Winter 12-15-2022

Emerging insights into a role for Topoisomerase IIB and Methyl-CpG binding protein 2 in regulating neuronal transcription

Sabin Nettles
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Neurosciences

Dissertation Examination Committee:
Harrison W. Gabel, Chair
Joseph R. Dougherty, co-Chair
Tychelle T Turner
Jason Yi
Andrew Yoo

Emerging insights into a role for Topoisomerase IIβ and Methyl-CpG binding protein 2 in regulating neuronal transcription

by

Sabin A. Nettles

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2022
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

This dissertation is the culmination of my doctoral research, made possible by invaluable support and guidance I have received over the years.

I would like to thank my PhD advisor Dr. Harrison Gabel. This work would not be possible without your tremendous support and mentorship. You are an endless source of knowledge and your enthusiasm and passion for science is magnetic. Thank you for fostering a lab environment that is committed to conducting rigorous scientific research. Thank you for elevating my critical thinking and for always being willing to take a deep dive into ideas for new experiments. Your guidance has helped me become a better scientist.

I would like to thank the members of my thesis committee: Dr. Joseph Dougherty, Dr. Andrew Yoo, Dr. Jason Yi and Dr. Tychele Turner. This is a group of exceptional researchers with broad expertise. You have provided considerable insight and support throughout every stage of my thesis research. Thank you for being so generous with your time.

I would like to thank the members of the Gabel lab. It has been such an honor to work alongside such a smart, hard-working and driven group. Thank you all for making our lab such a collaborative environment for sharing ideas, troubleshooting experiments, and discussing new data. Dennis Wu, thank you for your round-the-clock assistance for all of my programming question, constant encouragement and for always being willing to take a break for tacos. Adam Clemens, thank you for the quick-witted banter and for being the best corn hole partner during the annual lab BBQ games. Diana (formerly Christian) Beard, thank you for sharing your passion for teaching and for always providing the best constructive compliments. James Russell Moore, thank you for being a fellow lab night-owl and for spoiling me with your puppy Ziggy. Mati Nemera, thank you for always providing the comic-relief. Nicole Hamagami, thank you for your attention to detail and your

willingness to help, especially when it comes to revisions on a draft or data analysis. Ryan Liu, as the first member of the lab, thank you setting up the infrastructure of the lab, helping me with experiments and for always providing a calming presence. Jenna Martin, thank you for our long, thoughtful conversations on current events. Yasmin Razia, thank you for bringing a steady, nurturing presence to the lab. Your compassion and thoughtfulness are greatly appreciated. Samanda Velente-Soto, thank you for being the first member of "Team Topo" and my first mentee. I learned so much about mentorship through that experience. Chibueze Agwu, thank you for your daily optimism and tireless efforts on my projects. Your contributions helped me make significant progress. Alyssa Erickson and Ladan Abbasi, the newest members of the Gabel lab, thank you for asking great questions and adding new perspectives and expertise to the group.

I also would like to thank the members of the ZYGL (Zhao lab, Yi lab, Gabel lab and Li lab) group meetings for great discussions. A special thank you to Dr. Guoyan Zhao and Dr. Tristan Li, who have asked excellent questions and provided thoughtful feedback, which has been valuable for helping me improve my data presentation and communication skills.

I am fortunate to have received my scientific training through the DBBS Neuroscience program. The 2015 cohort is an extraordinary group of smart, highly motivated scientists. Thank you for sharing your knowledge. Many of you went above and beyond, always looking out for me (and my dietary restrictions) and for that I am eternally grateful! I will forever cherish your kindness. To the Neuroscience program coordinator, Sally Vogt, you are the backbone of the program. Thank you for all your time, effort, and support. Your impact is immeasurable. To the Neuroscience program directors, past and present – Dr. Erik Herzog, Dr. Larry Snyder, Dr. Tim Holy, Dr. Tammy Hershey, Dr. Daniel Kerschensteiner. Thank you for your leadership and support.

During my time at WashU, I have been able to learn from and engage with many incredible scientists in the broader neuroscience community. I am incredibly grateful to have been a trainee in

the Department of Neuroscience. It has been inspiring to be surrounded by outstanding and dedicated faculty who are consistently at the forefront and whose contributions are advancing the field. To the chairs of the Department of Neuroscience, past, interim, and present – Dr. Azad Bonni, Dr. Paul Taghert and Dr. Linda Richards, thank you for cultivating such a collaborative environment for scientist at all stages. I also appreciate the efforts from all those who work behind the scenes to keep the Department running smoothly. To Debbie Pfieffer, former executive assistant to the department chair, thank you for being the heart and soul of the department and for your unending encouragement. To Ann Olendorff, department coordinator, thank you for always making sure I had the resources I needed. To the custodial crew, especially Ms. Inez and Mr. Darrell, thank you for your kindness and for always looking out for me.

Outside of my dissertation research, my experience in graduate school has been enriched by participating in neuroscience outreach initiatives. I have had the privilege to serve as co-Director of Brain Discovery and Coordinator of St. Louis Neuroscience Outreach (STLNO). These opportunities have been incredible experiences and allowed me to share my passion for science with the next generation of students. These programs would not have been possible without the support from committed students, faculty, and administrators. To the Brain Discovery leadership team, past and present, thank you for your hard work and volunteering your time. A special thank you to my counterpart as co-Director, Katie Lefton, for stepping into these roles and continuing to work towards achieving our goals. To the STLNO Interest Group Coordinators – Dr. Erik Herzog, Dr. Linda Richards, Dr. Anneliese Schaefer and Kerry Grens, thank you for dedicating time to implement outreach initiatives. A very special thanks to Dr. Erik Herzog and Dr. Linda Richard, who have promoted and encouraged me during my tenure in leadership. It has been an honor and privilege to work with and learn from you.

I would also like to thank to the Neuroscience retreat planning committee, especially Dr.

Anneliese Schaefer, Kate Ruzicka, Carmen Horn, Laura William, Sally Vogt, and fellow neuroscience graduate student, Jenna Krizan. You are pillars in the community for organizing arguably the best event year after year. Thank you for welcoming me into the group and providing me the opportunity to run NeurOlympics. People are still talking about it!

I am incredibly thankful for my previous mentors – Dr. Marc Tetel, Dr. Annette Kim, Dr. Louis Muglia. Through your mentorship, you cultivated my interest in neuroscience and molecular biology, included me in your research groups and acknowledged my contributions. I am especially grateful to Dr. Kalpana Acharya and Dr. Liz Bless, thank you for bolstering my confidence and celebrating my successes. Collectively, you provided me with the tools, training and knowledge that prepared me for my research pursuits.

I am incredibly grateful to my friends and family! You have gone above and beyond to support me throughout my time in graduate school. Thank you for being a constant source of encouragement. You all are my motivation! To my sisters and brother-in-law, thank you for being my biggest “hypemen”. To my parents, you are my role models. Thank you for the unconditional love and support and for giving me the confidence to achieve my goals.

Sabin A. Nettles

Washington University in St. Louis

August 2022

Dedicated to the next generations of black girls who aspire to be a scientist.

ABSTRACT OF THE DISSERTATION

Elucidating a role for Topoisomerase IIβ and Methyl-CpG binding protein 2 in regulating the unique neuronal transcriptome.

by

Sabin A. Nettles

Doctor of Philosophy in Biology and Biomedical Sciences

Neurosciences

Washington University in St. Louis, 2022

Professor Harrison W. Gabel, Chair

The mammalian brain has unparalleled diversity of cell types with distinct molecular, morphological, connectional and functional properties. The specialization of these cells occurs during development through a series of spatially and temporally controlled changes in gene expression that are critical for proper assembly and function. Neurons in particular have complex and elaborate gene regulatory systems which allow specific combinations of genes to be expressed at distinct levels and at discrete developmental stages, giving rise to heterogenous neuronal subtypes. Identifying unique neuronal gene regulatory factors that contribute to the dynamic changes in gene expression are critical to understanding the molecular mechanism that underlie proper brain development and function. Recent progress has been made to characterize the gene regulatory programs that govern the establishment and organization of the neuronal transcriptome across development. A signature feature of the neuronal transcriptome that has provided considerable insight into the distinct gene regulatory events required for proper neuronal development is the expression of the very long genes (e.g. > 100 kilobases). These genes encode proteins enriched for

functions in neuronal physiology, including signaling molecules, receptors, ion channels and cell adhesion molecules, and many long genes have been identified as disrupted in neurodevelopmental disorders. Notably, two transcriptional regulators, the Methyl-CpG binding protein 2 (MeCP2) and Topoisomerase IIβ (TOP2β), were recently identified as being able to tune the expression of long genes in the neurons. MeCP2, the protein mutated in Rett syndrome, can function as a transcriptional repressor that binds methylated DNA within long genes and prevents the overexpression of long genes. In contrast, TOP2β functions to resolve topological constraints, such as supercoiling, by creating transient double-strand breaks in the DNA, and has been shown to facilitate the expression of very long genes. Therefore, defining the mechanism underlying how these seemingly antagonistic transcriptional regulators may co-operate to regulate the expression of long genes is paramount. Moreover, elucidating this mechanism may provide key insights towards understanding the unique gene-regulatory landscape in neurons during development and discerning how dysregulation may lead to neurodevelopmental disorders. Here, I investigate the role of MeCP2 and TOP2β during the transcription of long gene in neurons. Using biochemical approaches, I demonstrate a functional interaction between MeCP2 and TOP2β and map the sequences that mediate the TOP2β-MeCP2 interaction. The role of TOP2β during transcription and the precise mechanism by which TOP2β facilitates gene expression is largely unknown. To address the gap in knowledge, I adapted and implemented a unique genomic approach called etoposide-mediated topoisomerase immunoprecipitation sequencing (eTIP-seq) to profile TOP2β activity genome-wide. Through this analysis, I identify a length-dependent enrichment of TOP2β activity within long genes, including at key regulatory sites, downstream of promoters and at intragenic enhancers. To investigate the role of MeCP2 in regulating TOP2β activity, I conducted experiments in which I manipulated the levels of MeCP2 and assessed changes in TOP2β activity. I demonstrate that MeCP2 negatively regulates

TOP2β activity at long genes that are repressed by MeCP2. Finally, I highlight outstanding questions regarding the functional significance of the interaction between MeCP2 and TOP2β that may be relevant to Rett syndrome pathology. Future studies can address these open questions and build upon our current understanding of this neuronal transcriptional regulatory mechanism.

Chapter 1:

Emerging insights into the expression of long genes as a unique signature of the neuronal transcriptome

1.1 Expression of long genes as a hallmark of neural cell identity

The mammalian brain develops through a complex series of crucial events over a protracted period, including cell proliferation, differentiation, and migration, with temporal and spatial factors influencing the fate of each newborn neuron (Lim et al., 2018; Nowakowski et al., 2017). For example, in the mouse cerebral cortex, cortical progenitors rapidly divide between embryonic day 11.5 (E11.5) and birth (postnatal day 0, P0), giving rise to six cortical layers (Greig et al., 2013). Neuronal stem cells in the ventricular zone (VZ), intermediate progenitors of the subventricular zone (SVZ) and radial glia (RG) in the cerebral cortex undergo a series of divisions to produce additional intermediate progenitors or pyramidal neurons (Noctor et al., 2004). Terminally differentiated neurons undergo a long radial migration in an inside-out manner until they adopt their final position in a layer of the cortex. Abnormal cortical development causes prominent neurodevelopmental and neuropsychiatric disorders. While the basic structural and functional features of neocortical organization have been well studied, there remain several unknowns regarding the gene expression patterns and molecular regulatory programs that build neuronal diversity during corticogenesis (Lodato and Arlotta, 2015).

Cellular identity is defined by the expression of specific combinations of effector genes that give rise to a neurons biochemical, morphological and physiological properties (Closser et al., 2022). Technological advances in transcriptomics have provided the ability to profile gene expression patterns at distinct developmental stages and has allowed for the identification of key features underlying cortical development (Loo et al., 2019). Recent studies have discovered that neurons carry unique epigenetic and transcriptional features compared to other somatic cells,

that facilitate their complex transcriptomes. This includes the build up of high levels of methylated CA dinucleotides (Guo et al., 2014; Lister et al., 2013) and expression of long genes (e.g. >100 kb from the transcription start site to the transcription end site) (Gabel et al., 2015; Sugino et al., 2014; Zylka et al., 2015). Many long genes encode ion channels, receptors and cell adhesion molecules (Gabel et al., 2015), which may contribute to the extensive molecular diversity that exists among neuronal subtypes (Moore, Nemera, Gabel, unpublished results).

To gain insight into the gene length expression bias observed in the brain, several studies have analyzed gene expression profiles from various tissues, cell types and developmental stages at the single-cell and population level. Mouse and human brains express a greater proportion of long genes relative to non-neuronal tissues (Gabel et al., 2015). Furthermore, transcripts from longer genes are overrepresented in neural tissues and neurons (Gabel et al., 2015; Zylka et al., 2015). When evaluating the level of long gene expression from a single-cell RNA-seq study of neural cell subtypes from the mouse visual cortex (Tasic et al., 2016), the results showed preferential enrichment of long genes in neurons (i.e. Th, Pvalb, L4, L2.3, L5, Chodl, Sst, Vip, Ndnf, and L6) relative to glia (i.e. endothelial cells, microglia, astrocytes, oligodendrocytes and OPCs) (McCoy et al., 2018).

In a study analyzing the transcriptomes of the nine major cell types of the cerebral cortex and CA1 region of the hippocampus, the data showed each neuronal subtype was enriched for longer transcripts relative to non-neuronal cell types (Zylka et al., 2015). Furthermore, this analysis uncovered that the gene length expression bias could distinguish specific brain regions, as the transcripts enriched in hippocampal pyramidal neurons were longer than those expressed in the somatosensory cortex. When assessing post-mitotic neurons, the levels of long gene expression were high in the cortical and subcortical regions (i.e. intermediate zone, subcortical

plate, cortical plate, subpial granular zone and marginal zone) which are enriched for post-mitotic neurons (Miller et al., 2014) compared to the ganglionic eminences and ventricular zones (i.e. MGE, LGE, CGE, VZ, and SVZ) in which proliferating neural progenitor cells are prominent (McCoy et al., 2018). To pinpoint the timepoint during neuronal differentiation in which the level of long gene expression changes, one study used Div-seq to profile newborn neurons (immature neurons) in the neurogenic niche of the adult hippocampus, neural stem cells (NSCs), neural progenitor cells (NPCs) and neural blasts (NBs). The results show an increase in long gene expression as neural progenitor cells exit the cell cycle and differentiate into neurons and was maintained as the neurons continued to mature in vivo (Habib et al., 2016; McCoy et al., 2018).

Collectively, these findings demonstrate the expression of long genes as a hallmark of neuronal identity and maturity. Intriguingly, gene length and transcriptional mechanisms have been implicated in neurodevelopmental and neuropsychiatric disorders (Zylka et al., 2015). This unique feature of the neuronal transcriptome raises the question of how the transcriptional machinery "detects" gene size or is recruited to these genes. Genomic size and complexity have evolved from prokaryotes to multicellular eukaryotes. Increases in the size of introns as well as the number of introns gives rise to the variation in overall gene length. Long gene have more splice variants, as well as a greater number of regulatory elements, giving rise to the alternative isoforms with distinct molecular functions and expression patterns (McCoy and Fire, 2020; Sugino et al.).

Given that there are a finite number of protein coding-genes, the evolution of the complexity of the neuronal transcriptome suggests a sophisticated and elaborate gene regulatory system that facilitates unique gene expression patterns. Gene expression is typically controlled

by specific regulatory elements that are active in specific cell populations at distinct developmental stages, indicating the high level of specialization (Closser et al., 2022). Therefore, identifying the transcriptional regulators that control long gene expression is critical to better understand this unique signature of the neuronal transcriptome.

1.2 Role of methyl-CpG binding protein 2 (MeCP2) as a neuron enriched transcriptional repressor

In addition to expressing long genes, neurons are enriched for specialized regulatory factors that control the expression of long genes. MeCP2 and the specialized form of DNA methylation it reads out have emerged as important regulators of long genes. MeCP2 is ubiquitously expressed but is especially abundant in post-mitotic neurons (Jung et al., 2003; Mullaney et al., 2004; Shahbazian et al., 2002; Skene et al., 2010). MeCP2 protein was initially identified as a chromatin-associated nuclear protein that binds methylated CG-dinucleotides, a DNA modification associated with gene silencing (Lewis et al., 1992). In most mammalian cells, *de novo* DNA methylation is deposited in a CG dinucleotide (mCG) context by DNA methyltransferases (DNMT3A and DNMT3A) and is maintained throughout cell division (Reik, 2007), however in neuronal cells, non-CpG methylation (mCH, where H = A, T, C), the majority of which occur in a CA dinucleotide (mCA) context, and hydroxymethylation (hmC) are expressed at significant levels (Guo et al., 2014; Lister et al., 2013). The discovery of MeCP2 as an epigenetic reader of methylated DNA (Kinde et al., 2015) raised the possibility that MeCP2 binding mediates the repressive effects of methylation on both transcription and chromatin. These observations initiated a series of pioneering studies directly investigating the role of MeCP2 in transcriptional repression.

Additional interrogation of the protein structure of MeCP2 led to the identification of the

methyl-CpG binding domain (also known as the methyl-DNA binding domain; MBD) of MeCP2 (Nan et al., 1993a). DNA binding activity assays using deletion fragments, or truncated versions, of MeCP2, indicated that binding to methylated DNA requires a sequence of approximately 85 amino acids in the N-terminal region of MeCP2. In a follow up study, to determine whether localization of MeCP2 to chromatin depends on DNA methylation, MeCP2 was transiently expressed in wild-type cells and mutant cells deficient in DNA methylation. The result showed inefficient localization of MeCP2 in the absence of DNA methylation, therefore suggesting that proper localization of MeCP2 depends on the presence of DNA methylation. The MBD is the only domain of MeCP2 shown to bind DNA with sequence specificity (Ballestar et al., 2000; Yusufzai and Wolffe, 2000). Furthermore, deletions within the MBD of MeCP2 abolish binding to methylated DNA and prevent proper localization of MeCP2 to methylated heterochromatin, therefore indicating the MBD is necessary and sufficient for localization of MeCP2 to heterochromatin. These findings are consistent with chromatin immunoprecipitation analyses in which mutations in the MBD result in profound redistribution of MeCP2 genome-wide (Baubec et al., 2013). Collectively, this evidence demonstrates the MBD domain recruits MeCP2 to methylated DNA (Nan et al., 1996). With the identification of the MBD, subsequent studies explored whether this region of MeCP2 also represses transcription.

To determine whether the MBD is responsible for transcriptional repression, the MeCP2 MBD mutant, that was shown to impair binding to methylated DNA, was evaluated in transcription assays. The results showed the MeCP2 MBD mutant failed to inhibit transcription, thus indicating MeCP2 binding to methylated DNA alone is not sufficient to inhibit transcription and that a different region of MeCP2 protein mediates transcriptional repression. *In vitro* transcription reactions using deletion fragments of the C-terminal region of MeCP2 ultimately led to the identification of the

transcriptional repression domain (TRD) of MeCP2 (Nan et al., 1997). Subsequent studies demonstrate the mechanism underlying MeCP2-mediated transcriptional repression occurs through the recruitment of histone deacetylase complexes (HDACs) via the TRD of MeCP2 (Kokura et al., 2001; Nagy et al., 1997; Nan et al., 1998; Stancheva et al., 2003). Enzymatic activity assays using the HDAC inhibitor Trichostatin A reduced the transcriptional repressive behavior of MeCP2, thus suggesting that the deacetylation of histones or proteins, is an essential component of MeCP2-mediated repression (Nan et al., 1998). These early studies provided support for a model whereby MeCP2 inhibits transcription by recruiting an HDAC-corepressor complex to chromatin which in turn deacetylates chromatin (Jones et al., 1998; Nan et al., 1998).

More refined mapping of the TRD led to the identification of a discrete domain that serves as the site of interaction between MeCP2 and the HDAC-containing nuclear receptor co-repressor complex (NCoR) defined as the NCoR-interaction domain (NID). The recruitment of the NCoR co-repressor complex is alone sufficient for MeCP2-mediated transcriptional repression (Lyst et al., 2013). This finding provides evidence supporting a model in which the primary function of MeCP2 is to serve as a bridge between DNA and the NCoR co-repressor complex. Collectively, these studies provide substantial evidence in support of MeCP2 functioning as a transcriptional repressor.

1.3 Role of Topoisomerase IIβ (TOP2β) as a regulator of neuronal function

DNA topoisomerases are ubiquitous and essential enzymes that have critical roles in the fundamental biological processes of replication, transcription, and chromatin remodeling (Chen et al., 2013). Type II DNA topoisomerases disentangle topological problems that arise in DNA, such as supercoiling, by generating a transient double-strand break, which allows the passage of one DNA

duplex through another (Nitiss, 2009; Wang, 1998, 2002). While all cells require topoisomerases to maintain proper functioning, the nervous system in particular, given its complexity and diversity, has critical need for these enzymes. Indeed, aberrant topoisomerase function in neural cells has been implicated in neurodevelopmental and neuropsychiatric disorders (Mckinnon, 2016).

Extensive biochemical and structural analysis have provided key insights into biological contexts in which TOP2β functions. Eukaryotic TOP2 enzymes are highly conserved and structural studies have characterized the key domains of these enzymes. The amino-terminal region contains an ATP-binding domain, which allows for dimerization with other TOP2 monomers. The central region contains a TOPRIM domain and the active site tyrosine residue, which allows for the strand breakage and re-ligation. The active tyrosine of TOP2 binds the phosphodiester bond of DNA through nucleophilic attack thereby generating a strand break and a reversible enzyme-DNA covalent linkage. The carboxyl-terminus is not conserved between topoisomerase enzymes or the TOP2 paralogs. Thus, this region gives rise to the enzyme-specific functions of the enzymes and is implicated in nuclear localization as well as the regulation of enzyme activity by post-translational modifications and protein-protein interactions (Adachi et al., 1997; Austin and Marsh, 1998; Chen et al., 2013; Linka et al., 2007a).

Mammalian cells express two distinct type II topoisomerases (TOP2) enzymes, TOP2α and TOP2β. Initially it was thought that a single type II topoisomerase was present only in proliferating cells. However, evidence emerged for a second type II topoisomerase from two independent studies analyzing non-proliferating cells (Holden et al., 1990; Roca and Mezquita, 1989). This observation gave rise to questions as to whether the two topoisomerase enzymes, TOP2α and TOP2β, perform specialized functions or are redundant enzymes. This open question initiated a series of studies exploring the roles of TOP2α and TOP2β in the cell-cycle as well as the expression profiles of these

two enzymes across various mammalian tissues. TOP2α is expressed in proliferating cells and is essential for DNA replication. TOP2β is expressed in both proliferating and non-proliferating cells, however, it has been shown to be dispensable in proliferating cells, but required for gene regulation in differentiated cells (Akimitsu et al., 2003; Chaly et al., 1996; Li and Wang, 1998; Lyu et al., 2006; Meyer et al., 1997; Woessner et al., 1991).

Meanwhile, two separate studies carried out extensive expression profiling of TOP2 in the developing brain. The first study profiled TOP2α and TOP2β in the murine liver and brain before and after birth using cDNA clones and Northern blotting. Strikingly, *Top2b* was induced more than 6-fold in the brains of newborn mice compared to embryos, but not in the liver (Capranico et al., 1992). In a second study, the expression patterns of TOP2α and TOP2β during postnatal development in the rat cerebellum were assessed using immunohistochemistry. The results show a sharp transition in expression from TOP2α to TOP2β in the precursors that terminally differentiate into Purkinje or granule neurons (Tsutsui et al., 2001). The distinct expression patterns of TOP2α and TOP2β in the developing brain provided evidence that TOP2α and TOP2β have non-redundant roles and that additional studies are needed to understand the role of TOP2β in post-mitotic cells.

Initial indications of the importance of TOP2β function in the nervous system emerged from mouse models in which TOP2β was ablated from the nervous system. Constitutive knockout of *Top2b* (*Top2b*$^{-/-}$) in mice causes perinatal death, due to the inability of motor neurons to synapse and innervate the diaphragm, thus preventing newborn *Top2b*$^{-/-}$ mice from being able to breathe. Additionally, sensory neurons are unable to enter the spinal cord (Yang, 2000). Conditional knockout of *Top2b* in cortical structures, driven by Cre-Foxg1, causes aberrant lamination and broad disruption of neurogenesis (Lyu and Wang, 2003). Similar to the germline deletion, conditional knockout of *Top2b* in the mouse brain was perinatal lethal, indicating that TOP2β plays critical roles

in nervous system function.

Due to the perinatal death phenotype observed in the constitutive knockout and brain-specific knockout of *Top2b*, elucidating the precise roles and activities of TOP2β *in vivo* in the adult brain have been limited. One study sought to circumvent the perinatal death limitation by evaluating TOP2β in the developing retina, using Dkk3-Cre, which specifically targets retinal cells born embryonic day (E10), a time point prior to the birth of a majority of retinal cell type (i.e. photoreceptors, horizonal, amacrine and bipolar cells). The conditional knockout mice showed delayed neuronal differentiation, neurodegeneration of the outer layers of photoreceptors and significant reduction in retinal cell number. Gene expression analysis showed gene networks involved in neuronal survival and neurite outgrowth were preferentially dysregulated when *Top2b* is lost. In sum, these results suggest TOP2β plays an essential role in the survival and maintenance of post-mitotic neurons in the retina (Li et al., 2014).

Given that topoisomerases promote changes in DNA topology by resolving torsional stress during transcription, it is possible that altered gene expression may underlie the perinatal death phenotype of the *Top2b*$^{-/-}$ mice. Interestingly, gene expression analysis in *Top2b*$^{-/-}$ mouse embryos at the late gestation stage identified approximately four percent of actively expressed genes were downregulated. While only a moderate reduction in overall gene expression was observed, this analysis showed that nearly one-third of all developmentally relevant genes, such as genes involved in neurogenesis and neural differentiation, were altered (Lyu et al., 2006). Consistent with these observations, a study profiling neuronal differentiation in *Top2b*$^{-/-}$ embryonic stem cells (ESCs) observed normal growth, proliferation, and differentiation from ESC into neural progenitors, however, upon terminal differentiation, that is, the transition from progenitor into neuron, cells undergo apoptosis. Similar observations were made in cultured cortical neurons from *Top2b*$^{-/-}$

embryonic day 18.5 (E18.3) mice, coupled with a reduction in the expression of gene involved in neurogenesis (Tiwari et al., 2012). Therefore, it was suggested that dysregulation of TOP2β genes may lead to apoptosis of Top2b$^{-/-}$ cortical neurons. Altogether, these studies suggest TOP2β-mediated transcription regulation is largely dispensable during cellular proliferation, possibly due to presence of TOP2α, however, is necessary after neuronal differentiation and cannot be substituted for other topoisomerase enzymes expressed in this developmental window.

In parallel, seminal studies began profiling the binding of TOP2β across to genome in various tissues, including mouse cortical neurons, mouse cerebellar granule neurons, mouse liver cells, mouse spleen B-cells and human MCF7-cells, in order to elucidate the role of TOP2β in transcription. One of the initial studies using chromatin immunoprecipitation (ChIP) and tiling arrays to profile TOP2β binding was performed on whole brains from E18.5 mice. The investigators observed preferential binding of TOP2β at the 5' regions of many of TOP2β target genes, which was determined by gene expression data collected in *Top2b*$^{-/-}$ mice from the same study. These results support a model in which TOP2β regulates transcription via the promoter regions of developmentally regulated genes (Lyu et al., 2006).

