Identifying Protein Interaction Networks for MeCP2 using Proximity Labelling

Jessica Peixinho

Master of Science by Research (MRes) in Biology

University of York

Biology

January 2020

Abstract

Mutations in methyl-CpG binding protein 2 (MeCP2) have been attributed as the cause of Rett syndrome (RTT). RTT is an X-linked neurological disorder affecting approximately 1/10000 females. MeCP2 is a ubiquitously expressed protein and its main function is to act as a transcriptional repressor via association with NCoR/SMRT complex. However, MeCP2 has various reported functions and interacting partners. For this reason, the exact role of MeCP2 in RTT, its function and acting mechanism are not yet fully known.

In order to address this, the aim of this thesis was to shed light on the protein interactions of MeCP2 in a novel way, using proximity labelling (PL). Here, two promiscuous ligases, TurboID and miniTurbo, were fused to the C-Terminus of MeCP2 to enable biotinylation of proximal proteins. Mouse fibroblast cells were transfected with MeCP2-TurboID and MeCP2-miniTurbo and exposed to biotin. The high affinity of biotin to streptavidin allows biotinylated proteins to be isolated and identified.

Firstly, I have shown that fusion in the C-Terminus of MeCP2 does not affect its ability to localise to the nucleus. Additionally, supplementation with biotin allowed the visualisation of biotinylated proteins in mouse fibroblast cells. I have confirmed that protein biotinylation using the named ligases can occur in as little as 10 minutes. More exogenous biotinylation occurred in TurboID transfected cells than miniTurbo transfected cells. A cell line has been produced of MeCP2-TurboID. Overall, PL appears to be a promising avenue of research for identifying protein partners of MeCP2. In the future, this will aid more detailed insights into molecular mechanisms of MeCP2 and, indeed, implications on RTT.

Contents

List of Figures

List of Tables

Author’s declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as references.

Acknowledgements

Firstly, I would like to thank my family, my partner, and my friends for their continued support and encouragement throughout this project. Additionally, I also want to thank them for their understanding of my ongoing absence from London throughout my time in York completing my project.

Additionally, I would like to thank all the members of the Goffin lab and my very talented lab mates who taught me lots. My time at York would have not been the same without their input. I'd also like to thank my supervisor for sharing some of his knowledge and expertise.

1-Introduction

1.1. DNA methylation

The vertebrate genome is globally methylated (Bird and Taggart.,1980). DNA methylation is a common epigenetic mark whereby a methyl group is added to the 5' position of cytosine residues (Fig 1.1). DNA methylation is essential for controlling gene expression amongst various other functions (Bird., 2002; Baylin and Ohm., 2006). In mammals, this modification occurs predominantly at CpG dinucleotides. Approximately 70% of all CpG dinucleotides are methylated in mammals, with the exception of CpG islands (CGIs), CG-rich regions mostly coinciding with the promoter of protein-coding genes. CpG islands may become hypermethylated, leading to gene silencing. In this context, methyl-CpG binding proteins are vital in the process of DNA methylation-dependent gene silencing (Ballestar and Wolffe 2001).

DNA methylation has also been found to occur at non-CpG sites, particularly at CpA dinucleotides *in vitro* and the mouse brain (Lister et al.,2013; Guo et al.,2014; Xie et al.,2012), and contributes to the neuronal methylome (Kinde et al.,2016). The density of non-CpG methylation increases during neuronal maturation, at a time in development when MeCP2 levels are increasing, becoming the most predominant form of methylation in adult neurons (Lister et al.,2013; Guo et al.,2014; Xie et al.,2012). This indicates a possible role in MeCP2 recruitment to chromatin.

Another DNA modification, 5- hydroxymethylcytosine (5hmC) has gained considerable scientific attention. This modification arises from 5 methylcytosine by conversion catalysed by TET (ten eleven translocation) proteins (Tahiliani et al.,2009).. This DNA modification has been associated with gene enhancers, transcriptional start sites, and gene bodies (Stroud et al.,2011). Its roles in gene regulation and DNA methylation are still to be determined. Current suggestions depict that 5mC to 5hmC may facilitate DNA demethylation via the exclusion DNMT1- a maintenance DNA methyltransferase (Tahiliani et al.,2009).

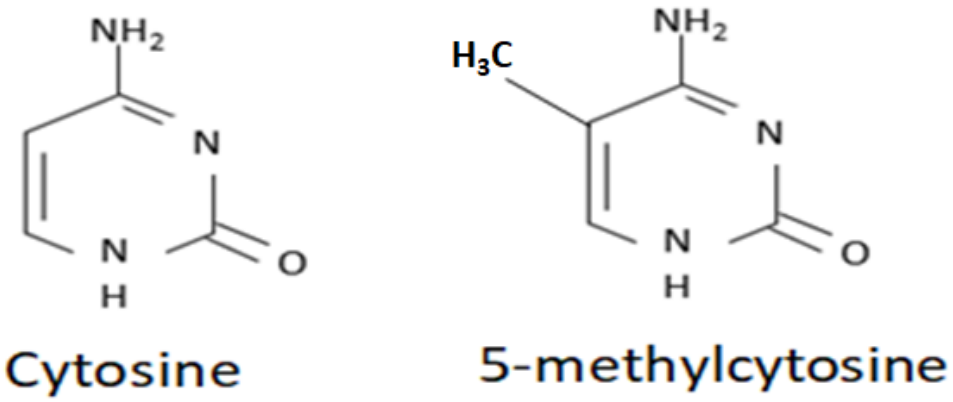


Figure 1.1 Chemical structure of cytosine and its modified form.

1.2. MeCP2

MeCP2 was characterised by its ability to bind to methylated DNA (Lewis et al.,1992; Meehan et al.,1992). In humans, *MeCP2*, a single gene on the X chromosome at the locus Xq28, codes for the 486 amino acid long, chromatin associated MeCP2 protein with established domains. Cytosine is frequently methylated and it has been noted that Methyl-CpG often interferes with transcription. While MeCP2's precise molecular function remains unclear, it is known that MeCP2 binds to methylated cytosine residues (Lewis et al.,1992) and that mutations in this gene underlie Rett syndrome (RTT).

MeCP2 is ubiquitously expressed, however, it is expressed at higher levels in the brain. Studies have determined the absolute abundance of MeCP2 in unsorted nuclei isolated from mature mouse brains as $16x10^{6}$ molecules per nucleus compared to $0.5x10^{6}$ molecules per hepatic nuclei (Skene et al.,2010). This is comparable to the number of nucleosomes, approximately $30x10^{6}$, suggesting MeCP2's fundamental role in neurons.

MeCP2 is known to occur in 2 isoforms, or variants, with the main difference between these residing in the N-terminus (Kriaucionis and Bird., 2004). One variant, MeCP2_e2, consists of 4 exons, with the translational domain present in exon 2, gives a protein of 486 amino acids (aa; Kriaucionis and Bird., 2004). The second variant, MeCP2_e1 omits exon 2 and gives a protein of 488 aa. Both isoforms of the protein are expressed in all tissues and function in the same way (Kriaucionis and Bird., 2004). Indeed, both protein isoforms are nuclear and colocalise with methylated heterochromatic foci in mice (Kriaucionis and Bird., 2004)

MeCP2 is believed to contain two main functional domains: the Methyl-CpG Binding Domain (MBD), responsible for binding to methylated DNA (Nan et al.,1993), as well as the NCoR interaction domain (NID; Lyst et al.,2013)

1.3 Rett syndrome

RTT is an X-linked neurological disorder, affecting approximately 1 in 10,000 live females at birth (Rett, 1966). Patients are characterised by the presence of a partial or complete loss of acquired purposeful hand skills, partial or complete loss of acquired spoken language, gait abnormalities, and development of stereotypic hand movements such as hand wringing or washing (Neul et al.,2010). Patients develop normally for the first 6-18 months of life, followed by regression in purposeful hand use and spoken language, gait abnormalities, and hand stereotypies (Neul et al.,2010). RTT is often associated with deceleration in head growth, although this is not required for diagnosis.

MeCP2 mutations are well established as being the main cause of RTT (Amir et al.,1999) and 80% of classic cases are caused by mutations in *MeCP2* (Cheadle et al.,2000). Sporadic mutations are of *MeCP2* are almost exclusively of paternal origin, explaining the high occurrence in females (Trappe et al.,2001; Cuddapah et al.,2014; Girard et al.,2001). It had been suspected that RTT causing mutations led to embryonic fatality in males, however, males with RTT have been reported (Schwartzman et al.,2001; Vorsanova et al.,2001; Neul et al.,2018). Mutations in males may arise from maternal inheritance or from mutations occurring in the oocytes (Schanen., 2001). Overall, there is greater clinical variation in males, complicating diagnosis. Furthermore, phenotypes are often more severe in males, leading to recent study including male encephalopathy as a diagnostic entity (Neul et al., 2018; Neul et al., 2014). RTT has been identified in a comorbid state, affecting boys who, generally, fall into two categories; those with 47XXY, also known as Klinefelter syndrome, or those with somatic mosaicism (Schwartzman et al.,2001).

1.4 Methyl-CpG-Binding Proteins

In 1984, the first evidence of a protein with specificity for methylated DNA was established (Huang et al.,1984). The first protein to be identified by its ability to bind to methylated DNA was named methyl-CpG binding protein 1 (MeCP1; Meehan et al.,1989). This was followed by the identification of MeCP2 (Lewis et al.,1992; Meehan et al.,1992).

The methyl-binding domain (MBD) was characterized by deletion studies of MeCP2 and was identified as the minimal domain responsible for interaction with methylated cytosines (Nan et al.,1993). The MBD family of proteins takes its definition from the MBD and include MBD1, MBD2, MBD3, and MBD4. This group of proteins is so named due to the fact that they contain a methyl-CpG-binding domain resembling that of MeCP2 (Hendrich and Bird, 1998). MBD1, MBD2, and MBD4, similarly to MeCP2, can bind to DNA containing only one symmetrically methylated CpG site and can localise to heterochromatin in transfected cells (Hendrich and Bird., 1998; Fujita et al.,1999; Ng et al.,1999). All MBD proteins, apart from MBD4, have been determined to be associated with histone deacetylase subunits as part of large complexes.

Mutations in the MBD domain have been associated with MeCP2's decreased time at heterochromatic foci, as assessed by fluorescence recovery after photobleaching (FRAP; Kumar et al.,2008, Schmiedeberg et al.,2009). This suggests that binding to DNA is impaired by mutations in the MBD *in vivo*.

1.5 MeCP2 as a transcriptional repressor

DNA methylation is typically associated with gene repression (Ballestar et al., 2001). The domain responsible for transcriptional repression, the transcriptional repressor domain (TRD) was mapped to a specific region of the protein between amino acids 207 and 310 (Nan et al.,1997). It was found that native and recombinant MeCP2 repress transcription *in vitro* from methylated promoters but do not repress transcription at unmethylated promoters (Nan et al.,1997). Moreover, the proposed mechanism by which MeCP2 acts as a transcriptional repressor was based on the fact that biochemical studies showed its interaction with histone deacetylase (HDAC) containing complexes Sin3A (Nan et al.,1998). Given the strong correlation between deacetylation of histone protein tails and transcriptional repression (Grunstein, 1997), these data suggest that MeCP2 represses gene transcription. In line with this finding it was observed that transcriptional repression of transporter genes was relieved by treatment with trichostatin A (TSA), a HDAC inhibitor (Yoshida et al.,1990). This suggested that this complex promoted deacetylation of chromatin, leading to gene repression (Nan et al.,1998). In studies of mice lacking *Mecp2* showed transcriptional alterations that could not be explained with this simple model of interaction.

Various mechanisms of how MeCP2 leads to transcriptional repression have been proposed. These are based on the detection of the binding of MeCP2 to various co-repressor proteins such as, nuclear receptor co-repressor (NCoR), and silencing mediator of retinoic acid and thyroid hormone receptor (SMRT; Kokura et al.,2001; Stancheva et al.,2003). Indeed, the NCoR/SMRT

co-repressor complex was shown to bind between amino acids 269-309- a region known as a hotspot for missense mutations, termed the NCoR/SMRT Interaction Domain (NID). Mutations in the NID abolished the interactions between MeCP2 and the NCoR-SMRT complex (Lyst et al.,2013). It is of note that in this study, the weak Sin3A binding was unaffected by NID mutations (Lyst et al , 2013) bringing into question the relevance of this co-repressor interaction in RTT and hence highlighting that the model of MeCP2 function is incomplete.

1.6 How MeCP2 mutations cause RTT

To properly explain RTT pathology, a model of MeCP2 function must explain how the same syndrome can arise from null and missense mutations in two main regions of the protein; Methyl-CpG Binding Domain (MBD) and the NID (Lyst et al.,2013). One model of MeCP2 function suggests that the main function of the protein is to act as a bridge between NCoR-SMRT and methylated DNA. Any of the three possible types of mutations (Null, MBD, and NID) would likely prevent the recruitment of NCoR-SMRT to chromatin by MeCP2 and hence inhibit transcriptional repression. Amongst the most frequent mutations are the following missense mutations; T158M and R306C, occurring in the MBD and the NID respectively (Figure 1.2) with T158M being more severe (Neul et al.,2008). An alternative model of MeCP2 function proposed that one of the main functions of MeCP2 is modulation of chromatin architecture via multiple interactions with DNA. The MBD is essential for MeCP2's ability to bind to 5-methylcytosine and mediates binding to methylated CpG dinucleotides found in heterochromatic regions of the chromosome (Lewis et al.,1992; Nan et al.,1993). Thus, mutations in the MBD domain impair binding to methylated DNA.

An alternative model of MeCP2 function is that the primary function of MeCP2 in chromatin remodelling. The NID has been proposed as having its primary function as a DNA binding domain (Heckman et al.,2014). Indeed, the R306C mutation present in this domain is reported to impair the interaction between MeCP2 and NCoR/SMRT in the brain (Lyst et al.,2013). If this model is correct, all disease causing mutations affecting the NID would abolish the interaction with methylated DNA.

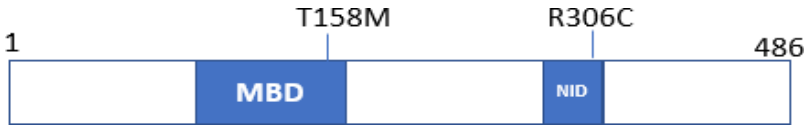


Figure 1.2. Rett syndrome mutations occur in the Methyl Binding Domain (MBD) and NCoR/SMRT Interaction Domain (NID). Schematic of the two main MeCP2 domains, the MBD and the NID.

1.7 Animal models of RTT

Animal models have proved invaluable for furthering the understanding of MeCP2 function. An early experiment conducted prior to the association of MeCP2 with RTT, aimed to establish the role of DNA methylation in mouse development. *MeCP2* was mutated in male mice embryonic stem (ES) cells using a lacZ reporter gene (Tate et al.,1996). Mutant ES cells lacking MeCP2 grew as well as parental lines, however, the creation of *Mecp2*-null mice from these ES cells was not possible due to embryonic lethality (Tate et al.,1996). Later, whole brain deletion of MeCP2 by Nestin-Cre led to a phenotype similar to that of a whole body knock-out (Chen et al.,2001; Guy et al.,2001). *Mecp2* deletion in post-mitotic neurons in the forebrain, hippocampus, and brainstem, but not in the cerebellum or glial cells, led to a delayed and less severe phenotype (Chen et al.,2001). This is suggestive of MeCP2's role in postmitotic, mature neurons rather than in early brain development.

Knockin mice recapitulating the T158A mutation exhibited various RTT like symptoms including developmental regression, motor dysfunction and learning deficits (Goffin et al., 2011). Interestingly, mice with MeCP2 duplication syndrome, overexpressing MeCP2, show similar RTT like symptoms as in models lacking MeCP2 (Na et al., 2012). Similarly, studies on mouse brains have revealed that protein depletion and gain of function both compromise cell morphology, neurotransmission, and processes supporting memory and learning (Na et al.,2013). These data suggest that both loss and gain of MeCP2 function result in the exhibition of similar symptoms and, thus, that the level of MeCP2 must be carefully regulated.

Considering that RTT is a neurological disorder and MeCP2's high abundance in the brain, it is important to acknowledge MeCP2 in a neural context. In a study that purified the MeCP2 from the brains of *Mecp2-EGFP* knockin mice, MeCP2 was found to interact with transducin-beta like factor 1 (TBL1; Lyst et al.,2013). A peptide comprising of MeCP2 residues bound directly to N-terminal regions of NCoR1 and SMRT and their shared homodimeric subunits of TBL1 and TBLR1 (Fig 1.3;Lyst et al.,2013), supporting the presence of MeCP2 binding sites on NCoR/SMRT complexes. Studies have shown that polymorphisms elsewhere in the protein are found in the general population, suggesting that other protein regions may be dispensable. Conversely, MeCP2 is almost abundant enough in the genome to coat the chromosome (Skene et al.,2010) hence the amount of MeCP2 supersedes the number of NCoR repressors, contradicting its proposed function.

Importantly, activation of a silenced *Mecp2* allele after phenotype presentations reverses these RTT-like symptoms (Guy et al., 2007; Robinson Ebert et al., 2013). This has raised the hope

that RTT is a curable disorder. But, we still have little understanding of the mechanisms by which loss of. Overall, these data suggest that RTT phenotypes can be rescued, at least in mice, highlighting the possibility for treatment and hence the need for further understanding of the etiology of RTT.

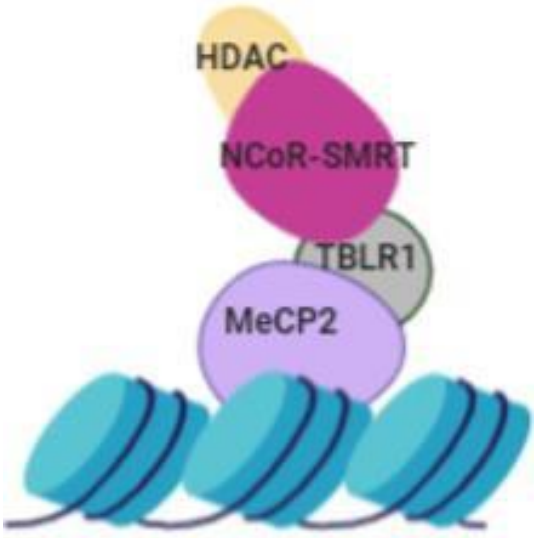


Figure1.3 The NCoR complex interacts with MeCP2 via a TBLR1 subunit. Illustrative schematic of depicting MeCP2 bound to heterochromatin and associated with NCoR/SMRT via TBLR1. This association leads to transcriptional repression (Lyst et al.,2013). Created with biorender.

1.8. MeCP2 as a transcriptional activator

MeCP2 has long been considered a transcriptional repressor (Nan et al*.,* 1997), however, recently its ability to act as both a transcriptional activator and repressor has been suggested. Initially, global expression profiling, using microarrays to measure gene activity throughout the genome, did not reveal significant differences in gene expression patterns in the brains of *MeCP2*- null mice compare to wild-types (Tudor et al*.,*2002). Subsequent analyses of specific brain regions, such as the hypothalamus, in *MeCP2*-null mice and mice overexpressing MeCP2 showed that numerous genes are consistently and reciprocally dysregulated, although subtly (Ben-Shachar et al*.,*2009; Chahrour et al*.,*2008). It was shown that more genes were downregulated than upregulated in the hypothalamus of *MeCP2* knock-out mice. The study showed that the majority of downregulated genes in *MeCP2*- null mice showed changes in patterns of expression and were rather upregulated and vice versa (Chahrour et al*.,*2008).

The mechanism proposed for MeCP2's role in transcriptional activation is based on MeCP2's recruitment of cyclic-AMP responsive element binding protein 1 (CREB1; Figure 1.4) to target gene promoters to activate transcription (Chahrour et al.,2008). However, the evidence for this is not compelling as experiments have not been replicated by others.

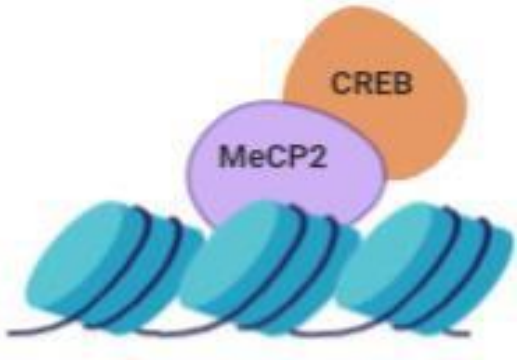


Figure 1.4. MeCP2 can function as a transcriptional activator when bound to CREB. Binding to CREB activates transcription. Adapted from Chahrour et al.,2008. Created with Biorender.

1.9. Other functions of MeCP2

1.9.1. MeCP2 in chromatin remodelling

Apart from the protein's well established role as a transcriptional repressor, MeCP2 has reported involvement in various other mechanisms. MeCP2 has been reported to be involved in controlling chromatin structure (Zlatanova ., 2005; Chadwick and Wade., 2007). MeCP2 is able to modulate chromatin architecture in a number of ways; by condensing DNA through regulating long range interactions (Horike et al.,2005), the formation of higher order chromatin structures (Georgel et al.,2003; Agarwal et al.,2011), and the formation of chromatin loops and DNA bridges (Georgel et al.,2003; Nikitina et al.,2007). MeCP2's chromatin associated role is evidenced through its localisation to DAPI stained heterochromatin regions of the nuclei (Zachariah et al.,2012; Craig et al.,2003).

1.9.2 Regulation of chromatin structure by replacement of histone H1

MeCP2's function in regulating chromatin structure may occur through displacement of histone H1. Experiments involving *in vitro* assembled protein highlighted that MeCP2 was able to compete with histone H1 (Nan et al., 1997). Furthermore, it has been observed that these proteins seem

to compete for the same substrates *in vivo* as microinjection of either MeCP2 or Histone H1 increased the FRAP kinetics of the other (Ghosh et al.,2010). There is evidence that this could be relevant in the brain; neurons from *Mecp2-null* mice have been found to express higher levels of H1 (Skene et al.,2010). Together, these data could present a mechanism by which MeCP2 regulates chromatin architecture.

1.9.3 MeCP2 in RNA splicing

It has been proposed that MeCP2 plays a role in RNA splicing. The main mechanism by which this occurs is by MeCP2 directly interacting with Y- box binding protein 1 (YB-1), a Y box transcription factor able to regulate alternative splicing. Additionally, expression of MeCP2 by transient transfection promotes the inclusion of variable exons in a reporter minigene. It has also been observed that depletion of MeCP2 in human cell lines leads to abnormal alternate splicing events. While MeCP2 has been identified as a methylation dependent transcriptional repressor, RNA profiling from mice lacking the protein did not reveal significant changes in gene expression (Young et al.,2005). Changes in the alternative splicing of genes, including *DLX5, Fgf2-5, Fut8,* and *Nf1* have been observed in mouse models of RTT (Mecp2308/Y) (Young et al.,2005). The latter, alongside reports of MeCP2 binding with RNA binding protein YB-1 (Young et al.,2005) support MeCP2's role in RNA splicing.

1.9.4 MeCP2 in microRNA regulation

MicroRNAs are non-coding RNAs that play key roles in gene expression regulation (Lee et al.,1993). These are short (22 nucleotides long), single stranded RNA molecules (Lee et al.,1993).

