lens and a fluorescent photon multiplier tube detector. Entire heterochromatin domains were bleached with a 488 Argon laser. 3 pre-bleach images were collected, and then time-lapse images were collected every second for the first 20 seconds and then every 5 seconds for the remainder of the FRAP experiment. During imaging, cells were maintained in a live cell chamber at 37℃ with 5% CO2. FRAP curves were normalized to the intensity of the whole nucleus throughout the FRAP experiment to account for the FRAP ROI and photobleaching of the entire cell as the result of repeated imaging.

2.4 Nuclei isolation

Whole P30 mouse brains previously frozen at -80°C or fresh nuclei of C2C12 cells were homogenized with a Dounce in 4 volumes of lysis buffer A (0.25 M sucrose, 60 KCl, 15mM NaCl, 10 mM MES pH 6.5, 5mM $MgCl_2$, 1mM $CaCl_2$, 0.5% triton X-100, protease inhibitor cocktail at 1:100 [cOmplete™ Mini, EDTA-free Protease Inhibitor Cocktail, Roche; 5056489001]), and phosphatase inhibitor at 1:50 (PhosSTOP ™ Roche; 4906837001) and incubated on ice for 3 - 6 minutes. Lysates were centrifuged at 600 x g for 5 minutes at 4°C. The cytoplasmic fraction was discarded, and the pellet was resuspended in 8 volumes of lysis buffer. The sample was pelleted at 600 x g for 5 minutes. Pellet was resuspended in 4 volumes of resuspension buffer (buffer B: 50mM NaCl, 10mM Pipes pH 6.8, 5mM $MgCl_2$, 1mM $CaCl_2$, protease inhibitor, phosphatase inhibitor). The same centrifugation time and speed were applied. Nuclei were resuspended in 2 volumes of buffer B. With cells, only one round of centrifugation with lysis buffer was completed. For OD measurement, nuclei were diluted with distilled water and vortexed briefly to create a hypotonic pressure and nuclei lysis. SDS was added to the final concentration of 0.2% in 1/200 dilution of nuclei. SDS binds core histones and breaks the DNA-histone interactions. OD_{260} was measured on CARY 1 BIO UV-VIS to estimate nuclei concentration. The dilutions were based on Beer-Lambert law that at 260nm, the extinction coefficient of double-stranded DNA is 0.020 $(\mu g/ml)^{-1}$ cm^{-1} to create an estimate of DNA concentration in nuclei isolates [147].

2.4.1 Alkaline phosphatase (AP) treatment

Nuclei were isolated as described with the exclusion of phosphatase inhibitor from buffers A and B. Isolated nuclei were treated with alkaline phosphatase (Antarctic Phosphatase, New England

Biolabs Inc; M0289S) at a concentration of 1U/5ug nuclei. AP treatment was performed following the manufacturer's protocol.

2.5 NaCl extraction of MeCP2 and H4

60ug of isolated nuclei were diluted in buffer B with protease inhibitor and aliquoted in 6 fractions. Equal volumes of 2x NaCl solutions ranging from 0.2M to 1M in 20 mM Tris-HCl pH7.5, 0.5 mM EDTA, 2 mM DTT, and protease inhibitor were added to fractions. Fractions intermittently vortexed for 10 seconds every 5 minutes for three rounds while on ice. Samples were incubated on ice for 30 minutes and centrifuged at 16'000 x g for 15 minutes at 4ºC. The supernatant was separated and mixed with 2x sample buffer (250 mM Tris HCl (pH 6.8), 4% SDS, 40% glycerol, 2.86 M β-mercaptoethanol, and 0.4% bromophenol blue). The pellet of each fraction was resuspended in 1x sample buffer (125 mM Tris HCl pH 6.8, 2% SDS, 20% glycerol, 1.43 M β-mercaptoethanol, and 0.2% bromophenol blue), sonicated for 2 minutes, and incubated at 100 ºC for 3 minutes before loading on SDS gel.

2.6 HCl extraction of nuclear proteins

Nuclei in buffer B were centrifuged at 600 x g. The pellet was homogenized with HCl to a final concentration of 0.65N with the Dounce. The sample was centrifuged at 13,000 x g for 10 minutes at 4ºC. The supernatant was added to 6 volumes of 100% acetone. The acetone-supernatant mixture was incubated at -20ºC overnight to precipitate nuclear proteins. Centrifugation at 13,000 x g for 10 minutes at 4ºC followed. Pellet was washed with another six volumes of fresh acetone. The pellet was scraped with a thin spatula to facilitate surface contact of nuclear proteins with freshly added acetone. The sample was centrifuged at 13,000 x g at 4 ºC for 10 minutes. The pellet was vacuum-dried with Jouan RC 1010 concentrator centrifuge for 15 minutes at room temperature and stored at -80ºC.

2.7 SDS-PAGE

Samples were loaded onto SDS gel (15% separating, 4% stacking) for 2 hours at 100V. Gels were stained with Coomassie stain (25% isopropanol, 10% acetic acid, 0.27% w/v Coomassie Brilliant Blue G-250) for at least one hour and distained with a solution containing 10% isopropanol and 10% acetic acid for 1-3 hours.

2.8 Western blotting

Proteins were transferred on the nitrocellulose membrane with Tris-glycine transfer buffer (40 Tris, 192mM glycine, 20% methanol) for 2 hours at 400mA or overnight (16 hours) at 100mA. The membrane was blocked with 5% skim milk in PBST (137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4 , 0.1% (w/v)Tween). The primary antibodies were incubated overnight at 4ºC and include anti-MeCP2 (Sigma Aldrich; M9317: 1:8000 dilution); H4 (in-house; 1:50'000); GFP clone 4B10(cell signalling; 2955S; 1:1,000 dilution). Secondary antibodies (IRDye 800 Anti-rabbit; Licor; 611-132-122; 1:10,000 dilution/ IRDye 680 anti-mouse; 926-68070; 1:5,000/ ECL Anti-Rabbit IgG, Horseradish peroxidase; GE Healthcare; NA934V; 1:2000) incubation was at room temperature for 1 hour with shaking. Images were analyzed with Li-Cor Odyssey (LI-COR Biosciences Lincoln, NE, USA) and Li-Cor Image Studio Lite 5.2.5 software.

2.9 Statistical Analysis

Results were expressed as mean ± standard error of the mean (s.e.m) as indicated with $\alpha < 0.05$. GraphPad Prism software performed one-way analysis of variation (ANOVA) with post hoc Dunnett's test, which compares values to the wildtype, and post hoc Tukey's test, which compares all values (https:/www.graphpad.com). Statistical power calculation was completed on the R-studio "pwr" plugin to calculate the number of observations needed for a power of 0.8.

Chapter 3: Biochemical characterization of MeCP2 with deleterious mutations

3.1 MeCP2 homeostasis is lost with several Rett mutations

Based on biophysical analysis, the categorization of MBD mutants yielded three clusters of MBD mutations based on protein stability and DNA binding. T158M is in group two, with moderate stability and DNA binding. The most destabilizing MBD mutation is P152R, yet it has not been experimentally verified [88,98,148]. P152 is not involved in direct DNA binding, and its mutation to arginine retains partial binding relative to T158M [98]. To confirm if this MBD mutant destabilizes MeCP2, I used an *in vitro* approach of C2C12 cells and transient transfection with MeCP2-GFP constructs to compare total MeCP2 levels in transfected cells (**Fig 3.1, A**). As a secondary control, I used the most common NID mutation, R306C **(Fig 3.1, B)**, which, *in vivo*, had indistinguishable protein levels relative to WT, suggesting that the mutation did not impact protein stability [55]. When comparing the levels of MeCP2-GFP in transiently transfected cells, there was no difference in protein levels between T158M and P152R, and contrary to the *in vivo* observation of R306C, this mutant had significant overexpression relative to WT at nearly 2x WT levels **(Fig. 3.1, C).**

Our lab had an intriguing observation in mouse models with the nonsense mutation R294X. The truncation results in MeCP2 missing the critical N-CoR1 interacting domain within residues 302 – 306 and the entire C-terminal domain encompassing the Ct-PEST motif. Unpublished *in vivo* work by our previous lab members on protein quantification by western blotting from four mice showed protein overexpression to varying levels with a mean of ~4x higher than WT. This observation was contradicted by published work, which showed the levels were indistinguishable from WT [102]. To address what could have caused the overexpression of R294X, we asked

whether the deletion of the Ct-PEST contributes to increased MeCP2 levels. Using the *in vitro* approach, we designed double mutants with the critical T158M mutation with modified PEST (mP+TM) or deleted PEST (nullP+TM). Modification entailed the substitution of proline residues with small hydrophobic amino acids, alanine, valine, and glycine, to prevent the formation of stabilizing polyproline type II, while deletion entailed complete removal of the Ct-PEST **(Fig 3.1, B; bottom).** Using transfection and western blotting, I compared MeCP2 levels in R294X, mPEST and nullPEST relative to WT and T158M. R294X showed a robust overexpression of ~4.4 times the WT.

NullPEST yielded ~25% more MeCP2 than T158M, while mPEST did not differ significantly from T158M, suggesting that the substitution of proline residues with small hydrophobic amino acids, alanine and valine, does not significantly change protein stability. The increase in protein levels in the double mutant nullP+T158M suggests that the Ct-PEST domain plays a partial role in the degradation of MeCP2; however, other mechanisms must be involved as the Ct-PEST deletion does not entirely restore MeCP2 levels **(Fig 3.1, E and F).** Altogether, these findings suggest that multiple factors mediate the mechanisms involved in degrading MeCP2, and the Ct-PEST is one contributor that expedites MeCP2 degradation. There have been reports from Dr. Adrian Bird's lab on non-canonical Rett mutations affecting the Ct-PEST. In these CTD truncations, referred to as CTD1 and CTD2, the truncation occurs at P389X and P388X, respectively, and in both cases, the truncation interrupts the stretch of polyproline residues [145]. When protein levels in humanized mouse models were measured, the CTD1 truncation caused a drop in MeCP2 levels. It is important to note that this truncation included two missense mutations, L384H and P385Q [145]. This observation suggests that the CTD has a vital role in protein stability, and specific residues within the Ct-PEST mediate this stability.

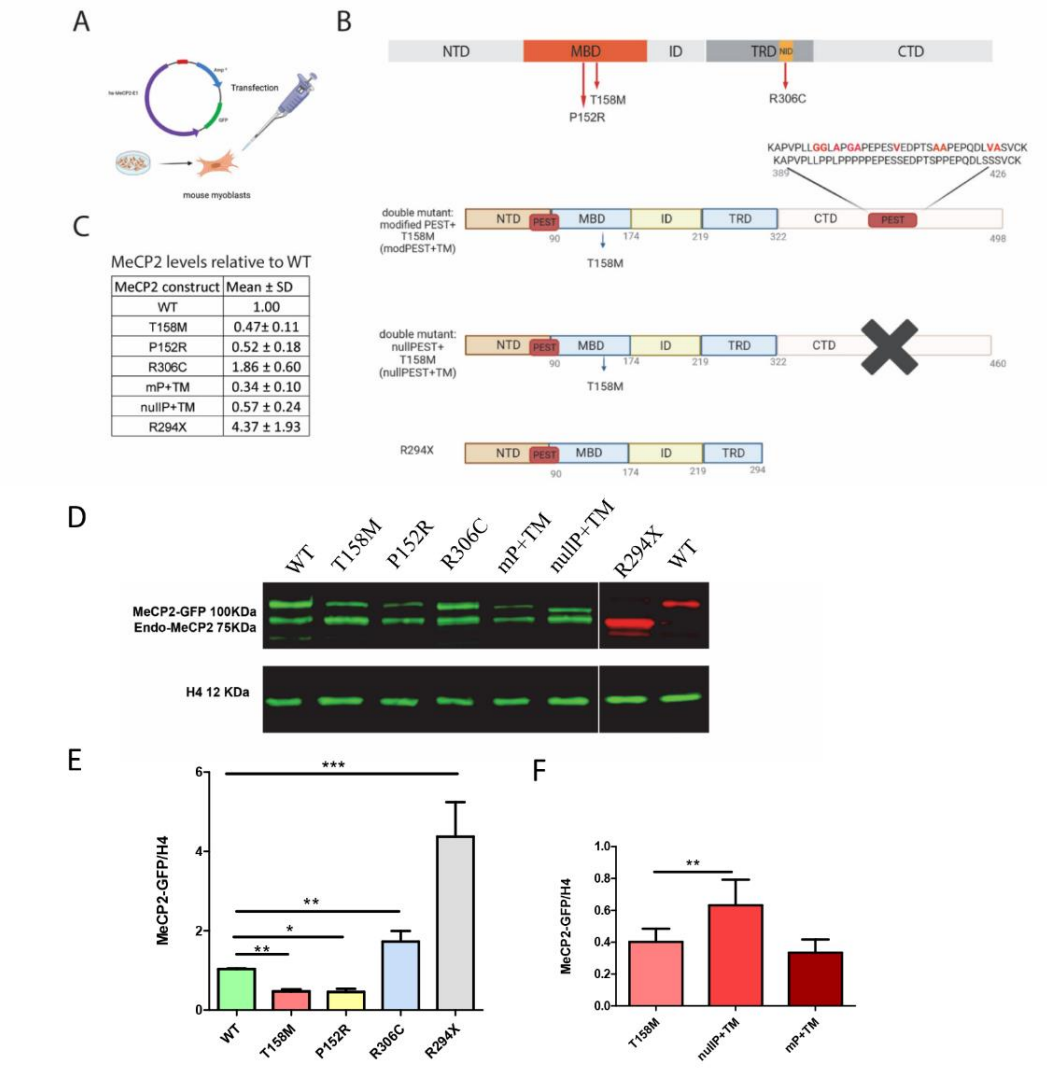


Figure 3.1. Comparison of MeCP2 levels *in vitro*. A. Cartoon diagram of transient transfection of C2C12 with GFP tagged MeCP2. **B.** MeCP2 primary structure depicting the N-terminal domain (NTD), methyl binding domain (MBD), intervening domain (ID), transcriptional repression domain (TRD), nuclear receptor co-repressor interaction domain (NID), and the C-terminal domain (CTD). The approximate position of the MBD mutations P152R and T158M are shown. R306C lies within the critical NID domain (top). Diagram of MeCP2 double mutants, modified PEST, and null PEST, both having the missense mutation T158M (middle). MeCP2 R294X diagram shows the absence of segments of the TRD, the entire NID, and the CTD due to the truncation (bottom). **C**. Mean values with the standard deviation of MeCP2 normalized by histone H4 from western blotting. **D.** Western blots of isolated nuclei of transfected myoblasts with MeCP2 constructs probed with MeCP2 antibody (green) that binds to the C-terminal epitope, which is absent in R294X; hence, this nonsense mutation is probed with GFP antibody (red). **E.** Bar graphs of normalized MeCP2 levels. **F.** Same as E, the double mutants nullPEST and modPEST with the missense mutation T158M are compared with MeCP2 T158M. Data are mean ± SEM; n = 5 biological replicates; one-way ANOVA with post hoc Dunnett's test; * $P \leq 0.05$, ** $P \leq 0.01$; *** $P \leq 0.001$. For quantification of MeCP2, only the upper band, MeCP2-GFP, was used. Endogenous (endo) MeCP2 was excluded.