In another study investigating TOP2β occupancy in cultured cerebellar neurons from postnatal day 8 (P8) rats, the investigators established a "functional" ChIP assay, termed etoposide-mediated topoisomerase immunoprecipitation (eTIP) in which a TOP2 inhibitor, etoposide, was used to trap TOP2β in a complex with DNA, forming a TOP2β-cleavage-complex, which allows for the identification of TOP2β action sites. After hybridization of the DNA to tiling arrays, this study showed developmental genes that are regulated by TOP2β are bias towards being long and encode membrane proteins, including ion channels and receptors. Moreover, these genes are adjacent to long

AT-rich intergenic regions (Sano et al., 2008). In a subsequent study using ChIP and tiling arrays to profile TOP2β in stem cell derived neurons, the investigators observed TOP2β binding at promoters enriched for histone H3 di-methylation at lysine 4 (H3K4me2), a modification associated with active chromatin (Tiwari et al., 2012). In mouse cortical neuron cultures, TOP2β occupancy was assessed using ChIP-seq in unstimulated and stimulated (with N-methyl-D-asparate, NMDA, which mimics the actions of the neurotransmitter glutamate, or, with potassium chloride, which triggers depolarization) conditions and the results showed increased binding of TOP2β upstream of actively transcribed genes, including at promoters and enhancers. The binding profile for TOP2β corresponds with the binding of transcription factors that regulate activity-dependent gene expression (e.g. CREB, SRF, CBP and CTCF) (Madabhushi et al., 2015). Collectively, these studies showed increased TOP2β binding at regulatory sites, predominantly the promoters of actively transcribed genes, as well as broadly distributed across regions of open chromatin. Moreover, they demonstrate contexts in which TOP2β influences the regulation of gene expression.

While the foundational studies provide insight into the expression and localization of the early stages of postnatal development, there remain many outstanding questions regarding the function of TOP2β in mature neurons of the adult brain. For example, how is TOP2β recruited to specific sites on the genome and how is TOP2β activity regulated at these sites. Along these lines, additional analyses are needed to determine which stages of transcription require TOP2β activity. One way to start addressing these questions would be to carry out proteomic studies to gain a better understanding of the protein complexes that include TOP2β. Identifying new protein interactors or protein complexes is a critical step to identifying relevant biological processes. Taken together, uncovering the role of TOP2β in regulating the neuronal genome in the mature brain is critical to

understanding its function.

1.4 Disruption of MeCP2 and TOP2β in disease pathology

Mutations in numerous transcriptional regulators has emerged as the cause of neurodevelopmental disorders and syndromes such as autism (Iossifov et al., 2014; de Rubeis et al., 2014). This raises the question of how do mutations of heterogenous and often ubiquitously expressed proteins lead to selective deficits in the nervous system. The prevalence of chromatin remodeler mutations underlying neurodevelopmental disorders points to an unappreciated role in neurons which is to produce stable levels of gene expression in order to maintain normal neuronal functioning (Ronan et al., 2013). A growing body of literature demonstrates the transcriptional regulators MeCP2 and TOP2β are required for maintaining proper expression of long genes, a population that encodes proteins associated neuronal physiology and synaptic activity (Gabel et al., 2015; King et al., 2013; Mabb et al., 2014; Sugino et al., 2014; Zylka et al., 2015).

Rett syndrome is a pervasive, postnatal neurodevelopmental disorder associated with synaptic dysfunction caused by mutations in the X-linked *MECP2* gene, which encodes the MeCP2 protein (Amir et al., 1999). Classical Rett syndrome affects approximately 1 in 10,000 females and is characterized by an initial period of normal development followed by progressive neurological dysfunction and developmental regression, including motor abnormalities, loss of speech, impaired breathing and repetitive hand wringing (Leonard et al., 2016; Neul et al., 2010; Percy et al., 2010; Zoghbi, 2016). Over two decades of research on Rett syndrome has provided insight into the function of the transcriptional regulator MeCP2. While MeCP2 is known to function as a transcriptional repressor, the precise mechanism underlying how transcription and gene expression

are dysregulated when MeCP2 is absent remains an open question. The development of loss-of-function (Chen et al., 2001; Guy et al., 2001) and gain-of-function (Collins et al., 2004; van Esch et al., 2005) mouse models of MeCP2 have been invaluable tools that enable the ability to connect protein biology and molecular mechanism with downstream pathology

One proposed model is that MeCP2 regulates transcription in a gene-length dependent manner, with long genes enriched for methylated CA-dinucleotides (mCA) being preferentially upregulated in the absence of MeCP2 (Gabel et al., 2015; Sugino et al., 2014). Indeed, the gene expression profile of the *Mecp2* knockout mouse (*Mecp2*-null) showed long genes (> 100kb) were modestly upregulated. Similar expression profiles were observed when examining two other Rett mice with nonsense mutations, R270X and G273X, that differ in prevalence and disease severity. Conversely, when profiling the mouse model of MeCP2 duplication syndrome (*MECP2*-Tg), the expression of long genes was markedly reduced (Gabel et al., 2015). Together, these studies provide direct evidence of the transcriptional repressor function of MeCP2 being biased towards longer transcripts. Collectively, these studies investigating loss or gain of MeCP2 function demonstrate that neurons are acutely sensitive to the dose of MeCP2. Too little or too much causes subtle changes in expression of a large number of genes and the cumulative effect of many changes may contribute to the underlying the pathology (Lombardi et al., 2015; Zoghbi, 2016).

This profound observation of a length-dependent impairment in gene transcription was also observed in studies examining inhibition or loss of function of topoisomerases, which have been implicated in neurodevelopmental disorders. Inhibitors of topoisomerase 1 (TOP1) enzyme have been shown to selectively reduce the expression of long genes associated with synaptic function. Notably, inhibitors to topoisomerase IIβ (TOP2β) enzymes also caused reduced expression of long genes (King et al., 2013; Mabb et al., 2014). Consistent with pharmacological inhibition, knockdown

of *Top1* and *Top2b* in cultured neurons using shRNA also resulted in the reduced expression of very long genes. Altogether, these data demonstrate topoisomerases are required for the expression of long neuronal genes, specifically genes encoding synaptic proteins and dysregulation of topoisomerase activity may contribute to the pathology of various neurological disorders (Zylka et al., 2015). Because topoisomerases play an essential role in regulating the expression of long genes, they have emerged as a potential therapeutic target for neurodevelopmental disorders, including Rett syndrome.

Notably, rare *de novo* missense mutations in *TOP2B* have recently been identified in two female individuals with neurodevelopmental disorder (NDD). The mutation identified in both individuals encodes c.187C>T, p(His63Tyr), missense mutation in the enzyme. The clinical presentation of the two individuals include microcephaly, hypotonia, impaired fine motor skills, no verbal communication skills, and autistic features, which partially overlap with Rett syndrome phenotypes (Hiraide et al., 2020; Lam et al., 2017). The rareness of the TOP2β mutations, coupled with the presumptive null, truncating mutations typically identified in individuals with NDD, suggests that this mutation may alter the activity of TOP2β in a specific manner rather than lead to loss-of-function. Indeed, in unpublished analysis, we have detected evidence that this mutation may lead to hyperactivation of the enzyme (Nettles, Soto, Gabel, unpublished results). This finding raises the possibility that the direct mutation of TOP2β may drive a disorder with Rett-like features. Future analysis can explore this possibility.

While dysregulation of MeCP2 and TOP2β represent a vulnerability for neurons that can give rise to neurodevelopmental disease, this neuronal gene-length dependent regulatory signature associated with both MeCP2 and TOP2β provides potential therapeutic targets for neurodevelopmental disorders (Zylka et al., 2015). For example, inhibition of TOP1 has been shown

to restore the expression of a subset of neural long genes to within normal range in neurons lacking MeCP2, demonstrating that at least in principle, a strategy involving targeting topoisomerases could be beneficial in the treatment of Rett syndrome (Gabel et al 2015).

1.5 Summary

Recent studies have implicated transcriptional regulators in neurodevelopmental disorders and genetic analysis in model organisms have provided considerable insights as to the function of these regulators during normal development. The findings from these studies allows speculation on underlying associated disease mechanism. MeCP2 and TOP2β are two chromatin regulators that have independently been shown to alter the expression of long genes in opposite directions. Several lines of evidence indicate MeCP2 can function as a transcriptional repressor and downregulate the expression of long, highly methylated genes. Loss or disruption of MeCP2 leads to the over-expression of long genes and this dysregulation may contribute to Rett pathology. In contrast, several studies suggest TOP2β plays an important role in transcription and that inhibition of TOP2β causes reduced expression of long genes.

The work I describe in this thesis uses molecular, biochemical and genomic approaches to expand our knowledge on a gene regulatory mechanism underlying the expression of long genes, a population enriched in neurons and critically important for neuronal activity and connectivity. In chapter 2, I describe studies in which I have employed molecular and biochemical approaches to identify regulatory proteins that interact with MeCP2 to determine how MeCP2 regulates the expression of long genes. Through this analysis, I identify an interaction between MeCP2 and TOP2β. Given that MeCP2 and TOP2β have opposing effects on gene expression, this raises the

possibility that MeCP2 may have a repressive effect on TOP2β activity.

To test this hypothesis, I carry out analysis interrogating the interaction between MeCP2 and TOP2β to determine the domains required for the interaction. *In vivo* TOP2β functional analysis in the brain is limited due to genetic manipulations of *Top2b* in mice being embryonic lethal. Furthermore, previous studies using conventional genomic approaches such as ChIP-seq to assess TOP2β occupancy show TOP2β is distributed broadly (Lyu et al., 2006; Madabhushi et al., 2015; Tiwari et al., 2012). To circumvent these limitations, I established a culture system which allows neurons to grow and mature. Moreover, using a modified ChIP-seq technique called eTIP-seq, I am able to profile TOP2β activity across the neuronal genome and gain insight into the mechanism by which TOP2β promotes gene expression. Through this analysis, I identified enriched TOP2β activity at regulatory regions associated with long genes as well as MeCP2-repressed genes. Recent findings from our lab have shown that MeCP2 downregulates the expression of long, highly methylated genes by repressing enhancers (Clemens et al., 2020). These observations prompted me to investigate how manipulating MeCP2 protein expression alters TOP2β activity at long genes and enhancers regulated by MeCP2. My data suggests that the interaction between MeCP2 and TOP2β is functionally relevant for regulating the expressions of long genes. This leads me to conclude that this interaction is necessary for neuronal maintenance and survival and that disruption of this interaction may contribute to Rett pathology.

Finally, in chapter 3, I discuss future studies that seek to identify the functional significance of the interaction between MeCP2 and TOP2β and how Rett-associated mutations may disrupt this interaction and contribute to disease pathology. Furthermore, I discuss future analyses that can be carried out to elucidate the role of TOP2β during transcription, including TOP2β recruitment to different sites across the genome and the role of post-translational modifications in regulating

TOP2β activity. These studies will give valuable insights into the normal function of TOP2β in mature neurons and will provide insight in the molecular mechanism that may be disrupted in the neurodevelopmental disorder associated with the *de novo* TOP2β mutation.

Chapter 2:

MeCP2 Represses the Activity of Topoisomerase IIβ in Long Neuronal Genes

This chapter is adapted from a manuscript submitted to *Cell Press*

Nettles, SA., Ikeuchi, Y, Agwu, C, Bonni A., Gabel, HW. MeCP2 Represses the Activity of Topoisomerase IIβ in Long Neuronal Genes. *Submitted*

2.1 Introduction

During development, diverse combinations of genes are expressed in order to establish the complex morphology, connectivity and function of neurons (Closser et al., 2022). The distinct complexity of the neuronal transcriptome is evidenced by the observation that the longest genes in the genome are expressed at uniquely high levels in the brain (Gabel et al., 2015). Many of these long pre-mRNAs (e.g. > 100 kilobases) encode proteins essential for neuronal function, including ion channels, synaptic receptors and cell adhesion molecules (Gabel et al., 2015), and have been implicated in neurological disorders (King et al., 2013). Long genes are expressed many fold higher in neurons than other cell types (Gabel et al., 2015; Sugino et al., 2014; Zylka et al., 2015) and analyses of transcript levels in developing mice and *in vitro*-derived human neurons demonstrate that the expression of long genes is a hallmark of functional maturity (McCoy et al., 2018). Pharmacological interventions that selectively downregulate long genes have been shown to lead to synaptic dysfunction and reduced neuronal transmission (Mabb et al., 2014). Thus, the expression of long genes is a distinctive and necessary feature of the neuronal transcriptome and aberrant expression of these genes represents a vulnerability in neurons that can lead to dysfunction.

Members of the topoisomerase gene family have been identified as essential for neural development and critical for the expression of long genes (Feng et al., 2017; King et al., 2013). Topoisomerases resolve DNA supercoiling and other topological constraints that arise during cellular processes including cell division, gene transcription and chromatin remodeling (Corbett and Berger, 2004; Nitiss, 2009; Roca, 2009). Type I topoisomerases (TOPI) relax DNA by transiently nicking and rejoining one strand of duplex DNA, whereas type II topoisomerases (TOP2) transiently break and rejoin both strands of duplex DNA simultaneously (Austin and

Marsh, 1998; Koster et al., 2005; Pommier et al., 2016; Stewart et al., 1998). Mammalian cells express two genetically distinct TOP2 enzymes, TOP2α and TOP2β. Although TOP2α and TOP2β have similar structures and biochemical activities, they have different expression patterns and biological roles (Austin et al., 1993; Jenkins et al., 1992). For example, TOP2α is highly expressed in proliferating cells but excluded from neurons, whereas TOP2β is highly expressed in both dividing and post-mitotic cells (Capranico et al., 1992; Harkin et al., 2016; Juenke and Holden, 1993; Kondapi et al., 2004; Tiwari et al., 2012; Tsutsui et al., 1993; Watanabe et al., 1994). Consistent with this expression pattern, TOP2α is essential for cell-cycle-related events, such as DNA replication, whereas TOP2β plays an essential role in DNA repair, transcription, and development. Germline or conditional deletion of *Top2b* leads to defective brain development and perinatal death (Lyu and Wang, 2003; Yang, 2000). Pharmacological inhibition or knockdown of TOP2β in cultured neurons also leads to reduced expression of long genes in neurons (King et al., 2013). While the exact mechanism by which TOP2β promotes gene expression is not clear, TOP2β may decatenate DNA to open chromatin and activate regulatory regions (Ju et al., 2006; Lyu et al., 2006) or unwind DNA to aid the progression of RNA polymerase though genes (King et al., 2013; Mabb et al., 2014; Zylka et al., 2015). Together, these findings indicate that TOP2β is critical for the robust expression of long genes, which are essential for normal brain function.

Balanced gene regulation is essential for neuronal function and recent studies have implicated methyl-CpG-binding protein 2 (MECP2) as an important regulator of transcription in neurons that preferentially represses long genes (Boxer et al., 2020; Chen et al., 2015; Gabel et al., 2015; Lagger et al., 2017; Renthal et al., 2018; Sugino et al., 2014). Loss or overexpression of MeCP2 causes the severe neurological disorders Rett syndrome (Amir et al., 1999) and MeCP2

duplication syndrome (van Esch et al., 2005), respectively, demonstrating its importance for nervous system function. Although multiple functions of MeCP2 have been described, molecular and genomic analyses support a model in which MeCP2 binds to methylated DNA within and around genes to promote a repressive chromatin structure and down-regulate gene transcription. Transcriptional repression by MeCP2 appears to be mediated in part by downregulation of enhancers within genes, leading to downregulation of transcription initiation (Clemens et al., 2020; Boxer et al., 2020; Chahrour et al., 2008; Lewis et al., 1992; Lyst et al., 2013; Nan et al., 1993a). Notably, because long genes are enriched for DNA methylation and contain many enhancers, these repressive effects most robustly affect long genes (Clemens et al., 2020). Consistent with the essential roles for long genes in neurons, MeCP2 represses genes that are critical for neuronal development and physiology (Gabel et al., 2015; Sugino et al., 2014). These studies suggest that an important function of MeCP2 is to preferentially tune down the expression of long genes and implicate disruption of long gene regulation in the pathology of MeCP2-associated neurodevelopmental disorders.

Mechanistically, biochemical studies of MeCP2 have identified protein interactors that associate with MeCP2 to mediate its functions. Extensive evidence demonstrates MeCP2 may function as a repressor by interacting with the NCoR/SMRT-corepressor complex (Lyst et al., 2013). In addition, interactors such as PSIP1/LEDGF and TCF-20 may also play important roles in MeCP2-mediated gene regulation (Jian et al., 2022; Li et al., 2016). However, whether additional protein interactors or mechanisms play a key role in MeCP2 functions, remains to be investigated.

In this study, we identify a physical interaction between MeCP2 and TOP2β with important functional implications for the regulation of long genes in neurons. Through unbiased proteomic analysis and targeted interaction mapping, we identify and interrogate the physical

association between TOP2β and MeCP2, implicating a key domain of MeCP2 required for this interaction. We profile the sites of TOP2β activity in neurons genome-wide and demonstrate that TOP2β is preferentially active at long genes, including genes repressed by MeCP2. We further demonstrate that altering MeCP2 levels in neurons alters TOP2β activity at these long, MeCP2-regulated genes. Altogether, our findings demonstrate an interaction between TOP2β and MeCP2 in the regulation of essential neuronal genes, that when disrupted, may contribute to MeCP2-related neurological disorders.

2.2 Results

2.2.1 MeCP2 selectively binds TOP2β in the mouse brain

To explore the molecular mechanism underlying MeCP2 transcriptional regulation of long genes, we sought to identify additional proteins that may interact with MeCP2 in neurons. We expressed and purified FLAG-tagged MeCP2 protein in cultured cortical neurons and performed unbiased mass spectrometry analysis of co-precipitating proteins (**Table S1**). In addition to previously identified interactors of MeCP2 (KPNA3, PSIP1/LEDGF), TOP2β was reproducibly identified in this analysis (**Figure 1A**). Motivated by previous findings from separate studies showing that inactivation of topoisomerase enzymes (King et al., 2013) and mutation of MeCP2 (Gabel et al., 2015) leads to opposing effects on long gene expression, we sought to further investigate this interaction. To independently test the association of TOP2β-MeCP2 and assess its functional relevance in the postnatal brain, we immunoprecipitated MeCP2 from forebrain extracts of wild-type 8-week mice. To eliminate the potential for nucleic acid bridging, nuclear extracts were treated with benzonase, a DNA and RNA nuclease, before the immunoprecipitation (IP) was

carried out. Through this endogenous co-immunoprecipitation (Co-IP) analysis, we found MeCP2 interacts with TOP2β but not TOP1, the other major topoisomerase that is robustly expressed in the brain (**Figure 1B**). Together, the detection and validation of this interaction between MeCP2 and TOP2β suggest that modulation of long gene expression by MeCP2 may occur through a mechanism that involves the activity of TOP2β.

The proteins MeCP2 and TOP2β are highly conserved across vertebrate species and are well-characterized, with specific domains having been defined as important for distinct protein functions (Austin et al., 1995; Baker et al., 2013; Caron and Wang, 1994; Heckman et al., 2014a; Lewis et al., 1992; Lyst et al., 2013; Nan et al., 1993a; Wang, 1996; Watt and Hickson, 1994). Thus, determining sequences in each protein required for the MeCP2-TOP2β association might provide insight into the nature and importance of the protein-protein interaction. We therefore established an *in vitro* co-IP assay in heterologous cells and asked which sequences within each protein are sufficient for the interaction. Co-expression of MYC-tagged full length (1-1612aa) TOP2β with FLAG-tagged full length (1-484aa) MeCP2 in HEK293 cells led to a detectable interaction upon co-IP (**Figure 2A**). Notably, this interaction was specific under these conditions as expression of MYC-tagged full length TOP2β with another nuclear protein, FLAG-tagged full length (1-908aa) DNMT3A, in these cells did not lead to a detectable co-IP signal. This validated the interaction we detected *in vivo* and allowed us to begin to dissect the amino acids that mediate the association between MeCP2 and TOP2β.

We next probed the sequences within TOP2β that are sufficient for the interaction with MeCP2. TOP2β contains three functional regions. The amino-terminal region contains an ATP-binding domain that allows dimerization with other TOP2 monomers. The central region of

TOP2β is the catalytic domain which contains the active tyrosine responsible for the strand breakage and re-ligation. The carboxyl-terminal region is suggested to be required for localization, regulation of enzyme activity by post-translational modification, and regulation of enzyme function by protein-protein interactions. (Austin and Marsh, 1998; Austin et al., 1993, 1995; Berger et al., 1996; Chang et al., 2013; Chen et al., 2013; Jenkins et al., 1992; Lindsley and Wang, 1991). To determine the region of TOP2β that binds MeCP2, we co-expressed exogenous MYC-tagged N-terminal (1-449aa), central catalytic (450-1198aa), and C-terminal (1199-1612aa) fragments with FLAG-tagged full length MeCP2 (1-484aa) in HEK293 cells. This analysis detected co-IP of MeCP2 specifically with the C-terminal region of TOP2β, whereas the N-terminal region and the catalytic region of TOP2β do not interact with MeCP2 (**Figure 2B**). The interaction between the MeCP2 and the C-terminal region of TOP2β is notable because the N-terminal ATPase and central breakage/reunion domains are very similar between TOP2α and TOP2β, whereas the C-terminal regions differ both in size and sequence and is considered the domain that gives rise to the enzyme-specific functions of TOP2α and TOP2β (Gilroy and Austin, 2011; Linka et al., 2007b; Meczes et al., 2008).

We next interrogated the region of MeCP2 required for the interaction with TOP2β. MeCP2 contains two critical domains: the methyl-DNA-binding domain (MBD), which is the region of MeCP2 that binds methylated DNA, and the NCoR-interaction domain (NID), required for the association with the NCoR corepressor complex (Guo et al., 2014; Heckman et al., 2014b; Lewis et al., 1992; Lister et al., 2013; Lyst et al., 2013; Nan et al., 1993a). Collectively, these two domains contribute to the repressive role of MeCP2 during transcription. A minimal protein containing the two domains is sufficient to rescue the multiple effects of MeCP2 loss in mice that

model the severe neurological disorder Rett syndrome, which is caused by MeCP2 mutation (Tillotson et al., 2017). To determine the region of MeCP2 that is required for interaction with TOP2β, we co-expressed a FLAG-tagged MeCP2 fragment that contains the MBD (1-167aa), a fragment spanning the MBD and NID (143-484aa), and a C-terminal fragment (384-484aa), together with MYC-tagged full length TOP2β (1-1612aa). The fragments containing the MBD of MeCP2, 1-167aa and 147-484aa, were efficiently precipitated by MYC-tagged full length TOP2β, while the C-terminus of MeCP2 did not interact (**Figure 2C**). These results suggest that TOP2β interacts with the MBD of MeCP2. Notably, this is one of the key domains for rescue of the Rett phenotype, raising the possibility that loss of TOP2β interaction could contribute to Rett syndrome.

2.2.2 TOP2β activity is enriched within long genes in neurons

Previous analyses have shown that pharmacological inhibition or knockdown of TOP2β in neurons causes reduced expression of long genes, but the site of TOP2β action during the transcription of long genes is not well understood. We therefore sought to identify where TOP2β acts in neurons in order to understand how this activity could be modulated by MeCP2. To gain insight into where on the genome TOP2β is active in neurons, we performed etoposide-mediated topoisomerase-immunoprecipitation (eTIP-seq) (Sano et al., 2008) in primary neurons isolated from mouse embryonic day 14.5 (E14.5) cerebral cortex. In this genomic profiling assay, etoposide treatment of cells is used to link topoisomerase to the DNA and profile its activity across the genome (**Figure 3A**). Etoposide is an inhibitor of TOP2 proteins that traps the enzyme in a covalently linked complex with the cleaved DNA at the intermediate step in its cutting and re-

ligation cycle (Osheroff, 1989; Wu et al., 2011). Thus, in contrast to a conventional chromatin immunoprecipitation (ChIP) assay, which uses formaldehyde to non-specifically crosslink protein to DNA and profile binding in the genome, employing eTIP allowed us to directly assess the sites of TOP2β activity in neurons and study how altering MeCP2 impacts TOP2β activity.

To assess TOP2β activity in neurons, we utilized cultured cortical neurons at day *in vitro* 12 (DIV12) where we could administer etoposide directly to the cells and perform eTIP-seq (**Figure 3A**). Isolation of DNA by eTIP from etoposide treated cultured neurons yielded high levels of immunoprecipitated DNA compared to vehicle (DMSO) treated controls, indicating the high specificity of the assay (Figure S1A, Table S5). Deep sequencing of eTIP-seq samples revealed enrichment of TOP2β activity compared to input controls at local sites as well as broad regions (Figure 3B). Visual inspection of the eTIP profile relative to genome annotations showed enrichment of TOP2β at promoters and gene bodies. Notably, the eTIP signal appeared to be higher at longer genes (>100kb) relative to shorter genes (**Figure 3B**). To quantify TOP2β activity at all genes in the genome, we performed aggregate analysis and quantification of the signal at the promoters and gene bodies. Aggregate plots of promoter regions showed TOP2β is consistently associated with these regions, with the peak of the signal located just downstream of the transcription start site (TSS), spanning the TSS +1kb to +3kb region (hereafter referred to as "promoter associated" eTIP signal) (**Figure 3C**). Aggregate plots of genes binned by gene length revealed a length-associated enrichment of TOP2β at promoter associated regions and gene bodies of long genes, with the highest TOP2β signal at the longest genes and lower signal at the shorter genes (Figure 3C). Quantitative analysis of TOP2β signal at the promoter associated regions and gene bodies of all genes supported the aggregate profiles, showing a robust correlation between

TOP2β enrichment and gene length (Figure 3D, Table S2, S3).

Long genes are robustly expressed in neurons, raising the possibility that TOP2β may be associated with these genes due to high expression in these cells. Therefore, we assessed whether the TOP2β signal at genes is correlated with gene expression, and if so, whether this association can explain the enrichment of TOP2β with long genes. We performed RNA-seq analysis on neuronal cultures at an identical timepoint as eTIP-seq and assessed TOP2β enrichment as a function of gene expression using aggregate plots, in which genes were binned by level of gene expression (**Figure S1B**). This analysis showed a modest association between expression level and enrichment of TOP2β at promoters and within genes. Notably, however, while long genes are robustly expressed in our cultured neurons, we find that long genes are not more highly expressed than shorter genes in these cultures (**Figure S1C, Table S3**). This analysis indicates that TOP2β activity is associated with mRNA levels, but that recruitment of TOP2β to long genes cannot be explained by high expression of these genes alone.

We noted that additional sites of local eTIP enrichment can be observed outside of gene promoter regions and these sites are reminiscent of peaks associated with enhancers (Figure 3B). We therefore sought to assess the locations of TOP2β relative to these distal regulatory regions across the genome. To define active regulatory regions in the genome, we performed ChIP-seq on the histone modification Histone H3 lysine 27 acetylation (H3K27ac), a mark associated with active promoters and enhancers. We then performed peak-calling on these H3K27ac chip data to define active promoters and putative enhancer elements and assessed the degree of overlap of eTIP signal with these sites (Table S4). Examination of eTIP profiles revealed clear enrichment of TOP2β activity at acetyl peaks corresponding to putative enhancer elements (**Figure 3B**).