MicroRNAs are involved in gene silencing. One study recently described role of MeCP2 in the regulation of MicroRNA (miRNA) expression in the prenatal brain. miR-132 and miR-212 have been reported to bind to 3' UTR of MeCP2 and post transcriptionally diminish translation (Wada et al.,2010). MeCP2 regulates miRNA processing is by interacting with DGCR8 in order to halt the formation of Drosha-DGCR8 complex thus leading to a halt in transcription and translation (Cheng et al.,2014). Despite this, retesting of the role of MeCP2 in microRNA regulation has not led to a consensus.

1.9.5 Silencing of retrotransposition

The mammalian genome is made up of more than 35% repetitive sequences (Yoder et al.,1997). Repetitive sequences, such as retrotransposons, have been found to be highly methylated in the mouse brain and, interestingly, MeCP2 associates with these regions, as highlighted by chromatin

immunoprecipitation (Muotri et al.,2010). Additionally, these elements are more abundant in the nuclei of *Mecp2-null,* suggesting that MeCP2 may act to reduce their expression (Skene et al.,2010). Considering this, increased transposition may be characteristic of MeCP2 deficiency. On par with the latter, increased L1 retrotransposition has been found in the brains of RTT patients as well as in *Mecp2-null* mice, compared to WT controls (Muotri et al.,2010). It has been proposed that the main function of DNA methylation is to protect the genome from retrotransposition events by transcriptional repression of these transposable elements (Yoder et al.,1997). These data thus reinforce MeCP2's role as a transcriptional repressor, however, further research into the protein's function is required to resolve controversies regarding its role as a transcriptional activator and repressor.

1.10. A single function for MeCP2?

Despite this evidence that MeCP2 acts as a multi-functional hub, a recent study suggests that MeCP2 has a single dominant function: to physically connect DNA with the NCoR/SMRT complex (Tillotson et al., 2017). In this study, the authors generated an MeCP2 construct whereby they removed almost all amino-acid sequences except the MBD and NCoR/SMRT interaction domain. They found that mice expressing this minimal protein survive for over one year with only mild symptoms. Moreover, they found that this minimal protein was able to prevent or reverse neurological symptoms when introduced into MeCP2-deficient mice by genetic activation or virus-mediated delivery to the brain. In contrast, mice lacking both the N- and C-terminal regions (approximately half of the native protein) are phenotypically near-normal. Together, these results suggest that the predominant function of MeCP2 is to bind methylated DNA and interact with the NCoR/SMRT complex. These results are supported by the finding that missense mutations that cause Rett syndrome are concentrated in two discrete clusters that coincide with the MBD and NCoR/SMRT interaction domain (Lyst et al., 2013).

1.11. MeCP2 interactions with importins

1.11.1 Nuclear Localisation Signal

MeCP2 is ubiquitously expressed throughout the genome and is present in the cell nuclei. The nucleus is separated from the cytoplasm by a nuclear membrane, which has an embedded nuclear pore complex (NPC). Small proteins are able to diffuse to the nucleus passively through the NPC, however, larger proteins are unable to do this. Larger proteins require a nuclear

localisation signal (NLS) in order to enter the nucleus. The NLS is a short motif that binds to proteins and mediates active transport through the NPC (Sorokin et al.,2007).

MeCP2's NLS was initially established by expression of protein fragments as beta-galactosidase fusion proteins and assaying their localization in mouse fibroblasts (Nan et al.,1996). The latter system required the NLS for the MeCP2's localisation to the nucleus. This is contradictory to the finding that the NLS is dispensable for MeCP2 localisation (Lyst et al.,2018). A likely explanation for the above is that beta galactosidase bound to MeCP2 may obligatory require active transport through the NPC. This hypothesis is hence suggestive of smaller MeCP2-EGFP fusions being able to enter the nucleus and be retained there by its affinity to DNA (Lyst et al.,2018) and that the only requirement for this to happen is the presence of an intact MBD.

Considering that the NLS appears not to be required for localisation of MeCP2 to the nucleus, a question is raised regarding the purpose of its interaction with KPNA3 and KPNA4. Disease progression in a mouse model of RTT was not impacted by a mutation abolishing the binding of MeCP2 NLS to importins KPNA3 and KPNA4 (Lyst et al.,2018). One possibility for the function of MeCP2's interaction with importins could be related to their proposed function as chaperones (Lyst et al.,2018).

While the general consensus remains that MeCP2 does not require KPNA3 and KPNA4 to localise to the nucleus, there is recent evidence implicating importin alpha 5 in MeCP2's nuclear localisation. Five importin alpha knockout mouse lines were produced (Panayotis et al.,2018). This study found that a knockout of importin alpha 5, but not importin alpha 3 and importin alpha 4, reduced MeCP2 nuclear localisation, particularly in hippocampal neurons (Panayotis et al.,2018). These data suggest it's premature to assume MeCP2's complete independence of this importin for localisation. Reinforcing the above, hippocampal neurons lacking importin alpha 5 revealed changes in presynaptic plasticity and modified expression of MeCP2 regulated genes (Panayotis et al.,2018).

1.12 TBLR1

The TBLR1 gene is located in 3q26.32 and is approximately 178bps long (Zhang et al.,2006). Human TBLR1 exists as two isoforms, differing at their carboxyl end (Zhang et al.,2006). The NCoR-SMRT complex, as isolated by immunoaffinity chromatography was shown to bind directly to TBLR1, via SMRT, which in turn binds directly to HDAC3. This suggests that TBLR1 has a role in the regulation of NCoR-SMRT.

1.12.1 TBLR1 and its importance in RTT aetiology

As mentioned previously, RTT mutations outside of the MBD are clustered in the NID (Lyst et al.,2013). A mutation in this domain, R306C, causes a severe RTT like phenotype in mice (Lyst et al.,2013). Additionally, models of MeCP2 duplication syndrome suggest that the MBD and NID must be intact for an adverse molecular pathology to develop (Lyst et al.,2013). The above facts suggest that the NID mediates critical functions of MeCP2. This could be through disruption of the recruitment of NCoR/SMRT corepressor complexes to chromatin. Alternatively, a major function of the NID could be to perform DNA binding. Previous data suggests that co-repressor recruitment to DNA is a core MeCP2 function that is disturbed in RTT (Lyst et al., 2013). It is hypothesised that this could be due to the loss of the DNA-MeCP2-NCoR/SMRT bridge compromising brain function by preventing transcriptional repression (Lyst et al., 2013).

When investigating the molecular basis of MeCP2-NCoR/SMRT interaction, previous studies have found that the TBL1 and TBLR1 subunits are direct MeCP2 binding partners. Furthermore, a crystal structure of the MeCP2-TBLR1 complex revealed that the four MeCP2 NID residues directly contacting TBLR1 are the same as those mutated in RTT (Kruusvee et al.,2017). The above suggests that one of the main functions of the NID could lie in the maintenance of this interaction. Furthermore, a study has shown that the NCoR subunit TBLR1 binds directly to MeCP2 via the C-terminus WD40 domain (Kruusvee et al.,2017). Reinforcing the main function of the NID as maintenance of MeCP2's interaction with TBLR1 is the fact that mutations in the TBLR1 WD40 domain were also found to prevent recruitment to the heterochromatic foci *in vitro*. The most common RTT-causing missense mutation in the NID is R306C, responsible for 5% of all cases. Mouse models of this mutation display the same phenotypic characteristics as other RTT mouse models, confirming pathogenicity (Guy et al.,2001; Brown et al.,2016; Lyst et al.,2013). Studies on transcriptional analyses have displayed the same patterns of dysregulation in R306C knock-in mice as the other models (Gabel et al.,2015). Overall, this data is evidence that disruption of the interaction between MeCP2 and the NCoR complex suffices in causing both the neurological defects and transcriptional changes characteristic of all RTT models

As mentioned previously, MeCP2 has been found to be nearly as abundant in the brain as the histone octamer (Skene et al.,2010). However, MeCP2 binds preferentially to methylated DNA (Skene et al.,2010). Additionally, it has been established that DNA methylation favours 2 MeCP2 molecules per nucleosome (Nikita et al.,2007; Skene et al.,2010) but there is enough MeCP2 in the brain to completely coat the entire nucleosome. Considering the above findings with regards to NCoR complex interaction with MeCP2 via a TBLR1 subunit, it may then be hypothesised that the excess presence of MeCP2, compared to nucleosomes, elucidates a specific function to bind to TBLR1.

1.13 Using biotinylation for identification of protein interaction networks

Despite decades of research and a well-documented interaction with NCoR, the interactome of MeCP2 is not yet established. Previous proteomic analysis methods, such as co-immunoprecipitation, have various technical limitations. Such techniques have the possibility of producing artefacts due to frail purification and contamination (Miernyk and Thelen., 2008). Also, co-immunoprecipitation also fails to identify weak or transient interactions. Currently used methods for the identification of protein-protein interactions involve using yeast 2 hybrids and affinity complex purification.

Biotin tagging of proteins is a useful tool for studying protein interactions. Biotin is a small, organic molecule with binds to avidin and streptavidin with high affinity (Weber et al.,1989). Biotinylation may be achieved by chemical or enzymatic means. Chemically, biotinylation occurs via the modification of protein amino groups with biotin-N-hydroxysuccinimide (Weber at al., 1989) Enzymatic biotinylation, however, is limited to proteins naturally containing this modification. Such proteins are mostly biotin-dependent carboxylases and decarboxylases (Chapman-Smith and Cronan., 1999). Biotinylation is one of the most commonly used protein labelling methods for facilitation of detection, immobilisation, and purification (Bayer and Wilchek., 1990).

Biotin protein ligases (BPLs) used to catalyse protein biotinylation are highly specific to their protein substrates and, thus, are not general protein modification enzymes (Choi-Rhee et al.,2004). BirA, *E*.coli's BPL, is an enzyme with the ability to catalyse covalent attachment of biotin to the lysine side chain within a 15 amino acid (aa) peptide termed the avi-tag (Barker and Campbell.,1981; Schatz 1993; Beckett et al.,1999; Cull and Schatz.,2000) E. coli B strain AVB101 contains an engineered pACYC184 plasmid with an inducible *BirA* gene. It is designed for co-expression of BirA and avi-tagged protein of interest for *in vivo* biotinylation. The 15 aa peptide has been widely used as a tag for protein biotinylation and purification. There are few other conjugation pairs that have the same high affinity and utility, and hence there are limitations on the development of applications that require more than one conjugation handle. By expanding the specificity of biotin ligase, it may prove possible to generate a wider array of peptide- and protein-labelling reagents (Lu et al.,2014).

Enzymatic biotinylation appears to be an ideal tag, however, the avi tag does not occur naturally and frequently in mammalian proteins. It was important to address this as biotinylated protein species are rare, meaning that endogenous biotinylation is reduced (Cronan., 2005). Additionally, the specificity and strong binding of biotin to streptavidin and avidin allow the sensitive detection

of biotinylated proteins (Cronan., 2005). To be able to use the above benefits of biotinylation, a mutant form of BirA was engineered to contain a R118G mutation. BirA-R118G, known as BirA*, showed a higher level of promiscuous biotinylation (Choi-Rhee et al.,2004). This method allows for identification of weak or transient protein interactions.

This approach for identification of neighbouring and interacting proteins is based on the use of a promiscuous prokaryotic biotin protein ligase. Based on BirA*, a biotin ligase was fused to a protein of interest, and then introduced into mammalian cells where it biotinylates vicinal proteins. Biotinylated proteins were then isolated and identified by methods such as mass spectrometry. This strategy, called BioID, was used to identify candidate proteins that are proximate to and/or interact with human lamin A (LaA), a well characterized component of the nuclear envelope (NE Roux et al.,2012). Aside from its use in mammalian cells, BioID has also been used for identification of interacting partners of the Asc1 proteins in yeast (Opitz et al.,2017), and for Identification of complexes mediating postsynaptic inhibition in mice neuronal cells (Uezu et al.,2016). Overall, data suggest that PL has widespread applications for shedding light on protein-protein interactions.

There are numerous limitations in the traditional methods used to study protein interactions, including limitations to isolation *in vitro*. Immunoprecipitation, one of the most commonly used techniques for studying protein-protein interactions, enriches proteins on the biochemical basis of their affinity to the bait protein (Branon et al.,2018). Proximity biotinylation, on the other hand, is based on through space biotinylation of proteins 1-10 nm from the bait (Figure 3A; Branon et al.,2017).

Despite extensive research, there is still a lack of understanding regarding MeCP2 target genes (Horvath and Monteggia., 2017). A main reason for this is the fact that *MeCP2* is an X linked gene, hence, a form of heterogeneity is added owing to X chromosome inactivation (Horvath and Monteggia., 2017). As a means of addressing the above, an *in vivo*, cell specific biotinylation system, combined with fluorescent activated cells to examine the effects of MeCP2 causing mutations on transcription was designed (Johnson et al.,2017).

1.13.1 Engineering MeCP2 to identify its targets

Biotin tagging has been recently used to reveal insights into the RTT transcriptome. Homologous recombination of MeCP2 to insert an affinity tag upstream of the protein's stop codon (Johnson et al.,2017) has been achieved. This tag comprised of a TEV protease cleavage site and a 15aa-

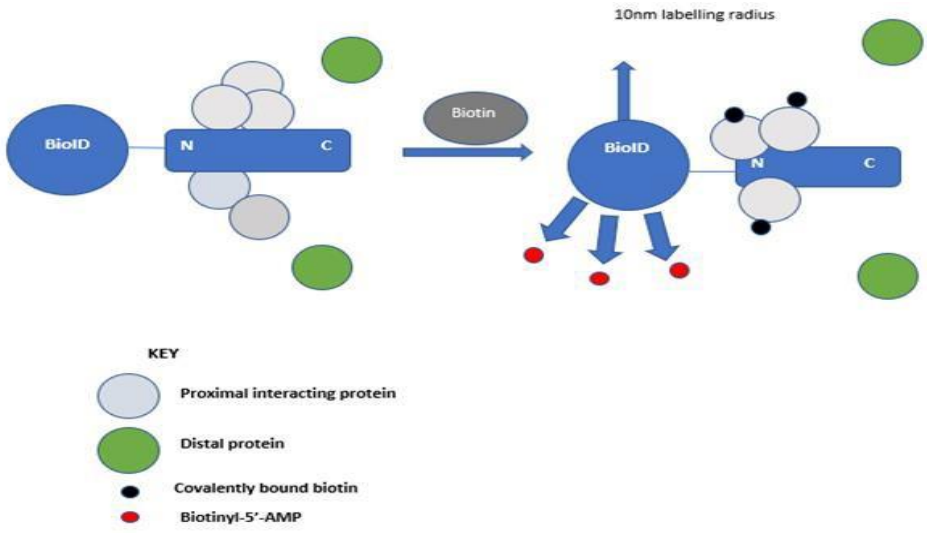
avi motif (Johnson et al.,2017). This motif can be post-translationally labelled with BirA (Johnson et al.,2017). The above approach was combined with an allelic series of Knock-in mice carrying frequent RTT mutations (Johnson et al.,2017). The above enabled the selective profiling of RTT associated nuclear transcriptomes (Johnson et al.,2017). Recently, more efficient mutant enzymes have been engineered, enabling the exploration of new avenues of identifying protein interactions.

1.14 Proximity labelling using TurboID and miniTurbo

The enzymes most commonly used for PL are; APEX2, and BioID (Roux et al.,2012; Branon et al.,2018). One of the major advantages of APEX labelling is its speed, enabling tagging of proteins in 1 min or less, and its versatility. On the other hand, APEX labelling requires H_2O_2, toxic to living samples (Branon et al.,2018). Contrastingly, BioID labelling is non-toxic and only requires the addition of biotin for the initiation of tagging. A major disadvantage of BioID is its slow kinetics requiring very long labelling times of approximately 18-24 hours (Branon et al.,2018). To address this, yeast display based directed evolution was used to engineer two BirA mutants (Figure 1.5; Branon et al.,2018).

Two promiscuous ligases, TurboID (35kD) and miniTurbo (28kD) were yielded. TurboID was yielded with 15 mutations relative to wild-type (WT) BirA, and miniTurbo with its N-terminal domain deleted and 13 mutations relative to WT BirA (Branon et al., 2018). These mutants were up to 26-fold more active than BioID, enabling labelling in as little as 10 minutes (Branon et al.,2018). Additionally, TurboID can produce more biotinylated material in 1 hour than BioID produces in 18 hours (Branon et al., 2018). These data suggest TurboID and miniTurbo enable proximity labelling (PL) to be performed efficiently, in new settings and thus broaden the scope for new discoveries.

A)



B)

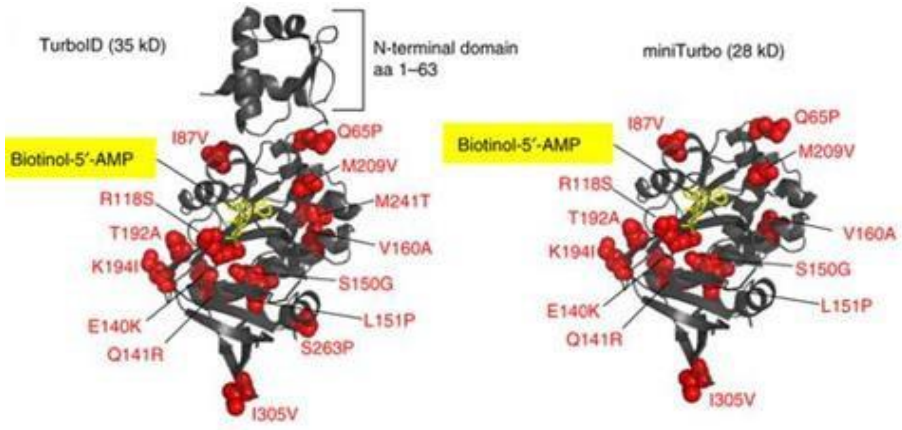


Figure 1.5. Previously used BioID is replaced in proximity labelling by TurboID and miniTurboID. a) Schematic of how BioID is used for proximity labelling (PL). b) Schematic of ligase structure of TurboID (left) and miniTurbo (right) highlighting differences resulting from mutations. N-Terminal domain is absent in miniTurbo. Other mutated sites mutated are shown in red and include the absence of S263P and M241T (Taken from Branon et al., 2018).

Two enzymes instead of one were engineered due to their unique properties. TurboID is more active and hence is most suitable for when biotinylation maximisation is the priority (Branon et al.,2017). It is of note that small amounts of promiscuous biotinylation have been witnessed to occur prior to the addition of biotin. This suggests that TurboID is able to utilise the small amounts of biotin in cells and organisms grown in media for example (Branon et al., 2017). Nearly all eukaryotes import biotin, as they cannot biosynthesize their own. Interestingly, bacteria, which can make their own biotin, did not give TurboID background prior to exogenous biotin addition. Overall, the use of TurboID and miniTurbo catalysed proximity labelling allows further understanding of the interacting partners of MeCP2 while overcoming the issues of toxicity and kinetics.

1.15 Hypotheses

We hypothesise that fusing a promiscuous biotin ligase to MeCP2 will allow for the efficient and unbiased labelling of MeCP2 interacting proteins (Figure 1.5).

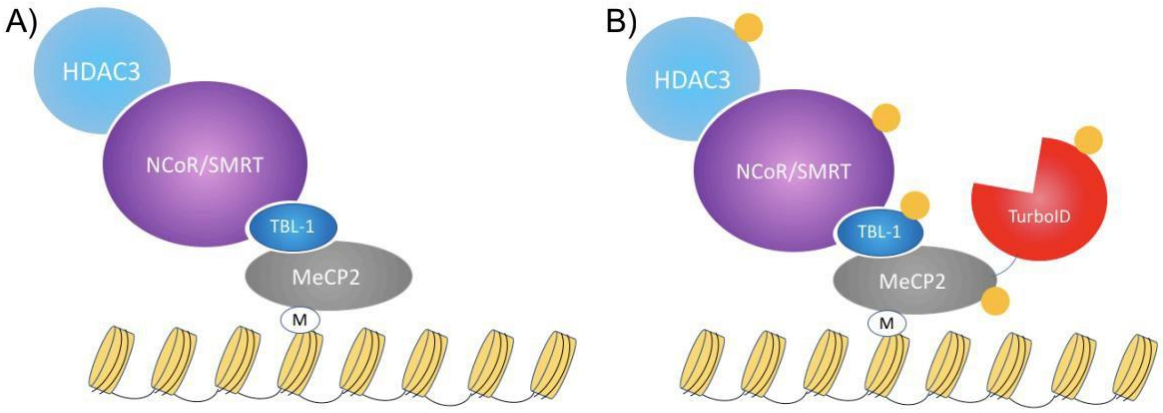


Figure 1. 6 Elucidating interacting proteins of MeCP2 using proximity labelling. Schematic showing our hypothesis for this project. (A) MeCP2 is known to interact with the NCoR/SMRT complex. (B) Fusion of the promiscuous biotin ligase, TurboID, to MeCP2 and addition of exogenous biotin causes the biotin ligase to biotinylate proximal proteins that interact with MeCP2. Biotinylated proteins can then be recognised by streptavidin. Streptavidin-coated beads can purify the biotinylated proteins from cell lysates for further analysis. miniTurbo acts in the same way.

1.16 Aims

Despite nearly three decades of research on MeCP2, its molecular function and the role it plays in RTT is still not fully elucidated. In addition to the above, the previous studies have attempted to shed light on the protein interactions of MeCP2 have done so using immunoprecipitation. In this project, I propose to generate recombinant constructs consisting of TurboID and miniTurbo and evaluate whether these biotinylate proteins in NIH-3T3 fibroblast cells. Overall, this study hoped to use novel PL to confirm and identify novel protein interactions of MeCP2.

The aims of this thesis were:

1: To search for novel proteins (using TurboID and miniTurbo) and confirm existing interactions with MeCP2

2: To compare biotinylating efficiency of conjugated TurboID and MiniTurbo.

3: To clone TBL1 into TurboID.

2- Materials and Methods

2.1 Generating recombinant plasmid constructs

Sequence analysis revealed that cutting out EGFP would conserve the N terminus of *MeCP2* and allow enough space for fusion of TurboID and miniTurbo. The restriction enzymes used were NOTI (NEB) and BamHI-HF (NEB). The next step was to amplify the insert DNA. As the insert DNA, did not have restriction sites in the correct orientation, I designed primers with the restriction site sequences included (Table1).

Table 1- Primers used for PCR and their sequence.

Primer	Sequence
TurboID (FWD)	5'-TAA GGA TCC G ATG TAC CCG TAT GAT GTT CCG-3'
TurboID (RVS)	5'-TTA GCG GCC GCT CAC ACC TTC CTC TTC TTC TTG-3'

*Primers for TurboID and miniTurbo were the same

2.2 Cloning of TurboID and miniTurbo plasmids

Sequence analysis enabled identification of antibiotic resistance genes of MeCP2-EGFP, TurboID and miniTurbo. It was revealed that MeCP2-EGFP displayed resistance to Kanamycin While TurboID and miniTurbo held resistance to ampicillin.

Agar (Miller) plates were prepared by adding 20 g of Agar to 500 ml of deionised water followed by autoclaving. The appropriate antibiotics were added to the molten Agar to a final concentration of 30 µg/ml and the agar was poured onto plates. A sterile pipette tip was used to make contact with the plasmid stab cultures and used to streak the plates containing appropriate antibiotic. Plates were incubated at 37 ℃ for 12-18 hours.

