The molecular basis of MeCP2 function in the brain

Citation for published version:

Digital Object Identifier (DOI):
10.1016/j.jmb.2019.10.004

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Journal of Molecular Biology

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
The Molecular Basis of MeCP2 Function in the Brain

Rebekah Tillotson¹,² and Adrian Bird³

¹ - Genetics and Genome Biology Program, The Hospital for Sick Children, The Peter Gilgan Centre for Research and Learning, Toronto, ON M5G 0A4, Canada
² - Medical Research Council (MRC) Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford, OX3 9DS, UK
³ - Wellcome Centre for Cell Biology, University of Edinburgh, The Michael Swann Building, King’s Buildings, Max Born Crescent, Edinburgh, EH9 3BF, UK

Correspondence to Adrian Bird: a.bird@ed.ac.uk
https://doi.org/10.1016/j.jmb.2019.10.004
Edited by Tuncay Baubec

Abstract

MeCP2 is a reader of the DNA methylome that occupies a large proportion of the genome due to its high abundance and the frequency of its target sites. It has been the subject of extensive study because of its link with ‘MECP2-related disorders’, of which Rett syndrome is the most prevalent. This review integrates evidence from patient mutation data with results of experimental studies using mouse models, cell lines and in vitro systems to critically evaluate our understanding of MeCP2 protein function. Recent evidence challenges the idea that MeCP2 is a multifunctional hub that integrates diverse processes to underpin neuronal function, suggesting instead that its primary role is to recruit the NCoR1/2 co-repressor complex to methylated sites in the genome, leading to dampening of gene expression.

© 2019 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Numerous chromatin-associated proteins and noncoding RNAs work together to establish cell type-specific epigenetic states that regulate gene expression. These factors are vital for normal mammalian development, with deletion of individual members often resulting in lethality in mutant mice [1]. Additionally, epigenetic factors often have links to neurological disease, caused by mutations affecting dosage such as haploinsufficiency, locus duplications or hypomorphic alleles [2]. Here, we discuss the role of MeCP2, a reader of the DNA methylome, encoded by the X-linked MECP2 gene. The gene has been implicated in several ‘MECP2-related disorders’ [3] prompting numerous studies of MeCP2 protein function. We discuss a spectrum of evidence that sheds light on the molecular mechanisms involved, including clinical genetic investigations of genotype-phenotype correlations and mouse models of the resulting human conditions.

Next-generation sequencing technologies have revealed where MeCP2 binds in the genome and how it interprets the DNA methylome to influence gene expression. Finally, structural studies have defined in molecular detail the interactions between MeCP2 and its key binding partners. Together, the findings provide a coherent picture of MeCP2 function as an essential reader of the DNA methylome in the brain that optimises neuronal transcription programmes.

MECP2-Related Disorders and Mouse Models

Loss of function mutations in the MECP2 gene in hemizygous male patients cause neonatal encephalopathy, which is usually fatal before the age of 2 years [4]. The same mutations cause the severe neurological disorder Rett syndrome (RTT) in heterozygous females [5]. RTT occurs in 1 in

0022-2836/© 2019 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
MeCP2 function in Brain

MeCP2 is an Essential Reader of DNA Methylation in the Brain

MeCP2 was initially discovered over quarter of a century ago because of its ability to bind DNA in a methylation-specific manner [23]. The ~90 amino acid region responsible for binding was called the methyl-CpG binding domain (MBD) [24]. Sequence similarity searches using the MBD sequence identified 10 other proteins that contain this domain: MBD1-4, MBD5 (alternative names TAM1, KIAA1461), MBD6 (alternative names TAM2, KIAA1887), BAZZA (alternative name TIP5), BAZZ2B, SETDB1 (alternative names ESET, KMT1E) and SETDB2 (alternative name CLLD8) [25–32]. Of these, only MeCP2, MBD1, MBD2 and MBD4 have been shown to bind DNA in vitro in a methylation-specific manner [23,26,32,33]. Specificity for the same binding sites in DNA raises the possibility of competition between these proteins, though analysis of their temporal-spatial expression patterns reveals differences between cell types. MeCP2 is expressed in all tissues but reaches near-histone abundance in neurons (~16 × 10^6 molecules per nucleus) [34]. Its levels increase during embryonic and postnatal development, plateauing at 10 years in humans [35] and 5 weeks in mice [34]. MBD1 is expressed during neurogenesis but is then downregulated [36]. MBD2 and MBD4 are more widely expressed across somatic tissues, and MBD4 is the only family member detected in embryonic stem cells [36,37]. Whereas complete deletion of Mecp2 in mice leads to severe neurological symptoms and death around 9 weeks of age [13,14], knocking out the other members results in minimal phenotypes [36,38–43]. Even though loss of MBD1 causes decreased neurogenesis, the animals have a normal lifespan and only mild behavioural defects [38,39]. These results suggest that MeCP2 is the MBD family member with the greatest role in interpreting the DNA methylome in the brain.

MeCP2 protein levels in human and mouse tissues correlate poorly with transcription of the MECP2/Mecp2 gene [35]. This may be explained by the activity of several posttranscriptional regulatory mechanisms, including alternative splicing, use of different polyadenylation sites and posttranslational modification. The gene spans four exons, which are transcribed and spliced to form two isoforms, e1 and e2, where only e2 includes exon 2 (Fig. 1) [44,45]. Isoform e1 is the ancestral form, conserved across vertebrates, whereas isoform e2 is only present in mammals. The two isoforms are very similar, differing only at the extreme N-terminus, and most evidence indicates that they are functionally interchangeable [46]. Importantly, e2 mRNA is translated at a much lower efficiency because of the presence of an upstream ATG [44], so the great majority of MeCP2 protein in the brain is derived from e1. The gene contains four alternative polyadenylation sites, producing 3’UTRs ranging from 0.1 to 8.6 kb, although only the longest and shortest mRNAs are detected in neurons. The use of different polyadenylation sites determines whether they contain binding sites for proteins and miRNAs that regulate mRNA stability and translation [47–49]. MeCP2 protein function and stability is also thought to be affected by multiple posttranslational modifications,
Table 1. The RTT-like phenotypic signature in Mecp2-mutant hemizygous male mice.