Aggregate plot analysis of eTIP at these putative enhancer elements revealed enriched TOP2β signal for enhancers found inside and outside of genes (**Figure 3E, Table S4**). Notably, enrichment of TOP2β was higher at enhancers located within genes (intragenic) and their surrounding regions relative to enhancers located outside of genes (extragenic) and their surrounding regions. This increased enrichment at intragenic enhancers may in part explain the increased signal inside long genes which can contain many enhancers.

To independently assess the profile of TOP2β activity genome-wide in an unbiased manner, we performed peak-calling on the TOP2β eTIP-seq signal. We then performed overlap analysis assessing where these peaks land in the genome, and compared them to resampled controls to assess if the distribution of TOP2β peaks differs from chance distributions. This analysis revealed that TOP2β peaks are significantly more prevalent within intragenic regions than extragenic regions compared to resampled controls (**Figure 3F, Table S4**). Analysis of TOP2β peak overlap with promoters and putative enhancers, as defined by H3K27ac peak calling analysis described above, indicated TOP2β peaks fall within promoters and enhancers significantly more than by chance (**Figure 3G, Table S4**). We next assessed whether TOP2β peaks are more likely to fall within intragenic enhancers or extragenic enhancers. Our analysis shows TOP2β peaks that overlap with enhancers are biased toward associating with intragenic enhancers compared to resampled controls (**Figure 3H, Table S4**). Together this peak distribution analysis is consistent with TOP2β enrichment within genes, including at intragenic enhancers, as described above (**Figure 3E, Table S4**). Our analysis of the distribution of TOP2β peaks revealed additional TOP2β peaks in unannotated regions (**Figure S1D, Table S4**). Initial evaluation of these unannotated peaks suggests many of these sites may indeed be enhancers that were just below the

level of detection of the H3K27ac peak calling analysis, however, many may not be enhancers. Most notable is the observation that nearly two-thirds of the unannotated TOP2β peaks are intragenic, suggesting TOP2β is engaged and active at many sites within these genes.

To further interrogate the profile of TOP2β activity across the genome and assess the validity of our findings, we performed eTIP-seq using a second, independently generated TOP2β antibody that targets a different region of the protein. Consistent with the initial observations, the second antibody showed TOP2β enrichment at promoter associated regions and gene bodies of long genes, as well as at enhancer regions (**Figure S2A – S2D, Table S4**). Peaks detected genome-wide were nearly three times more likely to fall within an intragenic region compared to an extragenic region (**Figure S2E, Table S4**), three times more likely to fall within an enhancer compared to a promoter (**Figure S2F, Table S4**), and two and a half times more likely fall within an intragenic enhancer compared to an extragenic enhancer (**Figure S2G, Table S4**). Furthermore, additional TOP2β peaks were observed in unannotated regions (**Figure S2H, Table S4**). Taken together, these analyses confirm TOP2β enrichment at promoters and enhancers within long genes, suggesting that the activity of TOP2β at these regulatory regions may play an important role in the transcription of these genes.

2.2.3 MeCP2 alters TOP2β activity at long genes repressed by MeCP2

The opposing effects of MeCP2 and TOP2β on long gene expression, along with our findings that MeCP2 and TOP2β interact in cells, raises the possibility that MeCP2 may inhibit TOP2β activity to affect transcription of long genes. We therefore sought to assess if TOP2β is active at MeCP2-regulated long genes and determine if MeCP2 modulates the activity of TOP2β at these genes.

Analysis of eTIP-seq signal at "MeCP2-repressed" genes, a set of long genes that have previously been shown to be consistently upregulated in expression when MeCP2 levels are low and downregulated in expression when MeCP2 levels are high across studies of MeCP2 mutations (Gabel et al., 2015), revealed TOP2β enrichment at the promoter associated regions and gene bodies of these genes. This signal was enriched relative to unchanged genes and "MeCP2-activated genes", a set of genes that are downregulated in expression when MeCP2 levels are low and upregulated in expression when MeCP2 levels are high (**Figure 4A**). Quantitative analysis of eTIP signal at MeCP2-regulated genes show TOP2β is significantly enriched at the promoter associated regions and genes bodies of MeCP2-repressed genes compared to all other genes (**Figure 4B, Table S3**). In this analysis, we also find that the MeCP2-activated genes displayed some TOP2β enrichment relative to unchanged genes, albeit, not to the levels of the MeCP2-repressed genes. This effect may be related to the relative expression level of this population of genes in our system. Altogether, these observations are consistent with our findings of a length-associated TOP2β signature and indicate that MeCP2-repressed genes, like other long genes, are targets of TOP2β activity.

Given that MeCP2 preferentially down-regulates a subset of neuronal long genes (e.g. MeCP2-repressed genes), we hypothesized TOP2β that is recruited to long genes may be negatively regulated by MeCP2, via the TOP2β-MeCP2 protein interaction, to restrict TOP2β activity and transcriptional activation of these genes. We therefore asked whether manipulating MeCP2 expression would alter TOP2β activity across the genome. To address this question, we carried out knockdown (KD) or overexpression (OE) of MeCP2 in our established neuronal culture system and performed eTIP-seq (**Figure 4C, Table S3**). Because the level of MeCP2 protein increases dramatically over postnatal development (Balmer et al., 2003; Skene et al., 2010), our culture system contains low levels of MeCP2 compared to the adult brain. Thus, we overexpressed MeCP2 in these neuronal cultures to

model the role of MeCP2 in the mature brain, while still providing access to the cells in order to perform eTIP-seq. We paired this with knockdown of MeCP2 to remove low levels of MeCP2 in the cultures and better recapitulate Rett syndrome-like conditions in which there is little to no MeCP2 activity present in neurons. We directly assessed MeCP2 RNA levels in the MeCP2 KD and MeCP2 OE cultured neurons using quantitative real-time PCR (RT-qPCR) (**Figure 4D, Table S5**). In the MeCP2 KD neurons, there was a robust reduction of endogenous MeCP2, while in the MeCP2 OE neurons, there was significantly higher MeCP2 expression.

To assess if our culture system mirrors aspects of the MeCP2-mediated gene regulation we observe in the brain, we performed total RNA-seq on the MeCP2 KD and MeCP2 OE cultured neurons in order to determine if the MeCP2-regulated genes identified in the *in vivo* studies of multiple brain tissues (Gabel et al., 2015) were misregulated in the same direction in the cultured neurons. We performed differential gene expression analysis and find that the MeCP2-repressed genes were significantly upregulated and the MeCP2-activated were significantly downregulated in MeCP2 KD neurons (**Figure 4E, Table S3**). This analysis suggests that our manipulation of MeCP2 in the cultured neurons elicits changes in gene expression that are consistent with the changes in gene expression observed in the brain, and therefore provides a suitable context to assess how levels of MeCP2 in neurons may affect TOP2β activity.

To investigate how MeCP2 expression alters TOP2β activity, we performed eTIP-seq in the MeCP2 KD and MeCP2 OE cultured neurons. Given that we observe TOP2β enrichment at the promoter associated regions and gene bodies of MeCP2-repressed genes, we hypothesized that loss of MeCP2 in the cultured neurons would induce an increase in TOP2β activity. Visualization of the eTIP signal in the MeCP2 KD and MeCP2 OE neurons at long, MeCP2-repressed genes suggested higher TOP2β activity in the MeCP2 KD neurons compared to MeCP2 OE neurons (**Figure 4F, Table S3**).

Differential analysis of eTIP signal from MeCP2 KD and MeCP2 OE neurons revealed increased TOP2β activity at the promoter associated regions and gene bodies of MeCP2-repressed genes in the MeCP2 KD neurons (**Figure 4G, Table S3**). At MeCP2-acitvated genes we find the TOP2β activity at promoter associated regions were unchanged while the gene bodies showed reduced signal in the MeCP2 KD neurons compared to MeCP2 OE neurons.

We previously showed one mechanism by which MeCP2 regulates long gene expression is by repressing the activity of intragenic enhancers associated with these genes (Clemens et al., 2020). To explore if TOP2β activity at these sites is affected by MeCP2, we evaluated the change in eTIP signal at intragenic enhancers of the MeCP2-regulated genes (Gabel et al., 2015) in the MeCP2 KD neurons compared to the MeCP2 OE neurons. Our analysis shows TOP2β activity is subtly, but significantly, increased at MeCP2-repressed enhancers (**Figure 4H, Table S4**) when MeCP2 protein expression is low compared to when it is high. Together, these findings raise the possibility that negative regulation of TOP2β by MeCP2 may contribute to the repression of intragenic enhancers.

2.3 Discussion

In this study, we have established a physical interaction in neurons between MeCP2 and TOP2β and demonstrated its importance for modulating TOP2β activity at long genes that are transcriptionally regulated by MeCP2. One model consistent with our findings is that when neurons are born in the developing brain, TOP2β is recruited to long genes and facilitates their expression. During postnatal development as neurons mature, MeCP2 becomes highly expressed and associates with methylated DNA in genes. At long, highly methylated genes, MeCP2 interacts with TOP2β to act as a molecular break on TOP2β activity within gene bodies and at enhancers, thereby leading to preferential restriction of TOP2β and tempering of gene expression. In the

absence of MeCP2, as in Rett syndrome, the molecular break on TOP2β has been removed thereby allowing for TOP2β activity to become unrestricted within long, MeCP2-regulated genes (**Figure 5**). This regulation may be key to the transcriptional regulatory role of MeCP2 to tune neuronal gene expression programs.

Type I and type II topoisomerases have been shown to promote the expression of long genes critical for neuronal development and synapses by cutting the DNA and resolving topological constraints. Through biochemical analysis, we show MeCP2 binds uniquely to the type II topoisomerase, TOP2β, but not the type I topoisomerase, TOP1, both of which are highly expressed in neurons. This is notable because TOP2 enzymes have been shown to have extended C-terminal sequences that play important roles in regulating these enzymes. Using fragments to map the regions sufficient for the interaction between MeCP2 and TOP2β, we show MeCP2 interacts with the C-terminal region (CTR) of TOP2β. Structural and functional studies have revealed that the CTR is the most divergent region between of TOP2α and TOP2β, and that removal of the CTRs of TOP2α and TOP2β did not affect the overall catalytic activity of either enzyme (Austin and Marsh, 1998; Linka et al., 2007a). However, truncation of the TOP2β CTR caused increased binding of TOP2β to DNA, whereas truncation of the TOP2α CTR had no effect on the binding of TOP2α to DNA, suggesting the CTR of TOP2β may have a negative regulatory function (Gilroy and Austin, 2011; Meczes et al., 2008). Collectively, these observations provide evidence of the CTR as a potentially important regulatory site of TOP2β activity, particularly in neurons. This specific interaction may have evolved in the vertebrate lineage, where the expression and specialization of TOP2β, TOP2α, and MeCP2 have emerged (Austin & Marsh, 1998, de Mendoza et al., 2021), and may contribute to the regulation of complex vertebrate

neuronal transcriptomes.

Our mapping of fragments of MeCP2 sufficient for the interaction with TOP2β implicates this interaction in Rett syndrome-related biology. Our study shows the TOP2β interaction occurs for MeCP2 fragments that contain the C-terminal portion of the MBD. In structural analysis, this portion of the MBD is found outside of the region that directly contacts the DNA (Nan et al., 1993b). This suggests that MeCP2 can interact with TOP2β and regulate its activity while bound to DNA. Previous studies have highlighted two critically important domains of MeCP2 for proper function, the MBD and the NID (Lyst et al., 2013; Nan et al., 1993a; Tillotson et al., 2017). The MBD domain of MeCP2 is required for the interaction with methylated DNA (Gabel et al., 2015; Guo et al., 2014; Lagger et al., 2017; Lister et al., 2013), while the NID domain of MeCP2 is required for the interaction with the NCoR-corepressor complex in cells (Lyst et al., 2013). Notably, our findings suggest the need to fully evaluate the effects of Rett syndrome causing missense mutations, which have been shown to be concentrated in the MBD, as some mutations may in fact disrupt the interaction with TOP2β rather than solely disrupting methyl-DNA binding capacity.

While the mechanism by which TOP2β facilitates long gene expression is not known, our study identifies TOP2β sites of action within long genes that may provide insight into its function in neurons. By using eTIP-seq to profile TOP2β activity across the genome, we specifically recovered TOP2β enzymes in the process of cutting double-stranded DNA, thereby associating functional relevance to each site of enrichment. Previous studies in primary mouse cortical neurons and mouse cerebellar neurons observed TOP2β enrichment by ChIP-seq at actively transcribed genes, promoters, across regions of open chromatin, and adjacent to genes that encode

ion channels and receptors (Madabhushi et al., 2015; Sano et al., 2008; Tiwari et al., 2012). To our knowledge, we are the first to report a genome-wide gene-length-associated enrichment of TOP2β activity in neurons. Globally, we find TOP2β is distributed throughout the gene body with TOP2β activity markedly higher in long genes relative to shorter genes. Interestingly, we also observe significant TOP2β enrichment just downstream of promoters and at intragenic enhancers. Future studies can assess how TOP2β activity is targeted specifically to longer genes, and what role it plays in facilitating the expression of these genes.

MeCP2 has been characterized as a repressor given its ability to bind methylated DNA and repress transcription via the recruitment of the NCoR-corepressor complex. Loss or disruption of MeCP2 has been shown to cause dysregulation of intragenic enhancers and up-regulation of long, highly-methylated genes which may contribute to Rett syndrome pathology (Boxer et al., 2020; Clemens et al., 2020). Our results provide evidence that TOP2β plays a role in MeCP2-mediated repression of long genes. By profiling TOP2β genome-wide in neurons that have low initial levels of MeCP2, we have identified TOP2β enrichment at the promoter associated regions and gene bodies of long, MeCP2-repressed genes (Gabel et al., 2015). This demonstrates that, like other long genes, MeCP2-repressed genes are preferential targets of TOP2β. Manipulating MeCP2 expression in cultured neurons to levels similar to adult wild-type (MeCP2 OE) or adult knockout (MeCP2 KD) alters the activity of TOP2β at MeCP2 regulated genes in a manner that is consistent with MeCP2 inhibiting TOP2β at these sites. Thus, our findings provide insight into the functional significance of the interaction between MeCP2 and TOP2β in the regulation of long genes. Because MeCP2 does not possess a catalytic domain, it is considered to function as a bridge between chromatin and the NCoR co-repressor complex (Lyst and Bird, 2015). One possible

model for this function is that MeCP2, through association with the NCoR complex, inhibits TOP2β activity to repress the expression of long genes. The NCoR complex contains a histone deacetylase (HDAC3) and may function to deacetylate TOP2β. Future studies can investigate whether post-translational modifications are altered on TOP2β in MeCP2 mutants.

Outside of a direct role in long gene expression, our findings, combined with evidence from recent studies, may also implicate MeCP2-TOP2β regulation as an important modulator of DNA damage of these genes in neurons. TOP2β cuts DNA to facilitate changes in DNA topology and if TOP2β activity is disrupted during this process, it can lead to an aborted cutting and ligation cycle that result in DNA damage in the form of a double-stranded break (DSB). Because long genes experience topological constraints that do not otherwise affect shorter genes, long genes may require higher levels of TOP2β activity to resolve these constraints. In the context of high TOP2β activity at long genes, MeCP2 may be an important modulator to down-regulate TOP2β activity within optimal levels. When MeCP2 is absent, TOP2β activity may be overactive, giving rise to additional errors and increased DSBs. Notably, previous studies have reported recurrent DSBs in long genes in neural progenitors (Wei et al., 2016) and detected increased levels of DSBs in neural progenitors from of *Mecp2* knockout mice (Alessio et al., 2018). Furthermore, a recent genetic suppressor screen identifying mutations in genes that ameliorate Rett-like phenotypes in mice, detected mutations that alter DNA damage response pathways. One suppressor mutation affects RBBP8, a protein that has been shown to play a role in regulating the repair of TOP2-induced DSBs (Enikanolaiye et al., 2020). This finding suggests that altering DSB pathways impacted by TOP2β dysfunction can modify Rett-like phenotypes. Thus, our findings of TOP2β regulation by MeCP2 may provide an underlying mechanism to explain these genetic findings.

Given this possibility, it will be of great interest to explore the presence of TOP2β-induced DSBs in long genes of wild-type and MeCP2 mutants.

Altogether, our study has identified a molecular interaction in neurons that regulates a subset of neuronally enriched long genes. Future studies can investigate the mechanism and functional importance of this interaction and may reveal potential avenues for future therapeutic targets for MeCP2-related disorders.

2.4 STAR Methods

2.4.1 Key Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-MeCP2	Millipore Sigma	Cat# 07-013; RRID:AB_2144004
Mouse monoclonal anti-MeCP2	Millipore Sigma	Cat# M7443; RRID:AB_477235
Mouse monoclonal anti-Topo I (H-5)	Santa Cruz	Cat# sc-271285; RRID:AB_10611597
Rabbit polyclonal anti-Topo II-beta (H-286) [secondary]	Santa Cruz	Cat# sc-13059; RRID:AB_2205866
Rabbit polyclonal anti-Topo II-beta [primary]	Bethyl Laboratories	Cat# A300-950A; RRID:AB_805860
Rabbit polyclonal anti-Myc	Abcam	Cat# ab9106; RRID:AB_307014

Goat polyclonal anti-Myc	Abcam	Cat# ab9132; RRID:AB_307033
Mouse monoclonal anti-FLAG	Millipore Sigma	Cat# F3165; RRID:AB_259529
Rabbit monoclonal anti-alpha-Tubulin (EP1332Y)	Abcam	Cat# ab52866; RRID:AB_869989
Rabbit polyclonal anti-Histone H3 (acetyl K27)	Abcam	Cat# ab4729; RRID:AB_2118291
Anti-mouse IgG, HRP-linked	Cell Signaling Technologies	Cat# 7076; RRID:AB_330924
Anti-Rabbit IgG, HRP-linked	Cell Signaling Technologies	Cat# 7074; RRID:AB_2099233
Donkey-anti-Goat IgG, HRP	Thermo Fisher	Cat# A16005; RRID:AB_2534679
IRDye 800CW Donkey anti-Mouse	LI-COR	Cat# 926-32212; RRID:AB_621847
IRDye 800CW Donkey anti-Rabbit	LI-COR	Cat# 926-32213; RRID:AB_621848
IRDye 800CW Donkey anti-Goat	LI-COR	Cat# 926-32214; RRID:AB_621846
Bacterial and Virus Strains		
One Shot TOP10 Chemically Competent E. Coli	Thermo Fisher	Cat# C404010
One Shot Stbl3 Chemically Competent E. Coli	Thermo Fisher	Cat# C737303
Chemicals, Peptides, and Recombinant Proteins		
Etoposide (VP-16)	Santa Cruz	Cat# sc-3512
Benzonase	Millipore Sigma	Cat# 71206

Dynabeads Protein A	Thermo Fisher	Cat# 10001D
Dynabeads Protein G	Thermo Fisher	Cat# 10003D
AMPure XP beads	Beckman Coulter	Cat# A63880
Critical Commercial Assays		
NotI (High Fidelity)	NEB	Cat# R3189
XhoI	NEB	Cat# R0146
Gibson Assembly	NEB	Cat# E5510
Q5 Hot Start High-Fidelity 2x Master Mix	NEB	Cat# M0494
RNeasy Mini Kit	QIAGEN	Cat# 74104
Accel-NGS 2S Plus DNA Library Kit (24 rxns)	Swift Biosciences	Ca# 21024
2S Indexing Kit (12 indices, Set A)	Swift Biosciences	Cat# 26148
2S Indexing Kit (12 indices, Set B)	Swift Biosciences	Cat# 26248
QuBit dsDNA HS Assay Kit	Thermo Fisher	Cat# Q32854
NEBNext Ultra Directional RNA Library Prep Kit for Illumina	NEB	Cat# E7420S
NEBNext Mulitplex Oligos for Illumina (Index Primers Set 1)	NEB	Cat# E7335S
NEBNext rRNA Depletion Kit (Human/Mouse/Rat)	NEB	Cat# E6310L
Ovation Ultralow Library System V2	NuGEN	Cat# 0344-32
Oligonucleotides		

qPCR primers for gene expression (Supplementary Table 5)		N/A
qPCR primers for eTIP & ChIP (Supplementary Table 5)		N/A
Recombinant DNA		
pBACh-TOP2B-Myc (full length; WT)	This paper	N/A
pBACh-TOP2B-Myc (1-449 aa)	This paper	N/A
pBACh-TOP2B-Myc (450-1198 aa)	This paper	N/A
pBACh-TOP2B-Myc (1199-1612 aa)	This paper	N/A
p3xFLAG-CMV10-MeCP2 (full length; WT)	(Lyst et al., 2013)	N/A
p3xFLAG-CMV10-MeCP2 (1-484 aa)	(Lyst et al., 2013)	N/A
p3xFLAG-CMV10-MeCP2 (1-167 aa)	(Lyst et al., 2013)	N/A
p3xFLAG-CMV10-MeCP2 (143-484 aa)	(Lyst et al., 2013)	N/A
p3xFLAG-CMV10-MeCP2 (308-484 aa)	(Lyst et al., 2013)	N/A
U6-shRNA-Ubiquitin-MeCP2-IRES-GFP-FUGW	(Zhou et al., 2006)	N/A
U6-MeCP2-shRNA-IRES-GFP-FUGW	(Zhou et al., 2006)	N/A
U6-non-targeting-shRNA-IRES-GFP-FUGW	(Zhou et al., 2006)	N/A
Ser312fs11x-FLAG-Dnmt3a-IRES-FUGW	(Christian et al., 2020)	N/A

Tyr324Phe-Cre-IRES-GFP-FUGW	This paper	N/A
Deposited data		
Raw and analyzed data	This paper	GEO: GSE201658
Mus musculus mm9 genome assembly	UCSC	http://hgdownload.soe.ucsc.edu/ goldenPath/mm9/
Ensembl gene models	UCSC	https://genome.ucsc.edu/cgi-bin/hgTables
Experimental models: organisms/strains		
C57BL/6J	The Jackson Laboratory	JAX:100012
Experimental models: cell lines		
Primary mouse cortical neurons	This paper	N/A
HEK293T	ATCC	N/A
Software and Algorithms		
DESeq2 (v1.14.1)	(Love et al., 2014)	http://www.bioconductor.org/packages/release/ bioc/html/DESeq2.html
edgeR (v3.16.5)	(Robinson et al., 2009)	https://bioconductor.org/packages/release/bioc/ html/edgeR.html
SAMtools (v1.3)	(Li and Durbin, 2009)	https://sourceforge.net/projects/samtools/files/
BEDtools2 (v2.25.0)	(Quinlan and Hall, 2010)	https://github.com/arq5x/bedtools2

Bowtie2 (v2.2.5)	(Langmead and Salzberg, 2012)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
STAR	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
fastQC		https://www.bioinformatics.babraham.ac.uk/ projects/fastqc/
MACS2 (v2.1.0)	(Zhang et al., 2008)	https://github.com/taoliu/MACS
Trim galore		https://www.bioinformatics.babraham.ac.uk/ projects/trim_galore/

Lead Contact and Materials Availability

Requests for reagents and resources should be directed toward the Lead Contact, Harrison Gabel (gabelh@wustl.edu).

2.5 Experimental Model and Subject Details

Mouse cortical cultures

Cortical neurons were cultured from C57BL/6J E14.5 mouse embryos as described in (King et al., 2013), with some modifications. Embryonic day 14.5 (E14.5) mouse cortices were dissected in 1X DPBS, dissociated and trypsinized with TrypLE express for two 6 min incubations at 37^{o}C, followed by DNAse treatment, to remove free-floating DNA and digest DNA from dead cells. Trituration of cells was performed with pipette to dissociate cells fully. Dissociated neurons were seeded onto 6-well plates pre-coated with poly-D-lysine (0.1mg/mL) at a density of 7.5x10^{5} cells per well (or at a

density of $3x10^5$ cells per well for 12-well plates). The plates were pre-coated with poly-D-lysine (0.1mg/mL) in water, washed three times with water and washed once with Neurobasal medium before use. Neurons were cultured with neurobasal medium with 5% fetal bovine serum, GlutaMAX, B27 Supplement and Antibiotic-Antimycotic and maintained at 37ºC with 5% CO_2. Neurons were grown *in vitro* for 3 days. At DIV3 and DIV9, cells were fed with one volume of neurobasal medium supplemented with 4.84 μg/mL uridine 5'-triphosphate, 2.46 μg/mL 5-fluoro-2'-deoxyuridine, GlutaMAX, B27 Supplement and Antibiotic-Antimycotic

Method Details

Mass Spectrometry

Immunoprecipitation followed by mass spectrometry (IP-MS) was performed as described in (Mejia et al., 2013; Sowa et al., 2009), with modifications. Mouse DIV5 primary cortical neurons were infected with lentiviruses encoding FLAG-tagged baits under the CMV or neuronal Synapsin1 promoter. Lysates of DIV10 cortical neurons were subjected to immunoprecipitation using FLAG resin (Sigma), followed by 3XFLAG peptide elution (Sigma) and trichloroacetic acid (TCA) precipitation. Proteins were trypsinized (Sequencing-Grade Trypsin, Promega) and washed (3M Empore C18 media), and tryptic peptides were loaded onto an LTQ linear ion trap mass spectrometer (ThermoFinnigan). Spectra were searched against target-decoy tryptic peptide databases by CompPASS analysis.

Plasmids

The mouse Top2b cDNA with a C-terminal MYC tag was amplified from MGC Mouse Top2b cDNA (Dharmacon) and cloned into pBACh, a modified pCAG (cytomegalovirus enhancer fused to chicken

beta-actin promoter) vector, using restriction enzyme cloning at XhoI (5') and NotI (3'). The modified pCAG vector was generated from a pCAG-mCherry, pIRES2-EGFP (Clontech) in which the IRES and EGFP were replaced by mCherry (a gift from Dr. Jason Yi). The truncated TOP2β constructs were generated by the insertion of PCR products into pBACh empty vector using Gibson Assembly cloning (NEB). All constructs were given consensus translation initiation sequences. Human full-length MeCP2 and MeCP2 fragments were cloned into p3xFLAG-CMV (a gift from Dr. Adrian Bird) and were the same as used previously (Lyst et al., 2013). Mouse full-length Dnmt3a was the same as used previously (Christian et al., 2020). The MeCP2 shRNA and MeCP2 overexpression constructs were the same as used previously (Zhou et al., 2006).