3.2 NaCl extractions of MeCP2 confirm Rett mutations impact its DNA binding

MeCP2, an IDP, gains its secondary structure by binding to its partners [149]. When unbound, it remains unstructured. A plausible explanation for the rapid proteolysis of MeCP2 with MBD mutations is that due to their reduced affinity to DNA, they are more often in suspension, unbound and exposed to proteolytic machinery; these conditions support rapid degradation [136]. To assess how the strength of DNA binding is affected in mutated MeCP2, NaCl extractions of MeCP2 from nuclei of transiently transfected C2C12 were carried out. The elution profile of MeCP2 in NaCl fractions from 0 to 0.5M was measured using densitometry analysis of the immunoblot **(Fig.3.2 A and B)**. At 0.2M NaCl, T158M MeCP2 showed a significant increase in MeCP2 elution relative to WT, suggesting that this mutation weakens the strength of DNA binding as low concentration of 0.2 M NaCl is enough to dissociate MeCP2 from its target DNA. Moreover, this increased elution of T158M was significantly higher than the double mutant nullP+TM, and to a lesser extent relative to modP+TM. A plausible explanation for this observation is that modifying or deleting the Ct-PEST containing the polyproline stretch enhances DNA binding, suggesting that the elimination of the secondary structure PPII that is the result of the several proline residues adjacent to each other may facilitate DNA binding, although more experimentation is needed to confirm this hypothesis.

At 0.3 and 0.4M NaCl, where most variation amongst mutants was observed, T158M did not show significant differences relative to WT or the double mutants, nullP+TM and mP+TM. However, R306C at both these salt concentrations had significantly lower elution of MeCP2 relative to WT, suggesting this mutant enhances the strength of DNA binding **(Fig. 3.2. B)**. At 0.3M NaCl, R294X also exhibited lower elution **(Fig 3.2. B, middle row),** and this observation corroborated Dr. Jeff Nuel's R294X mouse model, suggesting that this truncation enhances

DNA binding [102]. One can speculate that since this truncation creates a smaller protein, and since it occurs downstream the critical MBD, the protein's ability to associate with DNA is fortified as MeCP2 R294X no longer has the domains that enable protein-protein interactions, which could get in the way of DNA binding. Together, the NaCl extraction of MeCP2 constructs suggests that the mutations that strengthen the DNA binding of MeCP2, such as R306C and R294X, tend to delay its degradation by making the target lysine residues of MeCP2 inaccessible to the E3 ubiquitin ligases and the proteasomal machinery.

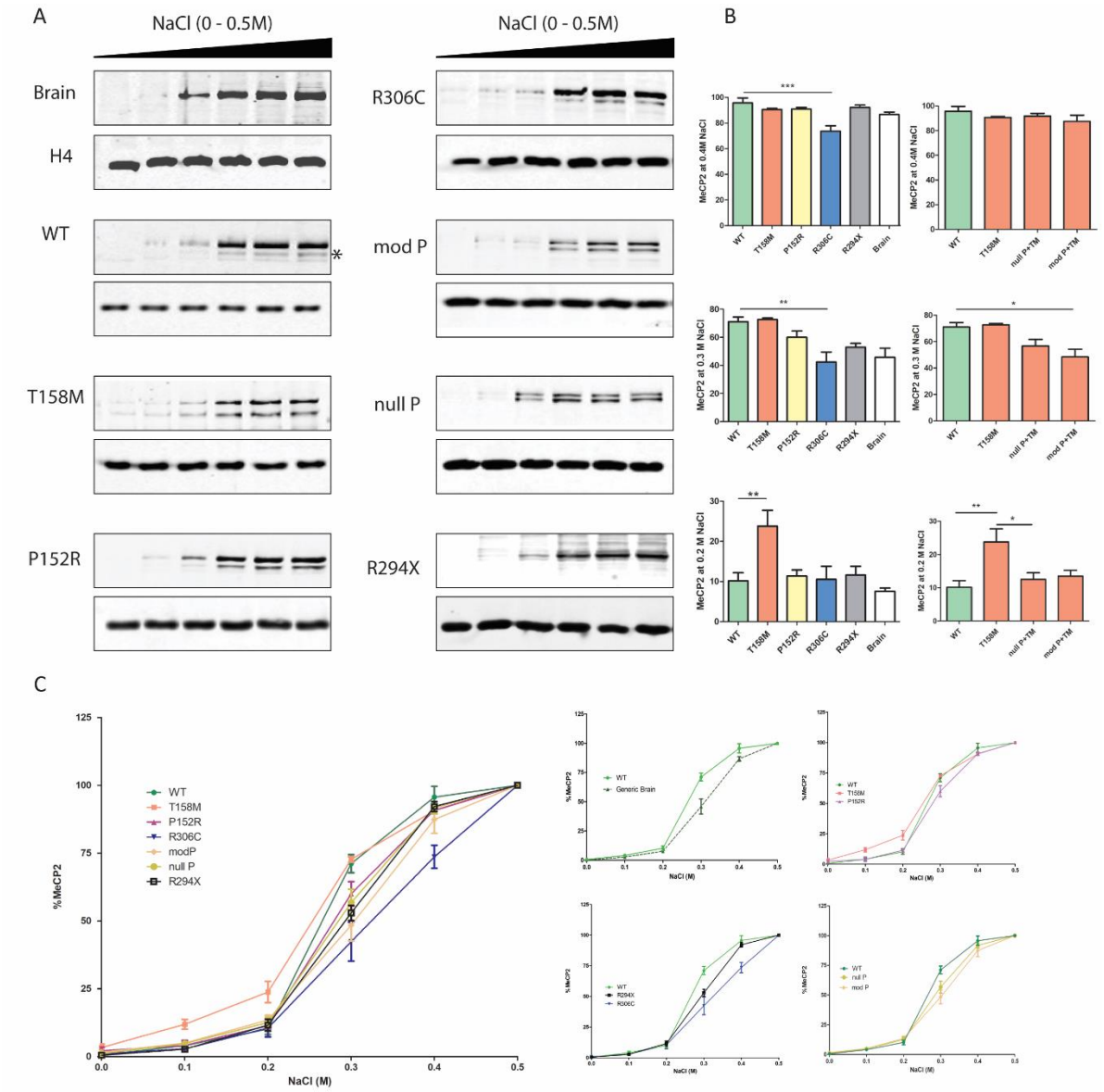


Figure 3.2. NaCl extractions of MeCP2. A. Western blot of the released MeCP2 in salt fractions from 0 to 0.5M. The first western is MeCP2 isolated from nuclei of P30 mouse, undetermined sex, brain (top), normalized by loading control histone H4 (bottom). MeCP2 was isolated from transiently transfected C2C12 normalized by H4 in the remaining westerns. The asterisk next to WT is the endogenous MeCP2. Only the upper band, MeCP2 with Ct-GFP, was used for quantification purposes. R294X was detected with an anti-GFP antibody since the epitope was absent due to truncation. **B.** Bar graphs quantification of MeCP2 released at 0.2 – 0.4 M NaCl. **C.** X, Y plots of released MeCP2 of all constructs in salt fractions (left). Individual comparisons were made for better data visualization, as follows: released MeCP2 in mouse brain and transfected myoblasts with WT MeCP2 (top left); WT with T158M and P152R, both located within the MBD (top right); WT with R306C, a NID mutation, and R294X (bottom left); and WT and T158M compared with the double mutants, nullPEST and modified PEST both having T158M (bottom right). One-way ANOVA with post hoc Tukey's test; n = 5 biological replicates; * $P \leq 0.05$, ** $P \leq 0.01$; *** $P \leq 0.001$.

3.3 MeCP2 phosphorylation at S80 yields a stronger signal in MBD mutants

One of the first identified post-translational modifications of MeCP2 was phosphorylation at Ser421[81] and Ser80 [80], amongst which Ser80 lies within the PEST motif. According to PEST-mediated degradation, PEST domains are conditional switches that induce proteolysis upon post-synthetic modifications; phosphorylation of a serine or threonine is what induces protein degradation as the phosphorylation impairs the stability of PEST regions within the IDP; hence it is followed by lysine ubiquitination [63,64] **(Fig 3.3, A).** To check if phosphorylation at S80 varies amongst mutants, and in particular, if the intensity of this PTM is highest in MeCP2 with destabilizing mutations, the nuclei of transfected cells were isolated, and S80 was detected using phospho-specific antibodies in western blots. The same membrane was stained with MeCP2 antibody for signal normalization **(Fig 3.3 B, C).** A slower migrating band was apparent for all constructs at 110kDa, just 10kDa higher than MeCP2-GFP. Interestingly, both bands at 100 and 110kDa were diminished upon alkaline phosphatase treatment of the nuclei **(Fig 3.3 B),** suggesting that both are due to phosphorylation. Both T158M and P152R exhibit the highest band ratio at 110kDa to the band at 100kDa. I postulated that the slower migrating band is because of MeCP2 ubiquitination, albeit it disappears upon alkaline phosphatase treatment. Ubiquitination at Lys82 (ubK82) has been identified by mass spectrometry [82]. It is highly probable that upon heat stress, which was part of alkaline phosphatase treatment to incubate the nuclei at 37 ºC, the ubiquitinated MeCP2 is quickly degraded because increased temperature stimulates the degradation of newly synthesized proteins via the proteasome system [150]. All phosphorylation signals were normalized by dividing the S80p signal by that of MeCP2-GFP, as shown in yellow bars. When comparing R294X levels, which would be approximately 25kDa lighter than full-length MeCP2, a band at 110kDa

appeared, which suggests non-specific binding at this region **(Fig 3.3 C; S80 western)**; hence, the R294X bar column is excluded from the pink bars corresponding to the band at 110kDa. When comparing the double mutants nullPEST and modified PEST, they both exhibit an intense slower migrating band at 110kDa; however, due to high variability amongst replicates, the lower band at 100kDa did not yield any significant result **(Fig3.3, C; navy bars)**. Nevertheless, all mutants with the missense mutation T158M show a higher ratio of the upper band to the lower band relative to WT **(Fig3.3, C; grey bars).**

Due to the lack of ubiquitin-specific antibodies, whether the slower migrating band is due to MeCP2 ubiquitination at K82 remains to be determined. I have tried eight in-house ubiquitin-specific antibodies, but they showed widely unspecific binding to all proteins (data not shown). While knowing that there is a background signal at 110kDa as identified by the S80p antibody in MeCP2 R294X mutant and also observed in non-transfected cells (data not shown), one way to assess if there is variability in this signal is to transfect the C2C12 cells with MeCP2 S80A, and compare the intensity of the signal. In future studies following this thesis, I plan to transfect C2C12 cells with mutants that include the S80A mutation and check for the presence of the band detected by the MeCP2 S80p antibody.

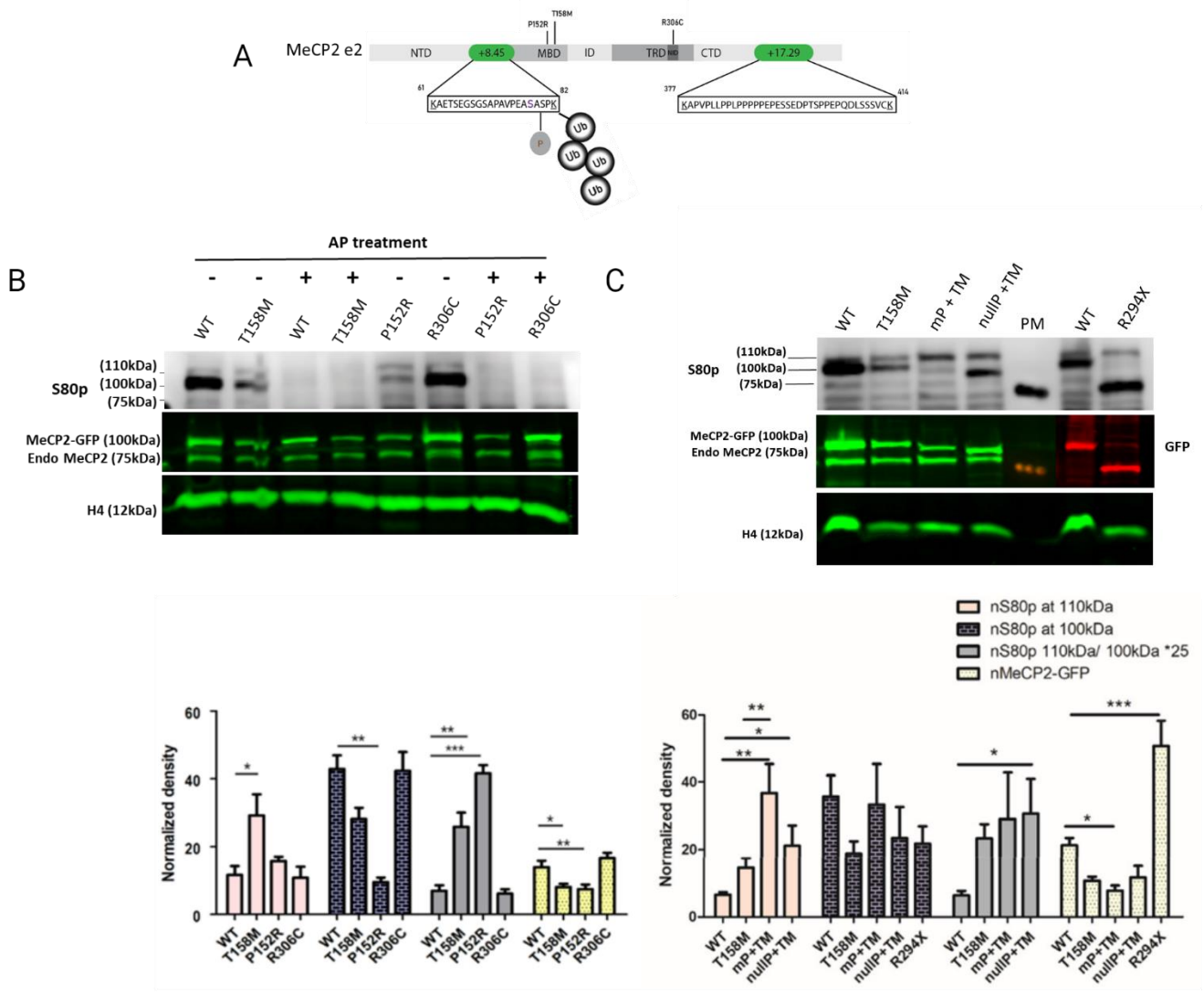


Figure 3.3. Phosphorylation at Ser-80 detected by phospho-specific antibody to MeCP2 S80. A. Cartoon diagram of MeCP2 e2 with the score, position, and sequence of its PEST domains. Phosphorylation of S80 precedes K82 ubiquitination, according to the PEST-mediated degradation hypothesis. **B**. Western blot of transfected C2C12 nuclei with and without treatment with alkaline phosphatase (AP), stained against S80p antibody; same membrane re-stained against MeCP2. Histone H4 was used as a loading control (top). Quantification of the phosphorylation signal (bottom): pink bars depict phosphorylation at S80 at 100KDa, equal migration as MeCP2-GFP. Navy bars are S80p signals of the slower migrating band at 110kDa, overlapping with ubiquitylated MeCP2. Grey bars are the ratio of the slower migrating band at 110kDa to the band at 100kDa multiplied by 25. Yellow bars depict the protein levels. **C.** Same as B, without AP treatment. The double mutants modified PEST (mP) and null PEST (nullP) have the missense mutation T158M. R294X was stained with GFP antibody as the epitope was removed due to truncation. The same densitometry analysis as B was performed. One-way ANOVA with post hoc Dunett's test; n = 3 biological replicate; * $P \leq 0.05$, ** $P \leq 0.01$; *** $P \leq 0.001$.