2.2.1 DNA extraction from bacterial colonies

2.2.2 Overnight culture

A sterile pipette was used to extract a single colony from the plates. The tip was ejected onto a 15 ml tube containing LB broth (10 g LB miller in 500 ml deionised water) and 5 µl of the appropriate antibiotic. Tubes were incubated in a shaking incubator for at least 12 hours at 225 rpm at 37 ℃.

2.2.3 Glycerol stocks

Glycerol stocks of every plasmid and construct were made using overnight bacterial growth. Briefly, 500 µl of overnight growth were added to 500 µl 50% glycerol in 2 ml screw top tubes. Stocks were kept at -80 ℃.

2.2.4 Miniprep DNA extraction

DNA extraction occurred according to manufacturer's (Qiagen) guidelines. Briefly, 4 ml of overnight culture underwent centrifugation at approximately 6,000 x *g* for 5 minutes at room temperature. The pellets were resuspended in 250 µl buffer P1 followed by addition of Buffer P2. Incubation occurred for 5 minutes prior to the addition of 350 µl of Buffer N3. Tubes were centrifuged for 10 minutes and the supernatant added to the spin column, followed by 60 seconds centrifugation. The spin column was then washed with 0.5 ml Buffer PB followed by centrifugation. The spin column was then washed with 0.75 ml buffer PE followed by centrifugation (60 s). This was followed by centrifugation (60 s) to remove residual wash buffer. DNA was eluted onto a clean microcentrifuge tube by adding 30 µl buffer EB to the spin column followed by centrifugation (1 min).

2.2.5 Midiprep DNA extraction

DNA to be used for transfections was extracted using a Midiprep Kit (Qiagen). Briefly, overnight culture (30 ml) was centrifuged at 6,000 x *g* for 15 minutes at 4 ℃. The pellet was resuspended in buffer P1 followed by lysis in buffer P2 While being incubated at room temp for no longer than 5 minutes. Lysis was followed by neutralisation with 4 ml buffer P3 and incubation on ice for 15 minutes. Neutralisation was followed by centrifugation at 20,000 x*g* for 30 min (4 ℃) and re-centrifugation for 15 min under the same conditions. Buffer QBT (4 ml) was used to equilibrate a column and supernatant added to the column. Column was then washed with 2 x 10 ml QC and DNA eluted with 5 ml Buffer QF.

2.2.6 PCR of insert DNA fragment

After I designed primers with restriction sites in the correct orientation (Table 1), a PCR was conducted (Figure 2.1). Each reaction was made to 20 µl and set up as follows; 4 µl 5X Reaction buffer, 1 µl FWD primer (10 µM), 1 µl (10 µM) RVS primer, template DNA (0.3 ng), 0.4 µl dNTPS (10 mM), 0.2 µl Q5 Taq polymerase (NEB). Reactions were performed according to the following programme; initial denaturation at 98 ℃ for 2 min followed by 32 cycles of 98 ℃ for 10 s, extension at 70 ℃ for 30 s followed by 70 ℃ for 45 s, and annealing at 72 ℃ for 2 min.

2.2.7 PCR DNA Purification and restriction digest

Following PCR, PCR products were purified according to manufacturer's (Qiagen) guidelines. Briefly, Buffer PB (5 volumes) were added to PCR reaction (1 vol), the mix was then added to the spin column and centrifuged for 60 s. Then, Buffer PE (0.75 ml) was added to the spin column and centrifugation occurred for 60 s. DNA was eluted with 30 µl Buffer EB.

A restriction digest of the purified PCR product (insert) as well as the vector was performed as a standard 20µL double digest with reagents: BamH1-HF(NEB) 0.4µL, Not1-HF (NEB) 0.4 µL, 10X CutSmart Buffer 2 µL (1X), DNA 0.3 µg, ddH_20 to 20 µL. This was carried out for at least 30 minutes. Digested DNA fragments were separated on a 1% TAE gel.

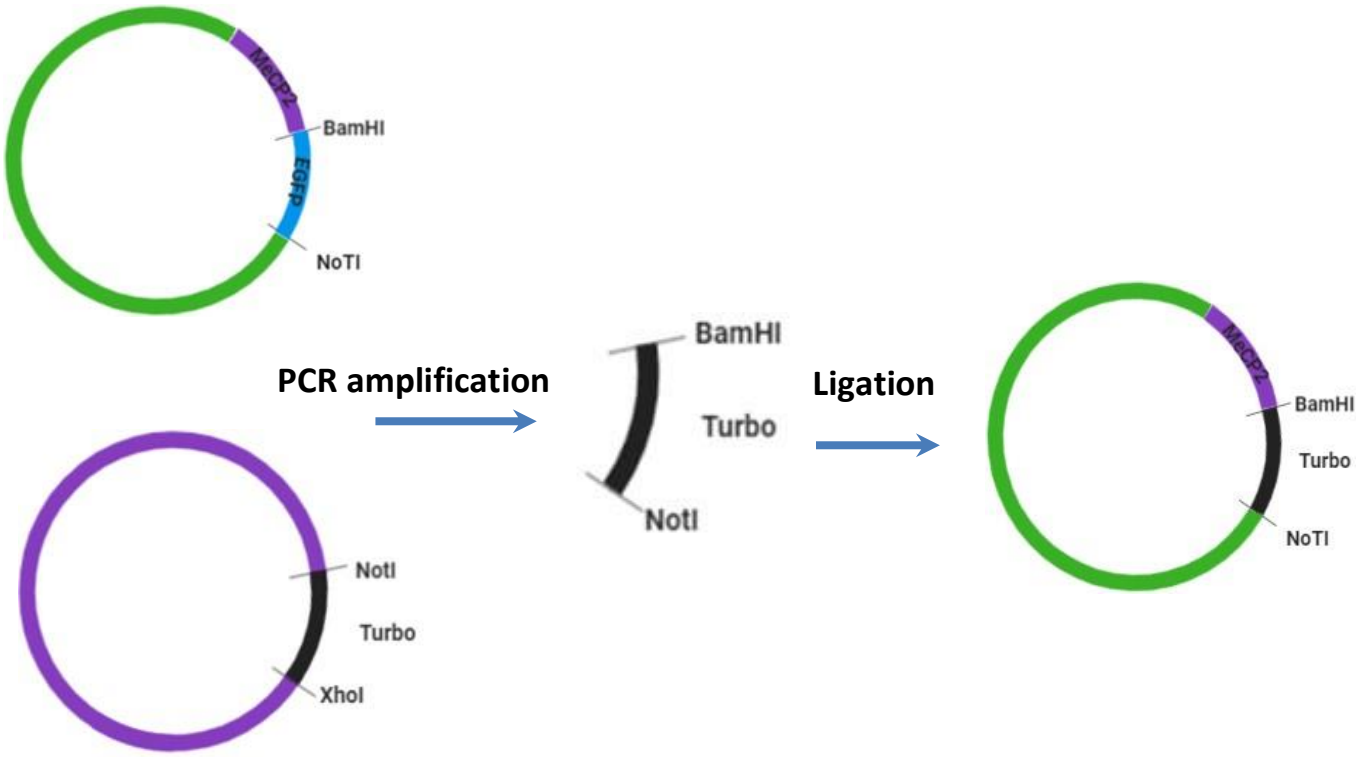
2.2.8 Gel extraction, DNA ligation and plasmid amplification

Post gel electrophoresis, the bands of interest were gel extracted according to manufacturer's (Qiagen) guidelines. Briefly, the bands of interest were excised, and Buffer QG (3 vols) were added to the gel fragment and the mix was incubated at 50 ºc. Isopropanol was added to the mix and this was applied to a spin column followed by centrifugation. Buffer PE (0.75 ml) was used to wash the DNA and Buffer EB was used to elute the DNA.

The concentration of purified DNA was measured using a nanodrop spectrophotometer and the vector (50 ng) was ligated with the insert DNA (30 ng) using T4 DNA ligase (NEB; Fig 2.1) and incubated at 16ºC overnight according to the manufacturer's protocol. Heat inaction of the ligation products occurred at 65 ºC for 10 minutes and the ligation product was used for transformation of DH5-α *E.Coli* competent cells. The cell and ligation product mix was heat shocked at 42 ºC for 45 seconds and incubated at 37 ºC for 1 hour at 225 rpm. Transformations were spread onto 30 µg/ml Kanamycin plates. The DNA from the resulting colonies was extracted by miniprep (Qiagen) according to manufacturer's guidelines and verified

by restriction digestion using NotI and BamHI enzymes followed by electrophoresis, to check for the presence of both plasmids (Fig 3.1). This was followed by sequencing.

A)



B)

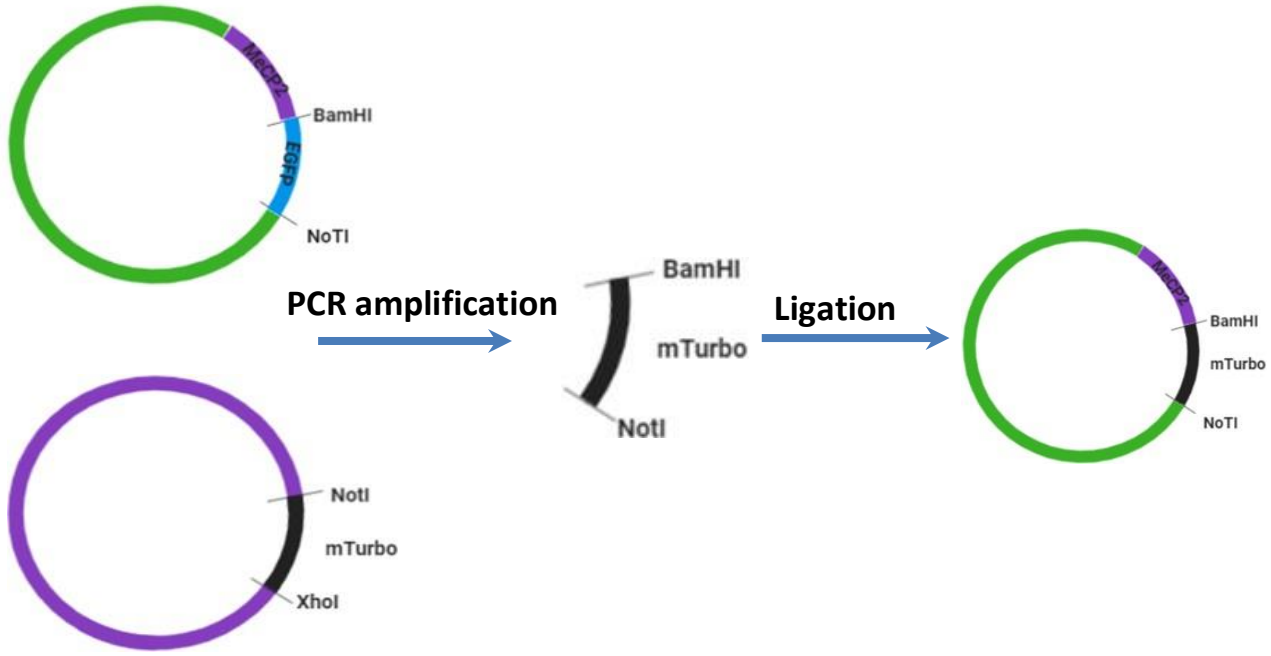


Figure 2.1 Schematic of cloning procedure of TurboID and miniTurbo into MeCP2. Compared to MeCP2, a) TurboID and b) miniTurbo had NotI and BamHI restriction sites in the wrong orientation. Primers were designed with restriction sites sequences for BamHI and NotI and PCR of the insert plasmids was conducted using these leading to the insertion of restriction sites in the correct orientation via ligation and transformation of competent cells. The transformation product was spread on Kanamycin plates and only plasmid containing bacteria would hence be able to grow.

2.3 Mammalian cell culture

NIH-3T3 cells were cultures as a monolayer in DMEM growth media (supplemented with 10% v/v foetal bovine serum (FBS; Gibco), 25 mM glucose, 1 mM sodium pyruvate, 50 units/mL penicillin, and 50 µg/mL streptomycin) at 37°C under 5% CO_2. 1% penicillin-streptomycin (Gibco). Cells were cultured in 75 cm^2 flasks (corning) in 15 ml of complete medium. When cells reached approximately 90% confluence, they were washed with PBS, trypsinised with 2 ml trypsin for approximately 5 minutes and seeded in a 1:10 dilution.

For imaging experiments, cells were counted in a haemocytometer and seeded at the desired density. For this purpose, cells were grown in 24 well plates with 500 µl growth medium. For western blotting, cells were grown on polystyrene 6-well plates with 2.5 mL growth medium.

2.4 Calcium phosphate transfection of NIH-3T3 Cells with MeCP2-TurboID and MeCP2-miniTurbo

NIH-3T3 cells ($1x10^4$) were calcium phosphate transfected in order to establish protein functionality and localisation in cells. Two tubes were prepared for each construct (MeCP2-EGFP, MeCP2-TurboID and MeCP2-miniTurbo); one containing $CaCl_2$ (2.5 M) and the other containing HEPES Buffered Saline (HBS). The appropriate DNA construct (1 µg) was diluted in $CaCl_2$.The HBS was added to the DNA-$CaCl_2$ solutions in a dropwise manner followed by 30 minutes incubation in the dark. This solution forms a calcium phosphate precipitate that is directly layered onto cells. The solution was then used for transfection and 20 µl of the DNA-$CaCl_2$ were added to the wells in a dropwise manner. This is followed by overnight incubation at 37 °C.

2.5 Immunocytochemistry of MeCP2-EGFP, MeCP2-TurboID, and MeCP2-miniTurbo

For fluorescent imaging experiments, NIH-3T3 cells were transfected for 48 hours and labelled with biotin (500 µM) for 10 mins or 1 hour. DMEM media was then aspirated, washed twice with PBS, and fixed with 4% paraformaldehyde in PBS for 15 mins. After washing twice with PBS, cells were permeabilised with 0.2% (v/v) Triton X-100 for 10 minutes. Non-specific binding was blocked with 3% Normal Goat Serum in PBS for 60 minutes. MeCP2 was visualised using an antibody directed against 6-His (1:1000 dilution; Bethyl) or Streptavidin (1:1000 dilution; Sigma) for 60 mins. This was followed by washing three times with PBS for 5 minutes, staining with DAPI (2 µg/ml) for 5 minutes, and two further washes in PBS. Cells were then mounted onto microscope slides.

Fluorescent microscopy was conducted on Leica DMIL LED. The software used to visualise was; Leica application suite af6000, version 4.0.0.11706. The pictures were analysed and processed in ImageJ (V 1.52a).

2.6 Transfection with TransIT-X2 reagent

Cells for further experiments were transfected with TransIT-X2. NIH-3T3 cells were plated onto coverslips and/ or 6 well plates containing DMEM supplemented with FBS (10%) and penicillin/streptomycin (1%) 48 hrs before transfection so they were 70-80% confluent at time of transfection. Transfections were carried out with TransIT-X2 dynamic delivery system (Mirus) according to manufacturer's instructions. Briefly, TransIT was left to warm to room temp and vortexed. Recommended concentration of DNA was diluted in recommended volume of serum and antibiotic free media. Recommended volume of TransIT was added to diluted DNA followed by gentle mixing and incubation at room temp for approximately 30 minutes. TransIT-DNA complexes were added to wells in a dropwise manner and plates incubated for approximately 30 hours.

2.7 Transfection efficiency of MeCP2-Turbo and MeCP2-miniTurbo in NIH-3T3 cells

Transfection efficiency- the percentage of cells were transfected successfully, was calculated from cells transfected with MeCP2-TurboID and MeCP2-miniTurbo. Fluorescent imaging occurred at low magnification and multiple images were taken from different areas of the cover slips in order to calculate an average. It is of note that these were taken from areas of highest signal. Transfection efficiency was calculated as follows:

TE= (number of true (transfected cells)/ number of live (DAPI stained) cells) x 100.

2.8 Biotin labelling with MeCP2-TurboID and MeCP2-miniTurbo in NIH-3T3 Cells

For labelling of transfected cells, I initiated biotin labelling approximately 36 hours following transfection. Biotin (Sigma) was diluted from a 100 mM stock (in dimethyl sulfoxide, DMSO), directly into serum containing cell culture medium to a final concentration 500 µM. Cells were treated with biotin and labelling was stopped after either 10 minutes and 1 hour by transferring the plates to ice and washing 4 times with ice cold PBS. For negative controls, I omitted the addition of biotin or used untransfected cells.

2.9 Protein separation by SDS PAGE

NIH-3T3 cells were plated, transfected and labelled with biotin as described above. Cells were subsequently detached from the flask through gently pipetting a stream of ice-cold PBS directly

onto the cells. Pellets were collected by centrifuging the resulting cell suspension at approximately 500 x *g* for 5 mins. The supernatant was removed, and the pellet was lysed with 100 µl RIPA buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.1% SDS, 0.5% Sodium Deoxycholate, 1% Triton X-100, 1 mM PMSF, and protease inhibitors (Sigma), by gentle pipetting and incubating at 4 ℃ for 5 minutes. Lysates were clarified by centrifugation at 16,000 x *g* for 10 minutes at 4 ℃. Protein concentration was estimated using the Bradford Assay, consisting of Bradford Reagent (Sigma) and BSA as a standard, according to manufacturer's guidelines.

Loading samples were made to 25 µl and prepared by addition of 15 µl of the protein sample to 6.25 µl of sample buffer (4x), 2.5 µl reducing agent (10x), followed by the addition of water to 25 µl. Samples were boiled at 70 ℃ for 10 min prior to loading. Precision plus protein kaleidoscope was used as a ladder.

SDS-Polyacrylamide gel electrophoresis (SDS-PAGE) was performed on home-made gels (10%) containing 8% acrylamide, 0.375 M Tris pH 8.8 (resolving gel) or 0.127 M Tris 6.8 (Stacking gel) 10% SDS, polymerised with 10% ammonium persulphate (APS) and TEMED. The electrophoresis occurred initially at 100V for 1.5 hours.

2.10 Western Blotting

Upon fractionation of proteins via SDS-PAGE, proteins were transferred to a nitrocellulose membrane using iBlot gel transfer device and iBlot gel transfer stacks (Invitrogen). The iBlot gel transfer stacks were constructed as per manufacturer's guidelines. Transfer occurred over 8 minutes. Blots were stained with Ponceau S solution (Sigma; 0.1% w/v in 0.5% acetic acid) for 5 minutes and washed with 0.1% Tween 20 in Tris-buffered saline (TBS-T). The membranes were then blocked with 5% BSA in TBS-T for 1hr at room temperature. Blots were then stained with Streptavidin-HRP (1/3000 v/v; Thermo Scientific, N504) in 3% BSA in TBS-T for 1-2 hours at 4 ℃. For visualisation, the membrane was washed with 0.1% TBS-T and developed with Pierce ECL reagent. Imaging was carried out with the Invitrogen iBright FL1000 imaging system.

2.11 Nuclear protein extraction

NIH-3T3 cells were plated as before onto a 6 well plate. Cells were counted, suspended in 500 µl NE10 buffer (20 mM HEPES (pH 7.6), 10 mM KCl, 1 mM MgCl2, 0.1% Triton X-100 (vol/vol), protease inhibitors (Roche), 15 mM β-mercaptoethanol) and transferred into cold Eppendorf tubes. Following 15 min incubation on ice, the cell suspension was passed through a 25 gauge needle 10 times and incubated on ice for 20 minutes with intermittent vortexing. Centrifugation

then occurred at 4 ℃ for 5 min at 500g. Nuclei were washed in 500µl NE10 buffer and the suspension was passed through a 25 gauge needle 10 times then incubated at 25 °C for 5 min with 250 units benzonase (Sigma) per 10^7 nuclei. Centrifugation occurred at 4 ℃ for 10 min at 500 x *g* before the nuclei were resuspended in 40 µl NE150 buffer (NE10 supplemented with 150 mM NaCl). Incubation on ice occurred for 20 minutes with intermittent vortexing. Lysates were cleared by centrifugation at 4 ℃ at 16,000 x *g* for 20 minutes and the supernatant was transferred to new tubes as a nuclear extract.

Samples were prepared for SDS-PAGE as determined previously and protein transfer onto a nitrocellulose membrane was carried out using an iBlot equipment and device according to manufacturer's instructions.

2.12 Probing for biotinylated proteins in the nuclear and cytoplasmic fractions

The membrane was blocked with 3% BSA in TBS-T for 1 hour. (1/1000) in TBS-T containing 3% BSA and left to incubate for at least 1 hour. After washing 3 times with TBS-T, streptavidin labelling was done as before for 1-2 hours.

2.13 Creating a stable NIH-3T3 cell line expressing MeCP2-TurboID and MeCP2-miniTurbo

Genomic sequence analysis revealed the presence of a G418 resistance gene in MeCP2 vector (Fig 2.2). The stable transfection of NIH-3T3 fibroblasts with recombinant plasmid containing MeCP2-TurboID and MeCP2-miniTurbo was established by using TransIT-X2 and G418 selection. Briefly, cells were split as before into 6 well plates 24 hours prior to transfection with TransIT-X2. Approximately 48 hours post transfection, media was replaced with media containing G418 (400 µg/ml) antibiotic. Media was replaced every 2-3 days and selection occurred over 1-2 weeks.

Once selection occurred, cells were plated onto 75 cm^2 flasks with 15 ml media containing G418 (400 µg/ml) and incubated at 37℃.

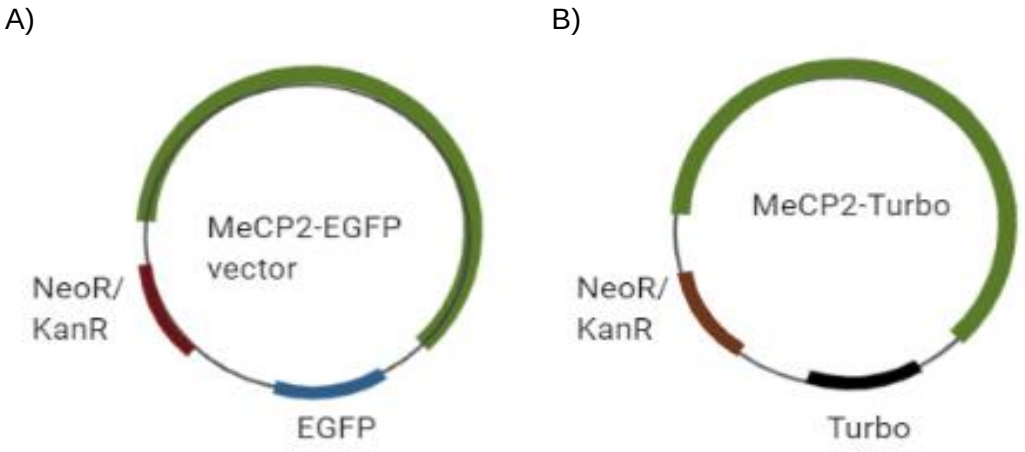


Figure 2.2 Vector and recombinant plasmid containing the TurboID ligase are resistant to G418. Schematic highlighting that a) the vector contains a neomycin coding gene and that this is conserved in b) the recombinant plasmid of interest. The neomycin gene confers resistance to G418. This will be used for selection of cells containing the plasmid of interest. Made with Biorender

2.14 Clonal stable cell line

For clonal isolation, wells containing MeCP2-TurboID transfected cells were washed wish PBS, trypsinised, counted and diluted to 10 cells/ml. This diluted suspension (100 µl) was plated onto 96 well plates. A few hours after plating, to allow for adhesion, wells containing only 1 cell were marked. Cells were left to grow from this single cell until >60% confluency was reached. Once adequate confluence was achieved in 96 well plates, cells in wells marked as having contained a single cell were trypsinised and cells plated onto cover slips in 24 well plates containing G418 and identically labelled 24 well plates. The latter was done as cells from a single well of the 96 well plate had to be imaged and some kept for maintaining the cell line. As the aim of this was to see if all cells in a particular well were expressing the proteins of interest, cells were left to adhere for 24 hrs prior to treatment with biotin (500 µM) and Streptavidin-488 for fluorescent imaging.