Endogenous co-immunoprecipitation

Endogenous co-immunoprecipitations were carried out as described previously (Ebert et al., 2013). Forebrains from 8-week-old C57B/J mice were isolated and lysed in NP-40 lysis buffer (10 mM HEPES, pH 7.9, 3 mM MgCl2, 10 mM KCl, 10 mM NaF, 1 mM Na3VO4, 0.5 mM DTT, 0.5% NP-40, 13 complete EDTA-free protease inhibitor cocktail (Roche)), dounced 15 times with a tight pestle, and pelleted at 1,000 g. Lysates were diluted 1:1 with benzonase buffer (10 mM HEPES, pH 7.9, 3 mM MgCl2, 280 mM NaCl, 0.2 mM EDTA, 10 mM NaF, 1 mM Na3VO4, 0.5 mM DTT, 0.5% NP-40, and 13 complete EDTA-free protease inhibitor cocktail (Roche)) and digested with 250 units benzonase (Millipore) for 1h rotating at 4ºC to release MeCP2 and its protein binding partners from the genome. Digested lysates were pelleted at 17,000g for 20 min at 4ºC and immunoprecipitation was carried out on the supernatant using the following antibodies: MeCP2 (07-013, Millipore) or TOP2β (H-286, sc-13059, Santa Cruz), in the presence of 150mM NaCl for 2hrs while rotating at 4ºC. The peptide-block control lysate was immunoprecipitated with MeCP2 antibody in the presence of a peptide to which the antibody was raised.

In vitro co-immunoprecipitation

HEK293T cells were transfected with TOP2β, MeCP2 or DNMT3A constructs using Lipofectamine 2000 (Invitrogen) and harvested after 24-48 hours. HEK293T cells were lysed in NE10 buffer (20 mM HEPES (pH 7.5), 10 mM KCl, 1 mM MgCl2, 0.1% Triton X-100 (v/v), protease inhibitors (Roche), 15 mM β-mercaptoethanol), dounced 15 times and pelleted 5min at 500 g. Nuclei were washed in NE10 buffer and then digested with 250 units benzonase (Millipore) for 30 min rotating at 25ºC. Nuclei were resuspended in NE150 buffer (NE10 supplemented with 150mM NaCl) and incubated for 20 min. Lysates were pelleted at 16,000 g for 20 min at 4ºC and supernatants were immunoprecipitated by incubating the Myc tag (ab9106, Abcam) antibody with Dynabeads Protein for 1hr at 4ºC. The IP fraction was recovered by magnetic separation followed by three washes with NE10 buffer containing 150mM-300mM NaCl. The IP was then eluted from the beads with 2X NuPage LDS buffer (Invitrogen) containing β-mercaptoethanol.

Immunoblotting

Nuclear extracts or IP eluates from brain tissues or HEK293T cells were resolved on 5-12% Tris-Glycine gels and transferred to nitrocellulose. Membranes were incubated overnight in the following primary antibodies: MeCP2 (Men-8, M7443, Sigma), TOP2β (H-286, sc-13059, Santa Cruz), TOP1 (H-5, sc-271285, Santa Cruz), Myc tag (ab9132, Abcam), Flag tag (F3165, Sigma), α-Tubulin (EP1332Y, ab52866, Abcam). Following washes, membranes were incubated with secondary antibodies conjugated to IRdye 800 and imaged with LiCOR Odyssey.

Virus production

For lentiviral-mediated overexpression and shRNA knockdown, virus was prepared as described in (Tiscornia et al., 2006) using the MeCP2 shRNA and MeCP2 overexpression plasmids previously

validated in (Zhou et al., 2006). To produce lentivirus, 10 μg of lentiviral plasmid (either shRNA-expressing or MeCP2 overexpression) was transfected into HEK293T cells along with third generating packaging plasmids pMDL (5 μg), RSV (2.5 μg), and VSVG (2.5 μg). HEK293T cells were maintained in complete DMEM media (DMEM (high glucose) media, 10% fetal bovine serum, 1% GlutaMAX, 1% penicillin/streptomycin). At 12-16 hrs following transfection, media was replaced with fresh complete DMEM media. Viruses were concentrated by ultracentrifugation 48-60 hrs after transfection and viral titers were determined by infection of HEK293T cells and were typically 0.5-1x10^5 IFU/μL. Cultured neurons were infected on DIV3 and harvested on DIV13 for eTIP-seq, ChIP-seq or RNA-seq. In some experiments (e.g. RNA-seq of unaltered neurons) cells were transduced with control virus to control for viral infection.

Etoposide-mediated Topoisomerase Immunoprecipitation (eTIP) protocol

eTIP experiments were performed as described in (Sano et al., 2008), with some modifications. Primary neuronal cultures from mouse cerebral cortex at 12 days in vitro (DIV12) were treated with 0.5mM etoposide (VP-16) or DMSO control in serum-free medium for 15min. The 2.1x10^6 treated cells were lysed with 300μL of TE buffer (10mM Tris and 1mM EDTA, pH 8.0) and 1% SDS. Before sonication, 3 volumes of a buffer containing TE buffer and protease inhibitor (Complete Mini Roche) was added to the lysates. 5% of the lysates was saved for protein evaluation. To fragment DNA, the lysates were sonicated with Covaris E220 sonicator (5% Duty Factory, 140 Peak Incidence Power, 200 cycles per burst, milliTUBE 1mL AFA Fiber). Under these conditions, length of DNA fragments, as determined by TapeStation, varied from 0.5kb-1.2kb. Before immunoprecipitation (IP), 3 volumes of a buffer containing 20mM Tris-HCl (pH 8.0), 3% Triton X-100, 450mM NaCl, 3mM EDTA and protease inhibitor mixture was added to 1 volume of lysates.

The immunoprecipitation was carried out as described in (Kim et al., 2010), with some modifications. The diluted lysates were pre-cleared with 15μL of Dynabeads Protein A by rotating the tubes for 2hrs @ 4ºC. After pre-clear, 3% of the lysate was saved and used as input for the IP reaction. The unbound fraction was recovered by magnetic separation. The reaction was initiated by the addition of 15μL of Dynabeads Protein A, which was pre-incubated with 2 μg of a TOP2β specific antibody. The beads suspension was incubated overnight at 4ºC and the IP fraction was recovered by magnetic separation followed by two washes with low salt buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl (pH 8.0), 150mM NaCl), two washes with high salt buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris- HCl (pH 8.0), 500mM NaCl), two washes with LiCl buffer (0.250mM LiCl, 1% NP40, 1% deoxycholic acid (sodium salt), 1mM EDTA, 10mM Tris (pH 8.0)) and one wash with TE buffer. The IP was then eluted from the beads twice by adding elution buffer containing TE buffer and 1% SDS and incubating the samples at 65ºC for 30min with brief vortexing every 10min. Elution buffer (1.5 volumes) was also added to the saved input material and this sample was processed together with the IP samples. Each eluate was treated with 10μg RNAse A and incubated for 1 hr at 37ºC and then with 20mg/mL Proteinase K and incubated for 2hrs for 55ºC. The IP DNA fragments were extracted with Phenol:Chloroform:Isoamyl alcohol (25:24:1, v/v). The resulting genomic DNA fragments were then purified using the QIAquick PCR purification kit (Qiagen) and DNA fragments were eluted from the columns twice.

Libraries were generated using ACCEL-NGS 2S PLUS DNA Library Kit (21024, Swift Biosciences), according to the manufacturer's instructions and PCR amplified for 12-16 cycles. Library quality was assessed using the Agilent 4200 TapeStation (Agilent Technologies). Libraries were pooled to a final concentration of 4-10nM and 50bp reads were generated on the Illumina HiSeq 3000 with the Genome Technology Access Center, or 75bp reads were generated on the Illumina

Nextseq 500 with the Center for Genome Science at Washington University in St. Louis, typically yielding 15-40 million single-end reads per sample.

Chromatin Immunoprecipitation protocol and library preparation (ChIP-seq)

ChIP experiments were performed on $2.1x10^6$ mouse cortical neurons cultured to *in vitro* day 12 as described in (Kim et al., 2010), with some modifications. To cross-link protein-DNA complexes, media was removed from primary neurons and cross-linking buffer (20 mM HEPES- NaOH, pH 8.0, 200mM NaCl, 2 mM EDTA, 2mM EGTA) containing 1% paraformaldehyde was added for 10 minutes at room temperature. Cross-linking was quenched by adding 125 mM glycine for 5 minutes at room temperature. Cells were then rinsed 2 times in ice-cold PBS containing PMSF protease inhibitor (36978, Thermo Fisher) and collected by scrapping. Cell were lysed and nuclei isolated by incubating in L1 buffer (100 mM HEPES- NaOH, pH 7.5, 280 mM NaCl, 2 mM EDTA, 2mM EGTA, 0.5% Triton X-100, 1% NP-40, 20% Glycerol, 10 mM sodium butyrate, protease inhibitor cocktail (Roche)) for 10 minutes at 4ºC. Nuclei were then pelleted by centrifugation at for 10 min at 4ºC. The isolated nuclei were resuspended in L2 buffer (100mM Tris-HCl, pH 8.0, 200mM NaCl, 10 mM sodium butyrate, protease inhibitor cocktail (Roche)) and re-pelleted. The isolated nuclei were resuspended in L3 buffer (20mM Tris-HCl, pH 8.0, 2 mM EDTA, 2mM EGTA, 10 mM sodium butyrate, protease inhibitor cocktail (Roche)). Nuclei were pelleted and either stored at -80ºC until use or immediately processed. Chromatin was sonicated using a Bioruptor (Diagenode) on high power mode for 50 cycles with 30 second pulses in sonication buffer (20mM Tris-HCl, pH 8.0, 2 mM EDTA, 2mM EGTA, 10 mM sodium butyrate, 0.1% Na-Deoxycholate, 0.5% SDS, protease inhibitor cocktail (Roche)). Following sonication, the immunoprecipitation was carried out as described for eTIP and the chromatin was incubated overnight at 4ºC with H3K27ac (0.025-0.1μg,

Abcam 4729) and Protein A Dynabeads. H3K27ac ChIPs were performed twice from independent neuronal cultures. Following the overnight incubation, the IP was washed and eluted as described for eTIP.

Libraries were generated using Ovation Ultralow Library System V2 (Tecan, 0344NB-32), according to the manufacturer's instructions and PCR amplified for 12-16 cycles. Library quality was assessed using the Agilent 4200 TapeStation (Agilent Technologies). Libraries were pooled to a final concentration of 4-10nM and 75bp reads were generated on the Illumina Nextseq 500 with the Center for Genome Science at Washington University in St. Louis, typically yielding 15-40 million single-end reads per sample.

Total RNA Isolation and library preparation (RNA-seq)

Neuronal cultures dissected and transduced (with MeCP2 KD, MeCP2 OE or control virus) on the same days, constituted technical replicates, whereas, Neuronal cultures dissected and transduced on independent days, constituted biological replicates. Cultured neurons were harvested directly using RLT buffer and homogenized in the QIAshredder spin column (Qiagen). To isolate RNA, the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. RNA libraries were generated from 250ng total RNA with rRNA depletion (NEBNext, E6310) and NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEBNext, E7420), using a modified amplification protocol (37^{o}C, 15 minutes; 98^{o}C, 30 s; (98^{o}C, 10 seconds; 65^{o}C, 30 seconds; 72^{o}C, 30 seconds) x 13 cycles; 72^{o}C, 5 minutes; 4^{o}C, hold). RNA libraries were pooled at a final concentration of 8-10nM and single end 75bp reads were generated on the Illumina Nextseq 500 with the Center for Genome Science at Washington University in St. Louis, typically yielding 20-40 million single-end reads per sample.

Quantification and Statistical Analysis

Etoposide-mediated Topoisomerase Immunoprecipitation sequencing analysis

eTIP-seq was analyzed as previously described for ChIP-seq data in (Clemens et al., 2020). Sequenced reads were mapped to the mm9 genome using bowtie2 alignment and reads were extended based on library sizes and deduplicated to consolidate PCR duplicate reads. Deduplicated reads were used to quantify read density normalized by the number of reads per sample and by read length in basepairs. Bedtools coverage -counts parameter was used to quantify eTIP signal and ChIP signal.

For eTIP, the signal was quantified at promoter associated region, defined as 1kb downstream to 3kb downstream of the TSS, and the gene body, defined as 3kb downstream of the TSS to the end of the transcript, based on our Ensembl gene models. edgeR was then used to determine differential eTIP-signal across conditions. Data were visualized using UCSC genome browser (http://genome.ucsc.edu).

Aggregate plots of eTIP signal at genes and enhancers were generated by calculating eTIP/Input for equally sized bins for the specified windows using Bedtools coverage -hist parameter and custom R and python scripts. For aggregate analysis at genes, the genes less than 3kb in length were filtered out in order to capture the signal at the promoter associated regions without conflating the ends of genes with this region. In some instances, the genes are further filtered such that the lengths of the genes plotted is equal to or greater than the aggregate length being plotted (Clemens et al., 2020) .

Resampling control analysis was performed with custom scripts, by moving true TOP2β peak randomly around the genome with the criteria that the shuffled peaks do not overlap each other however could at random overlap a true TOP2β peak.

Chromatin Immunoprecipitation sequencing analysis

ChIP-seq analysis was performed as previously described (Clemens et al., 2020). Sequenced reads were mapped to the mm9 genome using bowtie2 alignment and reads were extended based on library sizes and deduplicated to consolidate PCR duplicate reads. Deduplicated reads were used to quantify read density normalized by the number of reads per sample and by read length in basepairs. Bedtools coverage -counts parameter was used to quantify ChIP signal.

For ChIP, the signal was quantified at the TSS and enhancers. Enhancers in this study were defined by requiring the presence of H3K27ac peaks that occur outside of a known TSS region (TSS +/- 500bp). H3K27ac peaks were identified using MACS2 peak calling algorithm, in which the ChIP input was used as background signal, using the following parameters: macs2 call peak --nomodal -q 0.05. Bedtools intersect was used to identify H3K27ac peaks that did not overlap with gene promoter regions. These filtered H3K27ac peaks were defined as enhancers. As noted, this may have led to the exclusion of some subthreshold regions of H3K27ac enrichment that may represent true regulatory elements.

RNA sequencing analysis

RNA sequencing analysis was performed as previously described in (Clemens et al., 2020). Raw FASTQ files were trimmed with Trim Galore, using a quality filter of 20, followed by filtering out rRNA sequences using Bowtie2. Remaining reads were aligned to mm9 using STAR (Dobin et al.,

2013) with default parameters. Reads mapping to multiple regions in the genome were then filtered out, and uniquely mapped reads were converted to BED files and separated into intronic and exonic reads. Finally, reads were assigned to genes using bedtools coverage -counts parameter (Quinlan and Hall, 2010).

For gene annotations, we defined a "flattened" list of the longest transcript for each gene, generated on Ensgene annotations and obtained from the UCSC table browser. For each gene, Ensembl IDs were matched up to the MGI gene names. Then, for each unique MGI gene name, the most upstream Ensgene TSS and the most downstream TES were taken as the gene's start and stop. Based on these Ensembl gene models, we defined TSS regions, TSS-adjacent regions and gene bodies. DESeq2 was run using default parameters on exonic reads from MeCP2 KD and MeCP2 OE cultured neurons (n = 3 per condition), to identify differentially expressed genes.

Data and Software Availability

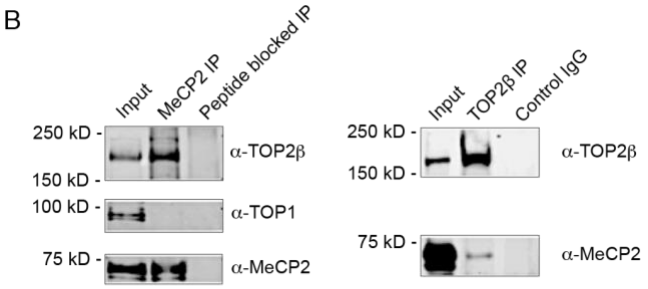
The accession number for the raw and processed data in this paper is Gene Expression Omnibus, GEO: GSE201658

2.6 Figures

A

Protein	IP1	IP2
MECP2	48	82
MTAP2	65	60
DHX30	10	11
JAKMIP1	15	6
CAMSAP3	6	10
TOP2B	**6**	**8**
SFRS14	6	6
SBF1	11	14
KPNA3	5	6
PSIP1	4	6

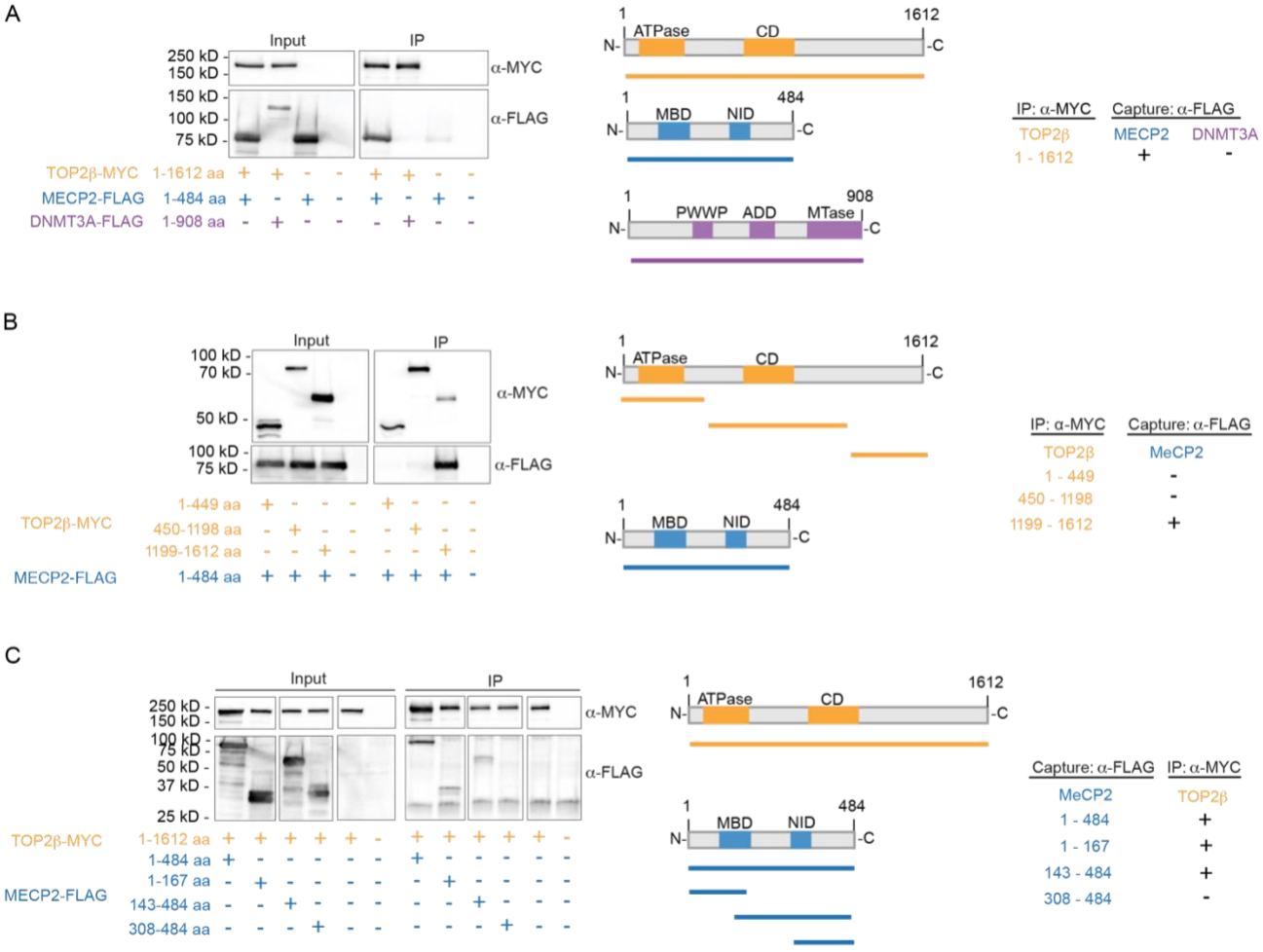
B



2.6.1 Figure 1. MeCP2 interacts with TOP2β in neurons

A. MeCP2 was isolated by FLAG immunoprecipitation from lysates of mouse cortical neurons infected with lentivirus expressing FLAG-MeCP2 from the CMV (IP1) or Synapsin1 (IP2) promoter. Interacting proteins were detected by liquid chromatography with tandem mass spectrometry (LC/MS/MS) and quantified by CompPASS analysis. The ten proteins with the highest average weighted D scores are shown for proteins detected in the two LC/MS/MS analyses (see Table S1). Total spectral count for each protein is shown for each analysis.

B. Left, immunoprecipitation (IP) of MeCP2 from nuclear extracts from the whole cortex of wild-type mice, using an antibody against MeCP2 shows co-IP of TOP2β but not TOP1. Peptide blocked IP indicates IP performed with antibody against MeCP2 that was pre-incubated with the peptide to which the antibody was raised. MeCP2-bound endogenous proteins were assayed on western blots. Right, IP of TOP2β with co-IP of MeCP2 from neuronal nuclear extracts from whole cortex of wild-type mice. Control IgG indicates IP performed with an antibody against another protein, CtBP.



2.6.2 Figure 2. Identification of protein regions sufficient for the MeCP2-TOP2β interaction

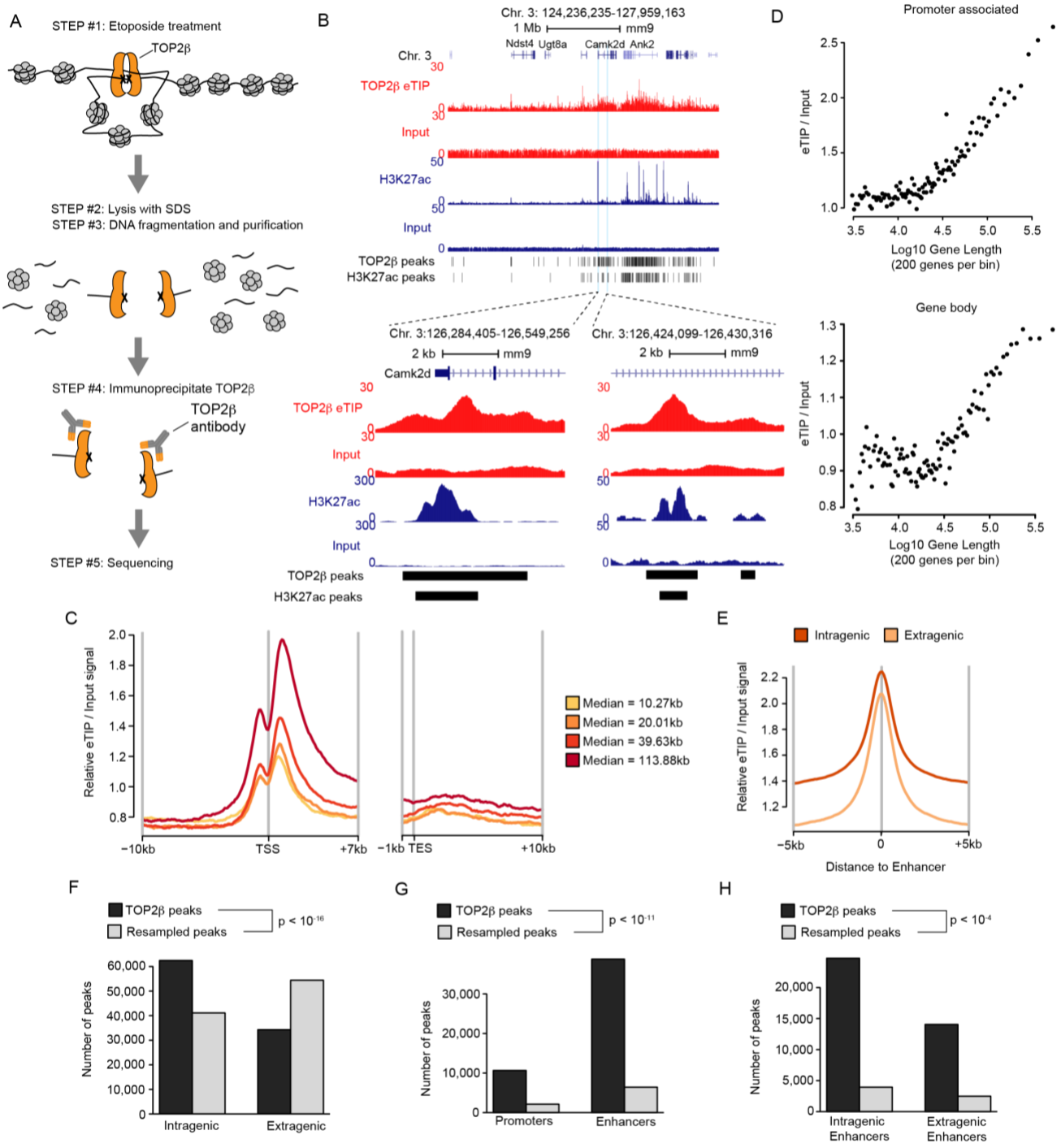
A. Co-immunoprecipitation analysis of MYC-tagged full-length TOP2β, FLAG-tagged full-length MeCP2 and FLAG-tagged full-length DNMT3A in HEK293T cells. Immunoprecipitation of full-length TOP2β binds full-length MeCP2 but failed to bind another nuclear protein, DNMT3A. Left, cell lysates were immunoprecipitated with an antibody against MYC, and co-precipitated FLAG was visualized by western blots. Center, overview of TOP2β (orange), MeCP2 (blue) and DNMT3A (purple) domain structure. Right, summary of results. Domains are annotated as follows: CD, catalytic domain containing the

active tyrosine; MBD, methyl-DNA-binding domain; NID, NCoR-interaction domain; PWWP, proline-tryptophan-tryptophan-proline domain; ADD, auto-inhibitory ATRX-DNMT3A-DNMT3L domain; MTase, methyltransferase domain.

B. Co-immunoprecipitation analysis of MYC-tagged fragments of TOP2β with full-length FLAG-tagged MeCP2 in HEK293T cells. Full-length MeCP2 interacts with the C-terminal region (1199-1612) of TOP2β and, but not with the N-terminal region (1-499) or the central region (450-1198). Left, cell lysates were immunoprecipitated with an antibody against MYC and co-precipitated FLAG was visualized by western blot. Lysates from untransfected cells were run in the last lane of the blot as a control. Center, overview of TOP2β (orange) and MeCP2 (blue) domain structure. Right, summary of results.

C. Co-immunoprecipitation analysis of FLAG-tagged fragments of MeCP2 with MYC-tagged full-length TOP2β. IP of full-length TOP2β recovers full-length (1-484) MeCP2, a fragment spanning the majority of the MBD (1-167), and a fragment spanning part of the MBD and the NID (143-484) but fails to detect a fragment containing the C-terminal region (308-484) of MeCP2. Left, cell lysates were immunoprecipitated with an antibody against MYC, and co-precipitated FLAG was visualized by western blot. Lysates from untransfected cells were run in the last lane of the blot as a control. Center, overview of TOP2β (orange) and MeCP2 (blue) domain structure. Blue lines show deletion fragments of MeCP2. Right, summary of results.

All blots in this figure are representative of at least two biological replicate experiments HEK293T cell transfections.