<table>
<thead>
<tr>
<th>Mutation (MGI)</th>
<th>Mutation type</th>
<th>Protein level</th>
<th>Median survival</th>
<th>Overt symptoms</th>
<th>Anxiety</th>
<th>Motor function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null [13, 95]</td>
<td>Mecp2&lt;sup&gt;tm1.1Bird&lt;/sup&gt;</td>
<td>Δex3-4</td>
<td>None</td>
<td>9 wks</td>
<td>↓ + + + + +</td>
<td>1/7 ↑</td>
</tr>
<tr>
<td>Null [14, 86, 179]</td>
<td>Mecp2&lt;sup&gt;tm1.1Jae&lt;/sup&gt;</td>
<td>Δex3</td>
<td>None</td>
<td>6–12 wks</td>
<td>↓ + + + + +</td>
<td>↓</td>
</tr>
<tr>
<td>Null [180]</td>
<td>Mecp2&lt;sup&gt;tm1Ppy&lt;/sup&gt;</td>
<td>Δex3 - part of ex4</td>
<td>None</td>
<td>6 wks</td>
<td>↓ + + + + +</td>
<td>↓</td>
</tr>
<tr>
<td>Mecp2/ly; Nestin-Cre [13]</td>
<td>Mecp2&lt;sup&gt;tm1Bird&lt;/sup&gt;; Tg(Nes-cre)1Kln</td>
<td>CNS KO (Δex3)</td>
<td>&gt; 90 % recombination in brain</td>
<td>6–12 wks</td>
<td>↓ ND ND + + ND</td>
<td>ND</td>
</tr>
<tr>
<td>Mecp2/ly; Nestin-Cre [14]</td>
<td>Mecp2&lt;sup&gt;tm1Jae&lt;/sup&gt;; Tg(Nes-cre)1Atp</td>
<td>CNS KO (Δex3)</td>
<td>&gt; 90 % recombination in brain</td>
<td>6–12 wks</td>
<td>↓ ND ND + + ND</td>
<td>ND</td>
</tr>
<tr>
<td>Mecp2 STOP/ly; Nestin-Cre [56]</td>
<td>Mecp2&lt;sup&gt;tm2Bird&lt;/sup&gt;; Tg(Nes-cre)1Kln</td>
<td>Peripheral KO (ΔSTOP in CNS)</td>
<td>&lt;0.1 x in periphery</td>
<td>&gt; 1 y</td>
<td>(1)</td>
<td>(1)</td>
</tr>
<tr>
<td>R106W-Tavi [87]</td>
<td>Mecp2&lt;sup&gt;tm4.1Joez&lt;/sup&gt;</td>
<td>RTT: MBD missense</td>
<td>~0.1 x (of WT-TAVI)</td>
<td>10 wks</td>
<td>Present, cumulative score ND</td>
<td>ND</td>
</tr>
<tr>
<td>Y120D [88]</td>
<td>Mecp2&lt;sup&gt;tm1.1Joez&lt;/sup&gt;</td>
<td>RTT: MBD missense</td>
<td>~0.1 x (of WT-EGFP)</td>
<td>11 wks</td>
<td>(1)</td>
<td>↓</td>
</tr>
<tr>
<td>T158M-EGFP [15]</td>
<td>Mecp2&lt;sup&gt;tm4.1Bird&lt;/sup&gt;</td>
<td>RTT: MBD missense</td>
<td>~0.1 x (of WT-EGFP)</td>
<td>14 wks</td>
<td>ND ND ND ND ND</td>
<td>ND</td>
</tr>
<tr>
<td>T158M-TAVI [87]</td>
<td>Mecp2&lt;sup&gt;tm4.1Joez&lt;/sup&gt;</td>
<td>RTT: MBD missense</td>
<td>~0.1 x (of WT-TAVI)</td>
<td>14 wks</td>
<td>ND ND ND ND ND</td>
<td>ND</td>
</tr>
<tr>
<td>T158A [86]</td>
<td>Mecp2&lt;sup&gt;tm1.1Joez&lt;/sup&gt;</td>
<td>RTT: MBD missense</td>
<td>~0.5 x</td>
<td>14 wks</td>
<td>ND ND ND ND</td>
<td>ND</td>
</tr>
<tr>
<td>R306C [171-181]</td>
<td>Mecp2&lt;sup&gt;tm2.1Meg&lt;/sup&gt;</td>
<td>RTT: MBD missense</td>
<td>~1 x (of WT-EGFP)</td>
<td>30 wks</td>
<td>ND ND ND ND</td>
<td>ND</td>
</tr>
<tr>
<td>R270X-EGFP (tg on null background)</td>
<td>Mecp2&lt;sup&gt;tm1Bird&lt;/sup&gt;</td>
<td>RTT: ID nonsense</td>
<td>Undetectable</td>
<td>12 wks</td>
<td>ND ND ND ND</td>
<td>ND</td>
</tr>
<tr>
<td>G273X-EGFP (tg on null background)</td>
<td>Mecp2&lt;sup&gt;tm1Bird&lt;/sup&gt;</td>
<td>RTT: ID nonsense</td>
<td>Undetectable</td>
<td>29 wks</td>
<td>ND ND ND ND</td>
<td>ND</td>
</tr>
<tr>
<td>G273X-EGFP</td>
<td>Mecp2&lt;sup&gt;tm1Bird&lt;/sup&gt;</td>
<td>RTT: ID nonsense</td>
<td>Undetectable</td>
<td>36 wks</td>
<td>ND ND ND ND</td>
<td>ND</td>
</tr>
<tr>
<td>S385PfsX6 [120]</td>
<td>Mecp2&lt;sup&gt;tm10.1Bird&lt;/sup&gt;</td>
<td>RTT: ID nonsense</td>
<td>Truncating (Patient mutation L386HfsX5)</td>
<td>20 wks</td>
<td>ND ND ND ND</td>
<td>ND</td>
</tr>
<tr>
<td>Floxed (exons 3 + 4) [13, 18, 91, 92]</td>
<td>Mecp2&lt;sup&gt;tm1Bird&lt;/sup&gt;</td>
<td>Hypomorphic</td>
<td>~1.8 x (of WT-EGFP)</td>
<td>41 wks</td>
<td>ND ND ND ND</td>
<td>ND</td>
</tr>
<tr>
<td>S80A [51]</td>
<td>Mecp2&lt;sup&gt;tm2.1Jae&lt;/sup&gt;</td>
<td>Mutation of phosphorylation site</td>
<td>~1 x (of WT-EGFP)</td>
<td>12 wks</td>
<td>ND ND ND ND</td>
<td>ND</td>
</tr>
<tr>
<td>T308A [52]</td>
<td>Mecp2&lt;sup&gt;tm3Moc&lt;/sup&gt;</td>
<td>Mutation of phosphorylation site</td>
<td>~1 x (of WT-EGFP)</td>
<td>12 wks</td>
<td>ND ND ND ND</td>
<td>ND</td>
</tr>
<tr>
<td>S421A/S424A [52]</td>
<td>Mecp2&lt;sup&gt;tm3Moc&lt;/sup&gt;</td>
<td>Mutation of phosphorylation sites</td>
<td>~1 x (of WT-EGFP)</td>
<td>12 wks</td>
<td>ND ND ND ND</td>
<td>ND</td>
</tr>
<tr>
<td>ΔAT-hook1 [122]</td>
<td>Mecp2&lt;sup&gt;tm1Smoc&lt;/sup&gt;</td>
<td>Deleted AT hook 1</td>
<td>~1.8 x (of WT-EGFP)</td>
<td>48 wks</td>
<td>ND ND ND ND</td>
<td>ND</td>
</tr>
<tr>
<td>G273X[NLS]-EGFP</td>
<td>Mecp2&lt;sup&gt;tm6Bird&lt;/sup&gt;</td>
<td>RTT ID nonsense + mutated NLS</td>
<td>ND</td>
<td>41 wks (NS vs G273X-EGFP)</td>
<td>ND ND ND ND</td>
<td>ND</td>
</tr>
</tbody>
</table>
ΔN-EGFP [130] (Mecp2em1Bird)  
ΔNC-EGFP [130] (Mecp2em1Bird)  
ΔN-EGFP [130] (Mecp2em2Bird)  
A140V [139,140] (Mecp2em2Bird)  