2.15 Cloning TBLR1 into MeCP2-TurboID

To clone TBLR1 into the vectors MeCP2-EGFP and MeCP2-TurboID, genomic sequence analysis was carried out and revealed that TBLR1 did not contain any of the restriction sites seen in MeCP2-EGFP or those incorporated into MeCP2-TurboID and MeCP2-miniTurbo in this study. To this end, primers were hence designed with restriction sites for XhoI and BamHI in the correct orientation when considering the vectors (Table 2). Integrating XhoI and BamHI restriction sites into the primers, which already present in the vectors to be used, means that the same set of primers may be used for cloning TBLR1 into both of the vectors.

Table 2- Primers with embedded restriction sites for cloning TBLR1

Primer	Sequence
F-Primer 1 (Too short)	5'-CGGCTCGAGATGAGTATAAGCAGTGATGAG-3'
R-Primer 2 (Too short/out of frame)	5'-CGGGGATCCTTTCCGAGGGTC-3'
F-Primer 2 (Correct length)	5'-CGGTAGCTCGAGATGAGTATAAGCAGTGATGAGG-3'
R-Primer 2 (In frame)	5'-CGACTAGGATCCACTTTCCGAGGGTCTAATACAC-3'

*The first set of primers used for cloning TBLR1 were too short and the reverse primer was out of frame. The second set of primers used for cloning have added bases and hence are long enough and in frame.

2.16 PCR of TBLR1

After designing primers with the chosen restriction sites, PCR of the insert was conducted according to standard protocol as highlighted at the beginning of this chapter. The conditions for the PCR amplification of TBLR1 were as follows:

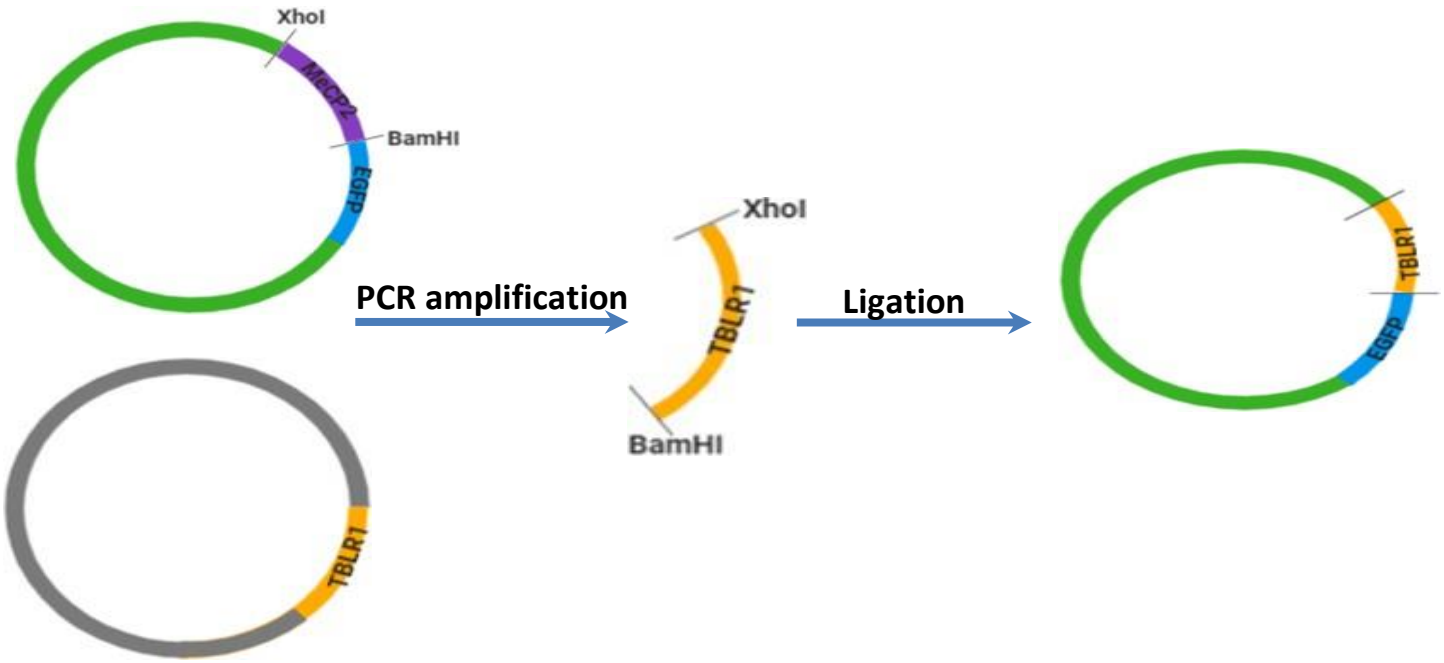
Table 3: Thermocycler settings for PCR

	1	2	3	4	Cycles- 34x	6	7
Temp ℃	98	98	53	72	///////////////	72	10
Time (min)	2:00	0:15	0:30	1:30	///////////////	2:00	∞

2.16.1 TBLR1 PCR Fragment purification

Following PCR, PCR products were purified according to manufacturer's (Qiagen) guidelines, and restriction digested using XhoI and BamHI. The bands of interest were gel extracted- MeCP2 at approximately 1500bp and TurboID and the rest of the plasmid at approximately 4900bp (Fig 2.3).

A)



B)

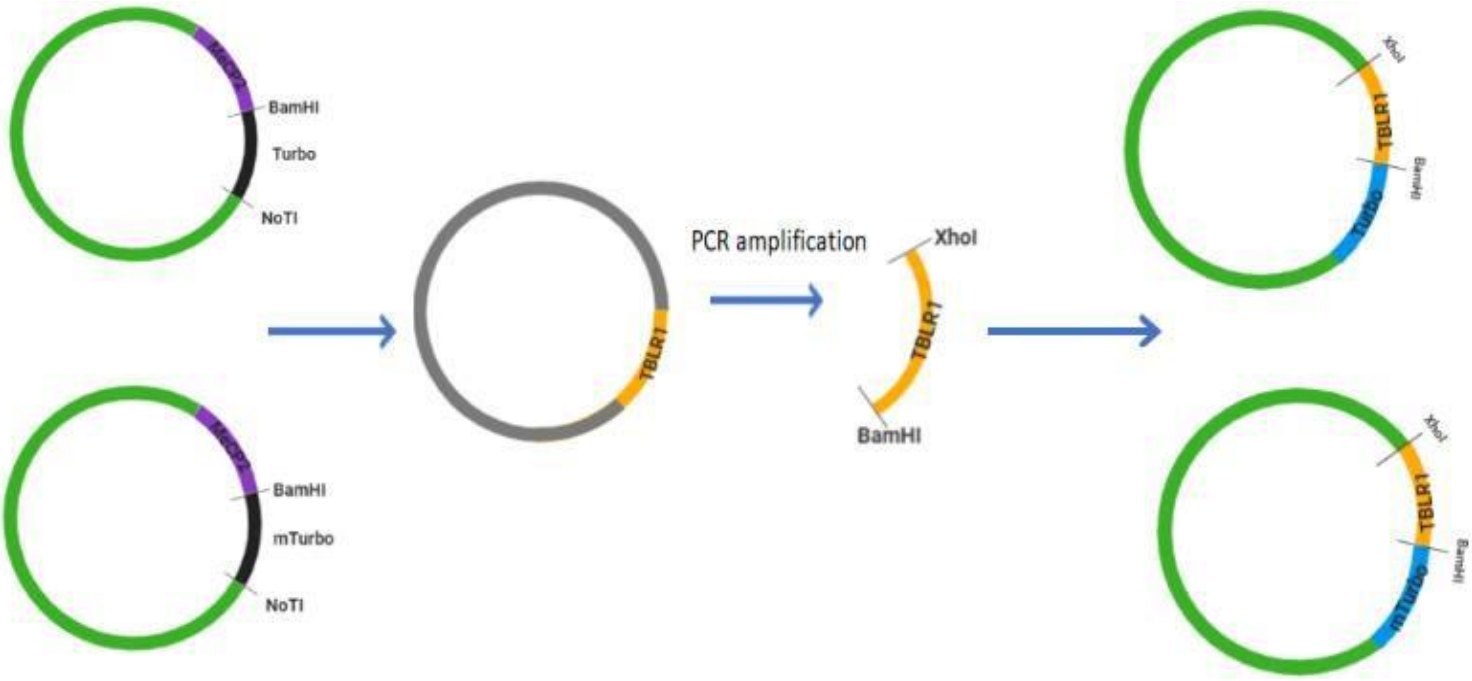


Figure 2.3 Schematic of cloning procedure of TBLR1. Genomic sequence analysis revealed that TBLR1 did not have appropriate restriction sites on the sequence of interest compared to MeCP2. PCR was conducted with primers consisting of restriction sites for XhoI and BamHI. This was a) followed by ligation of the vector (MeCP2-EGFP) and the insert (TBLR1). B) The same protocol may be used for the cloning procedure of the insert (TBLR1) into the vectors (MeCP2-TurboID and MeCP2-miniTurbo).

2.16.2 PCR purification of TBLI

The insert was amplified using PCR and the product was purified according to manufacturer's (Qiagen) guidelines. Brief details of protocol are highlighted earlier in this chapter.

2.16.3 Restriction digest and transformation

The insert and vector plasmids were digested with restriction enzymes XhoI (NEB) and BamHI (NEB) for up to 2 hours to ensure full digestion. DNA was separated on a 1% TAE gel and the bands of interest were excised according to manufacturer's (Qiagen) instructions.

The concentration of purified DNA was measured using a nanodrop spectrophotometer and the vector (50 ng) was ligated with the insert DNA (30 ng) using T4 DNA ligase (NEB) and incubated

at 16ºC overnight to produce sticky ends according to the manufacturer's protocol. XL Blue supercompetent cells (Agilent) were transformed with the ligated product according to manufacturer's guidelines. Briefly, 100 µl cells were thawed on ice, 1.7 µl of β-mercaptoethanol was added to the cells and incubated on ice. Ligated DNA was added to the mix followed by further incubation for 30 minutes. Cells underwent standard heatshock at 42 for precisely 45 s before a short incubation on ice. Cell-DNA mix was added to preheated medium and tubes were incubated at 37°C for 1 hour with shaking at 225–250 rpm Transformed cells were plated onto kanamycin plates (30 µg/ml).

3- Results

3.1 Generating MeCP2-TurboID and MeCP2-miniTurbo constructs

To identify protein interaction networks for MeCP2, I sought to develop new tools that enabled proximity-based labelling of interacting proteins. These tools were generated by the fusion of a promiscuous biotin ligase (Branon et al., 2018) to the C-terminus of MeCP2. These tools should allow the biotinylation of proteins bound to, or located, within 10 nm of MeCP2 (Branon et al., 2018). The biotin tag would then allow for the detection and purification of labelled proteins by taking advantage of the established strong biotin-streptavidin interaction. I generated two separate constructs whereby the two promiscuous biotin ligases, TurboID or miniTurbo (Branon et al., 2018) were fused to the C-terminus of MeCP2 using standard PCR-based cloning.

To this end, I excised EGFP from an MeCP2-EGFP plasmid using the restriction enzymes BamHI and NoTI (Fig 2.1). TurboID, or miniTurbo, was then PCR amplified from another plasmid using primers designed to incorporate the restriction sites matching those in the MeCP2-EGFP plasmid. I then ligated the cut plasmid and PCR insert together using T4 DNA ligase. I confirmed the addition of TurboID or miniTurbo onto the C-terminus using restriction digestion and sequencing (Fig 3.1).

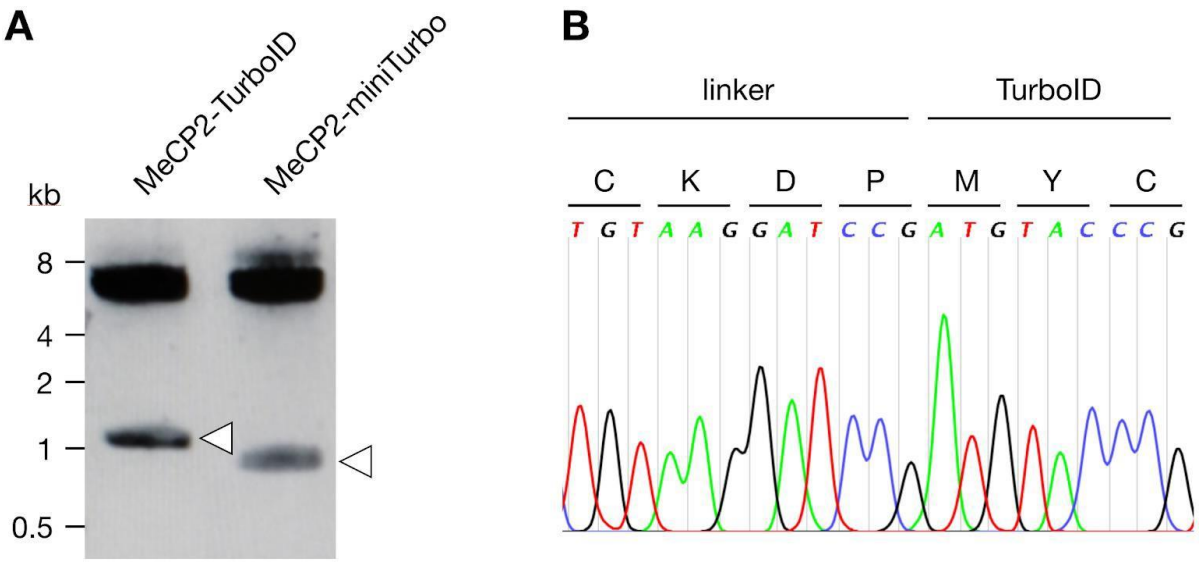


Figure 3.1 Generation of MeCP2-TurboID and MeCP2-miniTurbo constructs. A) Agarose gel electrophoresis of MeCP2-TurboID and MeCP2-miniTurbo after restriction digestion using NotI and BamHI. Arrows correspond to TurboID and miniTurbo, respectively. B) Sequencing chromatogram showing the 4 amino acid linker that joins MeCP2 and TurboID. The TurboID sequence is in-frame with the linker and MeCP2.

3.2 MeCP2-TurboID and MeCP2-miniTurbo localise to the nucleus of NIH-3T3 cells.

MeCP2 localises to the nucleus where it binds to methylated cytosines (Meehan et al.,1992). Previous studies have shown that the addition of EGFP to the C-terminus of MeCP2 does not affect it's tissue or cellular localisation (Lyst et al., 2018; Schmid et al., 2008). I therefore wanted to determine whether the addition of TurboID or miniTurbo would affect the nuclear localisation of MeCP2. To this end, we transiently transfected MeCP2-TurboID, MeCP2-miniTurbo or MeCP2-EGFP into NIH-3T3 cells. We visualised the localisation of these proteins using an antibody against 6-histidine residues, since this tag was incorporated into the MeCP2 sequence. In agreement with previous work, I found that MeCP2-EGFP was exclusively localised to the nucleus and was concentrated in heterochromatin foci (Fig 3.2), regions that are enriched for methylated cytosines (Nan et al., 1996). Moreover, I found that the localisation of MeCP2-TurboID and MeCP2-miniTurbo is indistinguishable from that of MeCP2-EGFP (Fig 3.2). Thus, the addition of TurboID or miniTurbo does not affect the nuclear localisation of MeCP2.

I next wanted to optimise the transfection protocol to ensure a high transfection efficiency, which would be required for future biochemical experiments. To this end, I transfected NIH-3T3 cells with increasing amounts of plasmid DNA. Overall, the TE ranged from 16-17% to 21-22% with TurboID and miniTurbo respectively. I found that the highest transfection efficiency (percentage

of cells expressing MeCP2) was obtained using 0.7µg DNA (Fig 3.3.2). Additionally, increasing the amount of DNA per well did not appear to significantly increase transfection efficiency, although the experiment must be repeated for adequate statistical analysis.

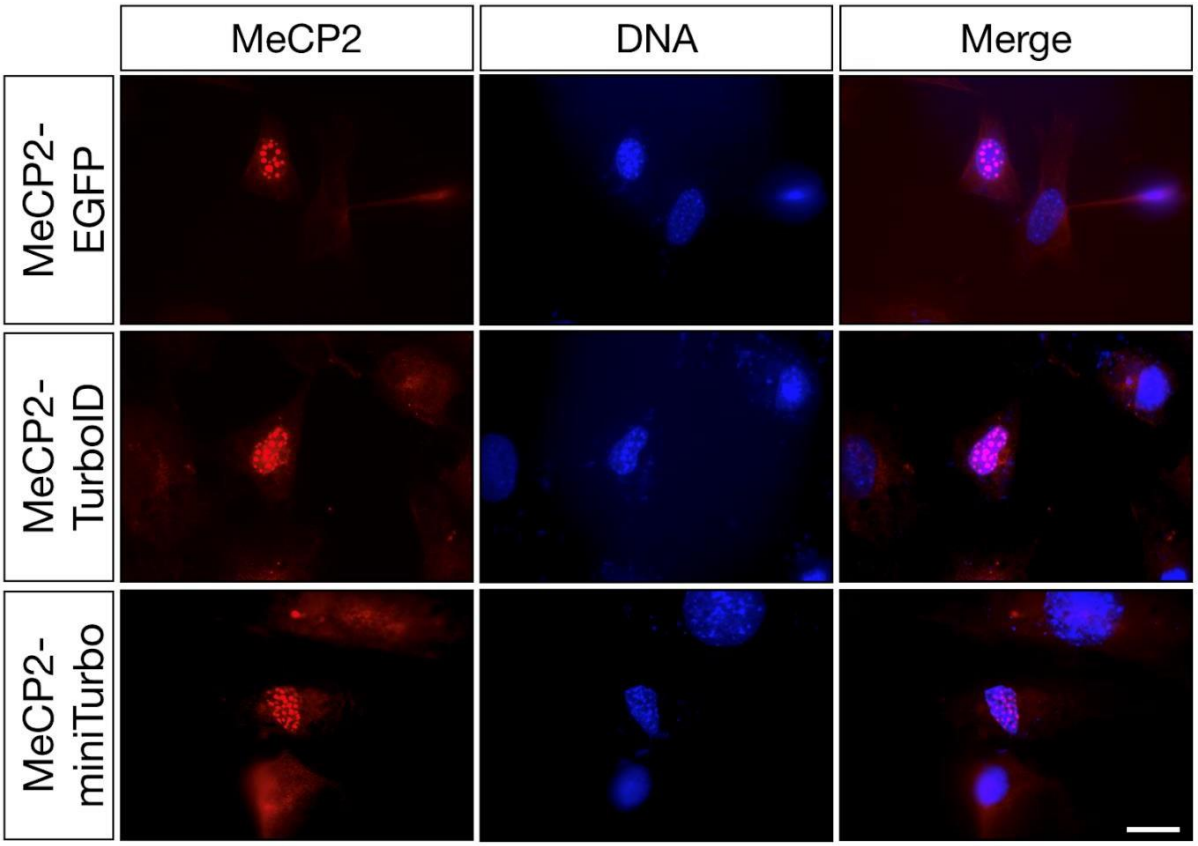


Figure 3.2 MeCP2-TurboID and MeCP2-miniTurbo locate to the nucleus of NIH-3T3 cells

NIH-3T3 cells were transfected with MeCP2-EGFP, MeCP2-TurboID, or MeCP2-miniTurbo. MeCP2 was visualised using an antibody directed against a 6-his tag incorporated into the MeCP2 sequence. DNA was visualised using DAPI. Scale bar represents 20µm.

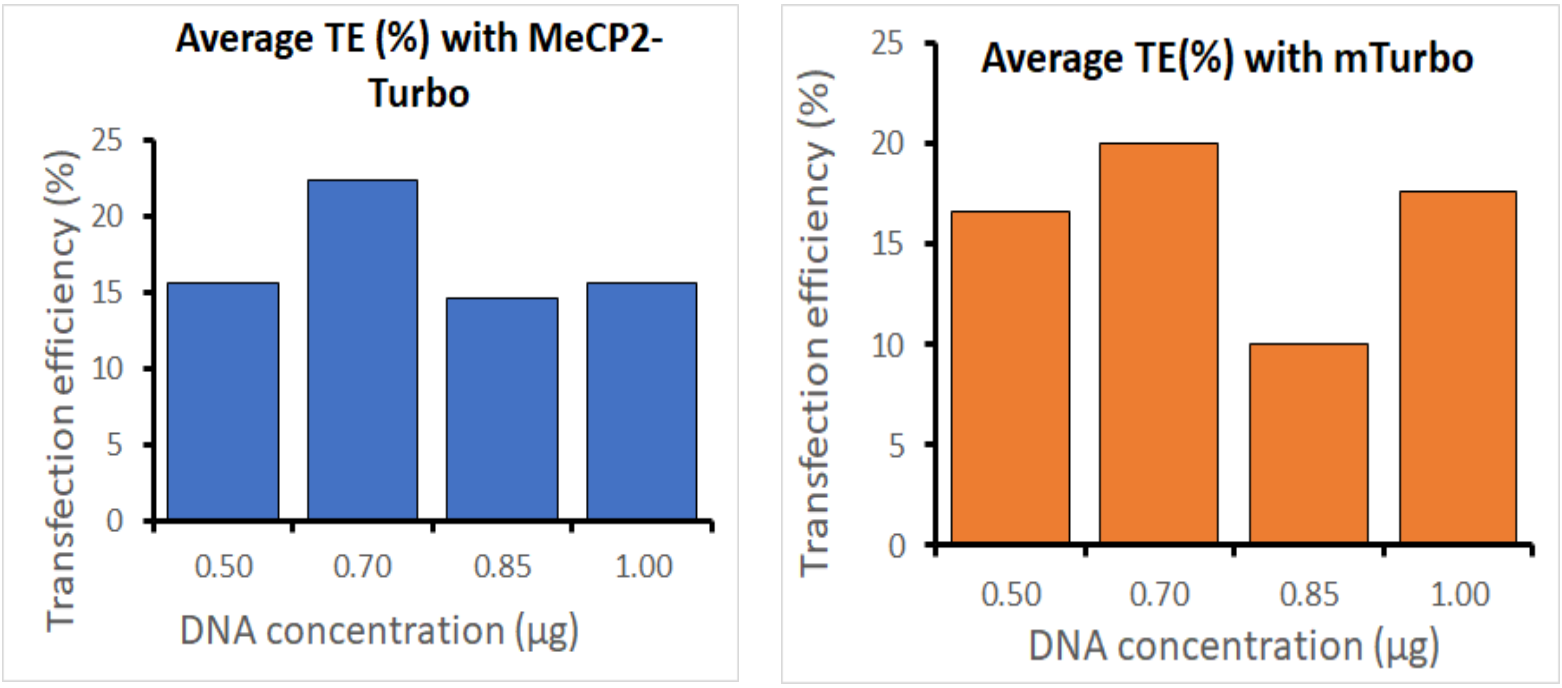


Figure 3.3. Average transfection efficiency was increased using similar amounts of MeCP2-TurboID and MeCP2-miniTurbo. Cells were Trans-IT X2 transfected with A) MeCP2-TurboID and B) MeCP2-miniTurbo. Cells were fixed with 4% paraformaldehyde, followed by incubation with rabbit anti-6 his primary antibody and anti-rabbit alexa-594 antibody. At least three images were obtained by fluorescent microscopy (40x) and an average transfection efficiency value calculated.