A
STEP #1: Etoposide treatment
TOP2β
STEP #2: Lysis with SDS
STEP #3: DNA fragmentation and purification
STEP #4: Immunoprecipitate TOP2β
TOP2β antibody
STEP #5: Sequencing
B
Chr. 3: 124,236,235-127,959,163
1 Mb
mm9
Ndst4 Ugt8a Camk2d Ank2
Chr. 3
TOP2β eTIP
Input
H3K27ac
Input
TOP2β peaks
H3K27ac peaks
Chr. 3:126,284,405-126,549,256
Chr. 3:126,424,099-126,430,316
2 kb
Camk2d
D
Promoter associated
eTIP / Input
Log10 Gene Length
(200 genes per bin)
Gene body
C
Relative eTIP / Input signal
−10kb
TSS
+7kb
−1kb TES
+10kb
Median = 10.27kb
Median = 20.01kb
Median = 39.63kb
Median = 113.88kb
E
Intragenic
Extragenic
Relative eTIP / Input signal
−5kb
0
+5kb
Distance to Enhancer
F
TOP2β peaks
Resampled peaks
p < 10-16
Number of peaks
Intragenic
Extragenic
G
p < 10-11
Promoters
Enhancers
H
p < 10-4
Intragenic Enhancers
Extragenic Enhancers

2.6.3 Figure 3. TOP2β is preferentially active at long genes in neurons

A. Schematic representation of the eTIP-seq procedure to directly map sites of TOP2β activity, adapted from (Sano et al., 2008). TOP2β cleaves a DNA duplex to make a gate for a second DNA duplex to pass through. In the eTIP assay, etoposide treatment freezes TOP2β in its cleavage complex, leaving it covalently linked to its site of action. Following denaturing and fragmentation under stringent conditions containing SDS, TOP2β-linked DNA is immunoprecipitated, purified and sequenced.

B. Genome browser snapshots of eTIP-seq and H3K27ac ChIP-seq from DIV12 cultured cortical neurons. Top, an ~2Mb region of the genome displaying the profile for TOP2β eTIP, H3K27ac and their respective inputs. Enrichment of eTIP-seq signal is more prominent at long genes (e.g. *Camk2d*) compared to shorter genes (e.g. *Ugt8a*). Bottom left, zoomed in view of the promoter region of the *Camk2d* gene. Bottom right, zoomed in view of an enhancer region of the *Camk2d* gene. Peaks of TOP2β eTIP and H3K27ac called by the MACS2 algorithm are indicated below the tracks. Gene annotations and scales are depicted above.

C. Aggregate plot of input-normalized eTIP signals at genes divided into quantiles of gene length (see methods). The median gene length for each group is indicated. Average signal around the transcription start site (TSS) and transcription end site (TES) is shown. Mean values plotted for 100bp bins.

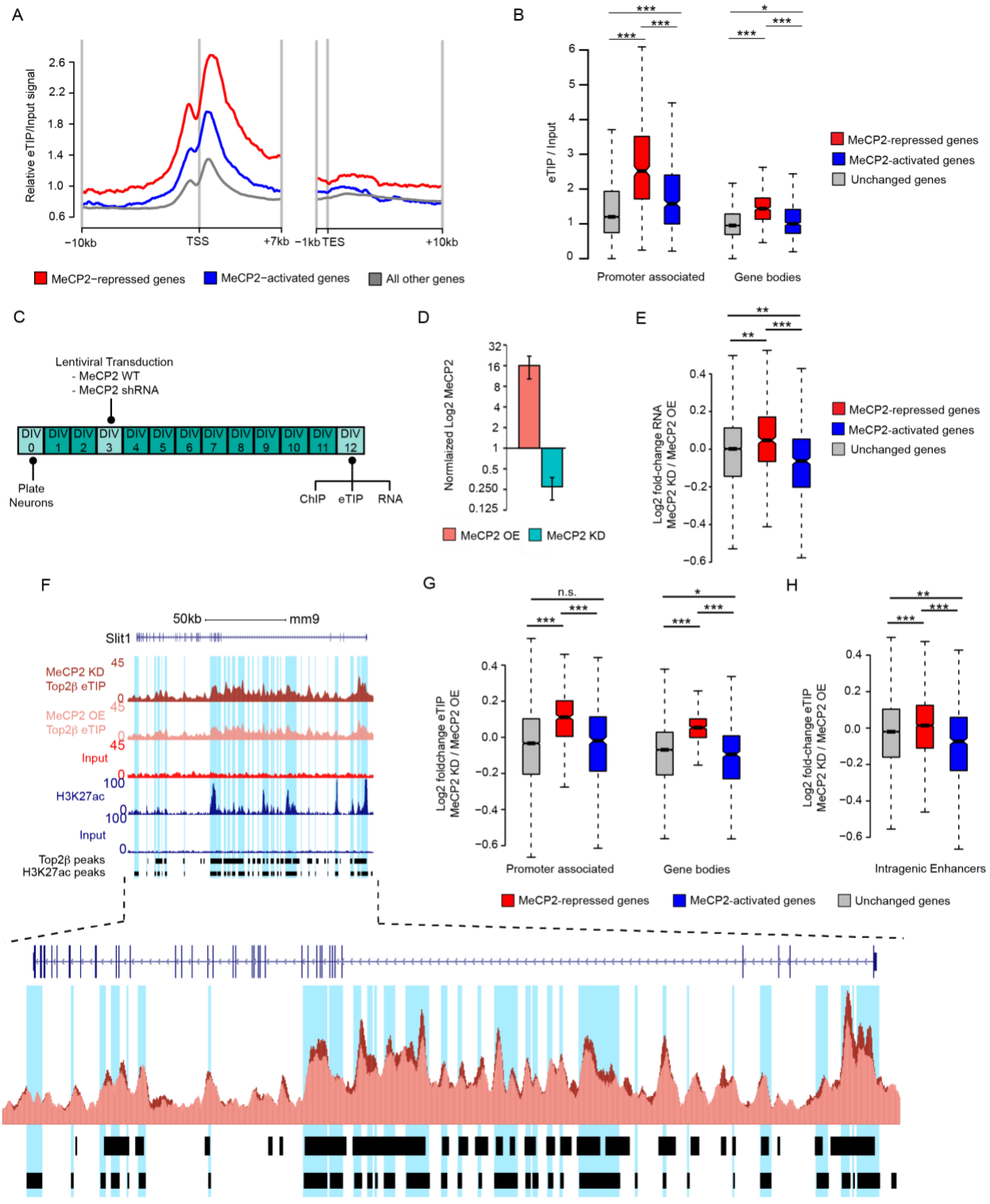
D. Running average plots of input-normalized eTIP signal at the promoter associated region (top) and gene bodies (bottom), from eTIP-seq. Averages are shown for bins of 200 genes ranked according to gene length.

E. Aggregate plot of input-normalized eTIP signal at putative intragenic and extragenic enhancers genome-wide in cultured cortical neurons. Putative enhancers were defined by H3K27ac peaks that did not overlap the TSS of a gene. The plot is centered at the midpoint for each enhancer and the surrounding 10kb region. Mean values plotted for 100bp bins.

F. Bar plots of the genomic distribution of TOP2β peaks and resampled TOP2β peaks at intragenic and extragenic regions. Resampled peaks were generated by shuffling TOP2β peak-sized regions randomly around the genome. A Chi-squared test was conducted between the frequencies of TOP2β peaks and resampled TOP2β peaks and resulting p-values are indicated.

G. Bar plots of the genomic distribution of TOP2β peaks and resampled TOP2β peaks at promoters and enhancers. Resampling and statistical analysis were performed as described in panel F.

H. Bar plots of the genomic distribution of TOP2β peaks and resampled TOP2β peaks at intragenic enhancers and extragenic enhancers. Resampling and statistical analysis were performed as described in panel F.


A
Relative eTIP/Input signal
2.6
2.2
1.8
1.4
1.0
0.6
−10kb
TSS
+7kb
−1kb
TES
+10kb
MeCP2−repressed genes
MeCP2−activated genes
All other genes
B
eTIP / Input
Promoter associated
Gene bodies
MeCP2-repressed genes
MeCP2-activated genes
Unchanged genes
C
Lentiviral Transduction
- MeCP2 WT
- MeCP2 shRNA
DIV 0
DIV 1
DIV 2
DIV 3
DIV 4
DIV 5
DIV 6
DIV 7
DIV 8
DIV 9
DIV 10
DIV 11
DIV 12
Plate Neurons
ChIP
eTIP
RNA
D
Normlaized Log2 MeCP2
MeCP2 OE
MeCP2 KD
E
Log2 fold-change RNA MeCP2 KD / MeCP2 OE
MeCP2-repressed genes
MeCP2-activated genes
Unchanged genes
F
50kb
mm9
Slit1
MeCP2 KD Top2β eTIP
MeCP2 OE Top2β eTIP
Input
H3K27ac
Input
Top2β peaks
H3K27ac peaks
G
Log2 foldchange eTIP MeCP2 KD / MeCP2 OE
Promoter associated
Gene bodies
H
Log2 fold-change eTIP MeCP2 KD / MeCP2 OE
Intragenic Enhancers
MeCP2-repressed genes
MeCP2-activated genes
Unchanged genes

2.6.4 Figure 4. Altering MeCP2 levels in neurons impacts TOP2β activity at long MeCP2-repressed genes

A. Aggregate plot of input-normalized eTIP signal in DIV12 primary culture cortical neurons at MeCP2-repressed, MeCP2-activated, and all other genes. Gene sets are those previously identified as consistently dysregulated across multiple MeCP2 mutant datasets and brain regions (Gabel et al., 2015).

B. Boxplot of input-normalized eTIP signal at the promoter associated regions (left) and gene bodies (right) of MeCP2-repressed, MeCP2-activated, and all other genes. *, $p < 0.01$; ***, $p < 10^{-15}$ Wilcoxon rank-sum test.

C. Schematic depicting experimental design to manipulate MeCP2 levels in primary culture cortical neurons. Embryonic neurons were isolated and cultured for 3 days before being transduced with exogenous wild-type MeCP2 (MeCP2 OE) or MeCP2 shRNA (MeCP2 KD) lentivirus. At DIV12, cells were harvested for downstream high-throughput genomic experiments (e.g. eTIP-seq, RNA-seq and ChIP-seq).

D. Relative MeCP2 RNA expression in DIV12 cortical primary neurons transduced with MeCP2 OE or MeCP2 KD lentivirus, as determined by RT-qPCR. n = 3 biological replicates per group. Graph shows *Actb* normalized mean ± SEM of MeCP2 expression relative to MeCP2 expression in control neurons in which MeCP2 expression was manipulated.

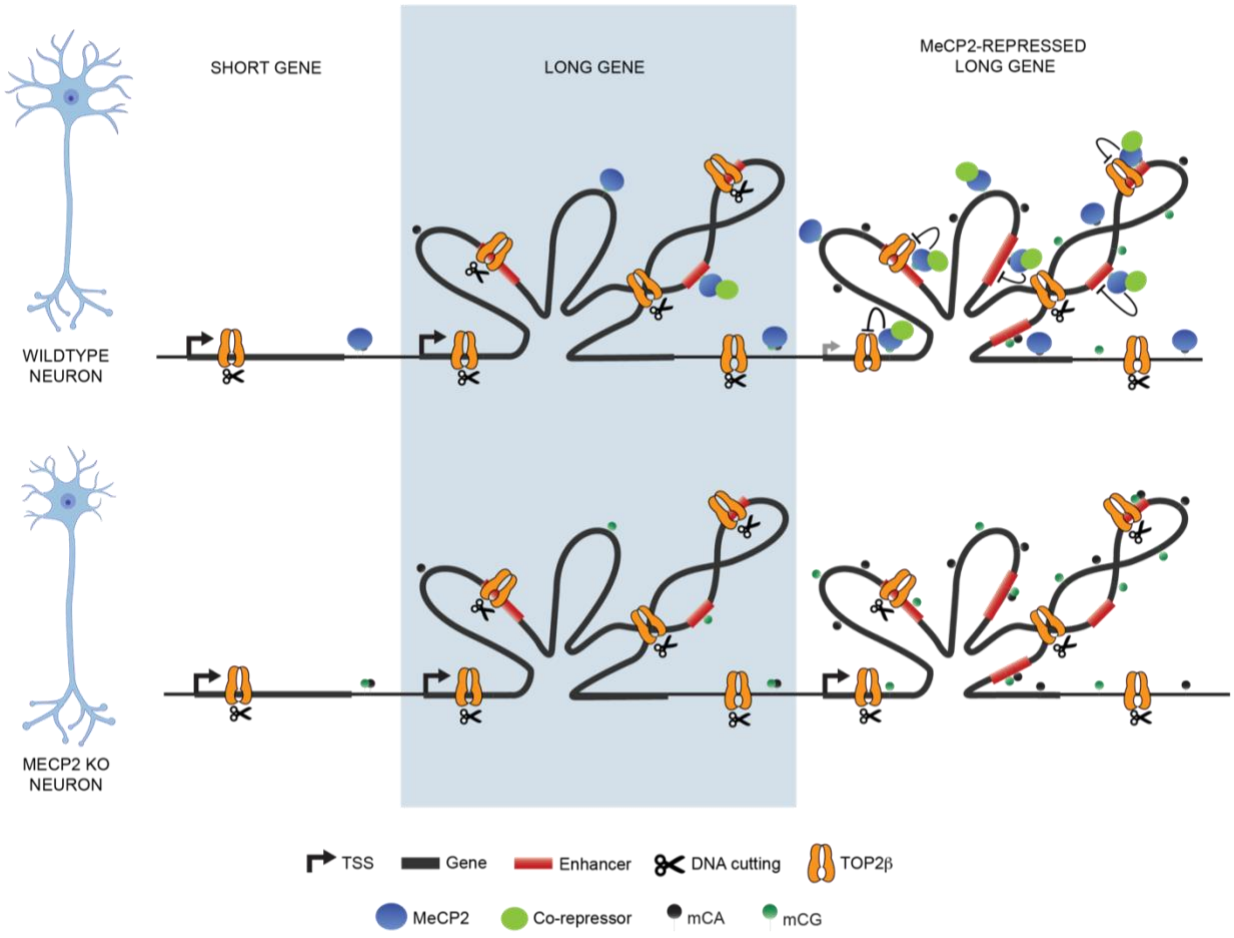
E. Boxplot of fold-changes in exonic RNA for MeCP2-repressed, MeCP2-actived and all other genes in DIV12 cortical neurons transduced with either MeCP2 OE or MeCP2 KD lentivirus. **, $p < 10^{-9}$; ***, $p < 10^{-16}$ Wilcoxon rank-sum test.

F. Top, genome browser snapshot of eTIP-seq signal at an example MeCP2-repressed gene, *Slit1*, in MeCP2 OE and MeCP2 KD DIV12 primary culture cortical neurons. Regions

corresponding to H3K27ac peaks (putative enhancers and promoters) are highlighted in blue. Bottom, superimposition of eTIP-seq tracks, illustrates higher signal at this gene in the MeCP2 KD condition compared to the MeCP2 OE condition.

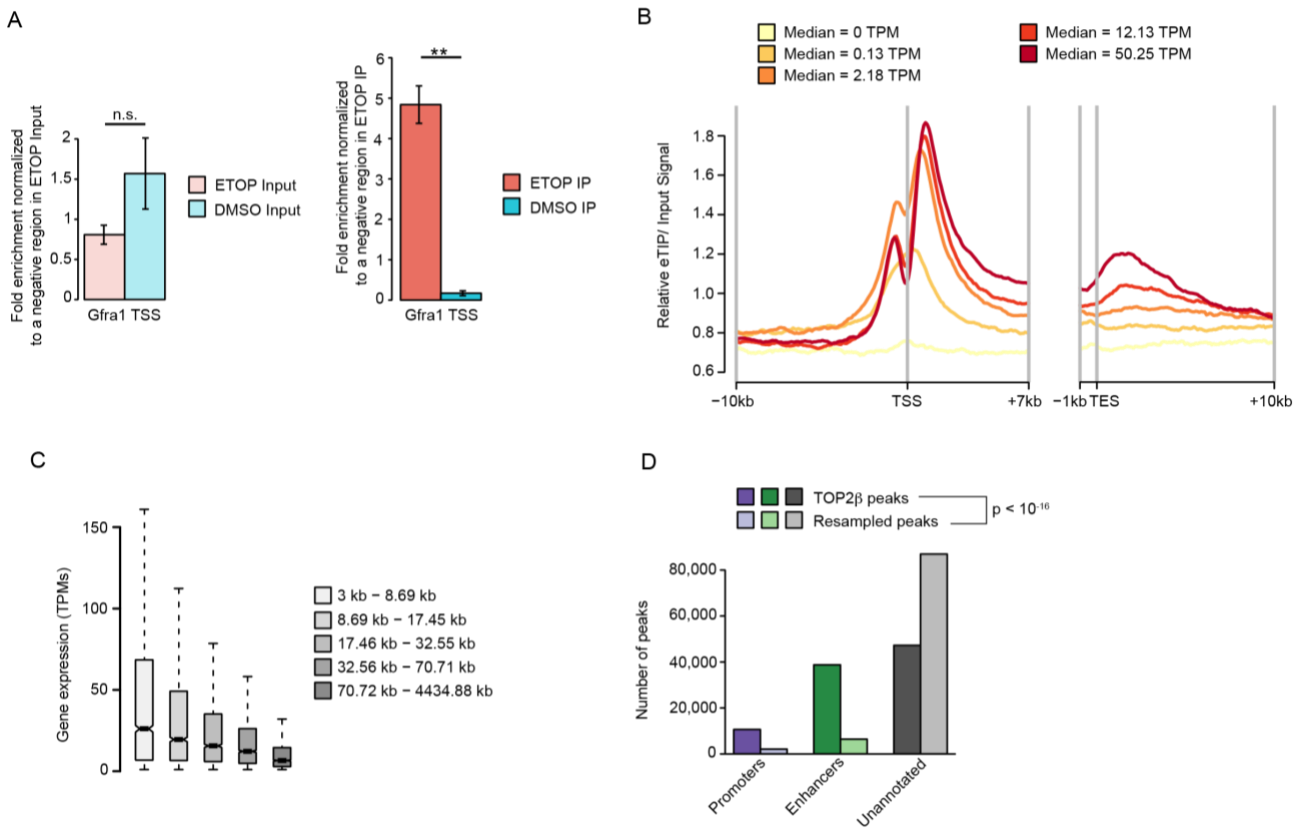
G. Boxplot of fold-changes in the promoter associated regions (left) and gene bodies (right) eTIP signal of MeCP2-repressed, MeCP2-actived and all other genes in DIV12 cortical neurons transduced with either MeCP2 OE or MeCP2 KD lentivirus. Fold-changes were calculated by edgeR analysis of eTIP signal at promoter associated regions and gene bodies. n.s., not significant; *, $p < 0.05$; ***, $p < 10^{-16}$ Wilcoxon rank-sum test.

H. Boxplot of fold-changes in eTIP signal in DIV12 cortical primary neurons transduced with either MeCP2 OE or MeCP2 KD lentivirus for enhancers located within MeCP2-repressed, MeCP2-actived and all other genes. Fold-changes were calculated by edgeR analysis of eTIP signal at enhancer regions. **, $p < 10^{-9}$; ***, $p < 10^{-16}$ Wilcoxon rank-sum test.



2.6.5 Figure 5. A model of repression of TOP2β activity in neurons by MeCP2

In neurons, TOP2β is preferentially recruited to promoter associated regions, gene bodies and enhancers of long genes, where it resolves topological constraints and facilitates gene expression. At a subset of long genes ("MeCP2-repressed genes"), DNA methylation (mCG and mCA dinucleotides) and MeCP2 binding are enriched. MeCP2 interacts with TOP2β, acting as a molecular break on activity of TOP2β to fine-tune the expression of these genes. When MeCP2 is absent (MeCP2 KO) the molecular break on TOP2β is removed, causing overactivity of TOP2β and overexpression of these long, MeCP2-repressed genes.



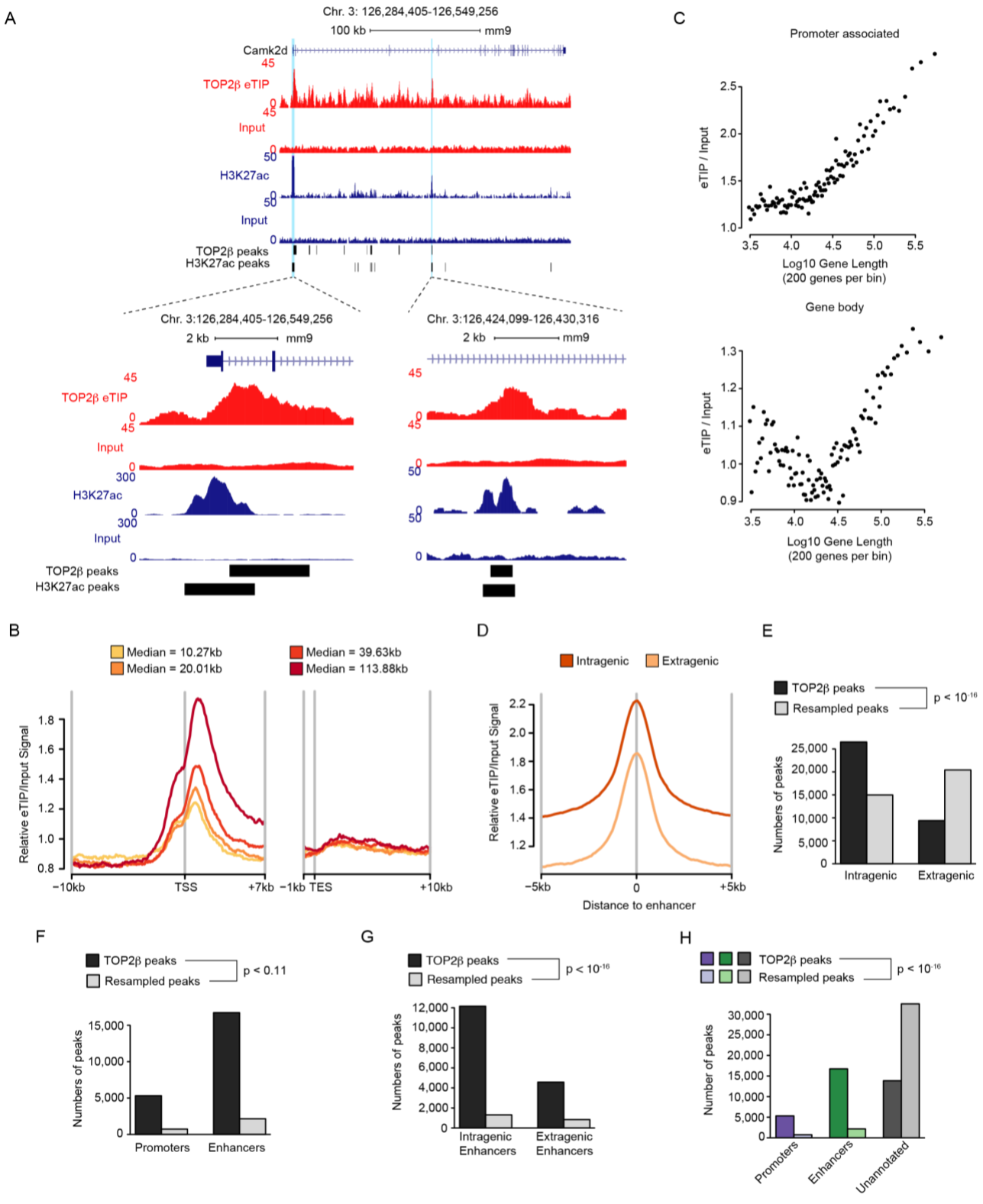
2.6.6 Figure S1. Analysis of TOP2β activity in neurons by eTIP-seq

A. Quantitative PCR (qPCR) analysis of input and IP DNA isolated from ETOP-treated vs. DMSO-treated DIV12 cultured cortical neurons. Data were normalized to a negative control intergenic region in the ETOP-treated condition (see *Methods*). Data are mean ± SEM for six (input) and seven (IP) independent experiments. Two-tailed unpaired t-test (n.s., not significant; **, $p < 10^{-5}$).

B. Aggregate plot of input-normalized eTIP signals at genes divided into quantiles of gene expression. The median gene expression (TPM) for each group is indicated. Average signal around the transcription start site (TSS) and transcription end site (TES) is shown.

C. Boxplot of exonic RNA expression from DIV12 wild-type cultured cortical neurons for genes divided in quantiles of gene length.

D. Bar plots of the genomic distribution of TOP2β peaks and resampled TOP2β peaks at promoters, enhancers and unannotated regions. Resampling and statistical analysis was performed as described in **Figure 3F**.


A
Chr. 3: 126,284,405-126,549,256
100 kb
mm9
Camk2d
TOP2β eTIP
Input
H3K27ac
Input
TOP2β peaks
H3K27ac peaks
Chr. 3:126,284,405-126,549,256
2 kb
mm9
TOP2β eTIP
Input
H3K27ac
Input
TOP2β peaks
H3K27ac peaks
Chr. 3:126,424,099-126,430,316
2 kb
mm9
C
Promoter associated
eTIP / Input
Log10 Gene Length
(200 genes per bin)
Gene body
eTIP / Input
Log10 Gene Length
(200 genes per bin)
B
Median = 10.27kb
Median = 20.01kb
Median = 39.63kb
Median = 113.88kb
Relative eTIP/Input Signal
−10kb
TSS
+7kb
−1kb TES
+10kb
D
Intragenic
Extragenic
Relative eTIP/Input Signal
−5kb
0
+5kb
Distance to enhancer
E
TOP2β peaks
Resampled peaks
p < 10-16
Numbers of peaks
Intragenic
Extragenic
F
TOP2β peaks
Resampled peaks
p < 0.11
Numbers of peaks
Promoters
Enhancers
G
TOP2β peaks
Resampled peaks
p < 10-16
Numbers of peaks
Intragenic Enhancers
Extragenic Enhancers
H
TOP2β peaks
Resampled peaks
p < 10-16
Number of peaks
Promoters
Enhancers
Unannotated

2.6.7 Figure S2. Validation of TOP2β activity in neurons by eTIP-seq with a second TOP2β antibody

A. Genome browser snapshots of eTIP-seq and H3K27ac ChIP-seq from DIV12 cultured cortical neurons at an example long gene, *Camk2d*. Regions corresponding to H3K27ac peaks (putative enhancers and promoters) are highlighted in blue. Bottom left, zoomed in view of the promoter region of the *Camk2d* gene. Bottom right, zoomed in view of an enhancer region of the *Camk2d* gene. MACS2 called-peaks of TOP2β eTIP and H3K27ac are indicated below the tracks. Gene annotation and scale are depicted above.

B. Aggregate plot of input-normalized eTIP signals at genes divided into quantiles of gene length. The median gene length for each group is indicated. Average signal around the transcription start site (TSS) and transcription end site (TES) is shown. Mean values plotted for 100bp bins.

C. Running average plots of input-normalized eTIP signal at the promoter associated region (top) and gene bodies (bottom), from eTIP-seq. Averages are shown for bins of 200 genes ranked according to gene length.

D. Aggregate plot of input-normalized eTIP signal at putative intragenic and extragenic enhancers genome-wide in cultured cortical neurons. Putative enhancers were defined by H3K27ac peaks that did not overlap the TSS of a gene. The plot is centered at the midpoint for each enhancer and the surrounding 10kb region. Mean values plotted for 100bp bins.

E. Bar plots of the genomic distribution of TOP2β peaks and resampled TOP2β peaks at intragenic and extragenic regions. Resampling and statistical analysis were performed as described in **Figure 3F**.

F. Bar plots of the genomic distribution of TOP2β peaks and resampled TOP2β peaks at promoters and enhancers. Resampling and statistical analysis were performed as described in **Figure 3F**.