<table>
<thead>
<tr>
<th>Strain</th>
<th>Testing dispensability</th>
<th>Intellectual disability</th>
<th>Normal lifespan</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔN-EGFP</td>
<td>-1x</td>
<td>ND (mRNA ~1x)</td>
<td>ND ND ND ND ND</td>
</tr>
<tr>
<td>ΔNC-EGFP</td>
<td>-1x</td>
<td>ND (mRNA ~1x)</td>
<td>ND ND ND ND ND</td>
</tr>
<tr>
<td>ΔN-EGFP</td>
<td>~0.5x</td>
<td>ND (mRNA ~1x)</td>
<td>ND ND ND ND ND</td>
</tr>
<tr>
<td>A140V</td>
<td>~1x</td>
<td>ND (mRNA ~1x)</td>
<td>ND ND ND ND ND</td>
</tr>
</tbody>
</table>

MeCP2 function in Brain

RTT models display six overt symptoms: Act. = reduced spontaneous activity; Gait = abnormal gait; HLC = hindlimb clasping (FP = fore paw clasping); Tr. = tremor; Br. = breathing irregularities; GC = deterioration of general condition. RTT models displayed impaired motor function, measured by Elevated Plus Maze or Elevated Zero Maze. Light/Dark test was used instead in these studies, showing increased or unaltered anxiety. RTT models displayed impaired motor function, measured by one or more the following: accelerating rotarod, running wheel, parallel rod footslip and hanging wire. RTT-like symptoms lead to premature death: median survival is given where reported. For non-RTT mutants, survival is stated as greater than the time point (e.g. ‘yr’) to which they were followed and ‘<50% animals died. Some studies simply state ‘normal lifespan’. ND = not determined; NS = not significantly different from WT controls; N/A = not applicable due to lethality; Bracketed results = weak phenotype. * = symptom not mentioned (only cumulative scores given); MGI = Mouse Genome Informatics; MBD = methyl-CpG binding domain; NID = NCoR1/2 interaction domain; CTD = C-terminal domain; NLS = nuclear localisation signal; EGFP = enhanced green fluorescent protein; RTT = Rett syndrome; ID = intellectual disability; CNS = central nervous system; WT = wild-type; KO = knock-out; OE = overexpression; fl = floxed; tg = transgene wks = year. The null background used to characterise transgenic mice was Mecp2tm1.1Bird.