3.4 Summary

The findings from the biochemical characterization of MeCP2 with loss of function mutations revealed three crucial aspects of MeCP2. Firstly, in MeCP2 with MBD mutations that destabilize the protein and expedite its degradation, the Ct-PEST domain was found to play a role in MeCP2's degradation as its deletion yielded ~25% higher protein levels. Secondly, the mutants that showed higher MeCP2 levels relative to WT were found to associate with DNA more strongly, and this increased binding, as detected by the NaCl elution profile of MeCP2, could play a role in delaying the degradation of MeCP2 by making the target residues inaccessible to proteasomal machinery. Lastly, the immunoblotting analysis of MeCP2 phosphorylation at S80 using a phospho-specific antibody showed a double banding pattern with a slower migrating band about 10kDa higher than MeCP2. The intensity of this band was highest in MeCP2 T158M, P152R and the double mutants nullPEST and modPEST. I speculated that the slower migrating band could be due to MeCP2 ubiquitination at K82, but this cannot be confirmed due to a lack of ubiquitin-specific antibodies. Future work using S80A MeCP2 will determine if the slower migrating band is due to MeCP2 ubiquitination, which, if it is, would agree with the PEST hypothesis.

Chapter 4: Cytological analysis of MeCP2 with deleterious mutations

4.1 MeCP2 clustering to chromocenters is severely impaired with MBD mutations

MeCP2 mutations, especially those within the MBD, are known to impair its binding to target DNA. To check how MeCP2 mutations affect its colocalization to target DNA, mouse myoblasts were transiently transfected with MeCP2 constructs. C2C12 cells show dense pericentromeric regions, and the heterochromatin DNA stains clearly with DAPI; hence, they are a great candidate for co-localization analysis. Pearson's correlation coefficient (PCC) analysis and the scatterplots were used as quantitative and qualitative methods of measuring colocalization, respectively. Not surprisingly, T158M, which has been shown to impair MeCP2's colocalization to methylated DNA [88], had the lowest PCC, whereas P152R, another MBD mutation, did not affect colocalization to the same extent **(Fig. 4.1, A and B).** However, when comparing chromocenter numbers, P152R did show the highest number of chromocenters **(Fig. 4.1, D).** In the truncation mutation R294X, which did not impact the MBD domain, the colocalization was indistinguishable from WT as the PCC values are around 0.7 **(Fig 4.1, B; see green and grey bars)**. An intriguing finding was that the chromocenter size for R294X was the largest **(Fig 4.1 A–C; see grey bars).** The larger chromocenters apparent in R294X may suggest how this truncation modifies the chromocenters and induces rearrangements; however, more work is needed to address why this enlargement of chromocenters occurs. R294X chromocenter numbers were similar to WT. The most common Rett-causing mutation outside the MBD, R306C, did not differ significantly from WT in colocalization, chromocenter size or number. The binding profile of MeCP2 follows methylated cytosine, although regions without DNA methylation show significant binding[151,152]. Beyond the MBD, other MeCP2 domains have been reported to associate with DNA. There is a basic cluster within MeCP2's primary sequence in the

TRD, and it is rich in positively charged amino acids that bind DNA *in vitro* and enhance MeCP2 binding to major satellite repeats *in vivo* [40,42,152]. R306C lies within this basic cluster. Although outside the MBD, this mutant influences MeCP2's DNA binding. R306H is one of the mutants that impairs the LLPS of MeCP2 when in complex with DNA, suggesting that this mutant impairs MeCP2's ability to condensate heterochromatin[87]; R306C appears to behave in an opposing manner to R306H as its ability to form condensates seems enhanced because what made R306C different from all other mutants was the appearance of several large chromocenters detected by both DAPI and the GFP channel **(Fig 4.1 A; top of column two).** We initially speculated that one reason for the enlarged appearance of chromocenters is the formation of disulfide bridges between MeCP2 proteins due to the introduction of the new cysteine residues.

To assess if the enlarged chromocenters were caused by MeCP2 binding to form dimers or higher number complexes, nuclei were isolated with WT, T158M, P152R and R306C MeCP2 and sample buffer with and without beta-mercaptoethanol (BME) was added. If there are any disulfide bridges present due to the introduced cysteine, they should remain intact as the reducing conditions are eliminated by excluding BME, allowing covalent linkages between cysteine residues to stay intact. Western blot of nuclei of MeCP2 constructs with and without BME did not exhibit any differences between the constructs (Appendix 1). We could not, however, rule out the possibility that there may be MeCP2 forming larger complexes with other MeCP2 molecules, and due to its higher molecular weight, it may not have been transferred to the nitrocellulose membrane in the 15% SDS gel. This phenomenon has been repeatedly observed in my western blots. To circumvent this issue, I ran isolated nuclei from a transgenic R306C mouse with or without BME on a 12% SDS for ~3x longer to allow band separation. I probed for MeCP2 in a western blot. While the separation between the E1 and E2 isoforms was

evident, there was no appearance of a higher molecular weight MeCP2 that could be due to MeCP2 dimers or oligomers (Appendix 2). Therefore, the chromocenters visibly larger than other MeCP2 contracts are not due to MeCP2 binding to itself via disulfide bridges; hence, the reason for the enlarged chromocenters in R306C requires further investigation.

The double mutants nullPEST and modPEST exhibited a diffuse distribution of MeCP2; however, some cells retained the defined chromocenters. The variability in the shape of chromocenters for the double mutants is shown in Appendix 3. The percentage of nuclei with defined boundaries between the chromocenters and the surrounding chromatin for nullPEST was 23%, and for modified PEST, it was 38%. In both cases, more than half of the nuclei of double mutants had a diffuse appearance where the chromocenter's boundaries were hardly noticeable.

Overall, the confocal fluorescence microscopy analysis of MeCP2 constructs confirms how the mutations within the MBD impair its binding by varying degrees. As expected, T158M causes the most severe impairment in the localization of MeCP2 to methylated DNA. To shed light on the complexity of Rett-causing mutations, chromocenter size and numbers of the transfected cells were analyzed, and it was confirmed that the MBD mutation P152R resulted in the highest chromocenter number, while R294X yielded the largest size of chromocenters. This finding suggests how Rett-causing mutations impair how MeCP2 remodels the heterochromatic regions and that while colocalization may not be affected, the recruitment of other repressive or activating complexes may not take place. Hence, there will be deleterious outcomes for neuronal cell functions where MeCP2 homeostasis is impaired. .

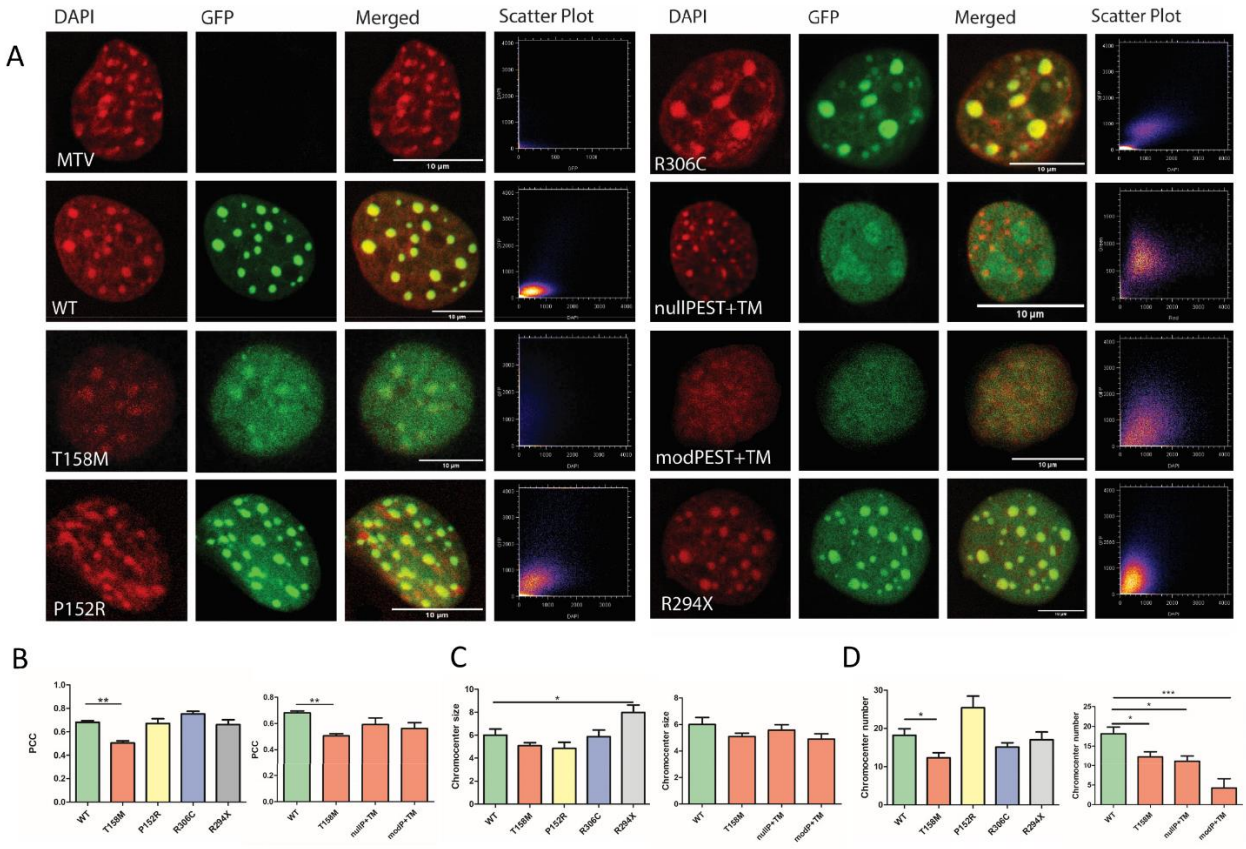


Figure 4.1. Confocal fluorescence microscopy of transiently transfected C2C12 with MeCP2 constructs. A. Confocal image stacks of DAPI (red), GFP (green), and merged channels. The scatter plot depicts the overlap of red and green pixels. MTV is the empty vector, pcDNA3.1-Ct-GFP. The missense mutations T158M and P152R are within the MDB, while R306C is part of the NID. The double mutants, nullPEST and modified PEST, contain the T158M mutation, and the nonsense mutation R294X is missing the critical NID and the entire Ct-domain. **B.** Pearson's correlation coefficient r_p (PCC) values of the DAPI and GFP within the nuclei for quantitative colocalization of MeCP2 to chromocenters. The bar graph to the right compares the double mutants nullPEST and modPEST with the missense mutation T158M to WT MeCP2 and T158M MeCP2. **C.** Chromocenter size of ~ 50 chromocenters measured from the DAPI channel in transfected cells; the bar graph at the right is the same as B. **D.** Chromocenter number of the nuclei of transfected C2C12. Data are mean ± SEM; n = 10 – 15 from three biological replicates. One-way ANOVA with post hoc Tukey's test; * $P \leq 0.05$, ** $P \leq 0.01$; *** $P \leq 0.001$.

4.2 FRAP analysis shows rapid recovery for mutations that enhance proteolysis

Fluorescence recovery after photobleaching (FRAP) experiments involve the bleaching of fluorescently labelled proteins at a predetermined region within a cell using a high-energy laser pulse following measuring the time it takes for the signal to recover at the same area; it is used as a kinetic model for investigating the dynamics of protein diffusion and binding [153,154].

To further assess the protein mobility and binding dynamics, Dr. Michael Henzel and Anastasia Roemer performed the FRAP analysis on transfected C2C12 cells using the same eight constructs **(Fig 4.2).** As shown by the blue arrows, the selected chromocenter was bleached with a 488nm argon laser, followed by time-lapse images from 0 – 60 seconds. Not surprisingly, all mutants with an MBD mutation, including the double mutants, exhibited a faster recovery, suggesting that the binding strength is weaker as it can quickly dissociate and rebound to DNA. In contrast, WT and R294X MeCP2 showed a much slower recovery. After 60 seconds, only partial recovery occurred, correlating with a tighter association of MeCP2 to DNA, restricting MeCP2's mobility**(Fig 4.2, A)**. As depicted in the relative intensity curves in **Fig 4.2, B**, which plot the relative intensity of the fluorophore to its surrounding chromatin as a function of time (s), recovery of P152R after 60 seconds is ~ 0.8. In comparison, the relative recovery of T158M is 0.9, suggesting its higher mobility and weaker binding, which correlates with the initial visual assessment of how T158M is more detrimental to the protein-DNA interaction of MeCP2 than P152R and other published work by Sheikh *et al*., focusing on the degree of binding impairment in MeCP2 MBD mutants[88]. The longer time it takes for the fluorescence signal to be recovered suggests that P152R allows a degree of MeCP2-DNA binding that is more so than T158M.

After 120 seconds, the NID mutant R306C relative intensity is ~ 0.7, while WT and R294X after 120 seconds have a slightly lower intensity of 0.6, suggesting how R306C may impair the

binding dynamics of MeCP2 more so than R294X. In other words, WT and R294X take longer to reach the same intensity of 0.7 than all other mutants, as the binding dynamics are at their minimum for these constructs due to the strength of protein-DNA binding.

As shown in the bar graphs, the half-time recovery better illustrates how WT and R294X are indistinguishable in their half-time recovery and have a value of ~70s. In contrast, all other missense mutations have significantly lower recovery times, with R306C at ~40s, P152R at 14s, and T158M at ~4s **(Fig 4.2, C).** When comparing the half-recovery time of the double mutants nullPEST and modPEST to T158M, the difference is within 2-3 seconds, yet both double mutants cause an increase in recovery time. The increase is significant when only comparing the three mutants **(Fig 4.2, D).** Although this increase is minimal, it suggests that the Ct-PEST domain may play a part in stabilizing the DNA binding.