3.3 MeCP2-TurboID and MeCP2-miniTurbo biotinylate nuclear proteins

To examine the ability of MeCP2-TurboID and MeCP2-miniTurbo to biotinylate endogenous proteins we performed biochemical analysis of biotinylated proteins. To this end, I transfected NIH-3T3 cells with MeCP2-TurboID and MeCP2-miniTurbo and treated cells with exogenous biotin as before. Biotinylated proteins were detected using Alexa488-conjugated Streptavidin. In non-transfected cells, or transfected cells without exogenous biotin, fluorescent staining was mostly restricted to the cytoplasm (Fig 3.4.1). It is of note that a shift in biotinylated proteins from the cytoplasm to the nucleus was seen in untransfected cells following the addition of exogenous biotin (Fig 3.4.1). In

transfected cells treated with exogenous biotin, however, I observed robust nuclear fluorescence localised to heterochromatin foci (Fig 3.4.2, and Fig 3.4.3).

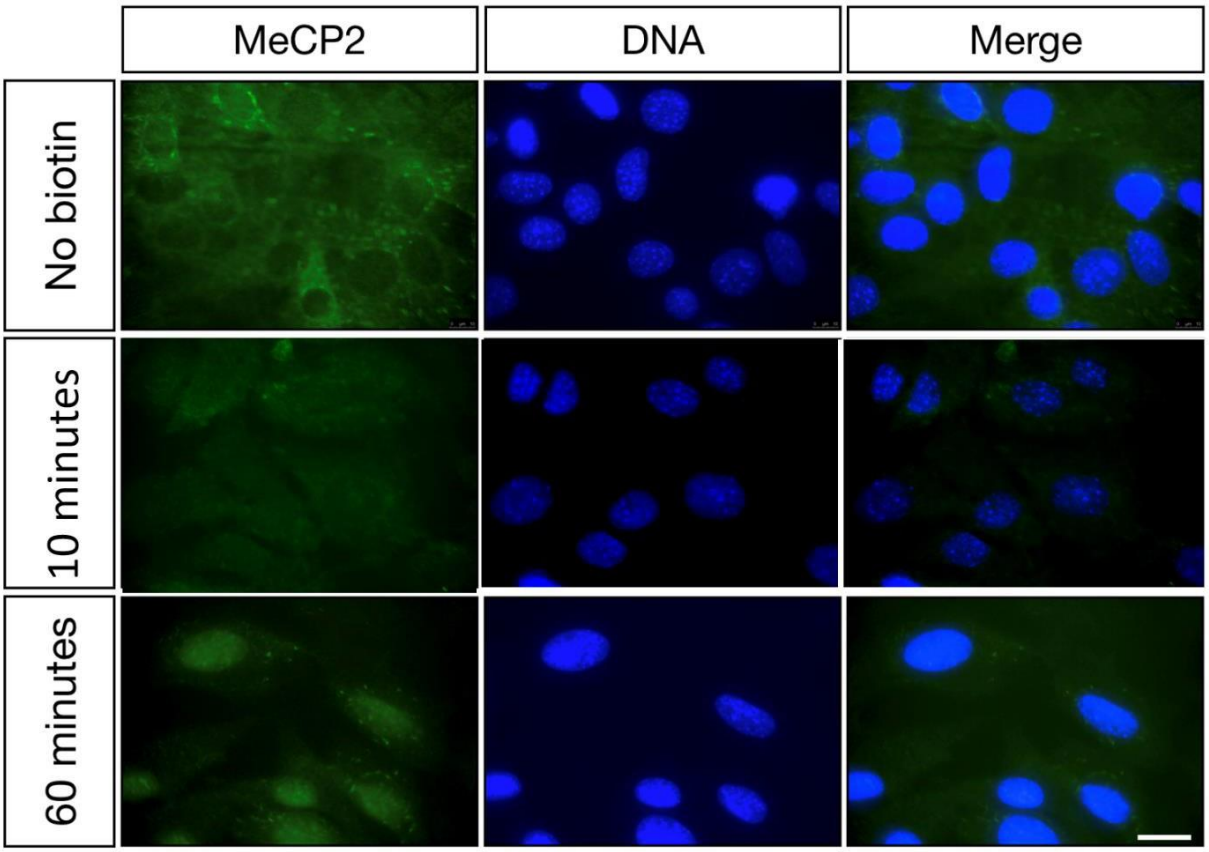


Figure 3.4.1. Endogenous biotinylation in NIH-3T3 cells. Non-transfected NIH-3T3 cells were treated with biotin (500 µM) for either 0, 10 or 60 minutes. Biotinylation was detected using Streptavidin conjugated to Alexa488. Scale bar represents 20 µm.

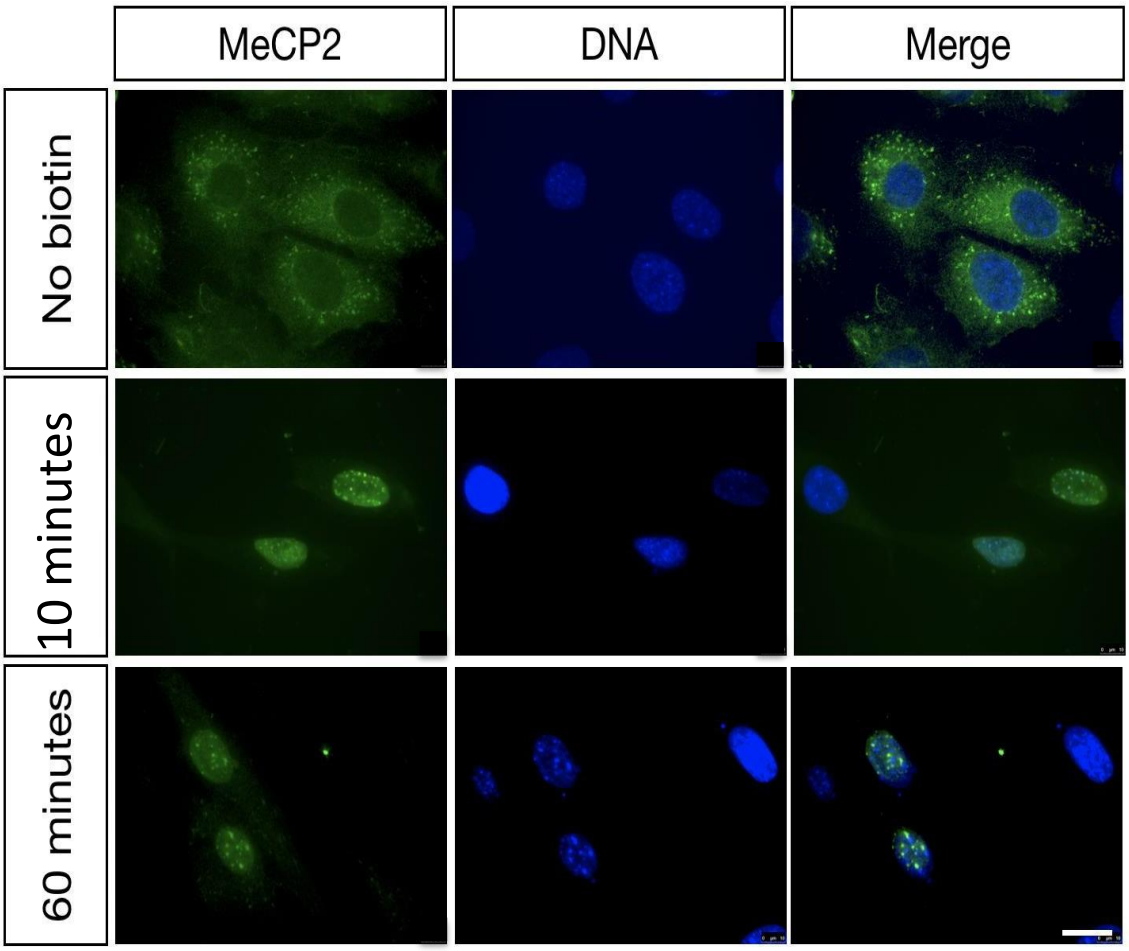


Figure 3.4.2. MeCP2-TurboID biotinylates nuclear proteins. NIH-3T3 cells were transfected with MeCP2-TurboID. After 48 hours, cells were treated with biotin (500 µM) for either 0, 10 or 60 minutes. Biotinylation was detected using Streptavidin conjugated to Alexa488. Scale bar represents 20 µm.

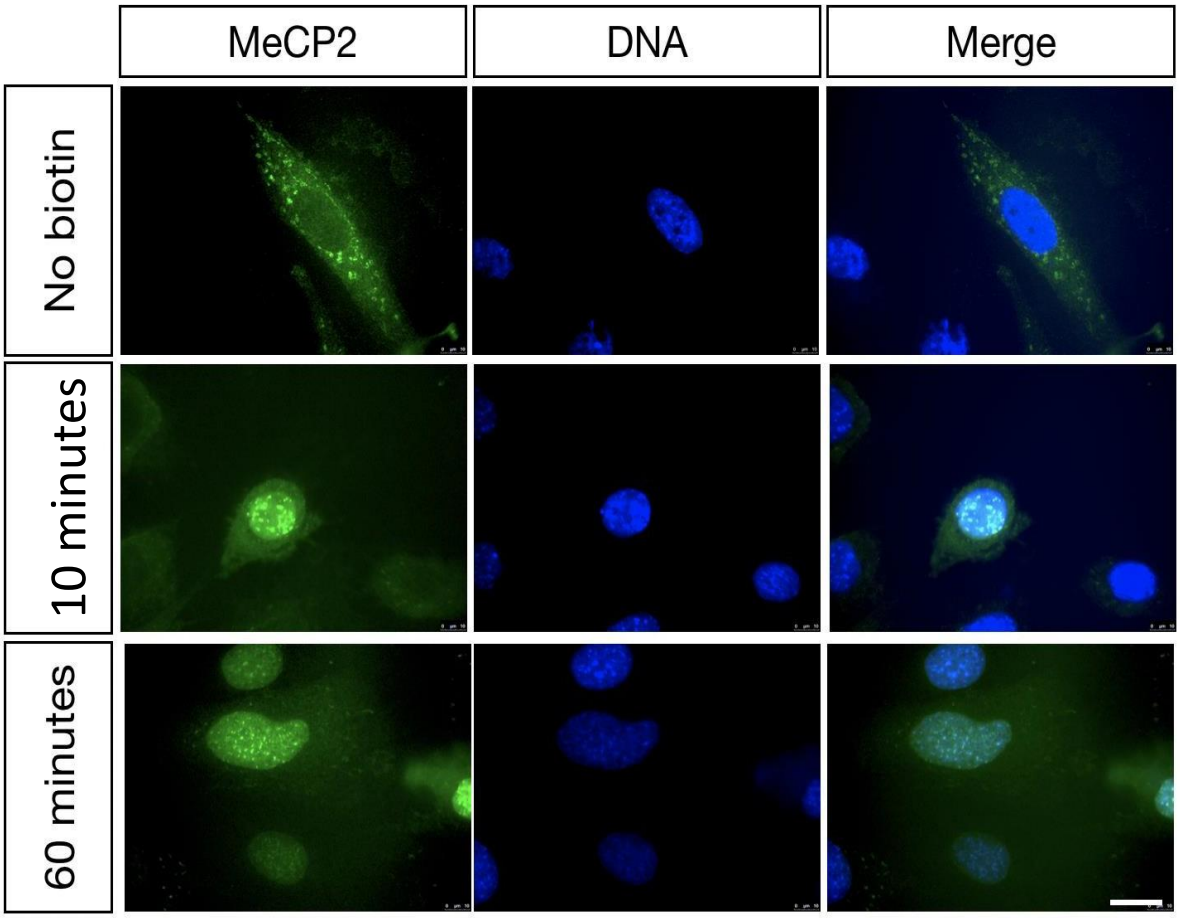


Figure 3.4.3. MeCP2-miniTurbo biotinylates nuclear proteins. Cells were transfected with MeCP2-miniTurbo. After 48 hours, cells were treated with biotin (500 µM) for either 0, 10 or 60 minutes. Biotinylation was detected using Streptavidin conjugated to Alexa488. Scale bar represents 20 µm.

3.4 Biotinylation by promiscuous ligases TurboID and miniTurbo

Studies have shown that both TurboID and miniTurbo are able to biotinylate proximal proteins, but, however, differ in efficiency. To establish whether conjugated TurboID and miniTurbo are able to induce robust biotinylation, I transfected cells with MeCP2-TurboID and MeCP2-miniTurbo. Cell lysates were separated on an SDS-PAGE. Subsequently, proteins were transferred to a nitrocellulose membrane, and detected with Streptavidin-HRP. In non-transfected cells, or transfected cells without exogenous biotin, I detected the presence of at least two bands (Fig 3.5). The intensity of these bands was unaffected by the presence of the presence of

exogenous biotin, suggesting these are endogenously biotinylated proteins. In transfected treated with exogenous biotin, I observed increased biotinylation of proteins (Figure 3.5b). The increase in biotinylation was similar for MeCP2-TurboID and MeCP2-miniTurbo. Together, these data demonstrate that MeCP2-TurboID and MeCP2-miniTurbo lead to robust biotinylation of proteins in NIH-3T3 cells

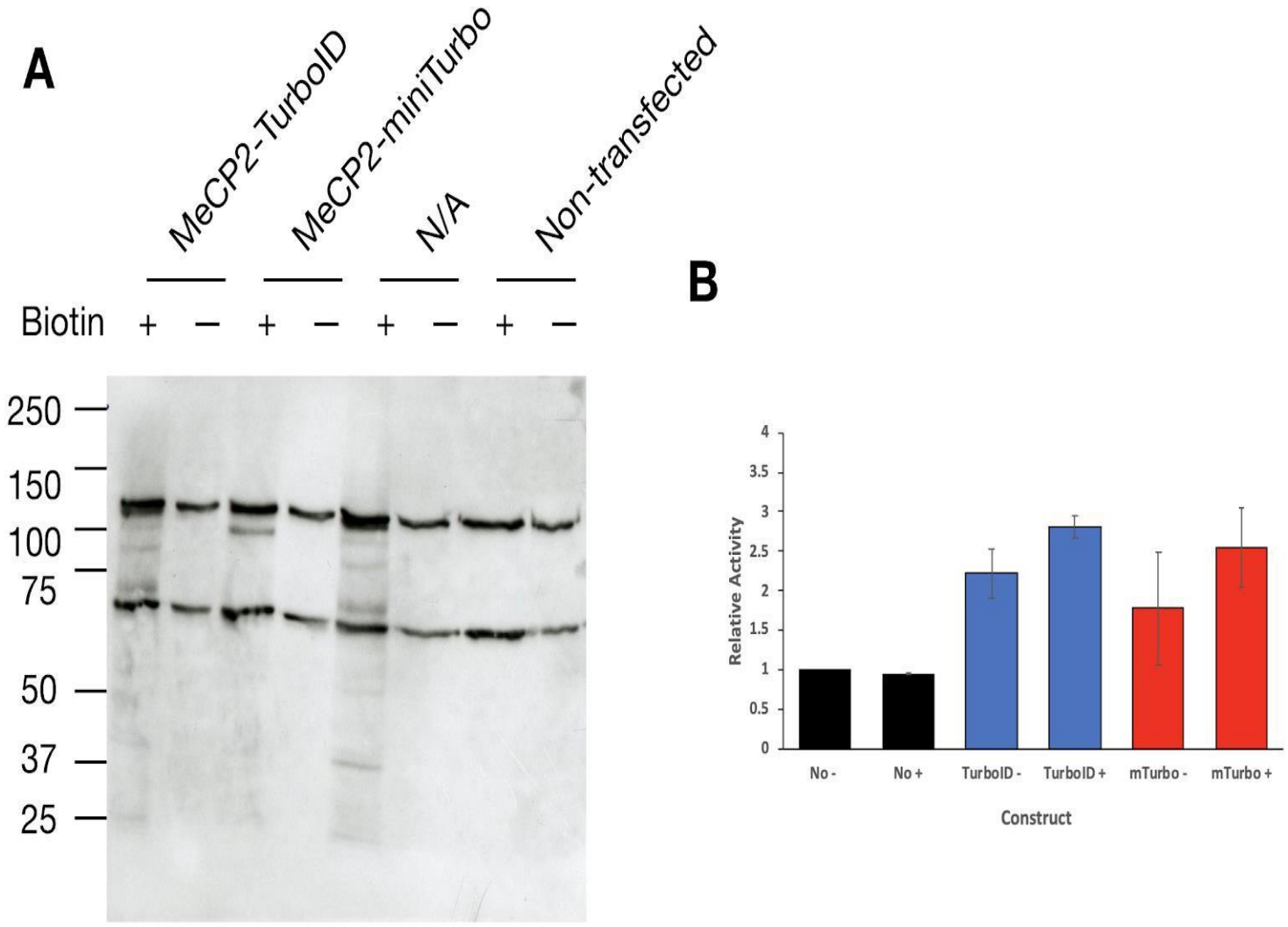


Figure 3.5. Biotinylation by MeCP2-TurboID and MeCP2-miniTurbo in NIH-3T3 cells. (A) Representative western blot showing biotinylated proteins in NIH-3T3 cell lysates from non-transfected cells and cells transfected with either MeCP2-TurboID or MeCP2-miniTurbo. Cells were incubated in the absence or presence of exogenous biotin (500 µM). Biotinylation was visualised using Streptavidin conjugated to HRP. (B) Quantification of Streptavidin-HRP signal intensity from western blots (n = 2). Bars represent mean ± sem.

3.5 Identifying protein distribution according to nuclear fractionization

To further examine the nuclear localisation of biotinylated proteins, I prepared nuclear and cytoplasmic extracts from NIH-3T3 cells transfected with MeCP2-TurboID and MeCP2-miniTurbo. In the absence of exogenous biotin, I observe 2 bands in the cytoplasmic fraction and a single band in the nuclear fraction in cells transfected with both MeCP2-Turbo and MeCP2-miniTurbo (Fig 3.6). Contrastingly, I observed robust biotinylation in both the cytoplasmic and nuclear fractions of MeCP2-Turbo transfected cells (Fig 3.6). Overall, it can be seen that, again, biotinylation is more robust with MeCP2-TurboID than with MeCP2-miniTurbo(Fig 3.6). Biotinylation appears to occur predominantly in the cytoplasm, as highlighted by the presence of biotinylated proteins in the cytoplasmic fraction, in the presence of biotin.

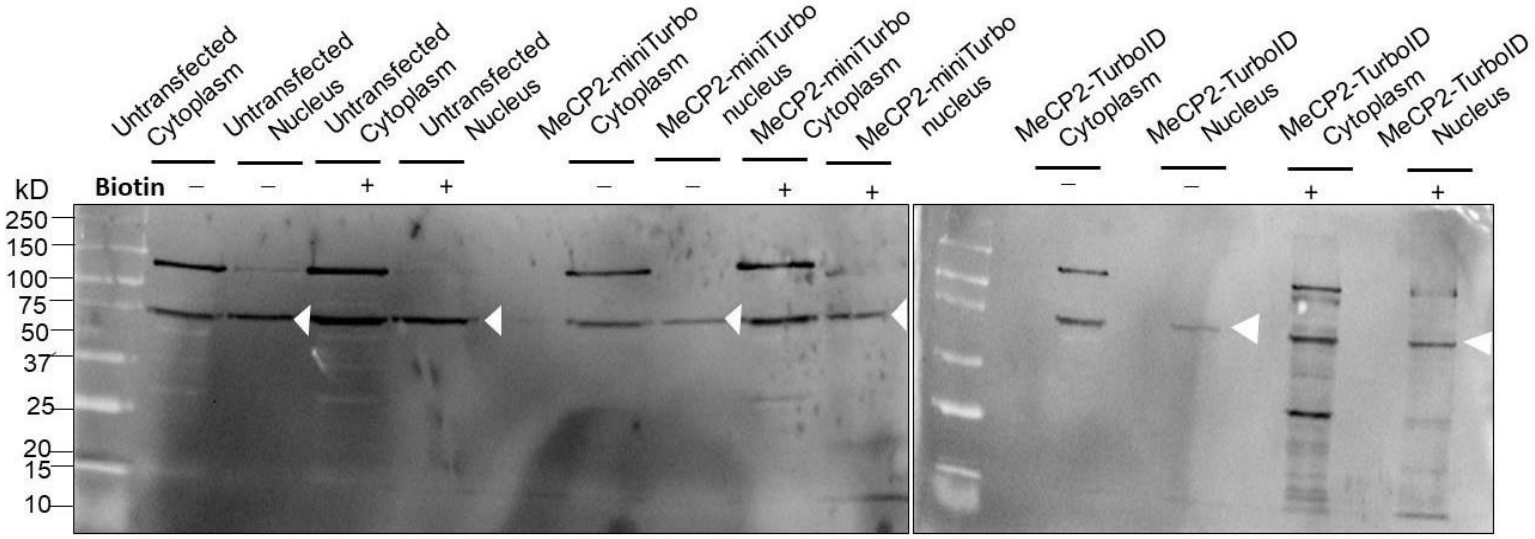


Figure 3.6 Nuclear extraction of transfected NIH-3T3 cells. NIH- 3T3 cells were TransIT-X2 transfected with MeCP2-TurboID and MeCP2-miniTurbo according to manufacturer's instructions and left to incubate for a minimum of 24 hrs at 37℃. Wells were either treated with biotin (500 µM) for 60 minutes or no biotin was added. Nuclear extraction was conducted to separate the cytoplasmic and nuclear fractions of cells and proteins were separated on a 9% SDS-PAGE gel. Proteins were transferred via a western blot transfer. Membrane was blocked in 3% BSA in TBS- T for 1 hr. Blocking was followed by incubation in 1/3000 streptavidin-HRP in 3% BSA TBS-T overnight at 4 ℃. Untransfected cells were used as a control. Arrows highlight proteins detected at the expected molecular weight of MeCP2.

3.6. Selection using G418

Cells were transfected with MeCP2-TurboID and MeCP2-miniTurbo, however, it was unclear whether biotinylation by transient transfection would provide sufficient signal for future pulldown assays. To address this, the next step was the production of a stable cell line using antibiotic selection. Sequence analysis highlighted that MeCP2 expresses a neomycin gene. In mammalian cells, when the neomycin gene is present, resistance to G418 is conferred by the neo gene in expression vectors which contain elements derived from transposons Tn601 or Tn5. The vector used in this project, pEGFP-N1_MeCP2 (WT; Addgene) contains aminoglycoside 3'-phosphotransferase from Tn5, conferring resistance to neomycin, Kanamycin and G418 (Fig 2.3a) This was also conserved in the conjugated plasmid (Fig 2.3b). Post transfection, cells were treated media containing G418. Only cells containing the plasmid which conferred resistance to G418 would hence be able to grow. After approximately 2 weeks, when selection was considered to be complete, cells were plated onto cover slips and treated with biotin followed by streptavidin-488 and visualised.