MeCP2 binds mCG dinucleotides and mCAC trinucleotides

The ability of MeCP2 to bind DNA containing methylated mCG dinucleotides was first identified by Southwestern assay [23] and has since been confirmed using a variety of in vitro techniques, including electrophoretic mobility shift assay (EMSA) [52], electron paramagnetic resonance (EPR) [62], and by fusion to a visible tag, taking advantage of nuclear magnetic resonance (NMR) [63]. MeCP2 binding has been visualised using immunofluorescence microscopy, including neuronal activity-dependent phosphor-mCAM (FRAP) [63]. The ability of MeCP2 to bind DNA containing methylated mCG dinucleotides was first identified by Southwestern assay [23] and has since been confirmed using a variety of in vitro techniques, including electrophoretic mobility shift assay (EMSA) [52], electron paramagnetic resonance (EPR) [62], and by fusion to a visible tag, taking advantage of nuclear magnetic resonance (NMR) [63]. MeCP2 binding has been visualised using immunofluorescence microscopy, including neuronal activity-dependent phosphor-mCAM (FRAP) [63]. The ability of MeCP2 to bind DNA containing methylated mCG dinucleotides was first identified by Southwestern assay [23] and has since been confirmed using a variety of in vitro techniques, including electrophoretic mobility shift assay (EMSA) [52], electron paramagnetic resonance (EPR) [62], and by fusion to a visible tag, taking advantage of nuclear magnetic resonance (NMR) [63]. MeCP2 binding has been visualised using immunofluorescence microscopy, including neuronal activity-dependent phosphor-mCAM (FRAP) [63]. The ability of MeCP2 to bind DNA containing methylated mCG dinucleotides was first identified by Southwestern assay [23] and has since been confirmed using a variety of in vitro techniques, including electrophoretic mobility shift assay (EMSA) [52], electron paramagnetic resonance (EPR) [62], and by fusion to a visible tag, taking advantage of nuclear magnetic resonance (NMR) [63]. MeCP2 binding has been visualised using immunofluorescence microscopy, including neuronal activity-dependent phosphor-mCAM (FRAP) [63]. The ability of MeCP2 to bind DNA containing methylated mCG dinucleotides was first identified by Southwestern assay [23] and has since been confirmed using a variety of in vitro techniques, including electrophoretic mobility shift assay (EMSA) [52], electron paramagnetic resonance (EPR) [62], and by fusion to a visible tag, taking advantage of nuclear magnetic resonance (NMR) [63]. MeCP2 binding has been visualised using immunofluorescence microscopy, including neuronal activity-dependent phosphor-mCAM (FRAP) [63]. The ability of MeCP2 to bind DNA containing methylated mCG dinucleotides was first identified by Southwestern assay [23] and has since been confirmed using a variety of in vitro techniques, including electrophoretic mobility shift assay (EMSA) [52], electron paramagnetic resonance (EPR) [62], and by fusion to a visible tag, taking advantage of nuclear magnetic resonance (NMR) [63]. MeCP2 binding has been visualised using immunofluorescence microscopy, including neuronal activity-dependent phosphor-mCAM (FRAP) [63]. The ability of MeCP2 to bind DNA containing methylated mCG dinucleotides was first identified by Southwestern assay [23] and has since been confirmed using a variety of in vitro techniques, including electrophoretic mobility shift assay (EMSA) [52], electron paramagnetic resonance (EPR) [62], and by fusion to a visible tag, taking advantage of nuclear magnetic resonance (NMR) [63]. MeCP2 binding has been visualised using immunofluorescence microscopy, including neuronal activity-dependent phosphor-mCAM (FRAP) [63]. The ability of MeCP2 to bind DNA containing methylated mCG dinucleotides was first identified by Southwestern assay [23] and has since been confirmed using a variety of in vitro techniques, including electrophoretic mobility shift assay (EMSA) [52], electron paramagnetic resonance (EPR) [62], and by fusion to a visible tag, taking advantage of nuclear magnetic resonance (NMR) [63]. MeCP2 binding has been visualised using immunofluorescence microscopy, including neuronal activity-dependent phosphor-mCAM (FRAP) [63]. The ability of MeCP2 to bind DNA containing methylated mCG dinucleotides was first identified by Southwestern assay [23] and has since been confirmed using a variety of in vitro techniques, including electrophoretic mobility shift assay (EMSA) [52], electron paramagnetic resonance (EPR) [62], and by fusion to a visible tag, taking advantage of nuclear magnetic resonance (NMR) [63]. MeCP2 binding has been visualised using immunofluorescence microscopy, including neuronal activity-dependent phosphor-mCAM (FRAP) [63]. The ability of MeCP2 to bind DNA containing methylated mCG dinucleotides was first identified by Southwestern assay [23] and has since been confirmed using a variety of in vitro techniques, including electrophoretic mobility shift assay (EMSA) [52], electron paramagnetic resonance (EPR) [62], and by fusion to a visible tag, taking advantage of nuclear magnetic resonance (NMR) [63]. MeCP2 binding has been visualised using immunofluorescence microscopy, including neuronal activity-dependent phosphor-mCAM (FRAP) [63]. The ability of MeCP2 to bind DNA containing methylated mCG dinucleotides was first identified by Southwestern assay [23] and has since been confirmed using a variety of in vitro techniques, including electrophoretic mobility shift assay (EMSA) [52], electron paramagnetic resonance (EPR) [62], and by fusion to a visible tag, taking advantage of nuclear magnetic resonance (NMR) [63]. MeCP2 binding has been visualised using immunofluorescence microscopy, including neuronal activity-dependent phosphor-mCAM (FRAP) [63].
protein and the large number of potential binding sites. In the bulk genome, CG dinucleotides occur on average every 100 bp and are highly methylated. In contrast, ‘CpG Islands’ are enriched in CG dinucleotides and tend to be unmethylated [66–68], but they account for only 1–2% of the genome. Consistent with MeCP2 occupying a large fraction of mCG sites, this analysis results in a relatively featureless ChIP-seq signal that follows methylation density, interrupted by dramatically reduced binding at unmethylated CpG islands [34,69,70]. In spite of technical limitations, peak-calling algorithms have been successfully employed for ‘summit analysis’ using ChIP-seq and bisulphite sequencing data from the mouse frontal cortex to produce aggregate plots showing an enrichment of peaks over mCG sites, but not at unmethylated CG dinucleotides [71]. To investigate the in vivo DNA binding footprint of MeCP2 further, assay for transposase-accessible chromatin using sequencing (ATAC-seq) was used to determine protected genomic regions around mCG sites in cultured human neurons [72]. The ATAC-seq signal from wild-type samples was divided by the equivalent signal in Mecp2 knock-out samples, revealing a striking DNA binding footprint corresponding to the region protected by MeCP2 over mCG [72]. This DNA methylation-dependent footprint, which is absent at nonmethylated CG sites, is 11 bp in size, consistent with in vitro DNaseI footprinting [24] and MNase protection [73].