Overall, the FRAP analysis of MeCP2 constructs provides insight into the binding dynamics and diffusion of MeCP2 to the densely methylated pericentromeric regions. Mutations that impair the DNA binding, such as those that are within the MBD, allow MeCP2 to quickly diffuse and, therefore, recover the fluorophore faster than other mutations that do not affect MeCP2's DNA binding, such as the R294X and R306C.

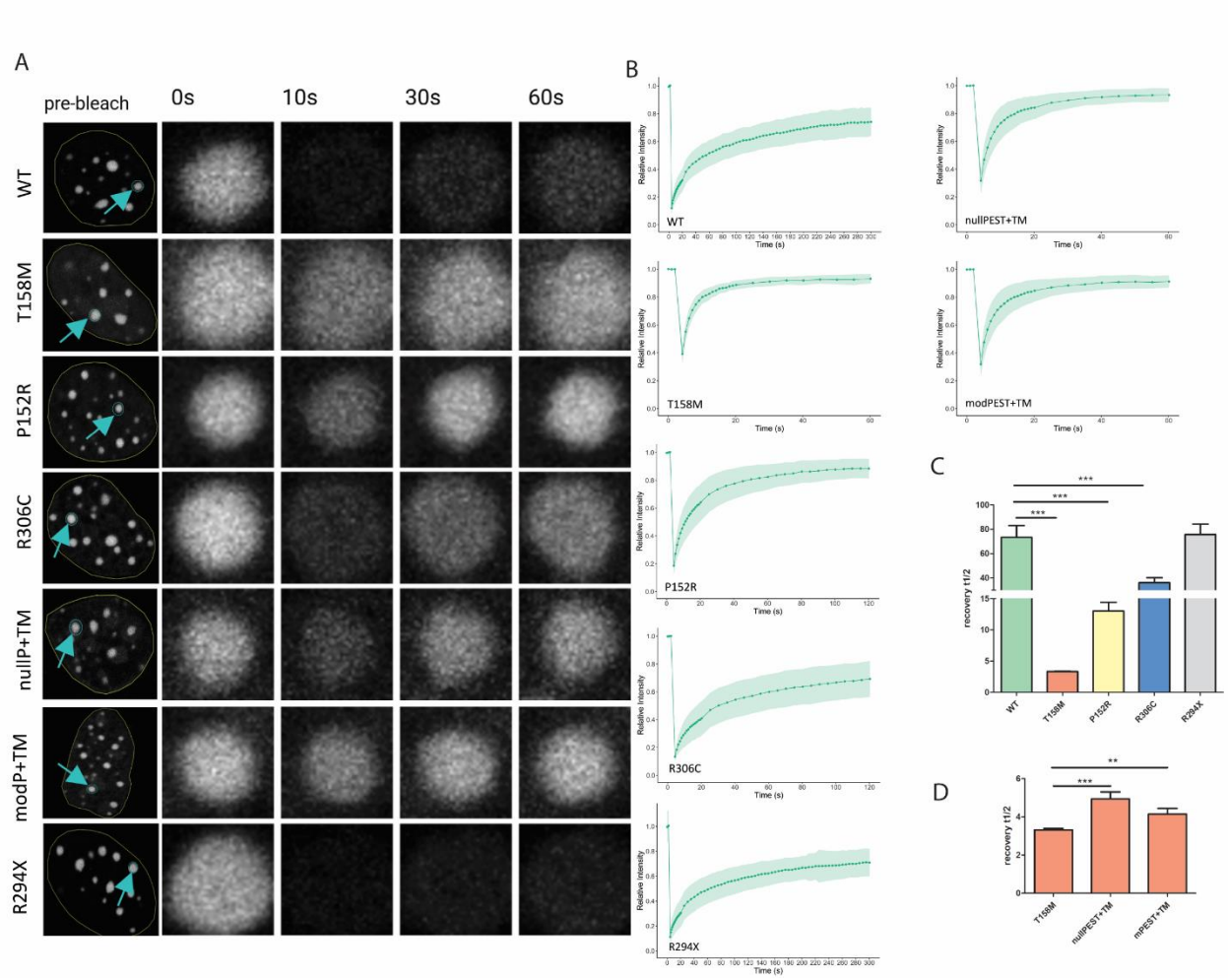


Figure 4.2. FRAP analysis of transfected C2C12 for comparing the mobility and binding dynamics of MeCP2 to chromocenters. A. Fluorescence recovery of the first 60 seconds of the GFP at the C-terminal of transfected MeCP2 after photobleaching. The blue arrow specifies the chromocenter that is targeted by the laser. The recovery for the first 60 seconds at 0s, 10s, 30s, and 60s is shown next to the pre-bleached full view of the cell. **B.** FRAP recovery curves of MeCP2 constructs at the chromocenters; n = 30-32. **C.** Bar graph comparisons of half recovery time of GFP in MeCP2 constructs. One-way ANOVA with post hoc Tukey's test. Comparisons were made only to WT and T158M. **D.** Same as C, except comparisons are made to T158M as both double mutants, nullPEST and modified PEST contain the missense mutation T158M. Unpaired t-test. P-values for nullPEST and mPEST are < 0.0001 and 0.0045, respectively. Data are mean ± SEM; ** $P \leq 0.01$ *** $P \leq 0.001$.

4.3 Summary

The cytological *in vitro* analysis of MeCP2 with loss of function mutations showed how colocalization of MeCP2 to densely methylated pericentromeric regions of myoblasts is most disturbed in T158M MeCP2, the most common MBD mutation [38,90]. In the confocal fluorescence microscopy images, the size and number of chromocenters of transfected cells were measured. The truncated protein R294X exhibited the largest size of chromocenters among all mutants. In addition, the apparent aggregation of chromocenters that appeared for R306C was not because of MeCP2 binding to itself via disulfide bridges. This enlarged appearance of chromocenters of R306C MeCP2 has never been addressed in the literature, and it is worthy of investigation. Moreover, protein diffusion and binding dynamics of MeCP2 analyzed by FRAP studies confirmed how T158M is a most disruptive mutation to protein-DNA binding as it had a half-time recovery of ~ 4 seconds, whereas that of WT was 70 seconds. All the missense mutations resulted in a significantly faster recovery of fluorophore. In contrast, R294X had an indistinguishable half-time recovery relative to WT, suggesting how this truncated MeCP2 is tightly associated with methylated DNA despite missing parts of the TRD and the entire CTD. When comparing the recovery time of the double mutants to T158M, the recovery for both mutants was significantly higher; however, this increase only accounted for 2-3 seconds, which, compared to WT, is a weak improvement. Overall, the colocalization and FRAP analysis of MeCP2 with deleterious mutations provides invaluable insight on a molecular level. Confocal fluorescence microscopy analysis showed how these mutations influence the chromocenter size, like the R294X, which had the largest chromocenter size and R306C, which had several very large but also very small chromocenters. When looking at chromocenter numbers, P152R exhibited the highest chromocenter number. The abolished colocalization, as seen in the T158M

and the double mutants, was observed in the FRAP and colocalization analysis. The diffusion of MeCP2 in the FRAP studies confirms the severity of the MBD mutations in disrupting protein-DNA binding. The mutants R306C and R294X that showed the highest colocalization, as assessed by the PCC values, also showed the longest half-recovery time in the FRAP studies, with values that were near or indistinguishable from WT.

Chapter 5: Discussion

A critical step in addressing treatments for RTT and other MeCP2-related disorders is to recover the MeCP2 dose to its normal levels, as many of the loss-of-function mutations impair the dose of this protein, or in uncommon occasions, duplication of *MECP2* as observed in MDS results in overexpression of MeCP2, yet it manifests many of the RTT phenotypes. I generated an *in vitro* system using mouse myoblasts C2C12 cells and six RTT-causing mutants, T158M, P152R, R306C, R294X, nullPEST, modPEST, and the WT of human MeCP2 E1 isomer to evaluate how they affect protein levels. This system was used for visual and biochemical assessment of the mutants and to analyze how they impair protein levels and their DNA-binding and dynamics. The C2C12 cell lines were a great model for visually evaluating colocalization and FRAP studies because of their defined heterochromatic regions, making the binding assessment simple and easily measurable; moreover, due to the striking similarities that C2C12 cells have with neurons, they are an ideal candidate for MeCP2 studies. One disadvantage, however, in using C2C12 cells is that they originated from a female mouse's thigh muscle after a muscle injury. Both being from a mouse and harvested following physical stress may create conditions that do not resemble human neurons. For comparison, human embryonic kidney cells, HEK293, were transfected with MeCP2 constructs and imaged. The DAPI staining of DNA is much more diffuse as HEK cells' chromatin structure does not have defined pericentromeric regions (Appendix 4). The evaluation of protein abundance for some mutants, such as P152R and the double mutants nullPEST and modPEST, has never been done before. Moreover, the generation of the two mutants that had their Ct-PEST domain modified or deleted was exclusive to our lab, and it was done with a primary focus on assessing the role of Ct-PEST domain in proteasomal degradation of MeCP2 with destabilizing mutations such as those within the MBD.

It has been over a decade since our lab proposed the PEST-mediated degradation of MeCP2 [62]. In 2017, Dr. Zhaolan Zhou and colleagues had an intriguing finding that rekindled the interest in the PEST hypothesis, which concerned a transgenic mouse model with the most common missense mutation in the MBD, T158M [38]. They found that the proteasome is responsible for the degradation of T158M MeCP2, in agreement with the PEST hypothesis. Beyond having two putative PEST domains, detected and scored by the PEST algorithm [78], numerous other molecular characteristics of MeCP2 also hinted at the possibility of PEST-mediated degradation, such as being intrinsically disordered, having a relatively short half-life, specific to the E1 isoform [53], and being constitutively phosphorylated at a Ser-80 [80] which resides within the Nt-PEST domain - a prerequisite of Lys-ubiquitination in PEST mediated degradation [64,79].

I addressed the abundance of MeCP2 with missense and nonsense mutations to first check if other MBD mutants behave similarly to T158M and assess if the reduced bioavailability of the protein can be recapitulated with other destabilizing mutations, such as P152R. In the following biochemical assessment of MeCP2 constructs, I focused on the level of Ser-80 phosphorylation. In the mutants that resulted in the fastest degradation of MeCP2, a more robust signal was detected right above transfected MeCP2. The slower migrating band is speculated to be ubiquitinated MeCP2 at Lys-82, yet this could not be confirmed due to a lack of ubiquitin antibodies with high specificity and affinity. Another phosphorylation event that is situated on MeCP2 Ct-PEST is at Ser-401 [82], albeit there were no phospho-specific antibodies for this epitope. Illustrated in **Fig 5.1[79]** is the predicted and experimentally confirmed MeCP2 PTM[80,82] that is situated with the two PEST motifs of MeCP2, including SUMOylation sites at K429 of mouse MeCP2[155], which is one of the flanking lysine residues of the Ct-PEST domain. The

alpha-fold prediction of polyproline type II (PPII), which in PEST-mediated degradation is predicted to be destabilized upon serine phosphorylation, is also shown [50,51].

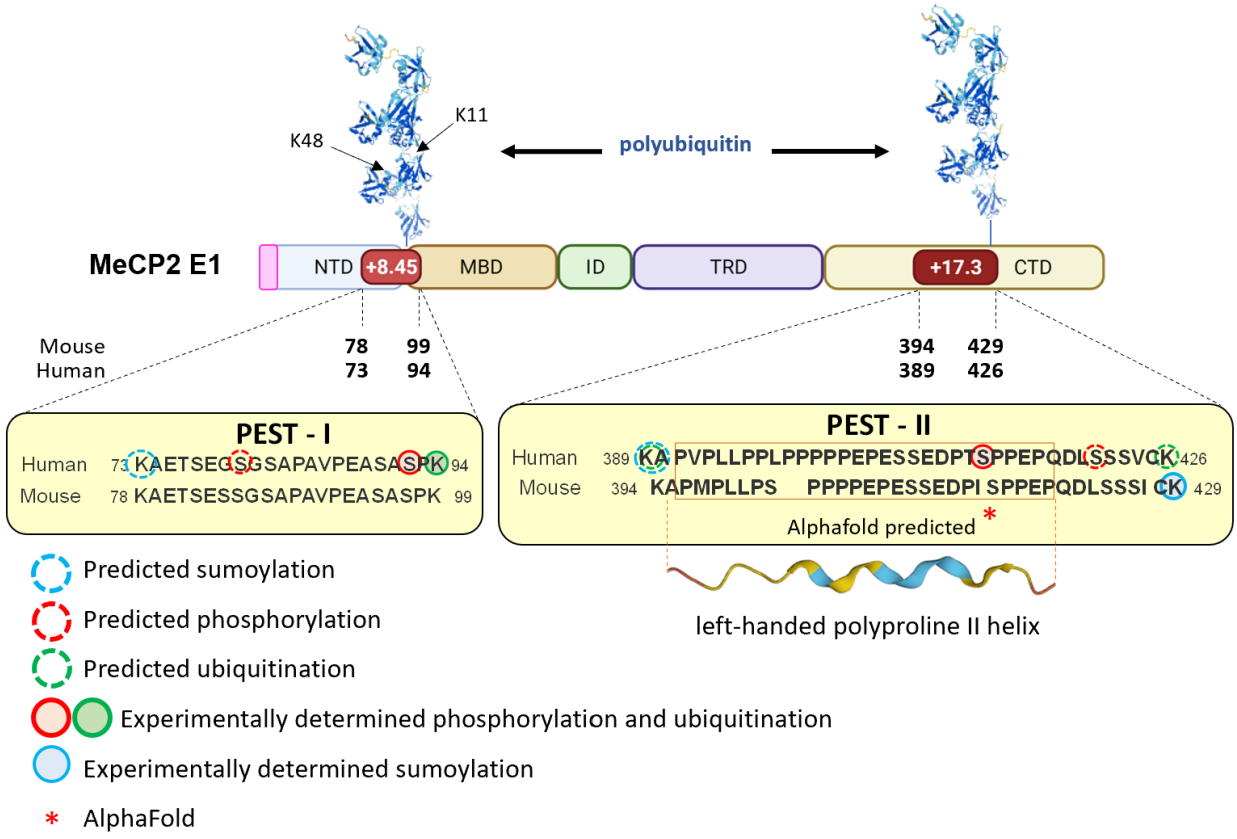


Figure 5.1. Cartoon diagram of MeCP2-E1 shows the location and PTM of its PEST domains. The sequence and location of the PEST domains of human and mouse MeCP2 are shown with the predicted and experimentally determined sites of sumoylation, phosphorylation and ubiquitination. The polyubiquitin chains are predicted to target the lysine residues that flank the PEST motifs through the K48/K11 linkage between the ubiquitin monomers. The alpha fold prediction of the left-handed polyproline type II helix is shown beneath the proline stretch of the Ct-PEST.