Selection appears to have been successful. Transfected NIH-3T3 cells were able to grow in the G418 containing media, indicating successful incorporation of the plasmid of interest (Fig 3.7). Additionally, fluorescent imaging showed no presence of live untransfected cells after approximately 14 days. Untransfected cells were unable to survive in G418 containing media when treated under the same conditions as transfected cells (Fig 3.7). This reinforces the fact that only cells containing the plasmid of interest would be present after the selection period.

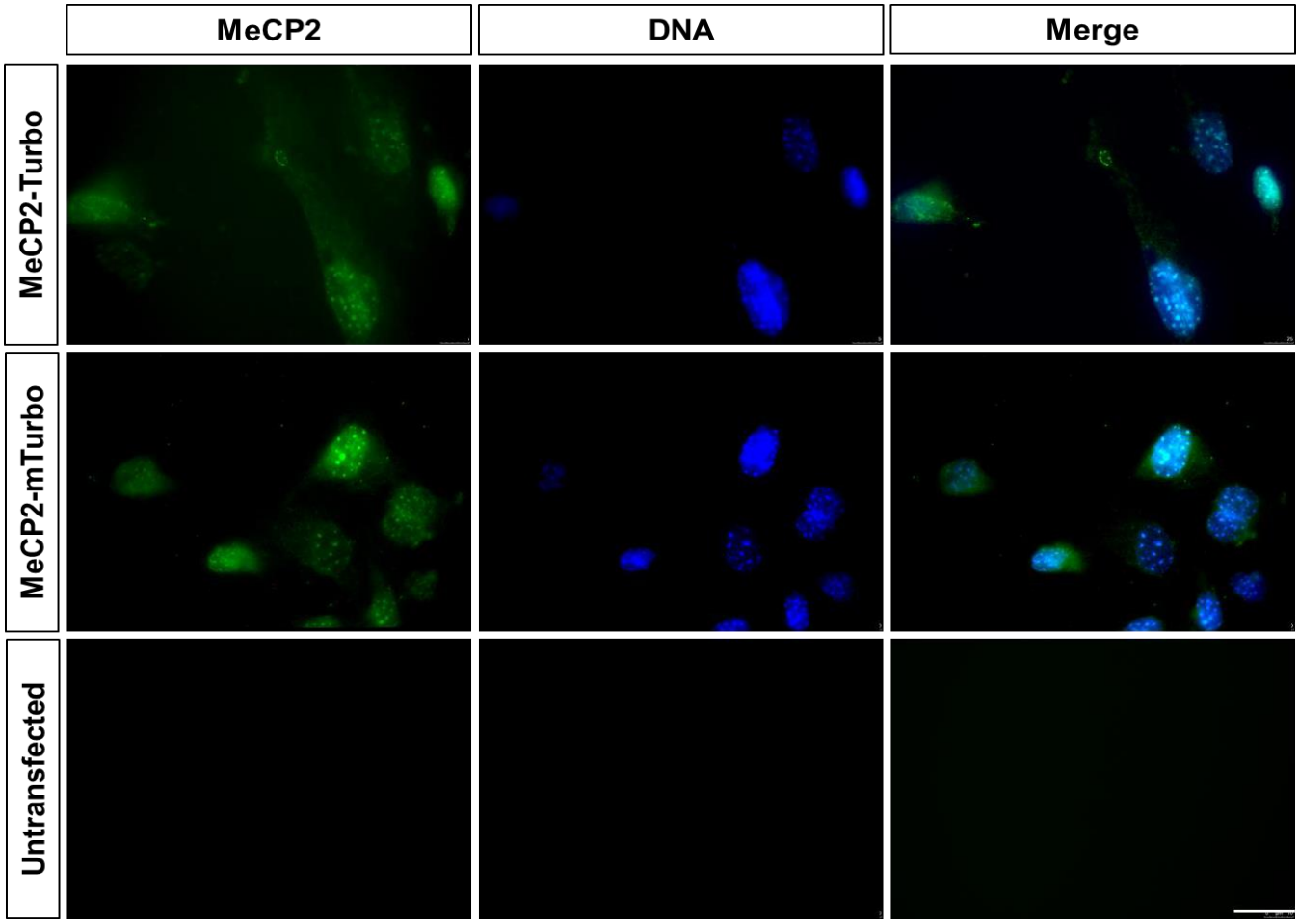


Figure 3.7 selection of cells containing MeCP2-TurboID and MeCP2-miniTurbo. Antibiotic selection of NIH-3T3 cells was created with cells TransIT-X2 transfected with MeCP2-TurboIDand MeCP2-miniTurbo. Cells were plated in G418 (400 µg/ml) containing media and media was replaced every 2 days. Cells were plated onto cover slips, treated with exogenous biotin (500 µM) for 1 hr followed by fixing with 4% parafolmadehyde and incubation with streptavidin-488 (1/1000). Visualisation occurred at 40x magnification. Scale bars represents 50 µm.

3.7 Generating a stable cell line containing MeCP2-TurboID

To overcome limitations caused by low transfection efficiencies, I generated a stably transfected cell line for MeCP2-TurboID derived from a single cell. The selection was performed using G418 (400 µg/ml).

The first round of clonal selection revealed a mix of cells containing the protein of interest and not containing the protein of interest. This was established by treatment with biotin followed by incubation with streptavidin. Cells containing MeCP2-TurboID would biotinylate nearby proteins

and travel to the nucleus, as seen in Figure 3.8, whereas those without MeCP2-TurboID incorporated would not be able to biotinylate proteins.

After considering which wells contained the highest percentage of protein expressing cells, these were left to grow and another round of monoclonal selection was conducted. To confirm stable cell line generation, we treated cells with exogenous biotin (500 µM) for 60 minutes and examined biotinylation using Streptavidin-488. I show that 100% of cells contained robust nuclear biotinylation (Figure 3.8). I thus confirmed the generation of a stable cell line.

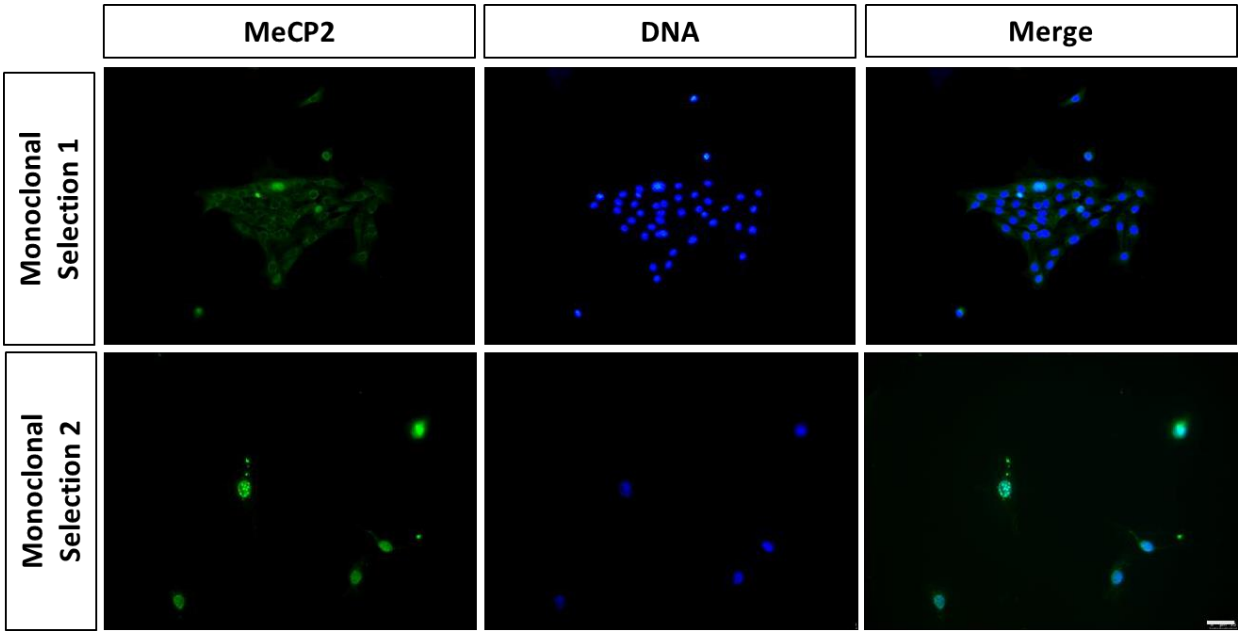


Figure 3.8 A stable cell line containing MeCP2-TurboID was created. NIH-3T3 cells transfected with MeCP2-TurboID. This was followed by treatment with biotin (500 µM) and Streptavidin-488 for fluorescent visualisation. Scale bar represents 25µm.

3.8 Cloning TBLR1 into MeCP2

3.8.1 Integration of restriction sites into TBLR1

Primers were designed containing XhoI and BamHI restriction sites on the forward and reverse primers respectively. Post PCR and PCR purification, a restriction digest using XhoI and NotI

restriction enzymes was conducted on the insert plasmid, TBLR1, as well as the vector plasmids MeCP2-EGFP and MeCP2-TurboID (Fig 3.9). The presence of the amplified insert DNA (TBLR1) band is seen at 1.5Kb and the bands present at approximately 5kb and 1.5kb in the vector highlight successful restriction digest.

Cloning of TBLR1 was attempted into WT MeCP2 (MeCP2-EGFP) so that, if successful, fluorescence microscopy of NIH-3T3 cells transfected with this construct would be used to confirm successful localisation of this protein.

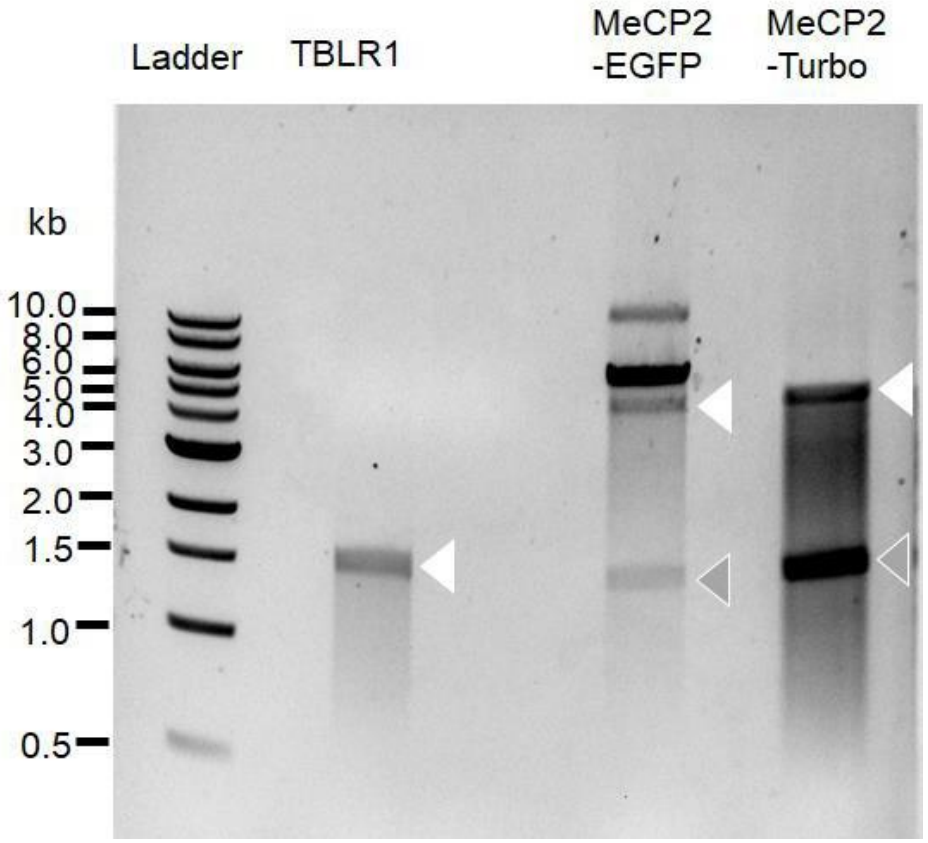


Figure 3.9 Restriction digest of TBLR1 PCR fragment and MeCP2-EGFP and MeCP2-TurboID plasmids. Representative agarose gel showing TBLR1 PCR fragment and the vector plasmids, MeCP2-EGFP and MeCP2-miniTurbo after restriction digestion with BamHI and XhoI. White arrow shows the cut plasmid; hollow arrow shows the cut MeCP2 fragment.

4-Discussion

4.1 Summary

This study applied a recently developed proximity labelling approach to the identification of MeCP2 protein interaction networks. Proximity labelling involves the covalent tagging of biotin to proximal interacting proteins of MeCP2 using a promiscuous biotin ligase. This methodology overcomes a number of issues with other protein-protein interaction approaches (Kim and Roux, 2016). Using two efficient promiscuous biotinylating enzymes derived from BirA, TurboID and miniTurbo, I attempted to provide evidence that proximity labelling is a feasible technique for finding interacting proteins of MeCP2.

Other aims of the study, which comprised confirming existing interactions and cloning TBLR1 into MeCP2-TurboID, were not achieved in the time frame of this project. This is further discussed in the limitations and future directions (see section 4.8 and 4.8.1).

4.2 MeCP2-TurboID and MeCP2-miniTurbo localise to heterochromatin

Prior studies have shown that biotin labelling may be carried out in different cell organelles (Branon et al.,2018). As MeCP2 is a nuclear protein, it was important to establish whether transfecting the protein's C-terminus hindered its ability to localise to heterochromatic foci. In this study I demonstrate that the promiscuous ligases, TurboID and miniTurbo (Branon et al., 2018) can be used to detect nuclear proteins in close proximity to MeCP2. Upon expression in NIH-3T3 cells, I show that tagging TurboID and miniTurbo to the C-terminus of MeCP2 do not affect the protein's nuclear localisation (Fig 3.2). This finding is in line with that of previous studies in mice, in which C and N termini alterations have been shown not to interfere with the functionality of the protein. When mice expressed the minimal protein, consisting of the two main domains, the MBD and NID, they survived for over a year with only mild symptoms (Tillotson et al.,2017). The critical regions of the protein, the MBD and NID were deliberately avoided when choosing insertion sites for TurboID and miniTurbo.

Additionally, previous studies have shown that fusing a Gal4 binding domain to the N-terminus of MeCP2 and co-transfecting this with a Gal4-binding site fused to a luciferase reporter gene, can quantify any changes in MeCP2-mediated repression of the reporter (Ebert et al., 2013; Lyst et al., 2013). I should perform a similar repression assay using MeCP2-TurboID and MeCP2-miniTurbo. Upon providing evidence that transcriptional repression is not affected, I can

confidently state that the function of MeCP2 is not abrogated upon fusion with TurboID or miniTurboID.

4.3 Transfection efficiency

Having established successful transfection, optimisation of (TE) was required in order to obtain enough signal on future western blots and pull downs. Previous studies have evaluated various commonly used non-viral transfection reagents with regards to their ability to deliver plasmid DNA and siRNA into cultured mouse cells (Tamm et al.,2016). TransIT-X2 has been previously found to be the most efficient reagent for transfection of adherent mouse cells (Tamm et al.,2016). Additionally, TransIT-X2 is widely considered an advanced, non-liposomal and low toxicity transfection reagent (Mirus protocol) and was hence chosen for transfections in this study. Transfections were optimised with differing concentrations of DNA. My data show that TE, using the TransIT-X2 system was highest when 0.7 µg DNA were present in each well (Fig 3.3). This is higher than that suggested by the manufacturers. This pattern can also be seen in cells transfected with MeCP2-miniTurbo where the highest TE was also achieved by transfection using 0.7 µg DNA per well (Fig 3.3). Despite this, efficiencies observed in this study were lower than expected when considering previous literature. TransIT transfection reagents have previously yielded efficiencies as high as 40% in NIH-3T3 cells, as determined by the number of EGFP expressing cells (Hayes., 2010). Considering this, the TE achieved in this study appears relatively low. It is of note that transfection in the C-terminus of MeCP2 has not been done before using constructs such as TurboID and miniTurbo, thus making direct comparisons difficult and thus the difference between this study's findings and previous study's findings may have occurred by chance.

Additionally, in this instance, only the DNA concentration was increased. While it is expected that increasing DNA concentration would increase transfection, it is surprising that the maximum TE was not achieved with the highest DNA concentration. This decrease in TE when using a higher DNA concentration could be due to inefficient DNA-TransIT complex formation. While DNA concentrations were increased, transfections were done with TransIT added to 2 µl/µg DNA. In future, to further maximise TE, transfections with differing amounts of TransIT from 2 µl/µg DNA to 6 µl/µg DNA should be done and TE recalculated. A higher TransIT concentration is likely to aid the formation of DNA-TransIT complexes and thus result in increased TE.

It is also likely that increased cell confluency could hinder transfection, as the densely packed cells could make the transfection reagent less able to access all cells. The TransIT reagent is recommended for transfection at 80% confluence (Mirus Protocol). The enzymatic dissociation involved in successful transfection induces the shaving of membrane-associated factors,

including glycoproteins with highly sulphated glycosaminoglycan (GAG) polysaccharide chains (Vogel., 1978). The large number of sulphate groups makes the chains negatively charged, anionic, and hence allows interaction with various positively charged substrates. Previous studies have shown that cell surface GAGs interact with transfection reagent polymers and liposomes and inhibit cation-mediated gene transfer (Ruponen et al.,2004). Attempting transfection at different levels of confluency would hence help determine the optimal density to maximise transfection reagent access to the cell surfaces.

DNA purity may be an additional factor hindering the attainment of maximum TE using TransIT-X2. A Qiagen-Midiprep kit was used to extract DNA for transfections. While this contains additional wash steps, the presence of endotoxins may still remain, which may hinder the formation of DNA-TransIT complexes, and hence reduce maximum TE. In the future, DNA extraction should be done using an endotoxin removal kit. Additionally, optimisation of TE could be continued by using different amounts of TransIT-X2 reagent. The latter is likely to aid achievement of optimal DNA-TransIT complexes and thus maximise TE achieved.

To summarise, my data reveal that at every concentration, the TE with both constructs was similar. A decrease in TE is seen in both of these constructs at higher concentrations. The similar pattern of TE for both constructs is expected as TurboID and miniTurbo have similar molecular weights of 35 kD and 28 kD respectively (Branon et al.,2018). It is important to note that this experiment was conducted once for each construct and a more accurate statistical comparison would be achieved with at least 2 more repeats and that, indeed, this finding may not be reproducible.

4.4 MeCP2-TurboID and MeCP2-miniTurbo biotinylate proteins in NIH-3T3 cells

To evaluate the effectiveness of MeCP2-TurboID and MeCP2-miniTurbo in biotinylating nuclear proteins, I transiently transfected NIH-3T3 cells with exogenous biotin. I then assessed biotinylation by isolating cell lysates separating protein lysates by SDS-PAGE and examining Streptavidin binding to Western blots. I found that found that both MeCP2-TurboID and MeCP2-miniTurbo led to a robust increase in the number of Streptavidin-reactive bands. No additional biotinylation occurred upon the addition of biotin in untransfected cells, suggesting any changes are solely a result of biotin ligase activity. MeCP2-TurboID displayed a higher efficiency to biotinylate than MeCP2-miniTurbo when supplemented with biotin. Therefore, TurboID is more likely to abundantly enrich interacting proteins of MeCP2 when labelling. Both biotin ligases displayed activity in the absence of exogenous biotin. This is likely due to endogenous biotin in

the culture media. We performed these experiments in duplicate. In future experiments, we should perform at least one more replicate to allow statistical comparison between MeCP2-TurboIDand MeCP2-miniTurbo.

However, TurboID displayed higher basal activity than MeCP2-miniTurbo without biotin addition. Thus, TurboID cannot be temporally controlled as easily as miniTurbo. These findings suggest that TurboID has higher efficiency but lacks the ability to be controlled, unlike miniTurbo, is concurrent with previous findings (Branon et al., 2018). The heightened activity of TurboID when using endogenous biotin, could sacrifice the specificity of labelled proteins once exogenous biotin is added. If TurboID has biotinylated a number of proteins prior to the addition of exogenous biotin, more proximal acceptor sites will be saturated when biotin is added. Therefore, upon exposure to exogenous biotin, more distal, potentially unspecific, acceptors will be labelled. This is similar to the previous demonstration that specificity is reduced if biotinylation is allowed for a prolonged period (Branon et al., 2018). Hence, miniTurbo may be a better enzyme in labelling specific interactors of MeCP2.

Here, I added biotin for 1 hour or 10 mins. When untransfected cells were exposed to biotin for either 10 minutes or 60 minutes, biotinylated proteins could be seen both in the nucleus and the cytoplasm. It is of note that when biotinylation occurred for 60 minutes, there was a shift in the biotinylation signal in untransfected cells. This suggests that longer exposure of untransfected cells to biotin leads to a reduction in biotinylation signal in the cytoplasm and an increase in the nucleus. For biotinylation labelling prior to western blotting, 60-minute labelling was used in order to come out with stronger blots and ensure the labelling of the largest quantity of MeCP2 interactions. It has been previously suggested that labelling for 60 minutes was less specific than that for 10 minutes (Branon et al.,2018). In future experiments, 10-minute biotin labelling would reduce the likelihood of unspecific interactors.

Indeed, miniTurbo is less stable than TurboID, which may be giving rise to the differing biotinylation activities found (Branon et al., 2018). This experiment should be repeated, and the western blot probed with an antibody specific to a haemagglutinin (HA) tag, found in the TurboID and miniTurbo constructs. The activity can then be quantified with respect to the expression of the ligase (HA-tag intensity), to understand the efficiency of each ligase. There are two strongly biotinylated bands at approximately 70 kDa and 110 kDa in every condition. These are likely endogenously biotinylated proteins, not interactors of MeCP2.

Our experiments revealed the presence of biotinylated proteins in untransfected NIH-3T3 cells. Using Alexa488-conjugated Streptavidin, we found that these endogenously biotinylated proteins

are predominantly localised to the cytoplasm in the absence or presence of exogenous biotin. In contrast, the addition of exogenous biotin led to a robust increase in nuclear biotinylation in cells transfected with MeCP2-TurboID or MeCP2-miniTurbo.

4.5 MeCP2-TurboID and MeCP2-miniTurbo biotinylate nuclear proteins

To summarise, MeCP2 has affinity for methylated DNA and hence I would expect to see it in the nucleus post successful transfection. On the other hand, the proteins detected in the nucleus after transfection with MeCP2TurboID/miniTurbo could indeed be nuclear proteins and be biotinylated and hence also transported to the nucleus. Both TurboID and miniTurbo were capable of biotinylating proteins after just 10 minutes (Fig 3.4.2-3.4.3). This result is expected as it has been previously found that biotinylation by unconjugated TurboID and miniTurbo can be seen after 10 minutes in the nucleus, mitochondrial matrix, ER lumen and ER membrane (Branon et al.,2018). These data suggest that biotinylation using the above constructs in the future can be done significantly more efficiently than before with BioID. These data also suggest that both ligases may be transfected into the C-terminus of the protein without significant decrease to biotinylating efficiency. The latter opens avenues for the investigation of molecular interactions of various proteins.