High levels of mCH (where H is A, C or T) have recently been discovered in the brain [74], primarily at CA dinucleotides. The highest level of this non-CG methylation is found at the trinucleotide CAC (~12%
of CAC sites) [70], as indicated by motif analysis of non-CG methylation in mouse and human brain tissue [75,76]. This is likely because of the relaxed enzymatic specificity of the de novo DNA methyltransferase, DNMT3A [77]. Although the percentage of methylation at individual CA dinucleotides is low, the higher abundance of CA compared with CG in the genome means that the total number of mCA sites approaches that of mCG sites [70,74]. Interestingly, non-CG methylation accumulates during neuronal maturation at the same time as the build-up of MeCP2 [34,35,74], raising the possibility that MeCP2 is able to bind methylation in a non-CG context. EMSA analysis demonstrated that MeCP2 could bind a probe containing mCH [76] and subsequent in vitro studies narrowed the sequence preference first to mCA [71] and then to mCAC (and to a lesser extent mCAT) trinucleotides [70]. The requirement for adenine in the second position suggested that MeCP2 could recognise the methyl group of thymine on the complementary strand. Indeed, removal of this group by replacement of thymine with uracil (mCAC/GTG to mCAC/GUG) abolished the interaction [70]. Binding to mCAC and mCAT but not mCAA or mCAG was confirmed in cultured cells co-transfected with MeCP2 and methylated oligonucleotides and ChIP-seq data from mouse hypothalamus indicated a strong peak of MeCP2 binding over mCAC (but not mCAT) sites in vivo [70]. Binding to mCAC dramatically increases the number of MeCP2 binding sites in the genome and alters their distribution. Comparison between neuronal subtypes found greater differences in the patterns of mCH than mCG, suggesting that mCAC is a major contributor to neuronal cell type-specific patterns of MeCP2 binding [78,79]. In mice, these patterns are established in the first few weeks of life when DNMT3A binding and mCH deposition occurs over the transcribed regions of lowly expressed genes [80].

The oxidised form of methyl-cytosine, hydroxymethylcytosine (hmC), is the product of the first step of the active demethylation pathway mediated by the ten-eleven translocation (Tet) enzymes [81]. Like mCH, hmC accumulates to uniquely high levels in postmitotic neurons [74,82], prompting the search for readers. In vitro analyses found that MeCP2 could bind to probes containing hmCAC (and to a lesser extent hmCAT) [60,70] but could not bind the predominant form of hydroxymethyl-cytosine, hmCG dinucleotides [57,59,70,71]. Co-transfection of cultured cells with MeCP2 and hydroxymethylated oligonucleotides confirmed the ability of MeCP2 to bind hmCAC sites [70]. Unlike hmCG, hmCAC sites are very rare in the mammalian brain [74], suggesting that the ability to bind hmCAC may not be of biological relevance. Interestingly, accumulation of hmCG in neurons by oxidation of mCG sites means that MeCP2 can no longer bind at these loci. This, plus the appearance of novel mCAC binding sites, will profoundly redistribute the profile of MeCP2 binding in the neuronal genome during postnatal maturation.

The crystal structure of the MBD of MeCP2 bound to a DNA molecule containing a central mCG site showed that binding is mediated by direct interactions between the nucleotides making up the mCG site and two arginine ‘fingers’, Arg111 and Arg133 [83]. The Arg111 sidechain is constrained by hydrogen bonding to Asp121, but Arg133 is relatively unconstrained. Computational modelling based on this structure predicted that the MBD can bind to mCAC or hmCAC-containing DNA in a single orientation by altering the position of Arg133 alone.

![Fig. 2. Posttranslational modifications of MeCP2.](image-url)
Recent structural studies have broadly confirmed this hypothesis for mCAG [84]. The minor structural change required to accommodate mCAG, together with the likelihood that the structure of MeCP2 outside the globular MBD is predicted to be largely disordered [85], suggests that MeCP2 does not interpret mCAG and mCAC sites differently.