Because I could not probe for other phosphorylation events of MeCP2, I took a different approach based on revisiting an observation in the transgenic mouse with the nonsense mutation R294X. This mutant resulted in overexpression of MeCP2, and since this mutant was missing the

entire CTD, including the Ct-PEST that had a higher score, we speculated the deletion of the Ct-PEST may have contributed to the increased levels of MeCP2. For this reason, I integrated the double mutants into my *in vitro* system. I found when the Ct-PEST is deleted, there are increased protein amounts, correlating with PEST-mediated degradation. This observation, for the first time in literature, had shown that the Ct-PEST does play a role in the degradation of MeCP2 with destabilizing mutations. However, whether this Ct-PEST influences MeCP2 degradation directly or indirectly remains to be addressed. If the Ct-PEST directly was responsible for the recruitment of the proteasome, then one or more of the serine residues within the Ct-PEST would be phosphorylated, and its flanking lysine residues at K377 and K414 **(Fig 1.6)** would have been ubiquitylated. However, the initial mass spectrometry analysis of MeCP2 did not find any ubiquitination in the CTD of MeCP2 [82]. It is important to note that the MS analysis focused on the PTMs in WT MeCP2; hence, looking at the PTM profile of MeCP2 T158M remains a critical step in further assessing where the ubiquitination occurs. Knowing that the E3 ubiquitin ligase that ubiquitinates MeCP2 is RNF4 [156] allows for a more targeted approach in deciphering the nature of MeCP2 ubiquitination, that is, in addition to finding which lysine is ubiquitinated, finding whether it is poly or mono ubiquitylated and via which linkage (homotypic or heterotypic and branched), remain essential questions amongst others. What is the outcome of ubiquitination? Does it act as a degradation or mislocalization signal, increasing its likelihood of degradation [157]? In other words, the direct or indirect role of the PEST motif in the proteasomal degradation of MeCP2 with destabilizing mutations such as T158M and P152R remains to be investigated. Depicted in **Fig 5.1** is the model proposed for MeCP2 ubiquitination in MeCP2 T158M via RNF4 ubiquitin ligase, which has better accessibility to MeCP2 due to its abolished DNA binding.

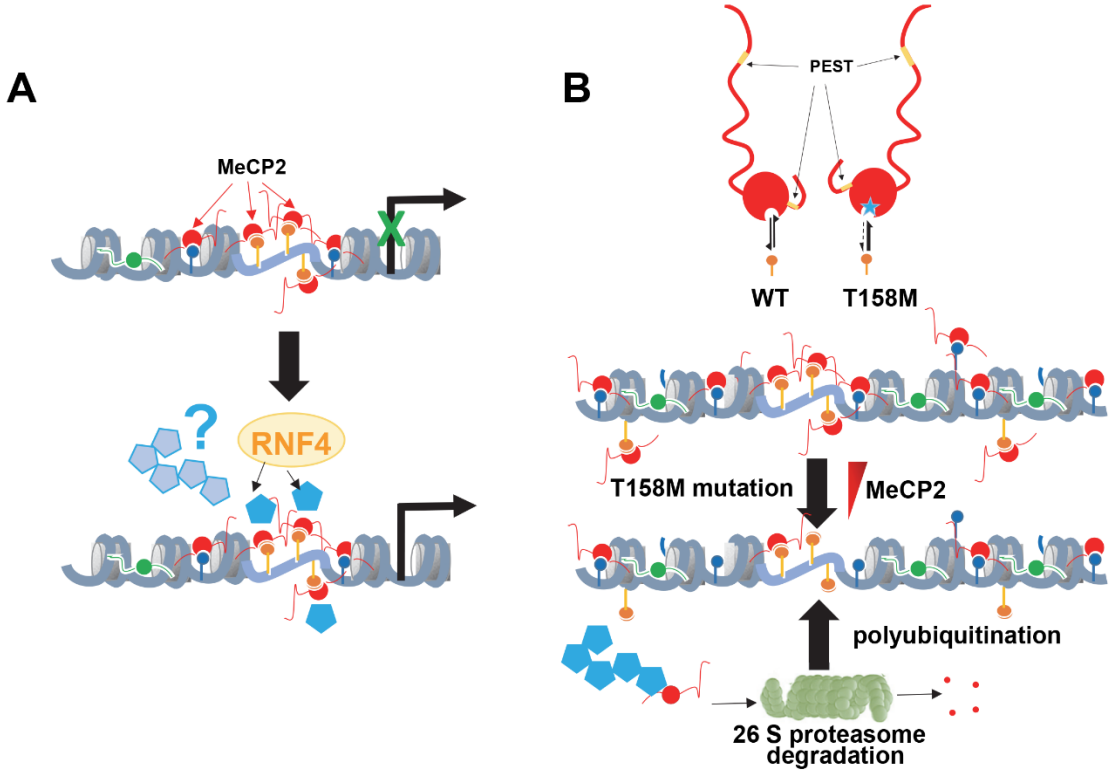


Figure 5.2. Model for MeCP2 ubiquitination in chromatin and PEST-mediated degradation.[79] **A.** RNF4 is a transcriptional co-activator, and the E3 ubiquitin ligase that ubiquitylates (blue pentagons) MeCP2 (red circles with lines), causing it to be released from the methylated promoter of genes (orange circles) [156], promoting transcription. **B.** Since T158M abolishes the DNA binding ability of MeCP2, it causes the protein to be in solution and more available for ubiquitination, which expedites its degradation.

A prerequisite of Lys-ubiquitination by RNF4 ubiquitin ligase is the accessibility of the target lysine residues of MeCP2[136]. The NaCl extractions of MeCP2 showed how in the mutants that resulted in higher MeCP2 levels than WT, such as R306C and R294X, less MeCP2 was being eluted at 300mM salt concentration which implies that MeCP2 R306C and R294X have tighter DNA binding. It would take higher ionic strength to dislodge them from DNA. Their tight binding can explain why the protein levels were higher, as when they are bound to DNA, they are protected from the 26S UPS proteasomal machinery by being inaccessible to ubiquitination.

In contrast, at lower salt concentrations where minimal MeCP2 was eluted in WT, T158M was easily dislodged from DNA, and while unbound and in solution, it will quickly become degraded, as observed in the protein quantification analysis; however, the same cannot be said about MeCP2 P152R which, in its NaCl elution, it did not significantly differ from WT despite having low protein levels similar to that of T158M. This contradiction in NaCl extraction of MeCP2 and protein levels sheds light on the complexity of the mechanisms responsible for MeCP2 degradation and how other mechanisms influence MeCP2 binding and its turnover. As observed in the colocalization and FRAP analysis, MeCP2's DNA binding and dynamics are most impaired by the mutations resulting in its rapid degradation. This evidence supports the accessibility of MeCP2 to the proteolytic machinery: when MeCP2 is free in solution, it is more likely to be degraded, as it is intrinsically disordered and exposed. While the NaCl extractions showed R306C having the highest binding strength than WT, and R294X was ranked as the second highest binding strength, in the FRAP analysis, the binding dynamics of R294X ranked as the mutant with the slowest recovery, similar to WT levels, and R306C was ranked the 2nd mutation that has the relative slowest recovery to WT. This observation suggests that DNA binding strength analysis, as observed in NaCl extractions of MeCP2, and diffusion and binding analysis, as observed in FRAP, may not always correlate, as FRAP analysis does not provide information on the binding strength. Deleting the Ct-PEST improved the protein amounts but did not enhance colocalization or recovery in the FRAP analysis to near WT conditions. This was not surprising because the Ct-PEST located with CTD of MeCP2 does not participate in DNA binding. Deleting the Ct-PEST delays MeCP2 degradation, thereby indirectly affecting its binding. Overall, my findings show that the Ct-PEST is one of the many players that regulate the turnover of MeCP2 with destabilizing mutations. When designing therapeutics and small

molecules that can stabilize the protein, it is essential to know by which mechanisms MeCP2 is degraded. Knowing that the Ct-PEST contributes to MeCP2 degradation opens a new pathway for drug discovery and therapeutics that can target the Ct-PEST to delay MeCP2 degradation.

5.1 Conclusions

My research findings shed light on a crucial prerequisite of MeCP2 degradation and the involvement of its Ct-PEST in its turnover when having destabilizing mutations; the findings can be summarized as follows:

1) The reduced protein availability observed in T158M is recapitulated in P152R, another destabilizing MBD mutant.
2) MeCP2's accessibility to the proteasomal machinery is hindered when MeCP2 is bound to target DNA with higher affinity, as seen in mutants that strengthen the DNA binding of MeCP2, such as R306C and R294X. In other words, MeCP2 is stabilized when it is bound to DNA.
3) Phosphorylation of a serine residue within the Nt-PEST domain of MeCP2, a prerequisite of PEST-mediated degradation, yields the strongest signal in MeCP2 with destabilizing mutations that expedite its degradation, such as P152R and T158M, corroborating the PEST-mediated degradation.
4) Deleting the Ct-PEST motif of MeCP2 with destabilizing mutation, T158M, recovers protein amounts to near WT levels. However, this deletion is insufficient to fully rescue the protein amounts, suggesting other mechanisms are at play, such as the Nt-PEST motif, which requires further investigation.

5.2 Future directions

A follow-up step in assessing PEST-mediated degradation of MeCP2 would be to design plasmids looking at both PEST motifs. While I have focused on the Ct-PEST in designing the double mutants, modified and nullPEST, whether my initial speculation about K82 ubiquitination was actual remains to be investigated. To address this question, one can integrate a plasmid expressing double mutant MeCP2 with two missense mutations, T158M and K82R, and probe the isolated nuclei from transfected cells with phospho-specific antibodies against MeCP2 S80 and check for the presence of the slower migrating band at 110kDa **(See Fig. 3.3).** If that was indeed due to ubiquitination, then while the slower migrating band would be present for T158M, this band in the double mutant T158M/K82R would cease to exist as this mutant cannot be ubiquitylated; arginine can become modified by methylation and not ubiquitination[158]. Moreover, suppose S80 phosphorylation is a prerequisite of K82 ubiquitination; in that case, one can integrate another double mutant with the missense mutations T158M/S80A and check for the slower migrating band relative to T158M when probed with a phospho-specific antibody to S80.

In finding possible therapeutics that can stabilize MeCP2 by targeting the Ct-PEST, it is critical to assess the biological significance of the Ct-PEST and design transgenic mouse models with nullPEST MeCP2. If the deletion of Ct-PEST does not have detrimental effects in mice, then targeting this domain can be explored as the pathway of drug development for Rett and other MeCP2-related disorders that result from MeCP2 dose dysregulation.

As mentioned earlier, the PTM profile of MeCP2 T158M remains to be studied to find the ubiquitylation sites, the linkage type and their outcome. Moreover, there may be more than one ubiquitin ligase of MeCP2. Other than RNF4, are there other E3 ubiquitin ligases that

ubiquitylate MeCP2, and are there deubiquitinases that can be used as a therapeutic approach to stabilize MeCP2?

These are some of the many questions that remain to be addressed about this relatively simple yet functionally pleiotropic and complex protein. Rett, to this day, remains an irreversible condition, yet with advances in CRISPR Cas-9 and other gene therapy approaches, and rigorous research funded by families whose children have been impacted, Rett may soon become a reversible condition in humans.

Bibliography

1. Noll, M. (1974) Subunit structure of chromatin. *Nature*, **251**, 249-251.
2. Luger K, R.T. (1998) DNA binding within the nucleosome core. . *Curr Opin Struct Biol.*, **8**, 33-40.
3. Baldi, S., Korber, P., Becker, P.B. (2020) Beads on a string-nucleosome array arrangements and folding of the chromatin fiber. *Nat. Struct. Mol. Biol.*, **27**, 109-118.
4. Arents, G., Moudrianakis, E.N. (1992) Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. *Proc.Natl.Acad.Sci.* , **90**.
5. Luger K, M.A., Richmond RK, Sargent DF, Richmond TJ. (1997) Crystal structure of the nucleosome core particle at 2.8 A˚ resolution. *Nature*, **389**, 251-260.
6. Wolffe AP, H.J. (1999) Chromatin disruption and modification. *Nucleic Acids Res.*, **27**, 711-720.
7. Hergeth, S.P., Schneider, R. (2015) The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO. Rep.*, **16**, 1439-1453.
8. Louters, L., Chalkley, R. (1985) Exchange of histones H1, H2A, and H2B in vivo. *Biochem.*, **24**, 3080-3085.
9. Thomas, J.O., Rees, C. (1983) Exchange of histones H1 and H5 between chromatin fragments. A preference of H5 for higher-order structures. *Eur J Biochem*, **134**, 109-115.
10. Wu LH, K.L., Rechsteiner M. (1986) Dynamic behavior of histone H1 microinjected into HeLa cells. . *J Cell Biol.*, **103**, 465-474.
11. Lever MA, T.n.J., Sun X, Hendzel MJ. . (2000) Rapid exchange of histone H1.1 on chromatin in living human cells. . *Nature*, **408**.
12. Millan-Zambrano, G., Burton, A., Bannister, A.J., Schneider, R. (2022) Histone post-translational modifications - cause and consequence of genome function. *Nat Rev Genet*, **23**, 563-580.
13. BM., T. (1993) Decoding the nucleosome. . *Cell.*, **75**, 5-8.
14. Lee Y, A.C., Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Rådmark O, Kim S, Kim VN. (2003) The nuclear RNase III Drosha initiates microRNA processing. . *Nature*, **425**, 415-419.
15. Hutvágner G, Z.P. (2002) A microRNA in a multiple-turnover RNAi enzyme complex. *Science*, **297**, 2056-2060.
16. Wightman B, H.I., Ruvkun G. . (1993) Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. . *Cell.* , **75**, 855-862. .
17. Oliveto, S., Mancino, M., Manfrini, N., Biffo, S. (2017) Role of microRNAs in translation regulation and cancer. *World J Biol Chem*, **8**, 45-56.
18. Roth, S.Y., Denu, J.M., Allis, C.D.,. (2001) Histone Acetyltransferases. *Annual Review of Biochemistry* **70**, 81-120.
19. Mona D. Shahbazian, B.A., Dawna L. Armstrong, Huda Y. Zoghbi. (2001) Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. *Oxford University Press*, **11**, 115–124.