G. Bar plots of the genomic distribution of TOP2β peaks and resampled TOP2β peaks at intragenic enhancers and extragenic enhancers. Resampling and statistical analysis were performed as described in **Figure 3F**.

Chapter 3:

Concluding Remarks and Future Directions

All experimental data discussed in this section is preliminary (i.e. based off a single replicate) and is presented to explain rationale and experimental design rather than draw conclusions. Future experiments should be optimized and contain additional controls. The experiments and data discussed in this chapter were conceived carried out by me (SAN).

3.1 Introduction

This dissertation begins to interrogate a transcriptional regulatory mechanism that gives rise to the unique length-associated gene expression signature in neurons. Recent studies have demonstrated that the level of expression of long genes must be precisely regulated for normal neuronal function. Our study in chapter 2, provides evidence supporting a model whereby MeCP2 functions as a transcriptional repressor and prevents the overexpression of long neuronal genes by affecting the activity of TOP2β at promoter associated regions and gene bodies to restrict the activity of TOP2β. These findings suggest the interaction between MeCP2 and TOP2β is a unique neuronal association necessary for maintaining the proper levels of expression of a distinct population of neuronally-enriched genes. In this chapter, I outline future studies that can be undertaken to probe the MeCP2-TOP2β interaction and explore its implications for the pathology of Rett syndrome and related neurodevelopmental disorders.

While our study in chapter 2 focused on prominent TOP2β activity at long, MeCP2-regulated genes, we observed TOP2β activity at additional sites that do not appear to be related to MeCP2-mediated transcriptional regulation, thus prompting important questions regarding the functional significance of TOP2β activity in the mature brain. For example, what is the mechanism underlying TOP2β recruitment to specific sites in the genome? What is the precise role of TOP2β activity during transcription (i.e. chromatin accessibility, transcription initiation, transcription elongation)? How is TOP2β activity regulated during transcription? Does TOP2β form interactions with other neuron-specific chromatin regulators in the mature brain? Here, I outline future studies that begin to address these topics.

3.2 Determine the specificity of the interaction between MeCP2 and TOP2β.

3.2.1 Determine whether MeCP2 interacts with the TOP2α

In our study in Chapter 2, we show MeCP2 preferentially interacts with the C-terminal region of TOP2β, however, whether MeCP2 interacts with TOP2α, remains to be determined. Given that the C-terminal region is where TOP2α and TOP2β diverge, one hypothesis is that MeCP2 will not interact with TOP2α. TOP2α, however, is not robustly expressed in neurons. Therefore, to determine whether MeCP2 and TOP2α interact, co-IP experiments can be performed in cells in which both TOP2α and TOP2β are expressed.

Previous studies have shown that TOP2α and TOP2β have distinct expression patterns and biological roles (Austin et al., 1993; Jenkins et al., 1992). TOP2α and TOP2β are paralogues that share considerable sequence identity. Eukaryotic TOP2 enzymes are large homodimers that possess an N-terminal ATPase domain (Dutta and Inouye, 2000), a TOPRIM (topoisomerase-primase) domain (Aravind et al., 1998), DNA-binding domain (Berger and Wang, 1996) and a less conserved and largely disordered C-terminal region (Austin et al., 1993; Crenshaw and Hsieh, 1993; Gadelle et al., 2003). This domain arrangement is conserved between all type II topoisomerases, from bacteria to human enzymes. However, despite the high degree of similarity between the N-terminal ATPase domain and central catalytic core, a key distinction between bacteria and eukaryotic enzymes resides in the C-terminal domain (Chen et al., 2013).

Structural studies have determined TOP2α and TOP2β share approximately 68% identity at the amino acid level, but this is not evenly distributed throughout the proteins. The N-terminal three-

quarters of the enzyme, consisting of the ATPase and central breakage and rejoining regions of TOP2α and TOP2β, share approximately 78% amino acid identity. In contrast, the least conserved portion of the TOP2 enzymes, the C-terminal regions, share approximately 34% amino acid identity (Austin and Marsh, 1998). The divergent C-terminal regions of TOP2α and TOP2β govern the enzyme-specific functions (Linka et al., 2007a). Notably, The C-terminal regions have been implicated as the site for the nuclear localization signals and many post-transcriptional modifications, including phosphorylation and acetylation (Austin and Marsh, 1998; Caron et al., 1994; Crenshaw and Hsieh, 1993; Jensen et al., 1996; McClendon et al., 2008).

The presence of two type II topoisomerases with divergent sequences in the C-terminal domain in vertebrate cells raised the question of whether these two enzymes perform specialized and non-redundant roles. TOP2α is expressed in proliferating cells and plays critical roles in mitosis. TOP2β is expressed in both proliferating and non-dividing cells (Juenke and Holden, 1993; Kondapi et al., 2004; Tsutsui et al., 1993; Zandvliet et al., 1996). Studies investigating the roles of TOP2 enzymes during cell-cycle observed oscillating levels of TOP2α, with increasing levels during S, G2 and M phases and decreasing levels during G1 and G0. In contrast, TOP2β levels remained consistent during cell cycle progression, however, increased considerably as cells entered quiescence (Woessner et al., 1991). In mammalian cells, the essential mitotic functions are performed by TOP2α, with mammalian cells lacking TOP2α terminating in the early stages of cell division, thus confirming that TOP2α is required (Akimitsu et al., 2003; Gonzalez et al., 2011; Grue et al., 1998). Mice lacking TOP2β die perinatally due to deficits in synaptic innervation at neuromuscular junctions (Yang, 2000). Conditional knockout of *Top2b* in brain tissue results in altered formation of brain structures and abnormal neuronal migration during corticogenesis (Lyu and Wang, 2003).

Recombinant proteins have been used to investigate the *in vitro* activities of the TOP2

enzymes. Both TOP2α and TOP2β exhibit strand passage activities as observed in decatenation and relaxation assays. Constructs in which the C-terminal regions have been swapped between TOP2α and TOP2β have been used to investigate the function of the C-terminal region in vitro (Gilroy and Austin, 2011; Meczes et al., 2008). A truncated form of TOP2β lacking the C-terminal region, bound to DNA most strongly, therefore suggesting the C-terminal region may have a negative regulatory role in DNA binding (Linka et al., 2007a). Collectively, these studies provide evidence implicating the C-terminal region in the isoform-specific functions of TOP2α and TOP2β. These studies demonstrate TOP2α and TOP2β display similar catalytic properties, therefore excluding a catalytic activity-based explanation for their divergent roles. Furthermore, the studies indicate the functional distinction between TOP2α and TOP2β is mediated through specific interactions with regulatory proteins and post-translational modifications, which occurs in the C-terminal domain.

These previous findings together with our results implicating the C-terminal domain as important for the MeCP2-TOP2β interaction, warrant further investigation into whether only TOP2β and not TOP2α interact with MeCP2. To carry out this analysis, we can perform co-IP experiments in which full-length MeCP2 is co-expressed with full-length TOP2α and the C-terminal region of TOP2α **(Figure 1)**. If the results show MeCP2 does not interact with TOP2α under conditions in which we successfully detect an interaction with TOP2β, this will provide evidence in support of the interaction between MeCP2 and TOP2β being a unique neuronal complex that may have evolved in order to regulate the expression of genes required to maintain proper neuronal function.

Moreover, determining whether MeCP2 interacts with TOP2α can help extrapolate the precise amino acids necessary for the association between MeCP2 and TOP2β. Along these lines, generating chimeric forms of TOP2α and TOP2β, in which the C-terminal regions of each gene are

swapped can help delineate which amino acids are necessary and sufficient for the interaction between MeCP2 and TOP2β **(Figure 1)**. For example, does expressing the N-terminal and central region of TOP2α with the C-terminal region of TOP2β establish an association between MeCP2 and TOP2α. Conversely, does expressing the N-terminal and central region of TOP2β with the C-terminal region of TOP2α abolish the interaction between MeCP2 and TOP2β. For future experiments, careful consideration should be made when designing chimeric TOP2α and TOP2β constructs. Historically, TOP2α and TOP2β have been quite challenging to clone, and this has limited our progress in this area to date.

3.2.2 Determine whether Rett mutations abolish the interaction between MeCP2 and TOP2β

One model of MeCP2 function in neurons is to serve as a bridge to link the NCoR corepressor complex to chromatin, thereby allowing histones to be deacetylated (see Chapter 1 for detail discussion). In support of this model, studies have shown that disrupting the bridging function, either by mutations that prevent MeCP2 from binding to methylated DNA or prevent MeCP2 from interacting with the NCoR corepressor complex, causes Rett syndrome (RTT) (Lyst et al., 2013).

Out of the hundreds of RTT-causing MeCP2 mutations that have been reported, there are eight mutations that are categorized as "hotspot" mutations (encoding the following missense and nonsense mutations: R106W, R133C, T158M, R168X, R255X, R270X, R294X and R306C), as they account for more than 60% of all documented cases (Ip et al., 2018; Neul et al., 2008). Additional studies interrogating RTT mutations in the MBD have shown that not all mutations alter the ability of MeCP2 bind methylated DNA (Yang et al., 2016). While some of these mutations in the MBD of MeCP2 result in an unstable protein due to improper folding, some mutations result in stable

proteins that maintain some ability to associate with methylated DNA. This observation suggests that the Rett phenotype may be driven by an alternative mechanism, that does not involve MeCP2 binding to methylated DNA. In our study in Chapter 2, we show TOP2β interacts with a region of MeCP2 containing the MBD domain. A proposed sequence sufficient for the interaction between MeCP2 and TOP2β is adjacent to the amino acid sequence identified as necessary for the interaction with methylated DNA (Nan et al., 1993a). Therefore, it is possible that disrupting the interaction between MeCP2 and TOP2β may contribute to the RTT phenotype.

Approximately 42% of cases are caused by frameshift mutations or truncation mutations, which are primarily located downstream of the MBD domain of MeCP2. Approximately 46% of RTT cases are identified as missense mutations, with nearly half of the mutations residing in the MBD domain of MeCP2 (Christodoulou et al., 2003). To test whether the RTT-causing mutation in the MBD domain of MePC2 disrupt the interaction with TOP2β, we can generate MeCP2 constructs containing the RTT-causing MBD mutations and perform co-IP experiments with TOP2β. This analysis will provide insight as to whether RTT mutations disrupt the interaction between MeCP2 and TOP2β and potentially implicate specific amino acids involved in this interaction. Validating these MBD mutations using a DNA binding assay would be informative, however, in the context of the *in vitro* co-IP, these reactions are carried out in the absence of nucleic acids as lysates are treated with benzonase.

Mutations in the MBD domain of MeCP2 can be further subdivided into three clusters (**Figure 1**). This categorization of mutations indicates the spatial location of the mutation rather the biological consequences of the specific mutations. In preliminary studies, we selected nine mutations that span the three clusters of the MBD of MeCP2: R111G, R133H, R133C, S134C, A140V. P152R, F155S, T158A, T158M. Each of these mutations has been previously assessed for the binding

affinity to methylated DNA (Yang et al., 2016). The R111G mutation results in a stable MeCP2 protein that has an 88-fold reduction in binding affinity to methylated DNA. While simulations do not predict the R133H mutation to affect the stability of the protein, reduced binding affinity to methylated DNA was observed. The R133C mutation is the second most frequent clinical mutation in the MBD of MeCP2 and causes reduced binding to methylated DNA. The S134C mutation does not fall within the sequence that binds directly to DNA, however simulations suggest this mutation will produce a relatively unstable protein. The A140V mutation results in a very stable protein with only moderately reduced binding affinity to methylated DNA. The P152R mutation produces an unstable protein, however, maintains relatively strong binding affinity to methylated DNA. The F155S and T158A mutations results in proteins that most resemble to wildtype MeCP2 protein in terms of protein stability and binding affinity to methylated DNA. The T158M mutation is the most frequent clinical mutation in the MBD of MeCP2. This mutation results in a moderately stable protein, similar folding stability to the wildtype MeCP2, and results in a 2-fold reduction in binding affinity to methylated DNA.

In addition to the known RTT-causing missense mutations in the MBD, we asked whether we could identify the specific amino acids involved in the interaction between MeCP2 and TOP2β. From our previous co-IP analysis, we observed TOP2β interacts with a fragment of MeCP2 containing amino acids 1-143 as well as a fragment of MeCP2 containing amino acids 167-484. Given these results, we tested whether removing amino acids 143-167 from wildtype MeCP2 (Δ143-167aa-FLAG-MeCP2) was sufficient to disrupt the interaction between MeCP2 and TOP2β.

Initial results from one replicate of this experiment indicated that we were able to express the mutant and deletion proteins (**Figure 2**). Western blot analysis of these MeCP2 mutants, in parallel with additional MeCP2 fragments previously tested, showed the mutations made were not sufficient

to disrupt the TOP2β interaction. This suggests that the disease-causing mutations examined may not disrupt this interaction. However, additional replicates and controls will need to be assessed to confirm these findings. The preservation of the interaction in the Δ143-167 MeCP2 construct suggests that while sufficient to mediate an interaction with TOP2β, the sequences contained in this region may not be strictly necessary.

Several considerations can be taken into account for future experiments probing this interaction with additional replicates and experimental constructs, such as, including additional amino acids and examining the C-terminal region of the MBD to determine the specific sequences involved in the interaction between MeCP2 and TOP2β. Moreover, careful examination of mutations in Cluster 3 is warranted, as they also reside in the C-terminal region of the MBD (**Figure 1**). It is also possible that RTT-causing mutations do not disrupt the interaction with TOP2β, however, additional co-IP experiments between TOP2β and the MeCP2 mutations are necessary to make this determination.

It is important to note that results from this initial experiment of co-IP analysis with MeCP2 mutants can be challenging to interpret. In parallel with this experiment, we generated an MeCP2 construct with the R306C mutation, which has been reported in the literature to disrupt the interaction between MeCP2 and the NCoR co-repressor complex (e.g. NCOR1, NCOR2, TBLR1, TBL1, HDAC3) (Kruusvee et al., 2017; Lyst et al., 2013). We co-expressed FLAG-tagged R306C MeCP2 with and without mCHERRY-tagged TBLR1 in HEK293T cells, to replicate the previous findings. In our studies, we find HEK293T cells innately express NCOR complex components at high levels. Expression of full-length FLAG-tagged MeCP2 in HEK293T cells followed by IP of FLAG showed an interaction with endogenous TBLR1 (data not shown), therefore, we hypothesized expressing FLAG-tagged R306C in HEK293T cells followed by IP FLAG would not interact with

TBLR1. Preliminary co-IP results of FLAG-tagged R306C MeCP2 showed the interaction with endogenous TBLR1 was maintained (data not shown). Importantly, this result is from a single experiment and additional replicates are necessary to draw conclusions, however this example highlights caveats associated with these experiments. Interestingly, studies in which FLAG-tagged MeCP2 and MCHERRY-tagged TBLR1 were co-expressed in HEK293T cells, resulted in reduced expression of FLAG-tagged MeCP2, indicating MCHERRY-tagged TBLR1was inducing some type of auto-inhibition. This effect on MeCP2 expression was observed in two separate experiments (data not shown).

3.2.3 Determine whether TOP2β interacts with other methyl-DNA binding protein family members

In chapter 2, we demonstrate MeCP2 and TOP2β interact *in vivo* in the brain. Using *in vitro* co-IP, we demonstrate the C-terminal region of TOP2β interacts with MeCP2 and the MBD of MeCP2 interacts with TOP2β. As previously discussed in Chapter 1, MeCP2 is an epigenetic reader of methylated DNA. In mammals, mCA is absent in proliferating cells, however, accumulates to high levels postnatally in neurons and is maintained into adulthood. The accumulation of mCA parallels the increase in MeCP2 protein levels. MeCP2 can bind mCG and mCA dinucleotides via the MBD domain, which distinguishes MeCP2 from other methyl-DNA binding proteins. Moreover, the unique ability for MeCP2 to bind mCA raises the possibility that this binding specificity may be important for neuronal gene regulation.

A recent study (Tillotson et al., 2017) examined the biological significance of the dual-binding ability of MeCP2 to mCG and mCA, by replacing the MBD of MeCP2 with an orthologous domain from a related methyl-DNA binding protein, MBD2, which exclusively binds mCG. This

resulted in a chimeric protein referred to as MeCP2-MBD2 (MM2). Knockin-mice expressing MM2 develop Rett-like phenotypes, which suggests binding mCG alone is insufficient for normal brain function. Moreover, gene expression analysis showed many of the genes upregulated in the MM2 mice are upregulated in the MeCP2 knockout mice, providing insight into the mechanisms underlying RTT-like phenotype.

Given these findings, coupled with the observation that mCA, MeCP2 and TOP2β are robustly expressed neurons, it raises the possibility that the MeCP2-TOP2β interaction requires an MBD that binds mCA. To test this hypothesis, we can perform *in vitro* co-IP experiments, by co-expressing MYC-tagged TOP2β with wild-type EGFP-tagged MeCP2 or EGFP-tagged MM2 followed by IP of MYC. If we detect an interaction between TOP2β and wildtype MeCP2, but fail to detect an interaction between TOP2β and MM2, this provides evidence supporting a functional significance for the interaction between MeCP2 and TOP2β. Furthermore, it would suggest the MeCP2-TOP2β interaction is a neuron-specific and occurs specifically with an MBD domain that can bind to mCA. Such a finding would further support disruption of this interaction contributes to RTT-like pathology.

While in vitro experiments are described here, these same experiments can be carried out in vivo. Moreover, additional *in vitro* and *in vivo* eTIP-seq analysis using MM2 can be carried out to profile changes in TOP2β activity.

3.2.4 Determine whether the minimal truncated version of MeCP2 interacts TOP2β

Several lines of evidence that suggest the primary function of MeCP2 is to serve as a bridge linking the NCoR co-repressor complex with methylated DNA (refer to Chapter 1 for detailed

discussion). A recent study (Tillotson et al., 2017) tested this hypothesis by generating a truncated version of MeCP2, containing only the MBD and NID domains (referred to as ΔNIC). They find ΔNIC is sufficient for MeCP2 function and observe mice expressing ΔNIC survive over a year with mild symptoms. Furthermore, expressing ΔNIC in MeCP2-deficient mice rescues the Rett-like phenotype.

In Chapter 2, we demonstrate TOP2β interacts with the MBD of MeCP2. To gain insight into the functional significance of the interaction between MeCP2 and TOP2β, we can perform in vitro co-IP experiments to test whether TOP2β interacts with the ΔNIC. Using HEK293T cells, we can co-express full-length MYC-tagged TOP2β with EGFP-tagged ΔNIC followed by co-IP of MYC. If we detect an interaction between TOP2β and ΔNIC, this provides evidence of an additional essential function for the MBD domain of MeCP2, which is to interact with TOP2β.

3.3 Investigating the regulation of TOP2β activity in neurons.

In Chapter 2, we demonstrate TOP2β activity is enriched at long genes, specifically at the promoter associated regions, gene bodies and intragenic enhancers of long genes. While many of these genes were also previously identified as MeCP2-regulated genes, we observed TOP2β activity at genes that were not regulated by MeCP2. This observation prompted many questions about recruitment of TOP2β to specific sites on the genome and the regulation of TOP2β activity.

As described in Chapter 1, one approach to address these questions is through proteomic analysis, specifically identifying protein interactors and post-translational modifications. As described in Chapter 2, proteomic analysis was carried out to identify MeCP2 protein interactors,

and through this analysis we identified TOP2β. This same approach can be used to identify TOP2β protein interactors. The C-terminal region of TOP2β is responsible for the enzyme-specific properties of TOP2β that are distinct from other topoisomerases, including the TOP2 paralog, TOP2α. Specifically, the C-terminal region is the site for nuclear localization and post-translational modifications. Several potential phosphorylation sites as well as putative sites for other posttranslational modifications, including acetylation, have been identified through proteomic and structural studies.

In Chapter 2, acetylation was highlighted as one potential mechanism underlying the regulation of long genes expression mediated by the interaction between MeCP2-TOP2β. As previously described, several lines of evidence suggest the primary function of MeCP2 is to serve as a bridge to link the NCoR corepressor complex with DNA. The NCoR corepressor complex contains the histone deacetylase (HDAC3), which functions to remove acetyl groups from histone and non-histone proteins, which in turn can have repressive effects transcription regulation.

It is possible MeCP2 associates with TOP2β, and through this proximity, the NCoR corepressor complex deacetylates TOP2β at an acetylation site in the C-terminal region of TOP2β. This deacetylation of TOP2β causes reduced TOP2β activity and downregulated long gene expression. In the absence of MeCP2, NCoR is unable to maintain proximity with TOP2β, therefore TOP2β does not become deacetylated. As a result, TOP2β activity is unregulated which results in the overexpression of long genes.

This model can be tested using proteomic approaches to identify acetylation sites on the TOP2β. To determine whether TOP2β activity is mediated by the deacetylase activity of the HDAC3-containing NCoR corepressor complex, the amino acid residues on TOP2β that are deacetylated can

be identified. Tandem mass spectrometry can be used to identify acetylation sites present on the endogenous TOP2β protein purified from 8-week mouse cerebral cortex nuclear extracts, as well as MYC-tagged TOP2β and MYC-tagged TOP2β C-terminal (1199-1612aa) fragments expressed in HEK293T cells.

3.4 Investigating the role of TOP2β in neuronal transcription.

Previous studies have shown that pharmacological inhibition or genetic knockdown of TOP2β causes reduced expression of long gene in neurons (King et al., 2013) (see chapter 1 for detailed discussion). These findings prompted questions about the role of TOP2β in neuronal transcription and how mechanistically does TOP2β facilitate the expression of long genes. As described in Chapter 1, genetic manipulation of TOP2β, either by constitutive or conditional knockout, in the brain of mice results in perinatal death. As a result, the ability to investigate TOP2β function *in vivo* is limited, especially in the context of the adult brain.

To circumvent these limitations and to query the role of TOP2β during transcription, I established a primary neuron culture system, using E14.5 Top2b$^{flx/flx}$ mice, in which I can ablate *Top2b* using lentivirus expressing Cre recombinase (active Cre), resulting inTop2bcKO, or a non-functional Cre recombinase (inactive Cre), resulting in wild-type. To assess the efficacy of this culture system, I profiled changes in gene expression between the Top2bcKO and wild-type neurons and observed significant reduction in the expression of long genes in the Top2bcKO, which is consistent with previous findings (King et al., 2013), however our system reported a larger number of misregulated genes as well as a larger effect size (*Nettles and Gabel unpublished*).

To investigate the role of TOP2β during transcription, we can test how loss of TOP2β effects transcription. In particular, we can assess changes in transcription at TOP2β active sites identified by eTIP-seq described in Chapter 2, including promoter associated regions, gene bodies and intragenic enhancers associated with long genes. One potential function of TOP2β is to promote acetylation of promoters and enhancers of long genes. Preliminary results, from a single replicate, of H3K27ac chip in Top2bcKO neurons suggests loss of TOP2β does not affect acetylation. We observed subtle changes acetylation at promoters, but do not observe changes in acetylation intragenic enhancers of long genes (data not shown). While additional replicates are necessary to fully understand the data, initial interpretations of these results suggest that acetylation is upstream of or in parallel with TOP2β activity during the transcription of long neuronal genes.

Several lines of evidence suggest TOP2β facilitates gene expression by promoting chromatin accessibility. To investigate this hypothesis, we can perform ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) in Top2bcKO and wildtype cultured neurons. If reduced chromatin accessibility is detected at long genes downregulated in Top2bcKO neurons, it is possible TOP2β facilitates gene expression by making chromatin accessible to transcriptional machinery or other chromatin regulators.

It is also possible that TOP2β facilitate transcription through another mechanism, and not directly though acetylation or chromatin accessibility. Previous studies suggest TOP2β plays a role in transcription elongation (King et al., 2013). Therefore, ChIP-seq for RNA polymerase II (RNA pol II) and its transcription-stage-specific modified forms (e.g. serine 5 phosphorylated RNA pol II – transcription initiation; serine 2 phosphorylated at RNA pol II – transcription elongation) can be assessed in the Top2bcKO and wild-type neurons using this system.

3.5 *In vivo* approaches for profiling TOP2β activity in neurons

As described in Chapter 1, there are limitations for evaluating TOP2β activity *in vivo*, including challenges generating a viable TOP2β loss-of-function mouse model. Another limitation that has emerged is the ability to profile TOP2β occupancy *in vivo* through conventional genomic approaches such as chromatin immunoprecipitation (ChIP). This limitation has made it especially challenging to evaluate TOP2β function in the adult brain.

In my own ChIP-seq data from the cerebral cortex of adult mice (*in vivo* ChIP), I observe broad occupancy of protein across the genome, with few prominent or distinct peaks, when compared to the input samples (data not shown). The broad distribution of ChIP signal is not uncommon for ubiquitously expressed proteins. Indeed, previous studies performing MeCP2 ChIP observed an extremely broad distributed signal (Chen et al., 2015; Cohen et al., 2011; Gabel et al., 2015; Kinde et al., 2016; Skene et al., 2010). The TOP2β ChIP and ChIP input profiles was especially notable when compared to the TOP2β eTIP and eTIP input profiles (refer to Chapter 2 for eTIP analysis). While this is not a direct comparison, as ChIP, which measures protein binding/occupancy across the genome, is distinct from eTIP, which measures TOP2β enzymatic activity (i.e. cut sites), and these two methods were performed in two different system, brain and neuron cortical cultures, respectively, there was a stark contrast in TOP2β profile between eTIP and ChIP. One interpretation of these results is that TOP2β is ubiquitously expressed and therefore binds broadly across the genome, however, TOP2β is active at distinct regions, which can only be assessed via eTIP. To test this hypothesis, TOP2β ChIP in wild-type and *Top2b* knockout mouse brain would need to be performed, however, as previously stated, there are no current viable brain-specific *Top2b*

knockout mouse models.

I previously carried out studies in which I attempted to perform eTIP from the cerebral cortex of adult mice (*in vivo* eTIP), however I encountered several technical challenges in the drug delivery. A brief description of the various approaches includes: (a) using fresh tissue and incubating cortex halves or quarters in solution containing etoposide or DMSO (vehicle control); (b) using fresh tissues, coarsely chopping the tissues and incubating in solution contain etoposide or DMSO (vehicle control); (c) using frozen tissue and incubating in solution contain etoposide or DMSO (vehicle control). For all paradigms tested, after the 15min incubation at 37ºC, the standard ChIP protocol was followed, excluding the crosslinking step. The results of the *in vivo* eTIP suggested that the drug did not permeate the tissue. In another set of *in vivo* eTIP studies, etoposide or DMSO was administered into the lateral ventricles via stereotaxic injections. Extensive evaluation of these samples was not carried out due to concerns with accounting for the precise amount of etoposide or DMSO that entered the brain (i.e. issues with the viscosity, concentration and volume of etoposide or DMSO administered). Altogether, there examples demonstrate the considerable challenges in previous *in vivo* eTIP studies.