### DNA Binding is Essential for MeCP2 Protein Function

A functional MBD is vital for the role of MeCP2 as a reader of the methylome. Its biological importance is indicated by the large number of RTT-causing missense mutations that map to this domain, almost all of which impair binding (Table 2). Further experimentation is needed to verify whether other, less well-characterised mutations also impact binding. For example, RTT-causing mutations affecting three prolines, P101, P127 and P152 were reported in vivo, because of their localisation at pericentromeric foci in fixed cells [62]. However, a DNA binding defect for these mutants cannot be ruled out as many mutants with defective DNA binding dynamics (quantified by FRAP) nevertheless showed localisation by this assay [65]. In addition, there is evidence that many RTT-causing mutations in the MBD reduce protein stability as measured by free energy changes of protein unfolding or dispersion of HSQC spectra (Table 2) [89,90]. It is likely that both disruption of DNA binding and instability contribute to the disease phenotype in MBD mutants, as reduced expression of the wild-type protein to similar levels caused a much milder phenotype (Table 1) [91,92], whereas increasing expression of MBD mutants improved health and survival [93,94].

Knock-in mice expressing MeCP2[R111G] at wild-type levels phenocopy MeCP2-null mice and show very mild RTT-like symptoms (Table 1) [51], consistent with a gain of function effect. However, analysis of the wild-type protein by genome-wide ChIP-seq failed to identify any regions of the genome with reduced MeCP2 binding alone is sufficient to cause Rett syndrome [95]. Further underlining the importance of DNA binding for the clinical phenotype, MeCP2[R133C] retains more DNA binding than other MBD mutants [15] and gives the mildest average clinical symptom score [96].

### The Key Interaction Partner of MeCP2 is the NCoR1/2 Co-repressor Complex

A prevalent view of MeCP2 molecular activity is that it is a multifunctional hub, involved in several cellular processes via its interactions with over 40 putative binding partners (Fig. 3) [100,101]. These roles include transcriptional repression, transcriptional activation, alternative splicing, chromatin remodelling and miRNA processing. An alternative view, based on studies of its effects on transcription, is that the most important role of MeCP2 is to inhibit gene expression. DNA methylation-dependent repression was demonstrated in early studies using transfected reporter constructs that were enzymatically modified at CG sites using a bacterial methyltransferase [13,102,103]. Similar experiments also demonstrated repression mediated by mCH sites [76]. Discovery of its repressive activity led to the expectation that MeCP2-deficiency would cause upregulation of a discrete set of target genes, but the failure to unambiguously identify these in MeCP2-deficient mice initially stymied the field. Loss of MeCP2 instead led to small expression changes in both directions at a large number of genes, leading to the belief that MeCP2 might function both to repress and activate transcription [104,105]. To distinguish between direct and indirect activity, recent studies have used ChIP-PCR and...
<table>
<thead>
<tr>
<th>RTT missense mutation</th>
<th>mCG binding \textit{in vitro}</th>
<th>mCG binding \textit{in vivo}</th>
<th>Position in structure (PDB: 3C2I [83])</th>
<th>Stability</th>
<th>Mouse model</th>
</tr>
</thead>
<tbody>
<tr>
<td>D97Y</td>
<td>ND</td>
<td>Reduced [62]</td>
<td>Interaction with R106</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>L100V</td>
<td>Reduced [90]</td>
<td>Partial [62,65,142]</td>
<td>ND</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>L100R</td>
<td>ND</td>
<td>ND</td>
<td>Loop</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>P101R</td>
<td>ND</td>
<td>Binding [62,142]</td>
<td>Loop</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>P101S</td>
<td>Binding [60]</td>
<td>Binding [62,142]</td>
<td>Loop</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>P101H</td>
<td>ND</td>
<td>Binding [62,142,186]</td>
<td>Loop</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>P101L</td>
<td>ND</td>
<td>Binding [62,142]</td>
<td>Loop</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>R106Q</td>
<td>Reduced [90]</td>
<td>Reduced [62,192]</td>
<td>Stabilises Asx-ST motif</td>
<td>Destabilised [192]</td>
<td>Not made</td>
</tr>
<tr>
<td>R106L</td>
<td>ND</td>
<td>ND</td>
<td>Stabilises Asx-ST motif</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>L108H</td>
<td>ND</td>
<td>ND</td>
<td>Stabilises Asx-ST motif</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>R111G</td>
<td>Reduced [84,90,189]</td>
<td>Reduced [62,65,95,137,142,186]</td>
<td>DNA interface</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>Y120D</td>
<td>Reduced [193]</td>
<td>Reduced [62,88]</td>
<td>Loop</td>
<td>Destabilised [88] RTT-like [88]</td>
<td></td>
</tr>
<tr>
<td>L124F</td>
<td>ND</td>
<td>Reduced [62]</td>
<td>Hydrophobic core</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>P127L</td>
<td>ND</td>
<td>Binding [62,142]</td>
<td>Loop</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>A131D</td>
<td>ND</td>
<td>ND</td>
<td>Hydrophobic core</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>R133P</td>
<td>ND</td>
<td>ND</td>
<td>DNA interface</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>R133L</td>
<td>ND</td>
<td>Reduced [62,142,186]</td>
<td>DNA interface</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>S134C</td>
<td>Reduced [90]</td>
<td>Binding [62,142]</td>
<td>DNA interface</td>
<td>Destabilised [89,90] Not made</td>
<td></td>
</tr>
<tr>
<td>S134F</td>
<td>Reduced [58]</td>
<td>ND</td>
<td>DNA interface</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>S134P</td>
<td>ND</td>
<td>ND</td>
<td>DNA interface</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>K135E</td>
<td>ND</td>
<td>Reduced [62]</td>
<td>DNA interface</td>
<td>Destabilised [89,90] Not made</td>
<td></td>
</tr>
<tr>
<td>L138S</td>
<td>ND</td>
<td>ND</td>
<td>Alpha helix</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>P152R</td>
<td>Reduced [90]</td>
<td>Binding [62,142,186]</td>
<td>Loop</td>
<td>Destabilised [89,90] Not made</td>
<td></td>
</tr>
<tr>
<td>F155S</td>
<td>Reduced [90,187,188]</td>
<td>Reduced [62,142,190]</td>
<td>Loop</td>
<td>Destabilised [194] Not made</td>
<td></td>
</tr>
<tr>
<td>D156E</td>
<td>Reduced [90]</td>
<td>Reduced [62]</td>
<td>Axx-ST motif</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>D156A</td>
<td>ND</td>
<td>ND</td>
<td>Axx-ST motif</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>F157L</td>
<td>ND</td>
<td>ND</td>
<td>Axx-ST motif</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>F157I</td>
<td>ND</td>
<td>ND</td>
<td>Axx-ST motif</td>
<td>ND</td>
<td>Not made</td>
</tr>
<tr>
<td>G161V</td>
<td>ND</td>
<td>ND</td>
<td>Axx-ST motif</td>
<td>ND</td>
<td>Not made</td>
</tr>
</tbody>
</table>