20. Ghosh, R.P., Horowitz-Scherer, R.A., Nikitina, T., Shlyakhtenko, L.S., Woodcock, C.L. (2010) MeCP2 binds cooperatively to its substrate and competes with histone H1 for chromatin binding sites. *Mol. Cell. Biol.*, **30**, 4656-4670.
21. Kokura, K., Kaul, S.C., Wadhwa, R., Nomura, T., Khan, M.M., Shinagawa, T., Yasukawa, T., Colmenares, C., Ishii, S. (2001) The Ski protein family is required for MeCP2-mediated transcriptional repression. *J. Biol. Chem.*, **276**, 34115-34121.
22. Agarwal, N., Hardt, T., Brero, A., Nowak, D., Rothbauer, U., Becker, A., Leonhardt, H., Cardoso, M.C. (2007) MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation. *Nucleic Acid. Res.*, **35**, 5402-5408.
23. Maria Chahrour1, S.Y.J., Chad Shaw1, Xiaobo Zhou3, Stephen T. C. Wong3, Jun Qin2,4, and Huda Y. Zoghbi. (2008) MeCP2, a Key Contributor to Neurological Disease, Activates and Represses Transcription. *Science*, **320**, 1224–1229.
24. Petriti, U., Dudman, D.C., Scosyrev, E., Lopez-Leon, S. (2023) Global prevalence of Rett syndrome: systematic review and meta-analysis. *Syst Rev*, **12**, 5.
25. Rett, A. (1966) *Über ein zerebral-atrophisches Syndrom bei Hyperammonämie*. Hollinek.
26. Richard H. Haas, M. (1988) The History and Challenge of Rett Syndrome. *Journal of Child Neurology* **3**.
27. Bengt Hagberg, M., ” Jean Aicardi, MD,? Karin Dias, MD,S and Ovidio Ramos, MDI-. (1982) A Progressive Syndrome of Autism, Dementia, Ataxia, and Loss of Purpos&l Hand Use in Girls: Rett’s Syndrome: Report of 35 Cases. *Annals of Neurology*
28. Percy, A. (2014) The American history of Rett syndrome. *Pediatr Neurol*, **50**, 1-3.
29. Foundation, I.R.S. (2022) *IRSF Research Finding Solutions for all*. https://www.rettsyndrome.org/ (accessed Feb 6).
30. Clinic, B.C.s. (2017) *Remembering Dr. Vanja Holm*. https://www.boyercc.org/family-stories/holm/ (accessed April, 2023).
31. Center, B.B.C.R. (2023) *About the Center*. Baylor College of Medicine https://www.bcm.edu/research/research-centers/blue-bird-circle-rett-center (accessed Jan 30 2023).
32. Coleman, M. (1990) Is classical Rett syndrome ever present in males? *Brain Dev*, **12**, 31-32.
33. Ruthie E. Amir, I.B.V.d.V., Mimi Wan, Charles Q. Tran, Uta Francke, Huda Y. Zoghbi. (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature*, **23**.
34. Joe D.Lewis, Ingrid Maurer-Fogy, Richard Ft. Meehan, William J. Henzel, Peter Jeppesen, Franz Klein, AdrianBird. (1992) Purification, Sequence, and Cellular Localization of a Novel Chromosomal Protein That Binds to Methylated DNA. *Cell Press*, **69**, 905-914.
35. Mnatzakanian, G.N., Lohi, H., Munteanu, I., Alfred, S.E., Yamada, T., MacLeod, P.J., Jones, J.R., Scherer, S.W., Schanen, N.C., Friez, M.J., et al. (2004) A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. *Nat. Genet.*, **36**, 339-341.
36. Guy, J., Gan, J., Selfridge, J., Cobb, S., Bird, A. (2007) Reversal of neurological defects in a mouse model of Rett syndrome. *Science*, **315**, 1143-1147.
37. Lombardi, L.M., Zaghlula, M., Sztainberg, Y., Baker, S.A., Klisch, T.J., Tang, A.A., Huang, E.J., Zoghbi, H.Y. (2017) An RNA interference screen identifies druggable regulators of MeCP2 stability. *Sci Transl Med*, **9**.

38. Lamonica, J.M., Kwon, D.Y., Goffin, D., Fenik, P., Johnson, B.S., Cui, Y., Guo, H., Veasey, S., Zhou, Z. (2017) Elevating expression of MeCP2 T158M rescues DNA binding and Rett syndrome-like phenotypes. *J. Clinic. Inv.*, **127**, 1889-1904.
39. Glaze, D.G., Neul, J.L., Kaufmann, W.E., Berry-Kravis, E., Condon, S., Stoms, G., Oosterholt, S., Della Pasqua, O., Glass, L., Jones, N.E., et al. (2019) Double-blind, randomized, placebo-controlled study of trofinetide in pediatric Rett syndrome. *Neurology*, **92**, e1912-e1925.
40. Adams, V.H., McBryant, S.J., Wade, P.A., Woodcock, C.L., Hansen, J.C. (2007) Intrinsic disorder and autonomous domain function in the multifunctional nuclear protein, MeCP2. *J. Biol. Chem.*, **282**, 15057-15064.
41. Wade, P.A. (2004) Dynamic regulation of DNA methylation coupled transcriptional repression: BDNF regulation by MeCP2. *Bioessays*, **26**, 217-220.
42. Georgel, P.T., Horowitz-Scherer, R.A., Adkins, N., Woodcock, C.L., Wade, P.A., Hansen, J.C. (2003) Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. *J. Biol. Chem.*, **278**, 32181-32188.
43. Mann, J., Oakley, F., Akiboye, F., Elsharkawy, A., Thorne, A.W., Mann, D.A. (2007) Regulation of myofibroblast transdifferentiation by DNA methylation and MeCP2: implications for wound healing and fibrogenesis. *Cell Death Differ*, **14**, 275-285.
44. Dragich J, H.-M.I., Schanen C. (2000) Rett syndrome: a surprising result of mutation in MECP2. . *Hum Mol Genet.* , **9**, 2365-2375. .
45. Thierry Bienvenu, Chelly, J. (2006) Molecular genetics of Rett syndrome: when DNA methylation goes unrecognized. *Nature Reviews Genetics*, **7**, 415–426.
46. Dragich, J.M., Kim, Y.H., Arnold, A.P., Schanen, N.C. (2007) Differential distribution of the MeCP2 splice variants in the postnatal mouse brain. *J Comp Neurol*, **501**, 526-542.
47. Liyanage, V.R.B., Olson, C.O., Zachariah, R.M., Davie, J.R., Rastegar, M. (2019) DNA Methylation Contributes to the Differential Expression Levels of Mecp2 in Male Mice Neurons and Astrocytes. *Int. J. Mol. Sci.*, **20**.
48. Pandey, S., Pruitt, K. (2017) Functional assessment of MeCP2 in Rett syndrome and cancers of breast, colon, and prostate. *Biochemistry and Cell Biology*, **95**, 368–378.
49. Ta, D., Downs, J., Baynam, G., Wilson, A., Richmond, P., Leonard, H. (2022) A brief history of MECP2 duplication syndrome: 20-years of clinical understanding. *Orphan. J. Rare. Disease.*, **17**, 131.
50. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583-589.
51. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., et al. (2022) AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acid. Res.*, **50**, D439-D444.
52. Claveria-Gimeno, R., Lanuza, P.M., Morales-Chueca, I., Jorge-Torres, O.C., Vega, S., Abian, O., Esteller, M., Velazquez-Campoy, A. (2017) The intervening domain from MeCP2 enhances the DNA affinity of the methyl binding domain and provides an independent DNA interaction site. *Sci. Rep.*, **7**, 41635.
53. Martinez de Paz, A., Khajavi, L., Martin, H., Claveria-Gimeno, R., Tom Dieck, S., Cheema, M.S., Sanchez-Mut, J.V., Moksa, M.M., Carles, A., Brodie, N.I., et al. (2019)

MeCP2-E1 isoform is a dynamically expressed, weakly DNA-bound protein with different protein and DNA interactions compared to MeCP2-E2. *Epigenetics Chromatin*, **12**, 63.

54. Brown, K., Selfridge, J., Lagger, S., Connelly, J., De Sousa, D., Kerr, A., Webb, S., Guy, J., Merusi, C., Koerner, M.V., et al. (2016) The molecular basis of variable phenotypic severity among common missense mutations causing Rett syndrome. *Hum. Mol. Genet.*, **25**, 558-570.
55. Lyst, M.J., Ekiert, R., Ebert, D.H., Merusi, C., Nowak, J., Selfridge, J., Guy, J., Kastan, N.R., Robinson, N.D., de Lima Alves, F., et al. (2013) Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. *Nat. Neurosci.*, **16**, 898-902.
56. Wakefield, R.I., Smith, B. O., Nan, X., Free, A., Soteriou, A., Uhrin, D., Bird, A. P., & Barlow, P. N. (1999) The solution structure of the domain from MeCP2 that binds to methylated DNA. *Journal of Molecular Biology*, **291**, 1055-1065.
57. Ho, K.L., McNae, I.W., Schmiedeberg, L., Klose, R.J., Bird, A.P., Walkinshaw, M.D. (2008) MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol. Cell.*, **29**, 525-531.
58. Kraulis, P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *Journal of Applied Crystallography*, **24**, 946-950.
59. Peter L. Jones1, G.J.C.V., Paul A. Wade1, Danielle Vermaak1, Stefan U. Kass2,Nicoletta Landsberger3, John Strouboulis1& Alan P. Wolffe1. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature America Inc.* , **19**.
60. Nan, X., Ng, H.-H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., Bird, A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. . *Nature*, **393**, 386–389.
61. Tillotson, R., Selfridge, J., Koerner, M.V., Kamal K. E. Gadalla, K.K.E., Guy, J., Dina De Sousa, R.D.H., Stuart R. cobb & Adrian bird. (2017) Radically truncated MeCP2 rescues Rett syndromelike neurological defects. *Nature*, **550**.
62. Thambirajah, A.A., Eubanks, J.H., Ausio, J. (2009) MeCP2 post-translational regulation through PEST domains: two novel hypotheses: potential relevance and implications for Rett syndrome. *Bioessays*, **31**, 561-569.
63. Rogers, S., Wells, R., & Rechsteiner, M. . (1986) Amino Acid Sequences Common to Rapidly Degraded Proteins: The PEST Hypothesis. *Science*
64. Singh, G.P., Ganapathi, M., Sandhu, K.S., Dash, D. (2006) Intrinsic unstructuredness and abundance of PEST motifs in eukaryotic proteomes. *Proteins*, **62**, 309-315.
65. Wright, P.E., & Dyson, H. (1999) Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. . *Journal of Molecular Biology*, **239**, 321-331.
66. Tompa, P. (2002) Intrinsically unstructured proteins. *Trends in Biochemical Sciences*, **27**.
67. García-Alai MM, G.M., Salame M, Wetzler DE, McBride AA, Paci M, Cicero DO, de Prat-Gay G. (2006) Molecular Basis for phosphorylation-dependent PEST mediated protein turnover. *Structure*, **14**, 309-319.
68. Rogers, M.R.a.S.W. (1996) PEST Sequences and Regulation by Proteolysis. *Elsevier Science Ltd*.
69. Chevaillier, P. (1993) Pest sequences in nuclear proteins. . *International Journal of Biochemistry*, **25**, 479-482.

70. Loetscher, P., Pratt, G., Rechsteiner, M. (1991) The C terminus of mouse ornithine decarboxylase confers rapid degradation on dihydrofolate reductase. Support for the pest hypothesis. *J. Biol. Chem.*, **266**, 11213-11220.
71. Barnes, J.A., Gomes, A.V. (1995) PEST sequences in calmodulin-binding proteins. *Molecular and Cellular Biochemistry*
72. Bordone, L., Campbell, C. (2002) DNA ligase III is degraded by calpain during cell death induced by DNA-damaging agents. *J. Biol. Chem.*, **277**, 26673-26680.
73. Ghoda, L., Wetters, D., Macrae, M., Ascherman, D., & Coffino, P. (1989) Prevention of Rapid Intracellular Degradation of ODC by a Carboxyl-Terminal Truncation. *Science*.
74. Ghoda, L., Phillips, M.A., Bass, K.E., Wang, C.C., Coffino, P. (1990) Trypanosome ornithine decarboxylase is stable because it lacks sequences found in the carboxyl terminus of the mouse enzyme which target the latter for intracellular degradation. *J. Biol. Chem.*, **265**, 11823-11826.
75. Bies, J., Nazarov, V., Wolff, L. (1999) Identification of Protein Instability Determinants in the Carboxy-Terminal Region of c-Myb Removed as a Result of Retroviral Integration in Murine Monocytic Leukemias. . *Journal of Virology* **73**, 2038–2044.
76. Roth, A.F., Davis, N.G. (2000) Ubiquitination of the PEST-like endocytosis signal of the yeast a-factor receptor. *J. Biol. Chem.*, **275**, 8143-8153.
77. Penrose, K.J., Mcbride, A.A. (2000) Proteasome-Mediated Degradation of the Papillomavirus E2-TA Protein Is Regulated by Phosphorylation and Can Modulate Viral Genome Copy Number. *Journal of Virology*, **74**, 6031–6038.
78. Schuster, M.K., Graber, M. (2002) *EMBOSS epestfind*. https://emboss.bioinformatics.nl/cgi-bin/emboss/help/epestfind (accessed June 28, 2023).
79. Kalani, L., Kim, B.H., Vincent, J.B., Ausio, J. (2023) MeCP2 ubiquitination and sumoylation, in search of a function (*). *Hum. Mol. Genet.*
80. Tao, J., Hu, K., Chang, Q., Wu, H., Sherman, N.E., Martinowich, K., Klose, R.J., Schanen, C., Jaenisch, R., Wang, W., et al. (2009) Phosphorylation of MeCP2 at Serine 80 regulates its chromatin association and neurological function. *Proc. Natl. Acad. Sci.*, **106**, 4882-4887.
81. Zhou, Z., Hong, E.J., Cohen, S., Zhao, W.N., Ho, H.Y., Schmidt, L., Chen, W.G., Lin, Y., Savner, E., Griffith, E.C., et al. (2006) Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription, dendritic growth, and spine maturation. *Neuron*, **52**, 255-269.
82. Gonzales, M.L., Adams, S., Dunaway, K.W., LaSalle, J.M. (2012) Phosphorylation of distinct sites in MeCP2 modifies cofactor associations and the dynamics of transcriptional regulation. *Mol. Cell. Biol.*, **32**, 2894-2903.
83. Weaving, L.S., Christodoulou, J., Williamson, S.L., Friend, K.L., McKenzie, O.L., Archer, H., Evans, J., Clarke, A., Pelka, G.J., Tam, P.P., et al. (2004) Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. *Am. J. Hum. Genet.*, **75**, 1079-1093.
84. Ariani, F., Hayek, G., Rondinella, D., Artuso, R., Mencarelli, M.A., Spanhol-Rosseto, A., Pollazzon, M., Buoni, S., Spiga, O., Ricciardi, S., et al. (2008) FOXG1 is responsible for the congenital variant of Rett syndrome. *Am. J. Hum. Genet.*, **83**, 89-93.
85. Alberti, S., Gladfelter, A., Mittag, T. (2019) Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell*, **176**, 419-434.