The most recent approach we have considered for *in vivo* eTIP would be to generate electrophysiology slice preparation style brain tissue. Administering the drug to the tissue and having the drug permeate the cells has been a consistent limitation in previous *in vivo* approaches. Preparing tissue in the same manner as slice preparation experiments would allow for the tissue slices to remain alive and be accessible to etoposide or DMSO treatment. Using this approach could allow for profiling TOP2β is a wide range of conditions. Of immediate interest is to perform eTIP in MeCP2 knockout and wildtype mice, in order to test how TOP2β activity changes in the absence of MeCP2. Additional mouse models include the MeCP2 duplication mouse, Dnmt3a knockout mouse,

MM2 knockin mouse, and any RTT-mutations that disrupt the interaction between MeCP2 and TOP2β

The MeCP2 duplication mouse would allow for profiling how overexpression of MeCP2 alters TOP2β activity. These studies would be a powerful complement to the neuronal cortical culture eTIP experiments in which MeCP2 levels were manipulated. Moreover, evaluating changes in TOP2β activity in the Dnmt3a knockout mouse would allow for profiling how DNA methylation influences TOP2β activity. If any Rett mutations discussed in the previous section (Chapter 3.2) disrupt the interaction between MeCP2 and TOP2β, it would be interesting to evaluate how TOP2β activity is altered in this context. For example, does it phenocopy the MeCP2 knockout or is the degree of effect different.

Finally, as discussed in Chapter 1, a rare *de novo* mutation in TOP2β was recently identified. Our preliminary studies suggest this mutation may be gain-of-function. Therefore, it is of considerable interest to investigate how TOP2β activity is altered in this model. While this would require the generation of a mouse model with this mutation, evaluating TOP2β activity in this *Top2b* mutant using this *in vivo* eTIP method may contribute to our understanding of the underlying disease pathology.

3.6 Investigating topoisomerase inhibition as a therapeutic approach

As described in chapter 1, previous studies demonstrate topoisomerases play an essential role in regulating the expression of long genes, specifically genes encoding synaptic proteins (King et al., 2013; Mabb et al., 2014). As a result, these pharmacological inhibitors of topoisomerase have

emerged as a potential therapeutic target for neurodevelopmental disorders, such as Angelman syndrome and Rett syndrome.

Angelman syndrome is a severe neurodevelopmental disorder caused by disruption of maternal allele of the ubiquitin-protein ligase E3A (UBE3A). *UBE3A* is biallelically expressed in most tissues, however, in neurons, only the maternal allele is expressed. One therapeutic strategy for treating Angelman syndrome is to activate the paternal *Ube3a* allele, which is epigenetically silenced in neurons, in order to compensate for the defective maternal allele (Mabb et al., 2011).

In a screen of candidate compounds, topoisomerase inhibitors were found to activate the paternal Ube3a gene. Administration of both TOP1 and TOP2 inhibitors in mouse cortical neurons showed reactivation of the silenced, paternal allele (Huang et al., 2012). Reactivation is thought to occur through inhibiting transcription of a long antisense *Ube3a* transcript, *Ube3a-ATS*, that normally blocks paternal Ube3a expression (Powell et al., 2013). These results suggest that a critical role for TOP1 and TOP2 is to eliminate negative supercoiling and that topoisomerase inhibitors that impede transcription elongation have unanticipated therapeutic implications.

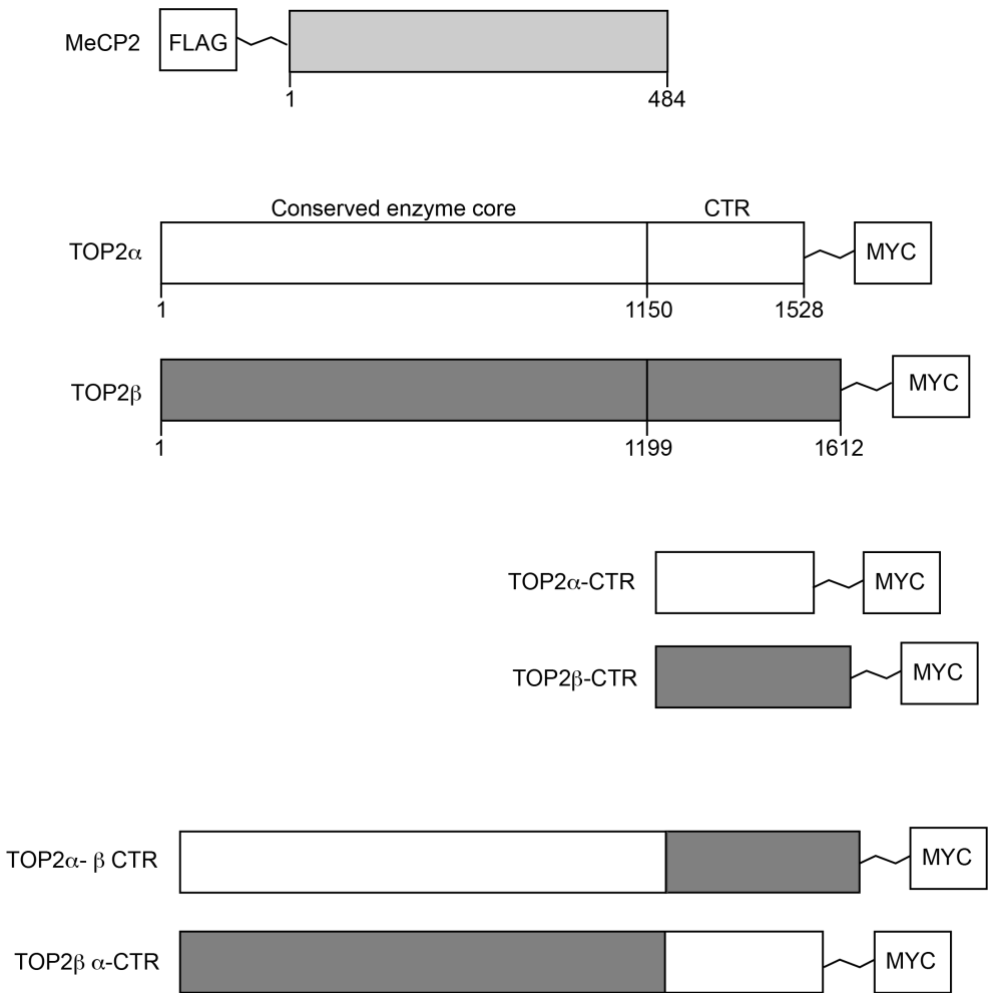
Intriguingly, UBE3A duplication is implicated in ASD. this suggests the dosages of UBE3A must be tightly regulated to maintain proper neuronal development. Moreover, these observations raise the question of whether there is aberrant expression of long genes in Angelman syndrome or ASD. It is possible topoisomerase inhibitors are downregulating the expression of long genes.
Rett syndrome is a severe neurodevelopmental disorder caused by the loss or disruption of MeCP2 (see chapter 1 for detailed discussion). Previous studies show inhibiting TOP1 leads to reduced expression of long gene in neurons, suggesting that topoisomerase inhibitors may function to prevent the overexpression of long genes when MeCP2 is lost or disrupted. One study tested this by knocking down MeCP2 using RNAi in neuronal cortical cultures and administering the TOP1

inhibitor, topotecan (Gabel et al., 2015). The results showed up-regulation of long genes in MeCP2 knockdown neurons, however, there was a dose-dependent down-regulation in MeCP2 knockdown neurons treated with topotecan.

These observations, coupled with our findings in chapter 2, where we show an interaction between MeCP2 and TOP2β in neurons and demonstrate MeCP2-mediated repression on TOP2β activity, raise the possibility that inhibiting TOP2β in the MeCP2 KO mouse may restore the expression of long gene to within normal range. Moreover, if the over-expression of long genes contributes to Rett syndrome pathology, then reducing the expression of long genes may rescue the Rett phenotype.

One approach to test this would require generating an MeCP2 knockout mouse that is heterozygous for *Top2b* ($Mecp2^{*/y}$;$Top2b^{+/-}$). Through gene expression profiling, we can determine whether heterozygous expression of *Top2b* is sufficient to reduce the expression of long genes. If we observe reduced long gene expression in the *Mecp2-null;Top2b-het* relative to the Mecp2-null, than, phenotypically, the *Mecp2-null;Top2b-het* may have less severe symptoms than the *Mecp2-null*. Moreover, extensive genomic profiling can be performed in these mice, including *in vivo* eTIP-seq, to examine changes in TOP2β activity, and H3K27ac ChIP-seq, to examine changes at MeCP2-regualted enhancers.

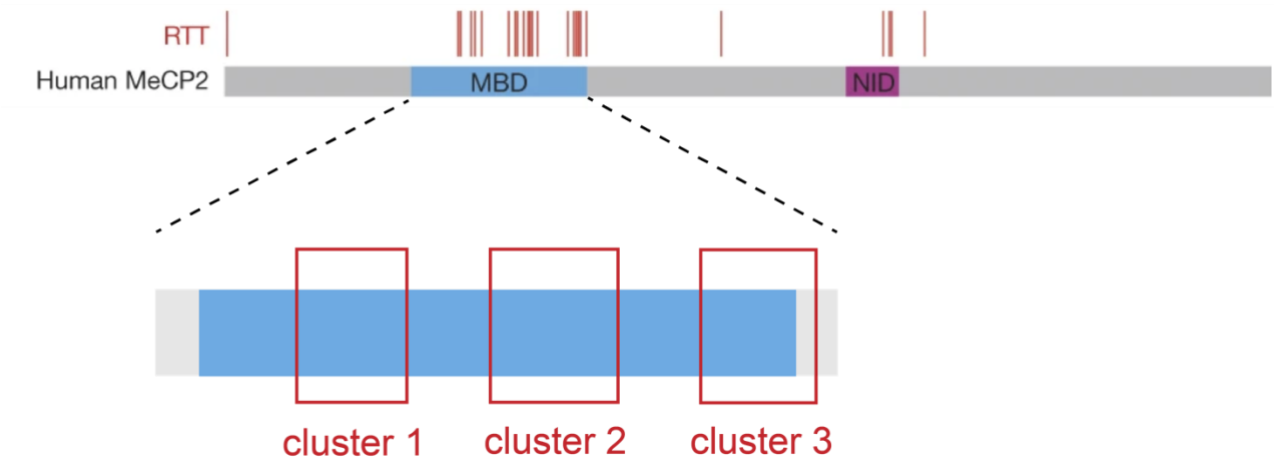
3.7 Figures



3.2.1 Figure 1. Schematic synopsis of proposed constructs to investigate the interaction between MeCP2 and TOP2α

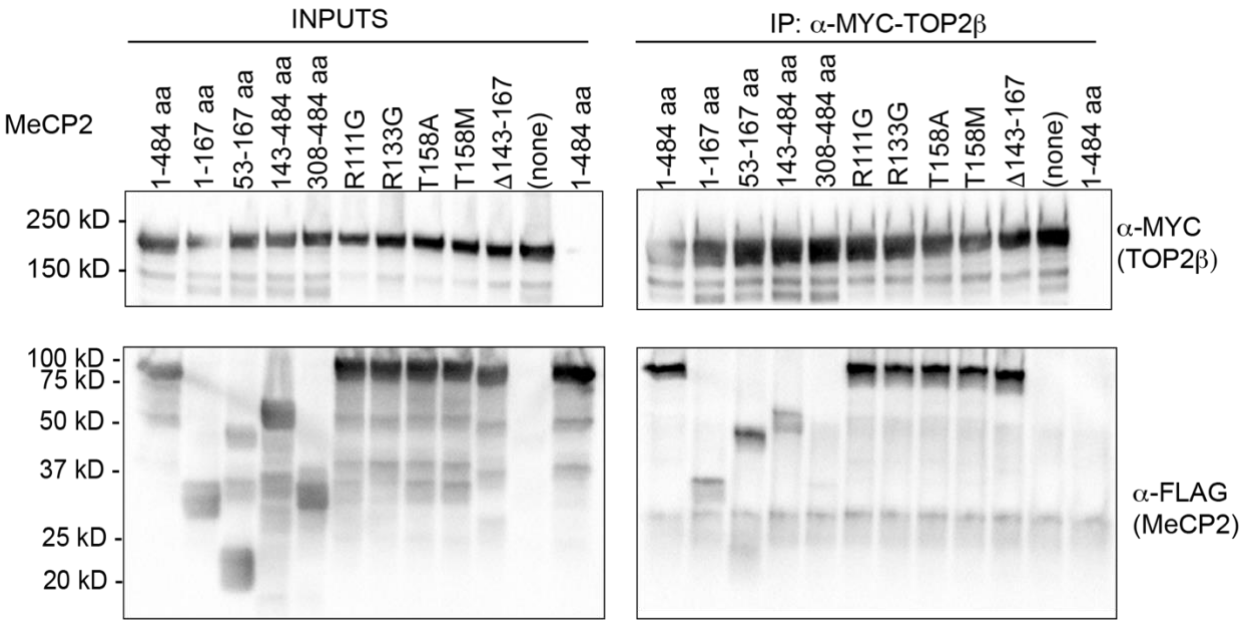
Schematic synopsis of the proposed constructs. MeCP2 constructs is fused to FLAG at its N-terminal end. TOP2 constructs are fused to MYC at their C-terminal ends. Construct description (top to bottom) – MeCP2: full-length MeCP2; TOP2α: full-length TOP2α; TOP2β: full-length TOP2β; TOP2α-CTR: truncated TOP2α (1150 – 1528aa); TOP2β-

CTR: truncated TOP2β (1199 – 1612aa); TOP2α-βCTR: chimeric TOP2α containing TOP2α (1 – 1150aa) and TOP2β-CTR (1199 – 1612aa); chimeric TOP2β containing TOP2β (1 – 1199aa) and TOP2α-CTR (1150 – 1528aa).



3.3.2 Figure 1. Schematic of missense Rett syndrome mutations in the human *MECP2* gene

This figure is adapted from (Tillotson et al., 2017). The human MeCP2 protein sequence with the MBD (blue), NID (magenta), and RTT-causing missense mutations (red lines) annotated. Missense mutations were extracted from RettBASE. Missense mutations in the MBD typically falling within one of three clusters. The domain implicated in the MeCP2-TOP2β interactions overlaps Cluster 3.



3.3.2 Figure 2. Candidate Rett syndrome mutations that disrupt the interaction between MeCP2 and TOP2β

Co-immunoprecipitation analysis of FLAG-tagged fragments of MeCP2 Rett syndrome mutants with MYC-tagged full length TOP2β. Lysates from cells transfected with only MYC-tagged TOP2β (lane labeled "none", indicating no MeCP2 construct) or FLAG-tagged MeCP2 (lane does not contain MYC-tagged TOP2β). ***NOTE: Preliminary results from single experiment***. This co-IP experiments would suggest that each construct tested is stably expressed. Initial interpretation of results would suggest that none of the tested RTT-causing MeCP2 mutations in the MBD domain was sufficient to disrupt the interaction between TOP2β and MeCP2. The deletion fragment (Δ143-167aa-FLAG-MeCP2) does not appear to disrupt the interaction between TOP2β and MeCP2.

References

Adachi, N., Miyaike, M., Kato, S., Kanamaru, R., Koyama, H., and Kikuchi, A. (1997). Cellular distribution of mammalian DNA topoisomerase II is determined by its catalytically dispensable C-terminal domain. Nucleic Acids Research *25*. https://doi.org/10.1093/nar/25.15.3135.

Akimitsu, N., Adachi, N., Hirai, H., Hossain, M.S., Hamamoto, H., Kobayashi, M., Aratani, Y., Koyama, H., and Sekimizu, K. (2003). Enforced cytokinesis without complete nuclear division in embryonic cells depleting the activity of DNA topoisomerase IIα. Genes to Cells *8*. https://doi.org/10.1046/j.1365-2443.2003.00643.x.

Alessio, N., Riccitiello, F., Squillaro, T., Capasso, S., del Gaudio, S., di Bernardo, G., Cipollaro, M., Melone, M.A.B., Peluso, G., and Galderisi, U. (2018). Neural stem cells from a mouse model of Rett syndrome are prone to senescence, show reduced capacity to cope with genotoxic stress, and are impaired in the differentiation process. Experimental and Molecular Medicine *50*. https://doi.org/10.1038/s12276-017-0005-x.

Amir, R.E., van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U., and Zoghbi, H.Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl- CpG-binding protein 2. Nature Genetics *23*, 185–188. https://doi.org/10.1038/13810.

Aravind, L., Leipe, D.D., and Koonin, E. v (1998). Toprim-a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins.

Austin, C.A., and Marsh, K.L. (1998). Eukaryotic DNA topoisomerase IIβ. BioEssays *20*. https://doi.org/10.1002/(SICI)1521-1878(199803)20:3<215::AID-BIES5>3.0.CO;2-Q.

Austin, C.A., Sng, J.H., Patel, S., and Fisher, L.M. (1993). Novel HeLa topoisomerase II is the IIβ isoform: complete coding sequence and homology with other type II topoisomerases. BBA - Gene Structure and Expression *1172*. https://doi.org/10.1016/0167-4781(93)90215-Y.

Austin, C.A., Marsh, K.L., Wasserman, R.A., Willmore, E., Sayer, P.J., Wang, J.C., and Fisher, L.M. (1995). Expression, domain structure, and enzymatic properties of an active recombinant human DNA topoisomerase IIβ. Journal of Biological Chemistry *270*. https://doi.org/10.1074/jbc.270.26.15739.

Baker, S.A., Chen, L., Wilkins, A.D., Yu, P., Lichtarge, O., and Zoghbi, H.Y. (2013). An AT-Hook Domain in MeCP2 Determines the Clinical Course of Rett Syndrome and Related Disorders. Cell *152*, 984–996. https://doi.org/10.1016/J.CELL.2013.01.038.

Ballestar, E., Yusufzai, T.M., and Wolffe, A.P. (2000). Effects of rett syndrome mutations of the Methyl-CpG binding domain of the transcriptional repressor MeCP2 on selectivity for association with methylated DNA. Biochemistry *39*. https://doi.org/10.1021/bi0001271.

Balmer, D., Goldstine, J., Rao, Y.M., and LaSalle, J.M. (2003). Elevated methyl-CpG-binding protein 2 expression is acquired during postnatal human brain development and is correlated with alternative polyadenylation. Journal of Molecular Medicine *81*. https://doi.org/10.1007/s00109-002-0396-5.

Baubec, T., Ivánek, R., Lienert, F., and Schübeler, D. (2013). Methylation-dependent and -independent genomic targeting principles of the mbd protein family. Cell *153*. https://doi.org/10.1016/j.cell.2013.03.011.

Berger, J.M., and Wang, J.C. (1996). Recent developments in DNA topoisomerase II structure and mechanism. Current Opinion in Structural Biology *6*. https://doi.org/10.1016/S0959-440X(96)80099-6.

Berger, J.M., Gamblin, S.J., Harrison, S.C., and Wang, J.C. (1996). Structure and mechanism of DNA topoisomerase II. Nature *379*. https://doi.org/10.1038/379225a0.
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. Molecular Cell *77*. https://doi.org/10.1016/j.molcel.2019.10.032.
Capranico, G., Tinelli, S., Austin, C.A., Fisher, M.L., and Zunino, F. (1992). Different patterns of gene expression of topoisomerase II isoforms in differentiated tissues during murine development. BBA - Gene Structure and Expression *1132*. https://doi.org/10.1016/0167-4781(92)90050-A.
Caron, P.R., and Wang, J.C. (1994). Appendix II: Alignment of Primary Sequences of DNA Topoisomerases. Advances in Pharmacology *29*. https://doi.org/10.1016/S1054-3589(08)61143-6.
Chahrour, M., Sung, Y.J., Shaw, C., Zhou, X., Wong, S.T.C., Qin, J., and Zoghbi, H.Y. (2008). MeCP2, a key contributor to neurological disease, activates and represses transcription. Science (1979) *320*. https://doi.org/10.1126/science.1153252.
Chaly, N., Chen, X., Dentry, J., and Brown, D.L. (1996). Organization of DNA topoisomerase II isotypes during the cell cycle of human lymphocytes and HeLa cells. Chromosome Research *4*. https://doi.org/10.1007/BF02265053.
Chang, C.C., Wang, Y.R., Chen, S.F., Wu, C.C., and Chan, N.L. (2013). New insights into DNA-binding by type IIA topoisomerases. Current Opinion in Structural Biology *23*. https://doi.org/10.1016/j.sbi.2012.11.011.
Chen, L., Chen, K., Lavery, L.A., Baker, S.A., Shaw, C.A., Li, W., and Zoghbi, H.Y. (2015). MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome. Proc Natl Acad Sci U S A *112*, 5509–5514. https://doi.org/10.1073/pnas.1505909112.
Chen, R.Z., Akbarian, S., Tudor, M., and Jaenisch, R. (2001). Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. Nature Genetics *27*. https://doi.org/10.1038/85906.
Chen, S.H., Chan, N.L., and Hsieh, T.S. (2013). New mechanistic and functional insights into DNA topoisomerases. Annual Review of Biochemistry *82*. https://doi.org/10.1146/annurev-biochem-061809-100002.
Christian, D.L., Wu, D.Y., Martin, J.R., Moore, J.R., Liu, Y.R., Clemens, A.W., Nettles, S.A., Kirkland, N.M., Papouin, T., Hill, C.A., et al. (2020). DNMT3A Haploinsufficiency Results in Behavioral Deficits and Global Epigenomic Dysregulation Shared across Neurodevelopmental Disorders. Cell Reports *33*, 2020.07.10.195859. https://doi.org/10.1016/j.celrep.2020.108416.
Christodoulou, J., Grimm, A., Maher, T., and Bennetts, B. (2003). RettBASE: The IRSA MECP2 variation database-a new mutation database in evolution. Hum Mutat *21*. https://doi.org/10.1002/humu.10194.
Clemens, A.W., Wu, D.Y., Moore, J.R., Christian, D.L., Zhao, G., and Gabel, H.W. (2020). MeCP2 Represses Enhancers through Chromosome Topology-Associated DNA Methylation. Molecular Cell *77*. https://doi.org/10.1016/j.molcel.2019.10.033.
Closser, M., Guo, Y., Wang, P., Patel, T., Jang, S., Hammelman, J., de Nooij, J.C., Kopunova, R., Mazzoni, E.O., Ruan, Y., et al. (2022). An expansion of the non-coding genome and its regulatory potential underlies vertebrate neuronal diversity. Neuron *110*. https://doi.org/10.1016/j.neuron.2021.10.014.
Cohen, S., Gabel, H.W., Hemberg, M., Hutchinson, A.N., Sadacca, L.A., Ebert, D.H., Harmin, D.A., Greenberg, R.S., Verdine, V.K., Zhou, Z., et al. (2011). Genome-Wide Activity-Dependent MeCP2 Phosphorylation Regulates Nervous System Development and Function. Neuron *72*.

https://doi.org/10.1016/j.neuron.2011.08.022.
Collins, A.L., Levenson, J.M., Vilaythong, A.P., Richman, R., Armstrong, D.L., Noebels, J.L., Sweatt, J.D., and Zoghbi, H.Y. (2004). Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. Human Molecular Genetics *13*. https://doi.org/10.1093/hmg/ddh282.
Corbett, K.D., and Berger, J.M. (2004). Structure, molecular mechanisms, and evolutionary relationships in DNA topoisomerases. Annual Review of Biophysics and Biomolecular Structure *33*. https://doi.org/10.1146/annurev.biophys.33.110502.140357.
Crenshaw, D.G., and Hsieh, T.S. (1993). Function of the hydrophilic carboxyl terminus of type II DNA topoisomerase from Drosophila melanogaster. II. In vivo studies. Journal of Biological Chemistry *268*. .
Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: Ultrafast universal RNA-seq aligner. Bioinformatics *29*. https://doi.org/10.1093/bioinformatics/bts635.
Dutta, R., and Inouye, M. (2000). GHKL, an emergent ATPase/kinase superfamily. Trends in Biochemical Sciences *25*. https://doi.org/10.1016/S0968-0004(99)01503-0.
Ebert, D.H., Gabel, H.W., Robinson, N.D., Kastan, N.R., Hu, L.S., Cohen, S., Navarro, A.J., Lyst, M.J., Ekiert, R., Bird, A.P., et al. (2013). Activity-dependent phosphorylation of MeCP2 threonine 308 regulates interaction with NCoR. Nature *499*. https://doi.org/10.1038/nature12348.
Enikanolaiye, A., Ruston, J., Zeng, R., Taylor, C., Schrock, M., Buchovecky, C.M., Shendure, J., Acar, E., and Justice, M.J. (2020). Suppressor mutations in Mecp2-null mice implicate the DNA damage response in Rett syndrome pathology. https://doi.org/10.1101/gr.258400.119.
van Esch, H., Bauters, M., Ignatius, J., Jansen, M., Raynaud, M., Hollanders, K., Lugtenberg, D., Bienvenu, T., Jensen, L.R., Gécz, J., et al. (2005). Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. American Journal of Human Genetics *77*. https://doi.org/10.1086/444549.
Feng, W., Kawauchi, D., Körkel-Qu, H., Deng, H., Serger, E., Sieber, L., Ariel Lieberman, J., Jimeno-Gonzalez, S., Lambo, S., Hanna, B.S., et al. (2017). Chd7 is indispensable for mammalian brain development through activation of a neuronal differentiation programme. Nature Communications *8*. https://doi.org/10.1038/ncomms14758.
Gabel, H.W., Kinde, B., Stroud, H., Gilbert, C.S., Harmin, D.A., Kastan, N.R., Hemberg, M., Ebert, D.H., and Greenberg, M.E. (2015). Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. Nature *522*, 89–93. https://doi.org/10.1038/nature14319.
Gadelle, D., Filée, J., Buhler, C., and Forterre, P. (2003). Phylogenomics of type II DNA topoisomerases. BioEssays *25*. https://doi.org/10.1002/bies.10245.
Gilroy, K.L., and Austin, C.A. (2011). The impact of the C-Terminal domain on the interaction of human DNA topoisomerase II α and β with DNA. PLoS ONE *6*. https://doi.org/10.1371/journal.pone.0014693.
Gonzalez, R.E., Lim, C.U., Cole, K., Bianchini, C.H., Schools, G.P., Davis, B.E., Wada, I., Roninson, I.B., and Broude, E. v. (2011). Effects of conditional depletion of topoisomerase II on cell cycle progression in mammalian cells. Cell Cycle *10*. https://doi.org/10.4161/cc.10.20.17778.
Greig, L.C., Woodworth, M.B., Galazo, M.J., Padmanabhan, H., and Macklis, J.D. (2013). Molecular logic of neocortical projection neuron specification, development and diversity. Nature Reviews Neuroscience *14*. https://doi.org/10.1038/nrn3586.
Grue, P., Gräßer, A., Sehested, M., Jensen, P.B., Uhse, A., Straub, T., Ness, W., and Boege, F. (1998). Essential mitotic functions of DNA topoisomerase IIα are not adopted by topoisomerase IIβ in human H69 cells. Journal of Biological Chemistry *273*. https://doi.org/10.1074/jbc.273.50.33660.