MBD = methyl-CpG binding domain; RTT = Rett syndrome; PDB = Protein Data Bank; ND = not determined.
Fig. 3. MeCP2 has been reported to interact with over 40 binding partners across its length. Schematic diagram showing the interaction sites of MeCP2 binding partners characterised by their proposed function when complexed with MeCP2. The minimal domain required for transcriptional repression of a reporter gene (the TRD 205–310 [102]) is indicated by the shaded red box. RTT-causing missense mutations (red) and neutral variants found in males on the ExAC database (black) are shown above. The hotspot in the C-terminus where RTT-causing truncations occur is indicated by a red arrow. Regions deleted in mice expressing truncated MeCP2 [130] (N = N-terminus, I = Intervening region and C = C-terminus) are shown by grey dotted lines. Interaction sites for: methylated DNA 78–162 [24]; AT-rich DNA 183–195 (hook 1) [121] and 257–272 (hook 2) [121,123]; DNA (‘basic patch’) 274–340 [95]; KPNA3 249–268 [128]; KPNA4 249–270 [128]; TBL1/TBLR1 (NCoR1/2 complex) 298–309 [119]; SIN3A 108–206 and 207–308 [103,107]; c-Ski 163–309 [106]; PU.1 1–162 and 163–309 [108]; YY1 202–255 [110]; TFIIIB 204–310 [117]; CBFI 109 (unmapped); PRMT6 109 (unmapped); SP3 109 (unmapped); SOX2 111 (unmapped); BRAHMA 109 (unmapped); G9a 114 (unmapped); HLCS [115] (unmapped); ATRX [137] 106–169; HP1α/β (unmapped), requires phospho-S229 [158]; HP1γ 1–55 [131], requires phospho-S229 [158]; SMC1/3 [167] (unmapped, requires phospho-S229 [158]); Lamin B 163–201 [168]; DNMT1 77–161, 162–206 and 207–310 [169]; TET1 170 (unmapped); CREB1 105 (unmapped); MYCN [171] (unmapped); RNA 160–200 [171]; YB-1 1195–329 [133], requires phospho-S80 [158]; FBPI 311-CT [132]; HYPC 311-CT [132]; PRPF3 1–140 and 207–308 [172]; LEDGF 263–293 [173]; DHX9 263–269 [173]; TDP-43 [173] (unmapped); FUS [173] (unmapped); DGCGR 380-CT, requires phospho-S80 [134]; FOXG1, requires 1–9 (unique to isoform e2) [174]; HIPK2 [155] (unmapped); CDKL5 202-CT [175]; IKKα [176] (unmapped); p300 [163] (unmapped); SIRT1 [163] (unmapped); HMG1 207-CT (motif 380–386) [134]; HTT [177] (unmapped); MB2 163–309 [178]; MeCP2 (dimerisation) 163–309 [178]. MBD = methyl-CpG binding domain; NID = NCoR1/2 interaction domain; RTT = Rett syndrome; SNP = single nucleotide polymorphism.
RNA-seq data to relate the level of wild-type MeCP2 binding in vivo to changes in transcription at genes in Mecp2-null mouse brain [69,70]. Based on the effect of MeCP2 deficiency on expression, genes were divided into three categories: upregulated/unchanged/downregulated, and MeCP2-enrichment at genes in each category was compared. Data collected from both cerebral cortex and hypothalamic brain showed that MeCP2-enrichment was highest over upregulated genes, consistent with the notion that the protein normally inhibits expression directly. Importantly, a reciprocal effect was observed in mice over-expressing MeCP2, as downregulated genes now had the highest MeCP2-enrichment. Lower MeCP2-enrichment over genes that are downregulated when MeCP2 is absent and upregulated when MeCP2 is overexpressed indicates that MeCP2 does not directly activate their transcription but suggests instead that these transcriptional changes are an indirect consequence of altered MeCP2 levels. It is notable that gene expression changes correlate with the density of both mCG and mCAC, suggesting MeCP2 interprets both sites in the same way [70–72]. The model of MeCP2 as a global repressor fits with the high frequency of MeCP2 binding sites in the genome [70] matched by the large number of MeCP2 molecules per cell [34] and with the robust association between DNA methylation and transcriptional inhibition [67].