86. Ghosh, R.P., Nikitina, T., Horowitz-Scherer, R.A., Gierasch, L.M., Uversky, V.N., Hite, K., Hansen, J.C., Woodcock, C.L. (2010) Unique physical properties and interactions of the domains of methylated DNA binding protein 2. *Biochem.*, **49**, 4395-4410.
87. Fan, C., Zhang, H., Fu, L., Li, Y., Du, Y., Qiu, Z., Lu, F. (2020) Rett mutations attenuate phase separation of MeCP2. *Cell Discov*, **6**, 38.
88. Sheikh, T.I., Ausio, J., Faghfoury, H., Silver, J., Lane, J.B., Eubanks, J.H., MacLeod, P., Percy, A.K., Vincent, J.B. (2016) From Function to Phenotype: Impaired DNA Binding and Clustering Correlates with Clinical Severity in Males with Missense Mutations in MECP2. *Sci. Rep.*, **6**, 38590.
89. Sheikh, T.I., de Paz, A.M., Akhtar, S., Ausio, J., Vincent, J.B. (2017) MeCP2_E1 N-terminal modifications affect its degradation rate and are disrupted by the Ala2Val Rett mutation. *Hum. Mol. Genet.*, **26**, 4132-4141.
90. Ehrhart, F., Jacobsen, A., Rigau, M., Bosio, M., Kaliyaperumal, R., Laros, J.F.J., Willighagen, E.L., Valencia, A., Roos, M., Capella-Gutierrez, S., et al. (2021) A catalogue of 863 Rett-syndrome-causing MECP2 mutations and lessons learned from data integration. *Sci. Data.*, **8**, 10.
91. Good, K.V., Vincent, J.B., Ausio, J. (2021) MeCP2: The Genetic Driver of Rett Syndrome Epigenetics. *Front. Genet.*, **12**, 620859.
92. Ballestar, E., Yusufzai, T., Wolffe, A. (2000) Effects of Rett Syndrome Mutations of the Methyl-CpG Binding Domain of the Transcriptional Repressor MeCP2 on Selectivity for Association with Methylated DNA. *Biochem.*, **39**, 7100-7106.
93. Christodoulou, J., Grimm, A., Maher, T., Bennetts, B. (2003) RettBASE: The IRSA MECP2 variation database-a new mutation database in evolution. *Hum. Mutat.*, **21**, 466-472.
94. Vermudez, S.A.D., Gogliotti, R.G., Arthur, B., Buch, A., Morales, C., Moxley, Y., Rajpal, H., Conn, P.J., Niswender, C.M. (2022) Profiling beneficial and potential adverse effects of MeCP2 overexpression in a hypomorphic Rett syndrome mouse model. *Gene. Brain. Behav.*, **21**, e12752.
95. Zhou, J., Hamdan, H., Yalamanchili, H.K., Pang, K., Pohodich, A.E., Lopez, J., Shao, Y., Oses-Prieto, J.A., Li, L., Kim, W., et al. (2022) Disruption of MeCP2-TCF20 complex underlies distinct neurodevelopmental disorders. *Proc. Natl. Acad. Sci.*, **119**, e2119078119.
96. Wang, L., Hu, M., Zuo, M.Q., Zhao, J., Wu, D., Huang, L., Wen, Y., Li, Y., Chen, P., Bao, X., et al. (2020) Rett syndrome-causing mutations compromise MeCP2-mediated liquid-liquid phase separation of chromatin. *Cell. Res.*, **30**, 393-407.
97. Heckman, L.D., Chahrour, M.H., Zoghbi, H.Y. (2014) Rett-causing mutations reveal two domains critical for MeCP2 function and for toxicity in MECP2 duplication syndrome mice. *Elife*, **3**, e02676.
98. Yang, Y., Kucukkal, T.G., Li, J., Alexov, E., Cao, W. (2016) Binding Analysis of Methyl-CpG Binding Domain of MeCP2 and Rett Syndrome Mutations. *ACS. Chem. Biol.*, **11**, 2706-2715.
99. Johnson, B.S., Zhao, Y.T., Fasolino, M., Lamonica, J.M., Kim, Y.J., Georgakilas, G., Wood, K.H., Bu, D., Cui, Y., Goffin, D., et al. (2017) Biotin tagging of MeCP2 in mice reveals contextual insights into the Rett syndrome transcriptome. *Nat. Med.*, **23**, 1203-1214.

100. Ortega-Alarcon, D., Claveria-Gimeno, R., Vega, S., Jorge-Torres, O.C., Esteller, M., Abian, O., Velazquez-Campoy, A. (2020) Molecular Context-Dependent Effects Induced by Rett Syndrome-Associated Mutations in MeCP2. *Biomol.*, **10**, 1533.
101. Stefanelli, G., Gandaglia, A., Costa, M., Cheema, M.S., Di Marino, D., Barbiero, I., Kilstrup-Nielsen, C., Ausio, J., Landsberger, N. (2016) Brain phosphorylation of MeCP2 at serine 164 is developmentally regulated and globally alters its chromatin association. *Sci. Rep.*, **6**, 28295.
102. Merritt, J.K., Collins, B.E., Erickson, K.R., Dong, H., Neul, J.L. (2020) Pharmacological read-through of R294X Mecp2 in a novel mouse model of Rett syndrome. *Hum. Mol. Genet.*, **29**, 2461-2470.
103. Neul, J.L., Fang, P., Barrish, J., Lane, J., Caeg, E.B., Smith, E.O., Zoghbi, H., Percy, A., Glaze, D.G. (2008) Specific mutations in methyl-CpG-binding protein 2 confer different severity in Rett syndrome. *Neurology*, **70**, 1313-1321.
104. Caffarelli, C., Gonnelli, S., Pitinca, M.D.T., Camarri, S., Al Refaie, A., Hayek, J., Nuti, R. (2020) Methyl-CpG-binding protein 2 (MECP2) mutation type is associated with bone disease severity in Rett syndrome. *BMC. Med. Genet.*, **21**, 21.
105. Kruusvee, V., Lyst, M.J., Taylor, C., Tarnauskaite, Z., Bird, A.P., Cook, A.G. (2017) Structure of the MeCP2-TBLR1 complex reveals a molecular basis for Rett syndrome and related disorders. *Proc. Natl. Acad. Sci.*, **114**, E3243-E3250.
106. Boxer, L.D., Renthal, W., Greben, A.W., Whitwam, T., Silberfeld, A., Stroud, H., Li, E., Yang, M.G., Kinde, B., Griffith, E.C., et al. (2020) MeCP2 Represses the Rate of Transcriptional Initiation of Highly Methylated Long Genes. *Mol. Cell.*, **77**, 294-309 e299.
107. Sandweiss, A.J., Brandt, V.L., Zoghbi, H.Y. (2020) Advances in understanding of Rett syndrome and MECP2 duplication syndrome: prospects for future therapies. *Lancet Neurol*, **19**, 689-698.
108. Ramocki, M.B., Peters, S.U., Tavyev, Y.J., Zhang, F., Carvalho, C.M., Schaaf, C.P., Richman, R., Fang, P., Glaze, D.G., Lupski, J.R., et al. (2009) Autism and other neuropsychiatric symptoms are prevalent in individuals with MeCP2 duplication syndrome. *Ann. Neurol.*, **66**, 771-782.
109. Collins, A.L., Levenson, J.M., Vilaythong, A.P., Richman, R., Armstrong, D.L., Noebels, J.L., David Sweatt, J., Zoghbi, H.Y. (2004) Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. *Hum. Mol. Genet.*, **13**, 2679-2689.
110. Yang, T., Ramocki, M.B., Neul, J.L., Lu, W., Roberts, L., Knight, J., Ward, C.S., Zoghbi, H.Y., Kheradmand, F., Corry, D.B. (2012) Overexpression of methyl-CpG binding protein 2 impairs T(H)1 responses. *Sci Transl Med*, **4**, 163ra158.
111. Sheen, V., Valencia, I.M., Torres, A.R. (2013) Atypical features in MECP2 P152R-associated Rett syndrome. *Pediatr Neurol*, **49**, 124-126.
112. Abuhatzira, L., Shamir, A., Schones, D.E., Schaffer, A.A., Bustin, M. (2011) The chromatin-binding protein HMGN1 regulates the expression of methyl CpG-binding protein 2 (MECP2) and affects the behavior of mice. *J. Biol. Chem.*, **286**, 42051-42062.
113. Klein, M.E., Lioy, D.T., Ma, L., Impey, S., Mandel, G., Goodman, R.H. (2007) Homeostatic regulation of MeCP2 expression by a CREB-induced microRNA. *Nat. Neurosci.*, **10**, 1513-1514.
114. Meur, N.L., Holder-Espinasse, M., Jaillard, S., Goldenberg, A., Joriot, S., Amati-Bonneau, P., Guichet, A., Barth, M., Charollais, A., Journel, H., Auvin, S., Boucher, C.,

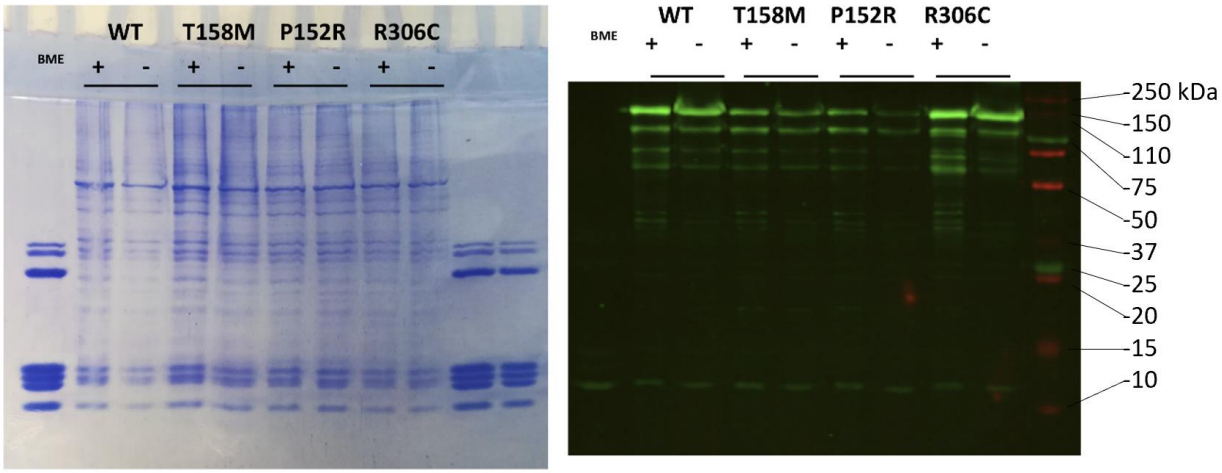
Kerckaert, P., David, V., Manouvrier-Hanu, S., Saugier-Veber, P., Frébourg, T., Dubourg, C., Andrieux, J., Bonneau, D. (2010) MEF2C haploinsufficiency caused by either microdeletion of the 5q14.3 region or mutation is responsible for severe mental retardation with stereotypic movements, epilepsy and/or cerebral malformations. *Journal of Medical Genetics*.

115. Trieschmann, L., Postnikov, Y.V., Rickers, A., Bustin, M. (1995) Modular structure of chromosomal proteins HMG-14 and HMG-17: definition of a transcriptional enhancement domain distinct from the nucleosomal binding domain. *Mol. Cell. Biol.*, **15**, 6663-6669.
116. Stros, M. (2010) HMGB proteins: interactions with DNA and chromatin. *Biochim. Biophys. Acta.*, **1799**, 101-113.
117. Nagarajan, R.P., Hogart, A.R., Gwye, Y., Martin, M.R., LaSalle, J.M. (2006) Reduced MeCP2 expression is frequent in autism frontal cortex and correlates with aberrant MECP2 promoter methylation. *Epigenetics*, **1**, e1-11.
118. Coussons-Read, M.E., Crnic, L.S. (1996) Behavioral assessment of the Ts65Dn mouse, a model for down syndrome: Altered behavior in the elevated plus maze and open field. *Behavior Genetics*, **26**.
119. Silverman, J.L., Yang, M., Lord, C., Crawley, J.N. (2010) Behavioural phenotyping assays for mouse models of autism. *Nat. Rev. Neurosci.*, **11**, 490-502.
120. Magill, S.T., Cambronne, X.A., Luikart, B.W., Lioy, D.T., Leighton, B.H., Westbrook, G.L., Mandel, G., Goodman, R.H. (2010) microRNA-132 regulates dendritic growth and arborization of newborn neurons in the adult hippocampus. *Proc. Natl. Acad. Sci.*, **107**, 20382-20387.
121. Wanet, A., Tacheny, A., Arnould, T., Renard, P. (2012) miR-212/132 expression and functions: within and beyond the neuronal compartment. *Nucleic Acid. Res.*, **40**, 4742-4753.
122. Chen, Y., Shin, B.C., Thamotharan, S., Devaskar, S.U. (2014) Differential methylation of the micro-RNA 7b gene targets postnatal maturation of murine neuronal Mecp2 gene expression. *Dev Neurobiol*, **74**, 407-425.
123. Su, M., Hong, J., Zhao, Y., Liu, S., Xue, X. (2015) MeCP2 controls hippocampal brain-derived neurotrophic factor expression via homeostatic interactions with microRNA-132 in rats with depression. *Mol Med Rep*, **12**, 5399-5406.
124. Rodrigues, D.C., Mufteev, M., Ellis, J. (2020) Regulation, diversity and function of MECP2 exon and 3'UTR isoforms. *Hum. Mol. Genet.*, **29**, R89-R99.
125. Chahrour M, J.S., Shaw C, Zhou X, Wong ST, Qin J, Zoghbi HY. MeCP2, a key contributor to neurological disease, activates and represses transcription. (2008) MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science*.
126. Li, H., Radford, J.C., Ragusa, M.J., Shea, K.L., Mckercher, S.R., Zaremba, J.D., Soussou, W., Nie, Z., Kang, Y.-J., Nakanishi, N., Okamoto, S.-I., Roberts, A.J., Schwarz, J.J., Lipton, S.A. (2008) Transcription factor MEF2C influences neural stem/progenitor cell differentiation and maturation in vivo. . *Proc. Natl. Acad. Sci.*
127. Zweier, M., Gregor, A., Zweier, C., Engels, H., Sticht, H., Wohlleber, E., Bijlsma, E.K., Holder, S.E., Zenker, M., Rossier, E., et al. (2010) Mutations in MEF2C from the 5q14.3q15 microdeletion syndrome region are a frequent cause of severe mental retardation and diminish MECP2 and CDKL5 expression. *Hum. Mutat.*, **31**, 722-733.