We showed that proteins biotinylated by MeCP2-TurboID and MeCP2-miniTurbo (when fused to MeCP2), are located in the nucleus, particularly at the heterochromatin, following the addition of biotin. The localisation of these proteins is in the same regions we found MeCP2 to localise. This suggests biotinylation of proteins is targeted within close proximity to MeCP2. We cannot be sure whether these biotinylated proteins are specific interactors of MeCP2, but it is promising that proteins appear to tightly co-localise to the MeCP2 position. We also showed that both 10 minutes and 1 hour of biotinylation is sufficient for MeCP2-TurboID and MeCP2-miniTurbo to produce enough biotinylated proteins for an intense immunofluorescent signal. Therefore, it may well be feasible to carry out biotin labelling for just 10 minutes.

In the absence of exogenous biotin, I noticed the presence of endogenously biotinylated proteins in the cytoplasm and note their absence from the nucleus. To remove these proteins during biochemical experiments, I prepared nuclear and cytoplasmic extracts from NIH-3T3 cells transfected with MeCP2-TurboID. I found the presence of two endogenously biotinylated protein bands in the cytoplasmic fraction. I also found a weak non-specific biotinylation band in the nuclear fraction. Whether this is a truly nuclear protein remains to be determined since I used a somewhat crude preparation of nuclei. Future experiments should use a more stringent preparation of neuronal nuclei.

Based on the nuclear localisation of MeCP2-TurboID and MeCP2-miniTurbo seen during earlier stages of this project, it was expected that nuclear extraction would also reveal the presence of biotinylated proteins in the nuclear fraction. Accordingly, nuclear extraction reveals a band at approximately 75kDa in the nuclear fraction of NIH-3T3 cells (Fig 3.6). Considering that endogenous MeCP2 is 72kDa, then the recombinant protein (MeCP2-Turbo or MeCP2-miniTurbo) should have also been biotinylated, which would give a MW weight band in transfected cells.Additionally, the fact that MeCP2 has affinity for methylated DNA, present in the nucleus, means a band at the above molecular weight is expected in the nuclear fraction. This affinity for methylated DNA would lead to the proteins localising to heterochromatic foci in the nucleus. Evidence suggests the redundancy in the role of a nuclear localisation signal in this process. MeCP2 may locate to the nucleus due to its affinity to DNA alone, without the need for importins such as KPNA3 and KPNA4 (Lyst et al*.,*2018), which reinforces the reason for a single band in the nuclear fraction. However, while a recent paper has reinforced the redundancy of KPNA3 and KPNA4, another importin, KPNA1, has been implicated in the process of MeCP2 nuclear localisation. Knockouts of KPNA1, reduced MeCP2 nuclear localisation (Panayotis et al*.,*2018). However, the same was not found for knockouts of KPNA3 and KPNA4 (Panayotis et al*.,* ,2018). More research is needed to solidify the role of KPNA1 and indeed other importins in MeCP2 nuclear localisation.

Additionally, a protein of approximately 75kDa is also detected in the nuclei of untransfected cells (Fig 3.6). There are two main possible explanations for this. This finding may have arisen if NIH-3T3 cells produce their own MeCP2. Despite being most abundantly in the brain, MeCP2 is present in most tissues and cell types (Coy et al*.,* 1999; Shahbazian et al*.,*2002; Akbarian et al*.,*2001). If this is the case, it is possible that the cell's own MeCP2 may be detected in a western blot analysis. Secondly, data may be depicting another protein of a similar molecular weight to that of MeCP2. While biotinylation detectable by western blot can be seen, which would likely only occur in the presence of MeCP2-TurboID and MeCP2-miniTurbo, probing the membrane with a specific antibody against MeCP2 would be shed light on the identification of the protein in the nuclear fraction of my analysis.

My data reveals a protein of approximately 75kDa is also present in the in the cytoplasmic fraction of transfected cells. This may arise due to the fact that proteins, including MeCP2, are produced in the endoplasmic reticulum, a cytoplasmic organelle. Additionally, biotinylated proteins can be seen in the cytoplasmic fraction of cells transfected with MeCP2-TurboID despite the fact that MeCP2 is a nuclear protein. This finding was unexpected as one of my previous experiments highlighted biotinylated proteins can be seen localised the nucleus of transfected NIH-3T3 cells post addition of exogenous biotin (Fig 3.4.2-3.4.3). This may occur due to numerous reasons. After MeCP2 production, it may still be present in the cytoplasm before entering the nucleus, explaining why proteins would be biotinylated in this fraction. Moreover, it is possible that MeCP2

interacts with importins in the cytoplasm but not in the nucleus. This explanation agrees with the hypothesis that importins may act as chaperones. Importins circulate between nucleus and cytoplasm and transfer cargo molecules from one side of the nuclear envelope to the other (Jäkel et al.,2002). At the nucleoplasmic side of the NPC, where Ran binds to the importin, the importin/substrate complex dissociates, leading to the importin being is re-exported from the nucleus to the cytoplasm (Jäkel et al., 2002). If this is the case, these interactions would be seen in the cytoplasmic fraction of transfected cells.

The slow diffusion kinetics of MeCP2 (Piccolo et al.,2019) may also shed light on the reason for the presence of biotinylation in the cytoplasm of transfected cells, particularly those transfected with MeCP2-TurboID. It has been identified that MeCP2 has strikingly slow diffusion dynamics compared to molecules such as Sox2 or TBP molecules (Piccolo et al.,2019). This slow rate of diffusion may hence mean that the protein is in the cytoplasm for a longer amount of time, allowing for biotinylation to occur here. It is important to consider that the new protein made containing the biotinylating ligases TurboID and miniTurbo may have an increased weight compared to the WT protein. Considering that the MeCP2-WT is at the limit for free diffusion through the nuclear pore (Wang and Brattain, 2007) due to its elongated shape (Klose and Bird, 2004) it may be that larger conjugated plasmids may have a further reduced rate of diffusion and hence are retained in the cytoplasm for longer. The possibility that the conjugated construct could weigh more and diffuse even slower cannot be excluded. Future experiments regarding the cellular kinetics of conjugated MeCP2 would be beneficial.

4.6 Stable cell line containing MeCP2-TurboID and MeCP2-miniTurbo

Plasmid vectors designed for stable expression of foreign and recombinant proteins often contain two distinct expression cassettes, or promoters; one for the protein of interest and another for the selection marker, such as G418 (Hobbs et al.,1998). Transfected cells are selected based on the acquired resistance to a specific drug, such as G418 (Hobbs et al.,1998). Using this basis for selection, in this study, I successfully selected for cells containing MeCP2-TurboID and MeCP2-miniTurbo. Figure 3.7 shows that no untransfected cells are present in wells which were in G418 media. The latter is expected due to the fact that untransfected cells would not have the G418 resistance gene present in the recombinant plasmids consisting of MeCP2 and hence would not survive in G418 containing media. Conversely, cells transfected with MeCP2-TurboID and MeCP2-miniTurbo are highlighted by their biotinylation ability, conferred by the presence of the

biotinylating ligases TurboID and miniTurbo in the plasmids of interest taken up by the transfected cells. Successful integration of the recombinant plasmid in transfected cells would confer resistance to G418 and, in theory, integration of the biotinylating ligase.

Non protein expressing cells can be seen in wells containing cells transfected with MeCP2-TurboID and MeCP2-miniTurbo (Fig 3.7). As the transfected cells express drug resistance but do not express the protein of interest, they are false positives (Hobbs et al.,1998). In this case, it is likely that cells may have successfully taken up the plasmids of interest, allowing survival and growth in media containing G418, but are not expressing the proteins of interest. It is possible that the transfected circular plasmid will be linearised by a random cut within the cell and it might be possible that sensitive parts of the plasmid such as the resistance gene or the gene of interest are destroyed upon linearisation. This means that if the resistance gene and the gene of difference are under different promoters, as is the case in the recombinant plasmid in this project, the resistance gene may be incorporated into the cell's genome without the incorporation of the protein of interest. Cells containing the resistance gene would survive in G418 media due to the G418 resistance gene present in the plasmid constructs without necessarily possessing the protein of interest. In the future, linearising the construct via digestion with BamHI and XhoI prior to transfection would aid more efficient incorporation of the protein of interest; the linearisation would mean that if the resistance gene is incorporated into the plasmid, so would the protein of interest. Additionally, the use of a bicistronic vector could also be employed for more efficient incorporation of the plasmid of interest. Bicistronic vectors are often used to express recombinant proteins at higher levels. The vital aspect of bicistronic cloning is the fact that an internal ribosome entry site (IRES) is involved. When an IRES segment is between two reporter reading frames in a eukaryotic mRNA molecule it is able to drive translation of the downstream protein coding region independent of the 5'-cap structure bound to the 5' end of the mRNA molecule (Pelletier and Sonenberg., 1988). In such a model, both proteins are produced and expressed in the cell.

4.7 Cloning TBLR1 into MeCP2

To the best of my knowledge, this study highlights the first attempt of cloning TBLR1 into MeCP2-EGFP and MeCP2-TurboID. Additionally, using the ligase containing vectors already designed as part of the project means that when carrying out biotin labelling in future experiments, protein interactions of MeCP2 and those of TBLR1 could be compared as they would have not been confounded by the use and hence protein interactions of a different vector. The latter would enable for more robust comparison of the interactors of MeCP2 and TBLR1 and potentially shed light on the function and importance of the MeCP2-TBLR1 interaction.

Sequence analysis highlighted the need to incorporate XhoI and BamHI restriction sites in the primers for amplification of the insert. The first set of primers designed were too short, with only 20 bases to anneal to the DNA fragment of interest (Table 2). The latter reduced is likely to reduce specificity and hence means this short primer is more likely to be able to bind elsewhere with complementary base pairs other than the insert fragment of interest. The latter would mean that the ligation would not be successful. To correct this, I designed primers that are longer and with more bases at the beginning to aid efficient annealing.

The reverse primer had an odd number of bases. This would lead to a lack of biotinylation when cloned with TurboID as the sequence coding for the biotinylating ligase would be out of frame due to the presence of odd number of bases in the linker between TBLR1 and TurboID. To correct this another base was added to the linker.

I was able to successfully amplify the insert DNA using the primers designed and use restriction digest with BamHI and XhoI to cut the insert and vector DNA. Bands at approximately 1.5kb, the size of the TBLR1 insert, indicate that the PCR reaction and the primers used in this successfully amplified an insert of the correct size. Additionally, restriction enzymes were able to use the complimentary restriction sites to cut the MeCP2-EGFP and the MeCP2-TurboID plasmids, resulting in bands of the expected length at the restriction digestion stage. The latter suggests success of the cloning procedure until the restriction digestion.

Despite the above amendments to the primers and thus cloning procedure, I was unable to successfully clone TBLR1 into MeCP2-EGFP and MeCP2-TurboID, as observed by lack of transformants post competent cell transformation and overnight incubation at 37^{o}C under antibiotic selection. There could be a number of reasons for this. One possibility is that the ligation was inefficient. Ligation was attempted for different durations of time- overnight at and for 2 hrs and this did not yield success.

One other cloning avenue to pursue could be to avoid the use of restriction sites altogether and use Sequence Ligation Independent Cloning (SLIC; see 4.8).

4.8 Limitations and future work

While the biotinylating efficiency of BioID has been compared to that of Turbo and miniTurbo, in this study the biotinylating efficiency of conjugated Turbo and miniTurbo have not been compared to conjugated BioID. It would be interesting to see if conjugating BioID has any effect on its biotinylating efficiency and to check if the differences in biotinylating efficiency seen by Branon et al in unconjugated forms of the ligases are reiterated in its conjugated forms in mammalian cells.

A stable cell line was generated for MeCP2-TurboID but not MeCP2-miniTurbo. MiniTurbo has been suggested as the primary ligase to use when highlighting specific interactions is a priority. A stable cell line of miniTurbo may be useful to produce and conduct biotinylation experiments in the future and can be achieved in an identical way as to that used for MeCP2- Turbo in this project; first by transfecting and seeding into individual cells followed by selection with G418.

Previous studies have highlighted various interactors of MeCP2, with some of the major ones being NCOR, SMRT, TBLR1 and HDAC3 (Kruusvee et al.,2017; Li et al.,1997; Codina et al.,2005). Post successful establishment of a stable cell line with both constructs, MeCP2-TurboID and MeCP2-miniTurbo, a pulldown assay could be performed. The interaction between biotin and streptavidin isolates the interacting proteins, thus, this interaction could be used to confirm if these proteins interact with MeCP2. To do this, streptavidin beads could be used to enrich the proteins biotinylated by TurboID or miniTurbo. Following this, peptides would be formed via on-bead digested on the protein prior to being isotopically tagged via a tandem mass tagging for the quantification of the relative amounts of proteins tagged by TurboID and miniTurbo. Liquid chromatography-tandem mass spectrometry analysis would be used to highlight the protein interactome for the proteins biotinylated by TurboID and miniTurbo conjugated MeCP2, aided by the stable nature of the proteins. This interactome would be analysed for the co-complexes known to be associated with MeCP2, and hence used to identify novel MeCP2 interactions. Comparisons of protein co-complexes would be done by using established databases such as RefSeq. Identifying novel interactions with, and confirming, rebuking, or quantifying established interactions, would be an initial step for a clearer picture of the biological purpose of MeCP2 to be constructed.

In this study, both MeCP2-TurboID and MeCP2-miniTurbo were able to label biotinylated proteins in as little as 10 minutes or 60 minutes. Conjugating Turbo and miniTurbo did not alter the biotinylation efficiency of the constructs. While I have shown that this short time frame continues to be sufficient for biotinylation to occur with a conjugated , an interesting avenue to pursue could be to enable biotinylation to occur for longer and various periods of time, such as 16 hrs, followed by a pull down assay of the biotinylated proteins for mass spectrometry analysis, as highlighted above. This would allow us to detect any differences in proteins biotinylated by the new constructs when biotinylation is allowed to continue for different amounts of time.

We used a mouse fibroblast cell line, NIH-3T3, for ease of use and the heterochromatin foci that form following DAPI staining. This cell line may not be particularly relevant for understanding interactions when considering RTT. MeCP2 is expressed very highly in the brain, therefore may have different functions and interactions with proteins in neurons. We should therefore repeat our

experiments using neuronal cell lines as this will likely elucidate interacting proteins of MeCP2 more relevant to brain functionality.

Changes in MeCP2 interactions in RTT can be assessed using this methodology. Fusing TurboID or miniTurbo to MeCP2 with an RTT-causing mutation, such as $MeCP2^{T158M}$ or $MeCP2^{R306C}$, can identify interactions lost due to mutation when compared to wild-type MeCP2. This may be applied to both neuronal cells and mouse models of RTT-syndrome. Creation of transgenic mice containing a mutant MeCP2 fused to TurboID or miniTurbo will be able elucidate pathologically relevant changes in interactions at an organismal context rather than just in cell culture. This approach may also be undertaken using cells from RTT patients. The ability to reprogramme somatic cells into iPSCs from RTT patients and differentiate them into neurons, can allow the use of proximity labelling to observe interactions in human cells carrying different RTT-causing mutations.

A potential constraint needing to be addressed in mouse models is loss of viability or health upon biotin ligase expression. It was found that Drosophila and C. elegans were less viable when TurboID and miniTurbo were expressed, most likely due to biotin starvation (Branon et al., 2018). This may present mice with health defects as a result of biotin deficiency. To overcome this, we could supplement the mouse diet with increased biotin as done previously (Johnson et al., 2017).

A SLIC method of cloning could be used to clone TBLR1 into TurboID and miniTurbo. SLIC mimics homologous recombination by using exonuclease, such as T4 DNA polymerase in the absence of dNTPs, for generation of single stranded overhangs are annealed *in vitro* either in the presence or absence or and transformed into cells (Figure 4.1; Li and Elledge., 2012). RecA may be used to promote homology searching and annealing *in vitro* but it need not be used of sufficient amounts of DNA are present (Li and Elledge., 2012). Importantly, the homology need not be perfect for SLIC, as stretches of nonhomology on the ends of annealed sequences are removed upon transformation in E. coli (Figure 4.1; Li and Elledge., 2012). In addition, excess excision is well tolerated, providing an increased margin of error in fragment preparation unless fragments are very small (Li and Elledge., 2012). SLIC is hence likely to be able to overcome limitations encountered in the restriction and ligation dependent method used in this project.

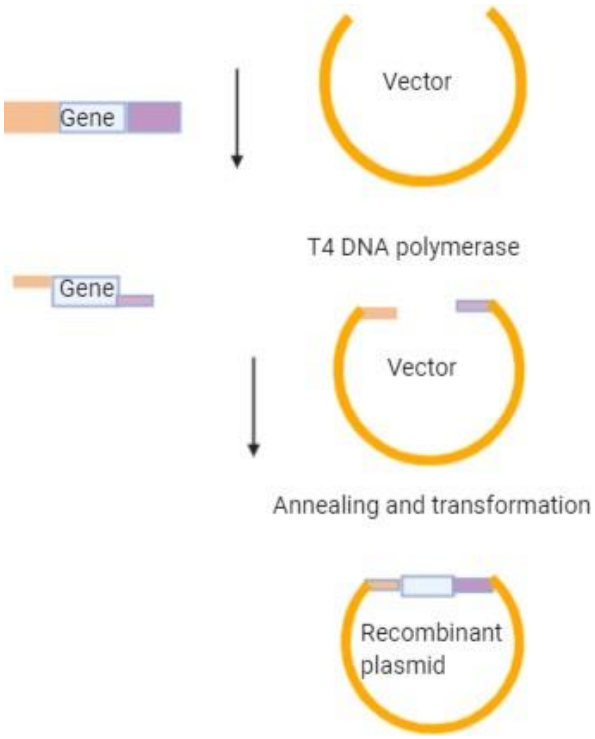


Figure 4.1 In vitro homologous recombination using SLIC. Schematic highlighting the process of producing recombinant DNA using SLIC. Created using Biorender and adapted from Li and Elledge, 2012.

Cloning of TBLR1 into the vector plasmid in order to establish its interaction with MeCP2 could provide an interesting avenue of research. Additionally, cloning TBLR1 into Turbo and miniTurbo containing plasmid would enable the determination of proximal proteins by biotinylation and thus possibly confirm the interaction between TBLR1 and MeCP2 as well as identify new interactions.

4.8.1 Alternative way of addressing and validating protein interactions found for MeCP2

Once mass spectrometry of proteins biotinylated by stable cell lines containing MeCP2-TurboID and MeCP2-miniTurbo is available, interactions found should be confirmed via another method. One way to confirm these interactions could be to use fluorescence resonance energy transfer (FRET). This term describes a distance dependent process by which energy is transferred from an excited fluorophore, known as the donor, to another fluorophore, known as the acceptor (Förster., 1965). FRET is an accurate measure of molecular proximity, and hence interaction, when proteins tagged with the donor and acceptor fluorophores are within 10nm of each other in live cells (Figure 4.2; Förster., 1965). This radius of 10 nm, measurable by FRET, is identical to that labelled by TurboID and miniTurbo, as described by Branon et al.,2018. Biotinylated proteins

identified by mass spectrometry can then be tagged with fluorescent FRET pairs, such as EGFP and mCherry and energy transfer between these established in live cells via confocal based FRET experiments (Bajar et al.,2016).

Additionally, the fact that FRET experiments are done on live cells means that FRET on proteins biotinylated by TurboID and miniTurbo could be used to identify their location. In turn, this could shed light on the localisation conundrum witnessed in this project where biotinylated proteins were seen in the nucleus by fluorescent imaging but in the cytoplasm when nuclear extraction was conducted.

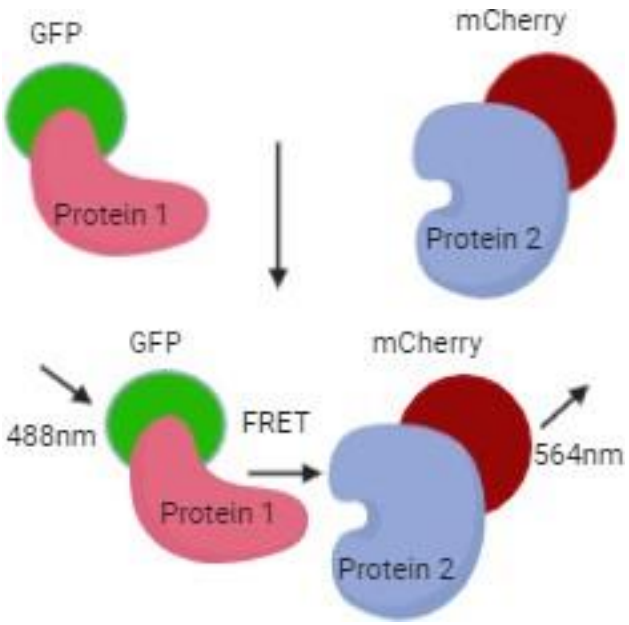


Figure 4.2 The principle of fluorescent energy transfer (FRET). When proteins tagged with an appropriate donor fluorophore, such as GFP, are interacting with (with 10nm protein) tagged with an acceptor fluorophore, FRET occurs.

4.8.2 Summary of limitations and contingencies

In this chapter I have discussed various limitations of this study as well as future directions to address these. Below, I summarise the main limitations and contingencies for this study (Table 4).

Table 4- Main limitations and contingencies for this study

Limitation	Possible cause	Contingency
Limited or reduced transfection efficiency	Endotoxins may have affected the formation of DNA-TransIT complexes:	Conduct DNA extraction using an endotoxin removal kit
		Attempt transfection optimisation using 2 µl/µg DNA to 6 µl/µg DNA
	Cell confluency may have been too high at the time of transfection	Attempt to transfect at different cell densities
Localisation of MeCP2 biotinylated proteins isn't consistent as seen via fluorescence and nuclear extraction	Protein expression may differ in non-fixed cells:	Conduct live cell imaging using FRET as this would allow the visualisation of the transportation of biotinylated proteins
Stable cell line contained non-protein expressing cells post antibiotic selection	Protein of interest and antibiotic marker are under different promoter sequences	Linearise the vector by digestion with BamHI and NotI as this would increase the likelihood that both insert and vector are incorporated together.
		Bicistronic vectors can be used to express a recombinant plasmid at higher level.
Cloning of TBLR1 was unsuccessful	Restriction sites were partially complimentary, leading to inefficient incorporation of the restriction sites into the insert plasmid	Use another set of restriction sites that are not complimentary
		Attempt a SLIC method to clone the insert into vectors

4.9 Implications of this project

Despite decades of research on MeCP2, there is still controversy regarding the role of MeCP2. Some controversy arises from 2 of various possible remaining questions:
-Is MeCP2 solely a transcriptional repressor?
-Can it truly act also as a transcriptional activator? If so, is this dependent on the tissue type where it is being expressed?

Further study of the molecular pathways and hence the protein interactions of this protein is required using novel approaches. One such approach is that used in this project. Using PL for depicting the protein interactions of MeCP2 may reveal additional targets which may thus reveal additional RTT therapeutic avenues to pursue.

5-Conclusions

In summary, we have shown that the promiscuous biotin ligases, TurboID and miniTurbo, can be effectively tagged to the C-terminus of MeCP2 to biotinylate local proteins in cells. This approach addresses some of the deficiencies in other methodologies, such as its ability to identify weak interactions and improved specificity. This study provides background work to apply a new approach to identifying interacting partners of MeCP2, with potential outcomes on understanding and treating RTT.