Guo, J.U., Su, Y., Shin, J.H., Shin, J., Li, H., Xie, B., Zhong, C., Hu, S., Le, T., Fan, G., et al. (2014). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. Nature Neuroscience *17*. https://doi.org/10.1038/nn.3607.
Guy, J., Hendrich, B., Holmes, M., Martin, J.E., and Bird, A. (2001). A mouse Mecp2-null mutation causes neurological symptoms that mimic rett syndrome. Nature Genetics *27*. https://doi.org/10.1038/85899.
Habib, N., Li, Y., Heidenreich, M., Swiech, L., Avraham-Davidi, I., Trombetta, J.J., Hession, C., Zhang, F., and Regev, A. (2016). Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. Science (1979) *353*. https://doi.org/10.1126/science.aad7038.
Harkin, L.F., Gerrelli, D., Gold Diaz, D.C., Santos, C., Alzu'bi, A., Austin, C.A., and Clowry, G.J. (2016). Distinct expression patterns for type II topoisomerases IIA and IIB in the early foetal human telencephalon. Journal of Anatomy *228*. https://doi.org/10.1111/joa.12416.
Heckman, L.D., Chahrour, M.H., and Zoghbi, H.Y. (2014a). Rett-causing mutations reveal two domains critical for MeCP2 function and for toxicity in MECP2 duplication syndrome mice. Elife *3*. https://doi.org/10.7554/elife.02676.
Heckman, L.D., Chahrour, M.H., and Zoghbi, H.Y. (2014b). Rett-causing mutations reveal two domains critical for MeCP2 function and for toxicity in MECP2 duplication syndrome mice. Elife *3*. https://doi.org/10.7554/elife.02676.
Hiraide, T., Watanabe, S., Matsubayashi, T., Yanagi, K., Nakashima, M., Ogata, T., and Saitsu, H. (2020). A de novo TOP2B variant associated with global developmental delay and autism spectrum disorder. Molecular Genetics and Genomic Medicine *8*. https://doi.org/10.1002/mgg3.1145.
Holden, J.A., Rolfson, D.H., and Wíttwer, G.T. (1990). Human DNA Topoisomerase II: Evaluation of Enzyme Activity in Normal and Neoplastic Tissues. Biochemistry *29*. https://doi.org/10.1021/bi00460a024.
Huang, H.S., Allen, J.A., Mabb, A.M., King, I.F., Miriyala, J., Taylor-Blake, B., Sciaky, N., Dutton, J.W., Lee, H.M., Chen, X., et al. (2012). Topoisomerase inhibitors unsilence the dormant allele of Ube3a in neurons. Nature *481*. https://doi.org/10.1038/nature10726.
Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. Nature *515*. https://doi.org/10.1038/nature13908.
Ip, J.P.K., Mellios, N., and Sur, M. (2018). Rett syndrome: Insights into genetic, molecular and circuit mechanisms. Nature Reviews Neuroscience *19*. https://doi.org/10.1038/s41583-018-0006-3.
Jenkins, J.R., Ayton, P., Jones, T., Davies, S.L., Simmons, D.L., Harris, A.L., Sheer, D., and Hickson, I.D. (1992). Isolation of cDNA clones encoding the β isozyme of human DNA topoisomerase II and localisation of the gene to chromosome 3p24. Nucleic Acids Research *20*. https://doi.org/10.1093/nar/20.21.5587.
Jian, Z., Hamdan, H., Krishna, Y.H., Kaifang, P., E, P.A., Joanna, L., Yingyao, S., A, O.-P.J., Lifang, L., Wonho, K., et al. (2022). Disruption of MeCP2–TCF20 complex underlies distinct neurodevelopmental disorders. Proceedings of the National Academy of Sciences *119*, e2119078119. https://doi.org/10.1073/pnas.2119078119.
Jones, P.L., Veenstra, G.J.C., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nature Genetics *19*. https://doi.org/10.1038/561.
Juenke, J.E.M., and Holden, J.A. (1993). The distribution of DNA topoisomerase II isoforms in differentiated adult mouse tissues. BBA - Gene Structure and Expression *1216*. https://doi.org/10.1016/0167-4781(93)90144-3.

Jung, B.P., Jugloff, D.G.M., Zhang, G., Logan, R., Brown, S., and Eubanks, J.H. (2003). The expression of methyl CpG binding factor MeCP2 correlates with cellular differentiation in the developing rat brain and in cultured cells. Journal of Neurobiology *55*. https://doi.org/10.1002/neu.10201.
Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. Nature *465*. https://doi.org/10.1038/nature09033.
Kinde, B., Gabel, H.W., Gilbert, C.S., Griffith, E.C., and Greenberg, M.E. (2015). Reading the unique DNA methylation landscape of the brain: Non-CpG methylation, hydroxymethylation, and MeCP2. Proc Natl Acad Sci U S A *112*. https://doi.org/10.1073/pnas.1411269112.
Kinde, B., Wu, D.Y., Greenberg, M.E., and Gabel, H.W. (2016). DNA methylation in the gene body influences MeCP2-mediated gene repression. Proceedings of the National Academy of Sciences *113*, 15114–15119. https://doi.org/10.1073/PNAS.1618737114.
King, I.F., Yandava, C.N., Mabb, A.M., Hsiao, J.S., Huang, H.S., Pearson, B.L., Calabrese, J.M., Starmer, J., Parker, J.S., Magnuson, T., et al. (2013). Topoisomerases facilitate transcription of long genes linked to autism. Nature *501*, 58–62. https://doi.org/10.1038/nature12504.
Kokura, K., Kaul, S.C., Wadhwa, R., Nomura, T., Khan, M.M., Shinagawa, T., Yasukawa, T., Colmenares, C., and Ishii, S. (2001). The Ski Protein Family Is Required for MeCP2-mediated Transcriptional Repression. Journal of Biological Chemistry *276*. https://doi.org/10.1074/jbc.M105747200.
Kondapi, A.K., Mulpuri, N., Mandraju, R.K., Sasikaran, B., and Rao, K.S. (2004). Analysis of age dependent changes of Topoisomerase II α and β in rat brain. International Journal of Developmental Neuroscience *22*. https://doi.org/10.1016/j.ijdevneu.2003.10.006.
Koster, D.A., Croquette, V., Dekker, C., Shuman, S., and Dekker, N.H. (2005). Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB. Nature *434*. https://doi.org/10.1038/nature03395.
Kruusvee, V., Lyst, M.J., Taylor, C., Tarnauskaite, Ž., Bird, A.P., and Cook, A.G. (2017). Structure of the MeCP2-TBLR1 complex reveals a molecular basis for Rett syndrome and related disorders. Proc Natl Acad Sci U S A *114*. https://doi.org/10.1073/pnas.1700731114.
Lagger, S., Connelly, J.C., Schweikert, G., Webb, S., Selfridge, J., Ramsahoye, B.H., Yu, M., He, C., Sanguinetti, G., Sowers, L.C., et al. (2017). MeCP2 recognizes cytosine methylated tri-nucleotide and di-nucleotide sequences to tune transcription in the mammalian brain. PLoS Genetics *13*. https://doi.org/10.1371/journal.pgen.1006793.
Lam, C. wan, Yeung, W. lan, and Law, C. yiu (2017). Global developmental delay and intellectual disability associated with a de novo TOP2B mutation. Clinica Chimica Acta *469*. https://doi.org/10.1016/j.cca.2017.03.022.
Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods *9*. https://doi.org/10.1038/nmeth.1923.
Leonard, H., Cobb, S., and Downs, J. (2016). Clinical and biological progress over 50 years in Rett syndrome. Nature Reviews Neurology *13*. https://doi.org/10.1038/nrneurol.2016.186.
Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F., and Bird, A. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to Methylated DNA. Cell *69*. https://doi.org/10.1016/0092-8674(92)90610-O.
Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*. https://doi.org/10.1093/bioinformatics/btp324.
Li, W., and Wang, J.C. (1998). Mammalian DNA topoisomerase IIIα is essential in early

embryogenesis. Proc Natl Acad Sci U S A *95*. https://doi.org/10.1073/pnas.95.3.1010.
Li, R., Dong, Q., Yuan, X., Zeng, X., Gao, Y., Chiao, C., Li, H., Zhao, X., Keles, S., Wang, Z., et al. (2016). Misregulation of Alternative Splicing in a Mouse Model of Rett Syndrome. PLoS Genetics *12*. https://doi.org/10.1371/journal.pgen.1006129.
Li, Y., Hao, H., Tzatzalos, E., Lin, R.K., Doh, S., Liu, L.F., Lyu, Y.L.L., and Cai, L. (2014). Topoisomerase IIbeta is required for proper retinal development and survival of postmitotic cells. Biology Open *3*. https://doi.org/10.1242/bio.20146767.
Lim, L., Mi, D., Llorca, A., and Marín, O. (2018). Development and Functional Diversification of Cortical Interneurons. Neuron *100*. https://doi.org/10.1016/j.neuron.2018.10.009.
Lindsley, J.E., and Wang, J.C. (1991). Proteolysis patterns of epitopically labeled yeast DNA topoisomerase II suggest an allosteric transition in the enzyme induced by ATP binding. Proc Natl Acad Sci U S A *88*. https://doi.org/10.1073/pnas.88.23.10485.
Linka, R.M., Porter, A.C.G., Volkov, A., Mielke, C., Boege, F., and Christensen, M.O. (2007a). C-Terminal regions of topoisomerase IIα and IIβ determine isoform-specific functioning of the enzymes in vivo. Nucleic Acids Research *35*. https://doi.org/10.1093/nar/gkm102.
Linka, R.M., Porter, A.C.G., Volkov, A., Mielke, C., Boege, F., and Christensen, M.O. (2007b). C-Terminal regions of topoisomerase IIα and IIβ determine isoform-specific functioning of the enzymes in vivo. Nucleic Acids Research *35*. https://doi.org/10.1093/nar/gkm102.
Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang, Y., Dwork, A.J., Schultz, M.D., et al. (2013). Global epigenomic reconfiguration during mammalian brain development. Science (1979) *341*. https://doi.org/10.1126/science.1237905.
Lodato, S., and Arlotta, P. (2015). Generating Neuronal Diversity in the Mammalian Cerebral Cortex. Annual Review of Cell and Developmental Biology *31*. https://doi.org/10.1146/annurev-cellbio-100814-125353.
Lombardi, L.M., Baker, S.A., and Zoghbi, H.Y. (2015). MECP2 disorders: From the clinic to mice and back. Journal of Clinical Investigation *125*. https://doi.org/10.1172/JCI78167.
Loo, L., Simon, J.M., Xing, L., McCoy, E.S., Niehaus, J.K., Guo, J., Anton, E.S., and Zylka, M.J. (2019). Single-cell transcriptomic analysis of mouse neocortical development. Nature Communications *10*. https://doi.org/10.1038/s41467-018-08079-9.
Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology *15*. https://doi.org/10.1186/s13059-014-0550-8.
Lyst, M.J., and Bird, A. (2015). Rett syndrome: A complex disorder with simple roots. Nature Reviews Genetics *16*, 261–274. https://doi.org/10.1038/nrg3897.
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. Nature Neuroscience *16*. https://doi.org/10.1038/nn.3434.
Lyu, Y.L., and Wang, J.C. (2003). Aberrant lamination in the cerebral cortex of mouse embryos lacking DNA topoisomerase IIβ. Proc Natl Acad Sci U S A *100*. https://doi.org/10.1073/pnas.1232376100.
Lyu, Y.L., Lin, C.-P., Azarova, A.M., Cai, L., Wang, J.C., and Liu, L.F. (2006). Role of Topoisomerase IIβ in the Expression of Developmentally RegulatedGenes. Molecular and Cellular Biology *26*. https://doi.org/10.1128/mcb.00617-06.
Mabb, A.M., Judson, M.C., Zylka, M.J., and Philpot, B.D. (2011). Angelman syndrome: Insights into genomic imprinting and neurodevelopmental phenotypes. Trends in Neurosciences *34*. https://doi.org/10.1016/j.tins.2011.04.001.

Mabb, A.M., Kullmann, P.H.M., Twomey, M.A., Miriyala, J., Philpot, B.D., and Zylka, M.J. (2014). Topoisomerase 1 inhibition reversibly impairs synaptic function. Proc Natl Acad Sci U S A *111*, 17290–17295. https://doi.org/10.1073/pnas.1413204111.
Madabhushi, R., Gao, F., Pfenning, A.R., Pan, L., Yamakawa, S., Seo, J., Rueda, R., Phan, T.X., Yamakawa, H., Pao, P.C., et al. (2015). Activity-Induced DNA Breaks Govern the Expression of Neuronal Early-Response Genes. Cell *161*. https://doi.org/10.1016/j.cell.2015.05.032.
McCoy, M.J., and Fire, A.Z. (2020). Intron and gene size expansion during nervous system evolution. BMC Genomics *21*. https://doi.org/10.1186/s12864-020-6760-4.
McCoy, M.J., Paul, A.J., Victor, M.B., Richner, M., Gabel, H.W., Gong, H., Yoo, A.S., and Ahn, T.H. (2018). LONGO: An R package for interactive gene length dependent analysis for neuronal identity. In Bioinformatics, p.
Mckinnon, P.J. (2016). Topoisomerases and the regulation of neural function. Nature Reviews Neuroscience *17*. https://doi.org/10.1038/nrn.2016.101.
Meczes, E.L., Gilroy, K.L., West, K.L., and Austin, C.A. (2008). The impact of the human DNA topoisomerse II C-terminal domain on activity. PLoS ONE *3*. https://doi.org/10.1371/journal.pone.0001754.
Mejia, L.A., Litterman, N., Ikeuchi, Y., de La Torre, L.U., Bennett, E.J., Zhang, C., Wade Harpe, J., and Bonni, A. (2013). A Novel Hap1-Tsc1 interaction regulates neuronal mTORC1 signaling and morphogenesis in the brain. Journal of Neuroscience *33*. https://doi.org/10.1523/JNEUROSCI.2290-13.2013.
de Mendoza, A., Poppe, D., Buckberry, S., Pflueger, J., Albertin, C.B., Daish, T., Bertrand, S., de la Calle-Mustienes, E., Gómez-Skarmeta, J.L., Nery, J.R., et al. (2021). The emergence of the brain non-CpG methylation system in vertebrates. Nature Ecology and Evolution *5*, 369–378. https://doi.org/10.1038/S41559-020-01371-2.
Meyer, K.N., Kjeldsen, E., Straub, T., Knudsen, B.R., Hickson, I.D., Kikuchi, A., Kreipe, H., and Boege, F. (1997). Cell cycle-coupled relocation of types I and II topoisomerases and modulation of catalytic enzyme activities. Journal of Cell Biology *136*. https://doi.org/10.1083/jcb.136.4.775.
Miller, J.A., Ding, S.L., Sunkin, S.M., Smith, K.A., Ng, L., Szafer, A., Ebbert, A., Riley, Z.L., Royall, J.J., Aiona, K., et al. (2014). Transcriptional landscape of the prenatal human brain. Nature *508*. https://doi.org/10.1038/nature13185.
Mullaney, B.C., Johnston, M. v., and Blue, M.E. (2004). Developmental expression of methyl-CpG binding protein 2 is dynamically regulated in the rodent brain. Neuroscience *123*. https://doi.org/10.1016/j.neuroscience.2003.11.025.
Nagy, L., Kao, H.Y., Chakravarti, D., Lin, R.J., Hassig, C.A., Ayer, D.E., Schreiber, S.L., and Evans, R.M. (1997). Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase. Cell *89*. https://doi.org/10.1016/S0092-8674(00)80218-4.
Nan, X., Meehan, R.R., and Bird, A. (1993a). Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. Nucleic Acids Research *21*. https://doi.org/10.1093/nar/21.21.4886.
Nan, X., Meehan, R.R., and Bird, A. (1993b). Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. Nucleic Acids Research *21*. https://doi.org/10.1093/nar/21.21.4886.
Nan, X., Tate, P., Li, E., and Bird, A. (1996). DNA methylation specifies chromosomal localization of MeCP2. Molecular and Cellular Biology *16*. https://doi.org/10.1128/mcb.16.1.414.
Nan, X., Campoy, F.J., and Bird, A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell *88*, 471–481. https://doi.org/10.1016/S0092-

8674(00)81887-5.
Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386–389. https://doi.org/10.1038/30764.
Neul, J.L., Fang, P., Barrish, J., Lane, J., Caeg, E.B., Smith, E.O., Zoghbi, H., Percy, A., and Glaze, D.G. (2008). Specific mutations in Methyl-CpG-Binding Protein 2 confer different severity in Rett syndrome. Neurology *70*. https://doi.org/10.1212/01.wnl.0000291011.54508.aa.
Neul, J.L., Kaufmann, W.E., Glaze, D.G., Christodoulou, J., Clarke, A.J., Bahi-Buisson, N., Leonard, H., Bailey, M.E.S., Schanen, N.C., Zappella, M., et al. (2010). Rett syndrome: Revised diagnostic criteria and nomenclature. Annals of Neurology *68*. https://doi.org/10.1002/ana.22124.
Nitiss, J.L. (2009). DNA topoisomerase II and its growing repertoire of biological functions. Nature Reviews Cancer *9*, 327–337. https://doi.org/10.1038/nrc2608.
Noctor, S.C., Martinez-Cerdeño, V., Ivic, L., and Kriegstein, A.R. (2004). Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. Nature Neuroscience *7*. https://doi.org/10.1038/nn1172.
Nowakowski, T.J., Bhaduri, A., Pollen, A.A., Alvarado, B., Mostajo-Radji, M.A., di Lullo, E., Haeussler, M., Sandoval-Espinosa, C., Liu, S.J., Velmeshev, D., et al. (2017). Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. Science (1979) *358*. https://doi.org/10.1126/science.aap8809.
Osheroff, N. (1989). Effect of Antineoplastic Agents on the DNA Cleavage/Religation Reaction of Eukaryotic Topoisomerase II: Inhibition of DNA Religation by Etoposide. Biochemistry *28*. https://doi.org/10.1021/bi00441a005.
Percy, A.K., Neul, J.L., Glaze, D.G., Motil, K.J., Skinner, S.A., Khwaja, O., Lee, H.S., Lane, J.B., Barrish, J.O., Annese, F., et al. (2010). Rett syndrome diagnostic criteria: Lessons from the Natural History Study. Annals of Neurology *68*. https://doi.org/10.1002/ana.22154.
Pommier, Y., Sun, Y., Huang, S.Y.N., and Nitiss, J.L. (2016). Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. Nature Reviews Molecular Cell Biology *17*. https://doi.org/10.1038/nrm.2016.111.
Powell, W.T., Coulson, R.L., Gonzales, M.L., Crary, F.K., Wong, S.S., Adams, S., Ach, R.A., Tsang, P., Yamada, N.A., Yasui, D.H., et al. (2013). R-loop formation at Snord116 mediates topotecan inhibition of Ube3a-antisense and allele-specific chromatin decondensation. Proc Natl Acad Sci U S A *110*. https://doi.org/10.1073/pnas.1305426110.
Quinlan, A.R., and Hall, I.M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics *26*. https://doi.org/10.1093/bioinformatics/btq033.
Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. Nature *447*. https://doi.org/10.1038/nature05918.
Renthal, W., Boxer, L.D., Hrvatin, S., Li, E., Silberfeld, A., Nagy, M.A., Griffith, E.C., Vierbuchen, T., and Greenberg, M.E. (2018). Characterization of human mosaic Rett syndrome brain tissue by single-nucleus RNA sequencing. Nature Neuroscience *21*. https://doi.org/10.1038/s41593-018-0270-6.
Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*. https://doi.org/10.1093/bioinformatics/btp616.
Roca, J. (2009). Topoisomerase II: A fitted mechanism for the chromatin landscape. Nucleic Acids Research *37*. https://doi.org/10.1093/nar/gkn994.
Roca, J., and Mezquita, C. (1989). DNA topoisomerase II activity in nonreplicating, transcriptionally

inactive, chicken late spermatids. EMBO Journal *8*. https://doi.org/10.1002/j.1460-2075.1989.tb03581.x.
Ronan, J.L., Wu, W., and Crabtree, G.R. (2013). From neural development to cognition: Unexpected roles for chromatin. Nature Reviews Genetics *14*. https://doi.org/10.1038/nrg3413.
de Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. Nature *515*. https://doi.org/10.1038/nature13772.
Sano, K., Miyaji-Yamaguchi, M., Tsutsui, K.M., and Tsutsui, K. (2008). Topoisomerase IIβ activates a subset of neuronal genes that are repressed in AT-rich genomic environment. PLoS ONE *3*. https://doi.org/10.1371/journal.pone.0004103.
Shahbazian, M.D., Antalffy, B., Armstrong, D.L., and 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*. https://doi.org/10.1093/hmg/11.2.115.
Skene, P.J., Illingworth, R.S., Webb, S., Kerr, A.R.W., James, K.D., Turner, D.J., Andrews, R., and Bird, A.P. (2010). Neuronal MeCP2 Is Expressed at Near Histone-Octamer Levels and Globally Alters the Chromatin State. Molecular Cell *37*. https://doi.org/10.1016/j.molcel.2010.01.030.
Sowa, M.E., Bennett, E.J., Gygi, S.P., and Harper, J.W. (2009). Defining the Human Deubiquitinating Enzyme Interaction Landscape. Cell *138*. https://doi.org/10.1016/j.cell.2009.04.042.
Stancheva, I., Collins, A.L., van den Veyver, I.B., Zoghbi, H., and Meehan, R.R. (2003). A mutant form of MeCP2 protein associated with human Rett syndrome cannot be displaced from methylated DNA by Notch in Xenopus embryos. Molecular Cell *12*. https://doi.org/10.1016/S1097-2765(03)00276-4.
Stewart, L., Redinbo, M.R., Qiu, X., Hol, W.G.J., and Champoux, J.J. (1998). A model for the mechanism of human topoisomerase I. Science (1979) *279*. https://doi.org/10.1126/science.279.5356.1534.
Sugino, K., Hempel, C.M., Okaty, B.W., Arnson, H.A., Kato, S., Dani, V.S., and Nelson, S.B. (2014). Cell-Type-Specific Repression by Methyl-CpG-Binding Protein 2 Is Biased toward Long Genes. The Journal of Neuroscience *34*, 12877 LP – 12883. https://doi.org/10.1523/JNEUROSCI.2674-14.2014.
Sugino, K., Clark, E., Schulmann, A., Shima, Y., Wang, L., Hunt, D.L., Hooks, B.M., Trä nkner, D., Chandrashekar, J., Picard, S., et al. Mapping the transcriptional diversity of genetically and anatomically defined cell populations in the mouse brain. https://doi.org/10.7554/eLife.38619.001.
Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nature Neuroscience *19*. https://doi.org/10.1038/nn.4216.
Tillotson, R., Selfridge, J., Koerner, M. v., Gadalla, K.K.E., Guy, J., de Sousa, D., Hector, R.D., Cobb, S.R., and Bird, A. (2017). Radically truncated MeCP2 rescues Rett syndrome-like neurological defects. Nature *550*. https://doi.org/10.1038/nature24058.
Tiscornia, G., Singer, O., and Verma, I.M. (2006). Production and purification of lentiviral vectors. Nature Protocols *1*. https://doi.org/10.1038/nprot.2006.37.
Tiwari, V.K., Burger, L., Nikoletopoulou, V., Deogracias, R., Thakurela, S., Wirbelauer, C., Kaut, J., Terranova, R., Hoerner, L., Mielke, C., et al. (2012). Target genes of Topoisomerase IIβ regulate neuronal survival and are defined by their chromatin state. Proc Natl Acad Sci U S A *109*. https://doi.org/10.1073/pnas.1119798109.
Tsutsui, K., Tsutsui, K., Okada, S., Watanabe, M., Shohmori, T., Seki, S., and Inoue, Y. (1993).

Molecular cloning of partial cDNAs for rat DNA topoisomerase II isoforms and their differential expression in brain development. Journal of Biological Chemistry *268*. https://doi.org/10.1016/s0021-9258(17)46737-8.
Tsutsui, K., Tsutsui, K., Sano, K., Kikuchi, A., and Tokunaga, A. (2001). Involvement of DNA Topoisomerase IIβ in Neuronal Differentiation. Journal of Biological Chemistry *276*. https://doi.org/10.1074/jbc.M008517200.
Wang, J.C. (1996). DNA topoisomerases. Annual Review of Biochemistry *65*. https://doi.org/10.1146/annurev.bi.65.070196.003223.
Wang, J.C. (1998). Moving one DNA double helix through another by a type II DNA topoisomerase: The story of a simple molecular machine. Quarterly Reviews of Biophysics *31*. https://doi.org/10.1017/S0033583598003424.
Wang, J.C. (2002). Cellular roles of DNA topoisomerases: A molecular perspective. Nature Reviews Molecular Cell Biology *3*. https://doi.org/10.1038/nrm831.
Watanabe, M., Tsutsui, K., Tsutsui, K., and Inoue, Y. (1994). Differential expressions of the topoisomerase IIα and IIβ mRNAs in developing rat brain. Neuroscience Research *19*. https://doi.org/10.1016/0168-0102(94)90007-8.
Watt, P.M., and Hickson, I.D. (1994). Structure and function of type II DNA topoisomerases. Biochemical Journal *303*. https://doi.org/10.1042/bj3030681.
Wei, P.C., Chang, A.N., Kao, J., Du, Z., Meyers, R.M., Alt, F.W., and Schwer, B. (2016). Long Neural Genes Harbor Recurrent DNA Break Clusters in Neural Stem/Progenitor Cells. Cell *164*. https://doi.org/10.1016/j.cell.2015.12.039.
Woessner, R.D., Mattern, M.R., Mirabelli, C.K., Johnson, R.K., and Drake, F.H. (1991). Proliferation- and cell cycle-dependent differences in expression of the 170 kilodalton and 180 kilodalton forms of topoisomerase II in NIH-3T3 cells. Cell Growth Differ *2*. .
Wu, C.C., Li, T.K., Farh, L., Lin, L.Y., Lin, T.S., Yu, Y.J., Yen, T.J., Chiang, C.W., and Chan, N.L. (2011). Structural basis of type II topoisomerase inhibition by the anticancer drug etoposide. Science (1979) *333*. https://doi.org/10.1126/science.1204117.
Yang, X. (2000). DNA topoisomerase IIβ and neural development. Science (1979) *287*. https://doi.org/10.1126/science.287.5450.131.
Yang, Y., Kucukkal, T.G., Li, J., Alexov, E., and Cao, W. (2016). Binding Analysis of Methyl-CpG Binding Domain of MeCP2 and Rett Syndrome Mutations. ACS Chemical Biology *11*. https://doi.org/10.1021/acschembio.6b00450.
Yusufzai, T.M., and Wolffe, A.P. (2000). Functional consequences of Rett syndrome mutations on human MeCP2. Nucleic Acids Research *28*. https://doi.org/10.1093/nar/28.21.4172.
Zandvliet, D.W.J., Hanby, A.M., Austin, C.A., Marsh, K.L., Clark, I.B.N., Wright, N.A., and Poulsom, R. (1996). Analysis of foetal expression sites of human type II DNA topoisomerase a and β mRNAs by in situ hybridisation. Biochimica et Biophysica Acta - Gene Structure and Expression *1307*. https://doi.org/10.1016/0167-4781(96)00063-2.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nussbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biology *9*. https://doi.org/10.1186/gb-2008-9-9-r137.
Zhou, Z., Hong, E.J., Cohen, S., Zhao, W. ning, Ho, H. yi H., 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*. https://doi.org/10.1016/j.neuron.2006.09.037.
Zoghbi, H.Y. (2016). Rett Syndrome and the Ongoing Legacy of Close Clinical Observation. Cell

167. https://doi.org/10.1016/j.cell.2016.09.039.
Zylka, M.J., Simon, J.M., and Philpot, B.D. (2015). Gene length matters in neurons. Neuron *86*, 353–355. https://doi.org/10.1016/j.neuron.2015.03.059.