128. Giglione, C., Boularot, A., Meinnel, T. (2004) Protein N-terminal methionine excision. *Cell Mol Life Sci*, **61**, 1455-1474.
129. Giglione, C., Fieulaine, S., Meinnel, T. (2015) N-terminal protein modifications: Bringing back into play the ribosome. *Biochimie*, **114**, 134-146.
130. Ree, R., Varland, S., Arnesen, T. (2018) Spotlight on protein N-terminal acetylation. *Exp Mol Med*, **50**, 1-13.
131. Nguyen, K.T., Mun, S.H., Lee, C.S., Hwang, C.S. (2018) Control of protein degradation by N-terminal acetylation and the N-end rule pathway. *Exp Mol Med*, **50**, 1-8.
132. Hwang, C.S., Shemorry, A., Varshavsky, A. (2010) N-terminal acetylation of cellular proteins creates specific degradation signals. *Science*, **327**, 973-977.
133. Gonda, D.K., Bachmair, A., Wünning, I., Tobias, J.W., Lane, W.S., Varshavsky, A. (1989) Universality and Structure of the N-end Rule. *J. Biol. Chem.*, **264**, 16700-16712.
134. Varshavsky, A. (2011) The N-end rule pathway and regulation by proteolysis. *Protein Sci*, **20**, 1298-1345.
135. Hwang CS, S.A., Varshavsky A. (2010) N-terminal acetylation of cellular proteins creates specific degradation signals. *Science*, **19**, 973-977.
136. Suskiewicz, M.J., Sussman, J.L., Silman, I., Shaul, Y. (2011) Context-dependent resistance to proteolysis of intrinsically disordered proteins. *Protein Sci*, **20**, 1285-1297.
137. Blau, H.M., Webster, C., Chiu, C., Guttman, S., & Chandler, F. (1983) Differentiation Properties of Pure Populations of Human Dystrophic Muscle Cells *Experimental Cell Research*, **144**, 495-503.
138. Yaffe, D., Saxel, O. . (1977) Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature*, **270**, 725-727.
139. Diokmetzidou, A., Tsikitis, M., Nikouli, S., Kloukina, I., Tsoupri, E., Papathanasiou, S., Psarras, S., Mavroidis, M., Capetanaki, Y. (2016) Strategies to Study Desmin in Cardiac Muscle and Culture Systems. *Methods Enzymol*, **568**, 427-459.
140. Manuelidis, L. (1985) Indications of centromere movement during interphase and differentiation. *Ann N Y Acad Sci*, **450**, 205-221.
141. Martou, G., De Boni, U. (2000) Nuclear topology of murine, cerebellar Purkinje neurons: changes as a function of development. *Exp Cell Res*, **256**, 131-139.
142. Brero, A., Easwaran, H.P., Nowak, D., Grunewald, I., Cremer, T., Leonhardt, H., Cardoso, M.C. (2005) Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation. *J Cell Biol*, **169**, 733-743.
143. Shahbazian, M.D., Antalffy, B., Armstrong, D. L., & Zoghbi, H. Y. (2002) Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. *Human molecular genetics*, **11**, 115-124.
144. Shahbazian, M.D., Antalffy, B., Armstrong, D.L., Zoghbi, H.Y. (2002) Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. *Hum. Mol. Genet.*, **11**, 115-124.
145. Guy, J., Alexander-Howden, B., FitzPatrick, L., DeSousa, D., Koerner, M.V., Selfridge, J., Bird, A. (2018) A mutation-led search for novel functional domains in MeCP2. *Hum. Mol. Genet.*, **27**, 2531-2545.
146. Akoglu, H. (2018) User's guide to correlation coefficients. *Turk J Emerg Med*, **18**, 91-93.
147. Ausio, J., Segar, D., Eisenberg, H. (1984) Nucleosome Core Particle Stability and Conformational Change. *J. Mol. Biol.*, **176**, 77-104.

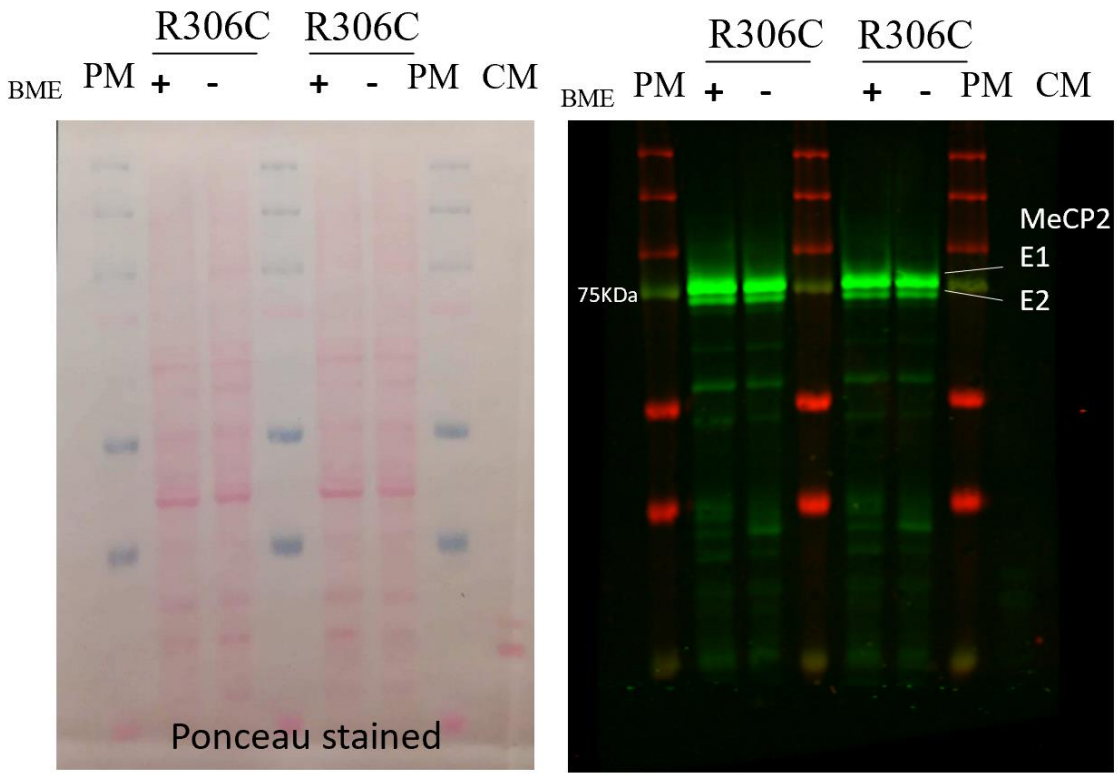
148. Franklin, D. (2019) P152R Mutation Within MeCP2 Can Cause Loss of DNA-Binding Selectivity. *Interdisciplinary Sciences: Computational Life Sciences*, **11**, 10-20.
149. Dunker AK, L.J., Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z. . (2001) Intrinsically disordered protein. *J. Mol. Graph. Model.*, **19**, 26-59.
150. Medicherla, B., Goldberg, A.L. (2008) Heat shock and oxygen radicals stimulate ubiquitin-dependent degradation mainly of newly synthesized proteins. *J Cell Biol*, **182**, 663-673.
151. Skene, P.J., Illingworth, R.S., Webb, S., Kerr, A.R., James, K.D., Turner, D.J., Andrews, R., Bird, A.P. (2010) Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Mol. Cell.*, **37**, 457-468.
152. Lyst, M.J., Bird, A. (2015) Rett syndrome: a complex disorder with simple roots. *Nat Rev Genet*, **16**, 261-275.
153. Mazza, D., Ganguly, S., McNally, J.G. (2013) Monitoring dynamic binding of chromatin proteins in vivo by single-molecule tracking. *Methods Mol Biol*, **1042**, 117-137.
154. Mueller, F., Mazza, D., Stasevich, T.J., McNally, J.G. (2010) FRAP and kinetic modeling in the analysis of nuclear protein dynamics: what do we really know? *Curr Opin Cell Biol*, **22**, 403-411.
155. Tai, D.J., Liu, Y.C., Hsu, W.L., Ma, Y.L., Cheng, S.J., Liu, S.Y., Lee, E.H. (2016) MeCP2 SUMOylation rescues Mecp2-mutant-induced behavioural deficits in a mouse model of Rett syndrome. *Nat. Commun.*, **7**, 10552.
156. Wang, Y. (2014) RING finger protein 4 (RNF4) derepresses gene expression from DNA methylation. *J. Biol. Chem.*, **289**, 33808-33813.
157. van Wijk, S.J., Fulda, S., Dikic, I., Heilemann, M. (2019) Visualizing ubiquitination in mammalian cells. *EMBO. Rep.*, **20**, e46520.
158. Sanchez-Bailon, M.P., Choi, S.Y., Dufficy, E.R., Sharma, K., McNee, G.S., Gunnell, E., Chiang, K., Sahay, D., Maslen, S., Stewart, G.S., et al. (2021) Arginine methylation and ubiquitylation crosstalk controls DNA end-resection and homologous recombination repair. *Nat. Commun.*, **12**, 6313.

Appendix 1. Comparison of MeCP2 in reducing and non-reducing conditions



Appendix 1. Nuclei isolated from transfected C2C12 run on SDS-gel and western blot probing for MeCP2. The nuclei of transfected cells with WT, T158M, P152R, and R306C MeCP2 with a Ct-GFP tag were run with and without beta-mercaptoethanol (BME) to check if MeCP2 dimers are present when BME is absent, especially for R306C where the extra cysteine is introduced. A band at ~200kDa would have been present if the dimerization occurred.

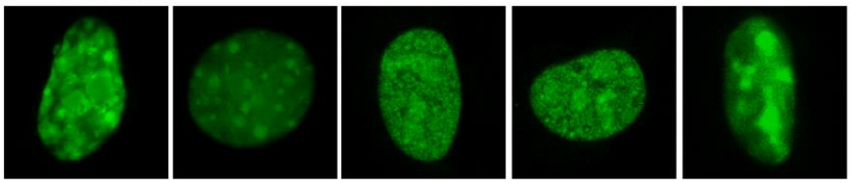
Appendix 2. MeCP2 R306C on 12% gel



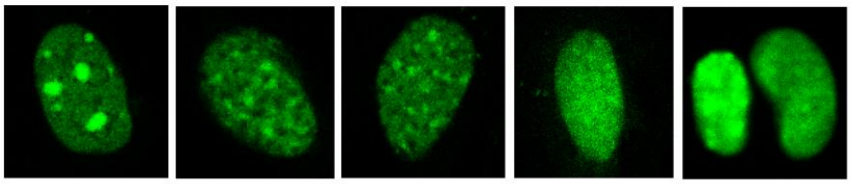
Appendix 2. Western blot of transgenic mouse brain nuclei with R306C MeCP2 with and without BME on 12% SDS. Right, ponceau stain of the nitrocellulose membrane of two technical replicates of R306C MeCP2. Left, western blot of R306C MeCP2 showing the two isoforms of MeCP2. As expected, the E1 is present at higher concentrations than the E2.

Appendix 3. Heterogeneity in the chromocenter appearance in confocal fluorescence microscopy of MeCP2-GFP

NullP+T158M



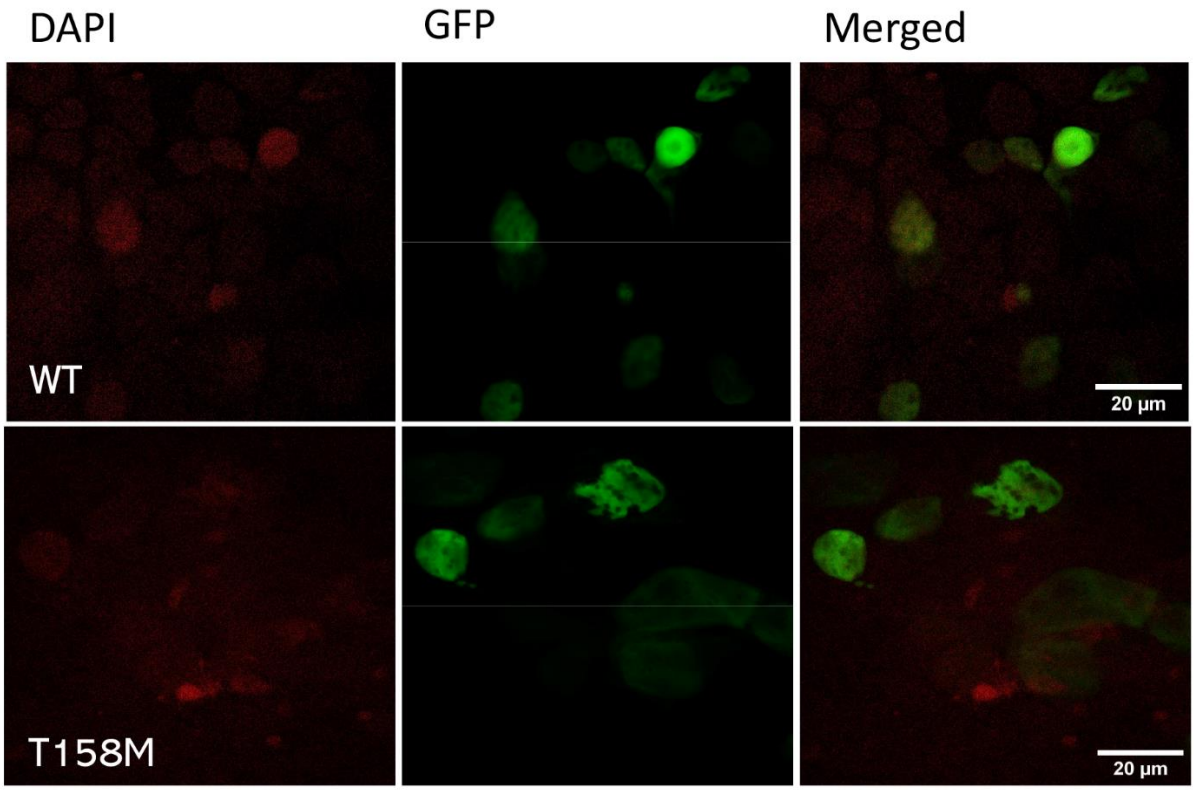
modP+T158M



% of cells with defined chromocenters	
NullP+T158M	24% (4/17)
modP+T158M	38% (6/16)

Appendix 3. Confocal fluorescence microscopy images of C2C12 nuclei showing the GFP channel and the ratio of the cells with defined chromocenters. The transient transfection of C2C12 with MeCP2 constructs nullPEST and modPEST with missense mutation T158M, causing diffuse chromocenters and loss of defined boundaries that separate the pericentromeric DNA from chromatin within the nuclei.

Appendix 4. MeCP2-GFP in transiently transfected HEK293 cells



Appendix 4. Confocal fluorescence microscopy images of HEK293 cells transiently transfected with MeCP2 WT and T158M with a GFP tag. HEK cells do not display the defined pericentromeric DNA as observed with C2C12 cells. The DAPI staining of HEK cells is diffuse and has high background binding.