Both TurboID and miniTurbo require the addition of exogenous biotin in order for biotin labelling to begin. The latter is highlighted by the fact that when transfected cells were not exposed to biotin, protein presence in the cells did not appear to be different to that of untransfected cells. Additionally, both TurboID and miniTurbo are able to biotinylate proteins in as little as 10 minutes. This is similar to that found when using the ligases in their unconjugated form and hence it is possible to conclude that conjugating the ligase did not to hinder its functionality.

Transfection efficiency was too low in this project to enable sufficient biotinylation for pull down experiments using streptavidin beads and mass spectrometry analysis. To address this, a stable cell line of MeCP2-TurboID and MeCP2-miniTurbo was attempted. Selection using G418 appeared to be successful as untransfected cells were no longer detected after 7 days, however, fluorescent imaging showed the presence on non-protein expressing cells after selection. Linearising the vector or employing different cloning methods may be used to address this.

I was unable to successfully clone TBLR1 into MeCP2-EGFP and MeCP2-TurboID. Without time constraints, the next step would have been to linearise the construct via digestion with BamHI and XhoI, as this has been shown to aid successful incorporation of genes of interest into mammalian cells or indeed design primers with a different combination of restriction sites. Additionally, bicistronic vectors could be employed for the cloning of TBLR1 into the vectors of interest.

Overall, the successful use of proximity labelling for the biotinylation of proximal proteins provides an interesting avenue for the pursuit of knowledge regarding protein interactions. In the future, these new constructs may be used for increasing knowledge on RTT by shedding light on specific protein interactions, which could, in turn, have significant clinical applications. Additionally, the use of biotinylating ligases may be used in constricts with other proteins and hence PL is likely to have general applications in the field of protein interactions.

Abbreviations

AA: Amino Acid

APS: Ammonium persulphate

BSA: Bovine standard albumin

BPLs: Biotin protein ligases

CGIs: CpG islands

CREB1: cyclic-AMP responsive element binding protein 1

DMEM: Dulbecco's modified Eagle's medium

DNA: Deoxyribonucleic acid

Dnmt1: DNA methyltransferase 1

Dnmt3a: DNA methyltransferase 3a

Dnmt3b: DNA methyltransferase 3b

FRET: Fluorescence resonance energy transfer

GAG: glycosaminoglycan

GFP: Green fluorescent protein

HA: Haemagglutinin

HDAC: Histone deacetylase

Hydroxymethylcytosine: hmC

MBD: Methyl binding domain

MBD1: Methyl binding domain protein 1

MBD2: Methyl binding domain protein 2

MBD3: Methyl binding domain protein 3

MBD4: Methyl binding domain protein 4

MeCP1: Methyl-CpG-binding protein 1

MeCP2: Methyl-CpG- binding protein 2

Methylcytosine: mC

NCoR: Nuclear receptor co-repressor

NE: Nuclear envelope

NID: NCoR/SMRT interaction domain

NGS: Normal goat serum

NLS: Nuclear localisation signal

NPC: Nuclear pore complex

PBS: Phosphate buffered saline

PCR: Polymerase chain reaction

PL: Proximity labelling

RIPA: Radioimmunoprecipitation assay buffer

RTT: Rett syndrome

SEM: Standard error of the mean

SDS: Sodium dodecyl sulphate

SDS-PAGE: SDS-Polyacrylamide gel electrophoresis

SLIC: Sequence ligation independent cloning

SMRT: Silencing mediator of retinoic acid and thyroid hormone receptor

TBL1: Tranducin (beta)-like 1

TBLR1: Transducin (beta)- like related 1

TBS: Tris buffered saline

TE: Transfection efficiency

TET: Ten eleven translocation

TRD: Transcriptional repressor domain TSA: Trichostatin AWT- Wild type

References

Abrahamyan, S. et al.,(2012). Measurement of the neutron radius of Pb 208 through parity violation in electron scattering. *Physical review letters*, 108(11), p.112502.

Agarwal, N., Becker, A., Jost, KL., Haase, S., Thakur, BK., Brero, A., Hardt, T., Kudo, S., Leonhardt, H., and Cardoso, MC. (2011) MeCP2 Rett mutations affect large scale chromatin organization. *Human molecular genetics*, 20(21), pp.4187–4195.

Amir, R.E. et al.,1999. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature genetics*, 23(2), pp.185–188.

Ballestar, E. and Wolffe, A.P. (2001) Methyl-CpG-binding proteins: Targeting specific gene repression. *European journal of biochemistry / FEBS*, 268(1), pp.1–6.

Bajar, B.T., Wang, E.S., Zhang, S., Lin, M.Z., and Chu, J. (2016) A guide to fluorescent protein FRET pairs. *Sensors,* 16(9), 1488.

Barker, D.F. and Campbell, A.M. (1981) The birA gene of Escherichia coli encodes a biotin holoenzyme synthetase. *Journal of Molecular Biology*, 146(4), pp.451–467.

Bayer, E.A. and Wilchek, M. (1990) Protein biotinylation. *Methods in Enzymology*. Academic Press, 184, pp. 138–160.

Baylin, S.B, and Ohm, J.E. (2006). Epigenetic gene silencing in cancer- a mechanism for early oncogenetic pathway addiction? *Nature Reviews Cancer*, 6, pp. 107-116

Beckett, D., Kovaleva, E. and Schatz, P.J. (1999) A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein science: a publication of the Protein Society*, 8(4), pp.921–92.

Ben-Shachar, S., Chahrour, M., Thaller, C., Shaw, CA., and. Zoghbi. (2009) Mouse models of MeCP2 disorders share gene expression changes in the cerebellum and hypothalamus. *Human Molecular Genetics*, 18(13), pp.2431–2442.

Bird, A.P. and Taggart, M.H., 1980. Variable patterns of total DNA and rDNA methylation in animals. *Nucleic acids research*, 8(7), pp.1485–1497.

Bird, A.P. (2002). DNA methylation patterns and epigenetic memory. *Genes and development*, 16, pp. 6-21.

Branon, TC., Bosch, JA., Sanchez, AD., Udeshi, ND., Svinkina, Tanya., Carr, SA., Feldman, J.L., Perrimon, N., and Ting, AY. (2018). Efficient proximity labeling in living cells and organisms with TurboID. *Nature biotechnology*, 36(9), pp.880–887.

Brown K, et al., (2016) The molecular basis of variable phenotypic severity among common missense mutations causing Rett syndrome. *Hum Mol Genet*, 25(3), pp.558–570

Chadwick, L.H. and Wade, P.A. (2007) MeCP2 in Rett syndrome: transcriptional repressor or chromatin architectural protein. *Current opinion in genetics and development*, 17(2), pp.121–125.

Chahrour, M., Jung, SY., Shaw, C., Zhou, X., Wong, STC., Qin., J., and Zoghbi, HY. (2008) MeCP2, a key contributor to neurological disease, activates and represses transcription.

Science, 320(5880), pp.1224–1229.

Chapman-Smith, A. and Cronan, J.E., Jr, (1999) The enzymatic biotinylation of proteins: a post translational modification of exceptional specificity. Trends in biochemical sciences, 24(9), pp.359–363.

Cheadle, J.P. (2000). Long-read sequence analysis of theMECP2gene in Rett syndrome patients: correlation of disease severity with mutation type and location. *Human Molecular Genetics*, 2000(9), pp.1119–1129

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(3), pp.327–331.

Cheng, T., Wang, Z.., Liao, Q., Zhu., Y., Zhou, W., Xu, W., and Qiu,' Z. (2014) MeCP2 Suppresses Nuclear MicroRNA Processing and Dendritic Growth by Regulating the DGCR8/Drosha Complex. Developmental cell, 28(5), pp. 547-560.

Choi-Rhee, E., Schulman, H. and Cronan, J.E., (2004). Promiscuous protein biotinylation by Escherichia coli biotin protein ligase. *Protein science: a publication of the Protein Society*, 13(11), pp.3043–3050.

Craig, JM., Earle, E., Canham, P., Wong, LH., Anderson, M., and Choo, KHA. (2003) Analysis of mammalian proteins involved in chromatin modification reveals new metaphase centromeric proteins and distinct chromosomal distribution patterns. *Human molecular genetics*, 12(23), pp.3109–3121.

Cronan, J.E. (2005) Targeted and proximity-dependent promiscuous protein biotinylation by a mutant Escherichia coli biotin protein ligase. *The Journal of nutritional biochemistry*, 16(7), pp.416–418.

Cuddapah, V.A., et al*.,* (2014) Methyl-CpG-binding protein 2 (*MEPC2*) mutation type is associated with disease severity in Rett Syndrome. *Journal of medical genetics*, 51(3), pp.152-158

Cull, M.G. and Schatz, P.J. (2000) Biotinylation of proteins *in vivo* and *in vitro* using small peptide tags. In *Methods in Enzymology*. Academic Press, pp. 430–440.

Ebert, D.H., et al. (2013) Activity-dependent phosphorylation of MeCP2 threonine regulates interaction with NCoR. *Nature,* 499, pp. 341-345.

Förster, T. (1965) Delocalized excitation and excitation transfer. In Modern Quantum Chemistry, 3, pp.93-137

Fujita, N., Takebayashi, S., Okumura, K., Kudo, S., Chiba, T., Saya, H., and Nakao, M. (1999) Methylation-mediated transcriptional silencing in euchromatin by methyl-CpG binding protein MBD1 isoforms. *Molecular and cellular biology*, 19(9), pp.6415–6426.

Gabel. et al*.,*(2015) Disruption of DNA-methylation-dependent long gene re-pression in Rett syndrome, Nature, 522(7554), 89-93

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

secondarychromatin structures in the absence of DNA methylation. The Journal of biological chemistry, 278(34), pp.32181–32188.

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

Girard, M., Couvert, P., Carrie, A., Tardieu, M., Chelly, J., Beldjord, C., and Bienvenu, T. (2001) Parental origin of de novo MECP2 mutations in Rett syndrome. *European Journal of Human Genetics*, 9, pp. 231-236.Li

Goffin, D. et al*.,* (2011) Rett syndrome mutation MeCP2 T158A disrupts DNA binding, protein stability and ERP responses. *Nature neuroscience*, 15(2), pp.274–283.

Guo, J.U. et al*.,* (2014) Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nature neuroscience*, 17(2), pp.215–222.

Grunstein, M. (1997) Histone acetylation in chromatin structure and transcription. *Nature,* 389, pp.349-352.

Guy, J., Hendrich, B., Holmes, M., Martin, J.E., 2001. A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nature genetics*, 27(3), pp.322–326.

Heckman, L.D., Chahrour, M.H. and Zoghbi, H.Y. (2014) Rett-causing mutations reveal two domains critical for MeCP2 function and for toxicity in MECP2 duplication syndrome mice. *eLife*, 3.

Hendrich, B. and Bird, A. (1998) Identification and characterization of a family of mammalian methyl CpG-binding proteins. *Genetical research*, 72(1), pp.59–72.

Horike, S., Cai, S., Miyano, M., Cheng, JF., Kohwi-Shigematsu, T. (2005) Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nature genetics*, 37(1), pp.31–40.

Horvath, P.M. and Monteggia, L.M. (2017) Engineering MeCP2 to spy on its targets. *Nature medicine*, 23(10), pp.1120–1122.

Huang, L.H., Wang, R., Gama-Sosa, MA., Shenoy, S., and Ehrlich, M. (1984) A protein from human placental nuclei binds preferentially to 5-methylcytosine-rich DNA. *Nature*, 308(5956), pp.293–295.

Ito-Ishida, A. et al*.,*2018. Genome-wide distribution of linker histone H1.0 is independent of MeCP2. *Nature neuroscience*, 21(6), pp.794–798.

Jäkel, S., Mingot, J.M., Schwarzmaier, P., Hartmann, E., and Görlich, D. (2002) Importins fulfil a dual function as nuclear import receptors and cytoplasmic chaperones for exposed basic domains. *The Embo journal*, 21(3), pp.377-386

Jellinger, K., Armstrong, D., Zoghbi, HY., Percy, AK. (1988) Neuropathology of Rett syndrome. *Acta neuropathologica*, 76(2), pp.142–158.

Johnson, B.S. et al*.,* (2017) Biotin tagging of MeCP2 in mice reveals contextual insights into the Rett syndrome transcriptome. *Nature medicine*, 23(10), pp.1203–1214.

Kim, D.I., and Roux, K.J. (2016). Filling the void; proximity-based labelling of proteins in living

Kimcells. *Trends Cell Biol.* 26, pp.804-817.

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 of the United States of America*, 113(52), pp.15114–15119.

Klose, R.J., and Bird, A.P. (2004) MeCP2 Behaves as an Elongated Monomer That Does Not Stably Associate with the Sin3a Chromatin Remodeling Complex. The journal of biological chemistry, 279(5), pp. 46490 –46496.

Kokura, K., et al*.,* (2001) The Ski protein family is required for MeCP2-mediated transcriptional repression. *Journal of Biological Chemistry*, 276(36), pp.34115–34121.

Kriaucionis, S. and Bird, A. (2004) The major form of MeCP2 has a novel N-terminus generated by alternative splicing. *Nucleic acids research*, 32(5), pp.1818–1823.

Kruusvee, V., Lyst, M.J., Taylor, C., Tarnauskaitė, 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. *Proceedings of the National Academy of Sciences of the United States ofAmerica*, 114(16), pp.E3243–E3250.

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

Lee ,R.C., Feinbaum, R.L., and Ambros, V. (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell,* 75(3), pp.843–854

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(6), pp.905–914.

Li, M.Z., and Elledge, S.J. (2012) SLIC: a method for sequence-and ligation-independent cloning. *Gene synthesis*, 852, pp.51-59.

Lister, R. et al*.,* (2013) Global epigenomic reconfiguration during mammalian brain development. *Science*, 341(6146), p.1237905.

Lu, W.C., Levy, M., Kincaid, R., and Ellington, A.D. (2014) Directed evolution of the substrate specificity of biotin ligase. *Biotechnology and bioengineering*, 111(6), pp.1071–1081.

Lyst, M.J. et al*.,* (2013) Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. *Nature neuroscience*, 16(7), pp.898–902.

Lyst, M.J., Ekiert, R., Guy, J., Selfridge, J., Koerner, M.V., Merusi, C., De Sousa, D., and Bird, A. (2018) Affinity for DNA Contributes to NLS Independent Nuclear Localization of MeCP2. *Cell reports*, 24(9), pp.2213–2220.

Meehan, R.R., Lewis, J.D. and Bird, A.P. (1992) Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic acids research*, 20(19), pp.5085–5092.

Meehan, R.R., Lewis, J.D., McKay, S., Kleiner, E.L., and Bird, A.P. (1989) Identification of a

mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell*, 58(3), pp.499–507.

Mellen, A., Ayata, P., Dewell, S., Kriaucionis, S., and Heintz, N. (2012) MeCP2 Binds to 5hmC Enriched within Active Genes and Accessible Chromatin in the Nervous System. *Cell*

Mellios, N. et al*.,* (2018) MeCP2-regulated miRNAs control early human neurogenesis through differential effects on ERK and AKT signalling. *Molecular psychiatry*, 23(4), pp.1051–1065.

Miernyk, J.A., and Thelen, J.J. (2008) Biochemical approaches for discovering protein–protein interactions. *The plant journal,* 53, pp. 597-609.

Muotri, A.R., Marchetto, M.C., Coufal, N.G., Oefner, R., Yeo, G., Nakashima, K., and Gage, F.H. (2010) L1 retrotransposition in neurons is modulated by MeCP2. *Nature,* 468(7332), pp.443-446

Na, E.S., Nelson, E.D., Adachi, M., Autry, A.E., Mahgoub, M.A., Kavalali, E.T., and Monteggia, L.M. (2012) A mouse model for MeCP2 duplication syndrome: MeCP2 overexpression impairs learning and memory and synaptic transmission. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 32(9), pp.3109–3117.

Na, E.S., Nelson, E.D., Kavalali, E.T., and Monteggia, L.M. (2013) The impact of MeCP2 loss- or gain-of-function on synaptic plasticity. *Neuropsychopharmacology,* 38*,* pp.212-219.

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

Nan, X., Meehan, R.R., and Bird, A. (1993) Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic acids research*, 21(21), pp.4886–4892.

Nan, X., Tate, P., Li, E., and Bird, A. (1996) DNA methylation specifies chromosomal localization of MeCP2. *Molecular and cellular biology*, 16(1), pp.414–421.

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(6683), pp.386–389.

Neul, J.L. et al*.,* (2010) Rett syndrome: revised diagnostic criteria and nomenclature. *Annals of neurology*, 68(6), pp.944–950.

Neul, J.L, et al*.,* (2014) Developmental delay in Rett syndrome: data from the natural history study. *Journal of neurodevelopmental disorders*, 6(20), pp. 1-9.

Neul, J.L. et al*.,* (2019). The array of clinical phenotypes of males with mutations in Methyl-CpG binding protein 2. *American journal of medical genetics*, 180(1), pp 55-67.

Ng, H.H. et al*.,* (1999) MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nature genetics*, 23(1), pp.58–61.

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

Opitz, N., Schmitt, K., Hofer-Pretz, V., Neumann, B., Krebber, H., Braus, G.H., and Valerius, O. (2017) Capturing the Asc1p/*R*eceptor for *A*ctivated *C K*inase *1* (RACK1) Microenvironment at the Head Region of the 40S Ribosome with Quantitative BioID in Yeast. *Molecular and cellular proteomics,* 16(12), pp.2199-2218.

Panayotis, N., et al*.,* (2018) Importin α5 Regulates Anxiety through MeCP2 and Sphingosine Kinase 1. *Cell reports,* 25(11), pp. 3169-3179.

Pelletier, J., and Sonenberg, Nahum. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature, 334, pp.320-325.

Piccolo, F. M., Liu, Z., Dong, P., Hsu, C., Stoyanova, E.I., Rao, A., Tijan, R., and Heintz, N. (2019) MeCP2 nuclear dynamics in live neurons results from low and high affinity chromatin interactions. http://dx.doi.org/10.1101/586867.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P., and Jaenisch, R. (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proceedings of the National Academy of Sciences of the United States of America*, 97(10), pp.5237–5242.

Rett, A. (1966) On a unusual brain atrophy syndrome in hyperammonemia in childhood. *Wiener medizinische Wochenschrift* , 116(37), pp.723–726.

2012, L., et al. (2012) Morphological and functional reversal of phenotypes in a mouse model of Rett syndrome. *Brain,* 135, pp.2699–2710

Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012) A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *The Journal of cell biology*, 196(6), pp.801–810.

Ruponen, M., Honkakoski, P., Tammi, M., and Urtti, A. (2004) Cell-surface glycosaminoglycans inhibit cation-mediated gene transfer. *The Journal of Gene Medicine, (6)*, 405–414

Schatz, P.J. (1993) Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in Escherichia coli. *Biotechnology,* 11(10), pp.1138–1143.

Schmiedeberg, L., Skene, P., Deaton, A., and Bird, A. (2009) A temporal threshold for formaldehyde crosslinking and fixation. *PloS one*, 4(2), p.e4636.

Schwartzman, J.S., Bernardino, A., Nishimura, A., Gomes, R.R., Zatz, M. (2001). Rett syndrome in a boy with a 47,XXY karyotype confirmed by a rare mutation in the MECP2 gene. *Neuropediatrics*, 32(3), pp.162–164.

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

Schmid, R.S., Tsujimoto, N., Qu, Q., Lei, H., Li., E., Chen, T., and Blaustein, C.S. (2008) A methyl-CpG-binding protein 2- enhanced green fluorescent protein reporter mouse model provides a new tool for studying the neuronal basis of Rett Syndrome. *Neuroreport,* 19, pp. 393.

Sorokin, A.V., Kim, E.R., and Ovchinnikov, L.P. (2007) Nucleocytoplasmic transport of proteins. *Biochemistry. Biokhimiia*, 72(13), pp.1439–1457.

Stroud, H., Feng, S., Kinney, S.M., Pradhan, S., and Jacobsen., S.E. (2011) 5 - Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome biology*, 12, pp.1-8

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), pp.425–435.

Tamm, C., Kadekar, S., Pijuan-Galitó, S., and Annerén, C. 2016. Fast and efficient transfection of mouse embryonic stem cells using non-viral reagents. *Stem cell reviews and reports*, 12, 584-591.

Tate, P., Skarnes, W., and Bird, A. (1996a) The methyl-CpG binding protein MeCP2 is essential for embryonic development in the mouse. *Nature genetics*, 12(2), pp.205–208.

Tahiliani , M., et al*.,* (2009) Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science,* 324(5929), pp. 930-935.

Tillotson, R. et al*.,* (2017) Radically truncated MeCP2 rescues Rett syndrome-like neurological defects. *Nature*, 550(7676), pp.398–401.

Tudor, M., Akbarian, S., Chen, R.Z., and Jaenisch, R. (2002) Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain. *Proceedings of the National Academy of Sciences of the United States of America*, 99(24), pp.15536–15541.

Uezu, A., Kanak, D.J., Bradshaw, T.W.A., Soderblom, E.J., Catavero, C.M., and Burette, A.C. (2016) Identification of an elaborate complex mediating postsynaptic inhibition. *Science,* 353(6304), pp. 1123-1129.

Vogel, KG. (1978) Effects of hyaluronidase, trypsin, and EDTA on surface composition and topography during detachment of cells in culture. *Experimental Cell Research, 113*, 345–357

Wada, R., Akiyama,Y., Hashimoto, Y., Fukamachi, H., and Yuasa, Y. (2010) miR-212 is downregulated and suppresses methyl-CpG-binding protein MeCP2 in human gastric cancer. *International Journal of cancer*, 127(5), pp.1106-1114.

Wang, R., and Brattain, M.G. (2007) The maximal size of protein to diffuse through the nuclear pore is larger than 60 kDa. *FEBS Letters*, 581, pp. 3164–3170

Weber, P.C., Ohlendorf, D.H., Wendoloski, J.J., and Salemme, F.R. (1989). Structural origins of high-affinity biotin binding to streptavidin. *Science*, 243(4887), pp.85–88.

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

Yoder, J., Walsh, C.P., and Bestor, T.H. (1997). Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics*, 13, pp.335–340.

Yoshida, M., Kijima, M., Akita, M., and Beppu, T. (1990) Potent and specific inhibition of

mammalian histone deacetylase both *in vivo* and *in vitro* by trichostatin A. *The Journal of biological chemistry*, 265(28), pp.17174–17179.

Young, J.I. et al*.,* (2005). Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proceedings of the National Academy of Sciences of the United States of America*, 102(49), pp.17551–17558.

Zachariah, R.M., Olson, C.O., Ezeonwuka, C., and Rastegar, M. (2012) Novel MeCP2 Isoform-Specific Antibody Reveals the Endogenous MeCP2E1 Expression in Murine Brain, Primary Neurons and Astrocytes. *PLoS ONE*, 7(11), p.e49763.

Zhang, X.M., Chang, Q., Zeng, L., Gu, J., Brown, S., and Basch, R.S. (2006) TBLR1 regulates the expression of nuclear hormone receptor co-repressors. *BMC cell biology,* 7(1), pp. 1-31

Ziller, M.J. et al*.,* (2011) Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS genetics*, 7(12), p.e1002389.

Zlatanova, J. (2005) MeCP2: the chromatin connection and beyond. *Biochemistry and Cell Biology*, 83(3), pp.251–262.