Of the large number of proposed MeCP2 interaction partners, 13 proteins have been suggested to mediate its role in transcriptional silencing. The first group of partners comprises the HDAC-containing NCoR1/2 [106,107] and SIN3A complexes [103] and three co-repressors that interact with these complexes (cSki [106], PU.1 [108] and CBF1 [109]). The second group is made up of the transcription factors, YY1 [110], SOX2 [111] or SP3 [112]. Third, evidence has been presented that MeCP2 recruits the histone methyltransferases PRMT6 [113], G9a [114] or HLCS [115] or the chromatin remodeller BRM [116]. Finally, MeCP2 was reported to inhibit transcription by binding TFIIB, preventing its incorporation into the preinitiation complex [117]. To date, the binding sites of only six of these (the NCoR1/2 complex, SIN3A, c-Ski, PU.1, YY1 and TFIIB) have been mapped to regions of MeCP2 (Fig. 3). To identify the domain required for MeCP2-mediated transcriptional repression, fragments were fused to the DNA binding domain of the yeast transcription factor GAL4, and residues 207–310 were found to be sufficient to repress a reporter gene regulated by GAL4 binding elements [102]. This region was named the transcriptional repression domain (TRD). As none of the mapped interaction sites lies wholly outside the TRD, knowledge of this domain does not intrinsically exclude responsibility of any of the 13 interaction partners for repression. A common approach to predicting the biological importance of protein regions or residues relies on evolutionary conservation. MeCP2 is 95% identical in amino acid sequence between human and mouse. This level of conservation is greater than the average for all proteins (86.4% identity [116]) and therefore suggests that the entire amino acid sequence is functionally important. A strikingly different picture is seen when the missense mutations causing Rett syndrome are mapped onto the protein, as most are confined to two discrete domains [107]: the MBD and a second smaller cluster at the C-terminal end of the TRD (Fig. 3). The importance of these two domains for MeCP2 protein function is further supported by the pattern of variants identified in the human population, as recorded in the ExAC database. Mapping these variants showed changes to the protein sequence could be tolerated throughout the protein sequence but rarely in these two domains (Fig. 3). Interestingly, the C terminal cluster of RTT-causing mutations coincides with the region responsible for the interaction between the NCoR1/2 complex and MeCP2: residues 298–309, named the NCoR1/2 interaction domain (NID) [107,119]. RTT-causing mutations in this cluster (P302R, K304E, K305R and R306C) destroy this interaction and disrupt repressive activity [107]. The NID interacts with the WD40 domains of transducin-beta like 1 (TBL1) and TBL1-related (TBLR1), two paralogs that are core components of the NCoR1/2 complex. Co-crystallisation of the NID peptide with the WD40 domain of TBLR1 showed that all four MeCP2 residues mutated in RTT make extensive contacts with TBLR1 [119].

The most common RTT-causing missense mutation in the NID is R306C, responsible for 5% of all cases. Knock-in mice carrying this mutation display the same phenotypic signature as other RTT mouse models (Table 1) [13,15,107], confirming pathogenicity. In addition, transcriptional analysis shows the same patterns of dysregulation in R306C knock-in mice as the other models [71]. Taken together, this evidence shows that disruption of the interaction between MeCP2 and the NCoR1/2 complex is sufficient to cause both the neurological defects and transcriptional changes common to all RTT models. The presence of an activity-dependent phosphorylation site, T308, in the NID provides a potential mechanism by which NCoR1/2 complex recruitment by MeCP2 could be regulated by neuronal activity [52]. The importance of this site was demonstrated in vivo by the production of T308A knock-in mice. Cortical neurons derived from T308A mice had reduced induction of activity-dependent genes upon, consistent with retaining the co-repressor complex. These mice displayed very mild RTT-like symptoms including impaired motor function, rather than the gain of function phenotype that might be predicted (Table 1) [52].
Testing the Bridge Hypothesis: MeCP2 Recruits the NCoR1/2 Co-repressor Complex to Methylated DNA

The weight of evidence discussed so far supports a model where MeCP2 recruits the NCoR1/2 co-repressor complex to methylated sites on chromatin, requiring two functional domains: the MBD and the NID. In other words, MeCP2 acts as a bridge between DNA and the co-repressor. Although most RTT-causing missense mutations affect these two domains, three lie elsewhere in the protein: A2V (e1 isoform only, Fig. 1), P225R and P322L. There is also a major cluster of truncating mutations downstream of the NID, making up ~10% of cases of RTT (Fig. 3) [120]. Considering first the A2V mutation, evidence indicates that it impairs posttranslational processing of MeCP2. Normally, the N-terminal methionine and up to five downstream alanine residues are excised from newly synthesised MeCP2, followed by acetylation of the alanine now at the N-terminus. Replacement of the alanine in position 2 with valine reduces the efficiency of methionine aminopeptidase (the enzyme responsible for cleavage), leading to greatly reduced MeCP2 stability [54]. Mouse models of the two proline mutations, P225R and P322L, reproduce the RTT phenotype (Table 1). Both mutant proteins have reduced abundance in the brain and both are deficient in their ability to recruit TBL1 to pericentromeric heterochromatin [120]. Two of the most common patient mutations in the C terminal cluster were chosen for characterisation. The first, L386HfsX5, drastically reduced MeCP2 abundance leading to RTT-like symptoms in knock-in mice (equivalent mutation in mice is S385PfsX6). P389X, the second, was initially recapitulated in the mouse model (equivalent mutation in mice is P384_S385insPL; P387X), but introduction of this patient mutation in the context of the human protein sequence demonstrated that it too leads to greatly reduced MeCP2 protein levels. Intriguingly, the C-terminal truncating mutations are the first RTT models found to have lower mRNA levels, suggesting transcript stability is affected when this region is mutated [120]. In summary, all RTT-causing mutations adversely affect the level of bridge function and do not therefore uncover unanticipated functional domains in MeCP2.

A more stringent test of the bridge hypothesis is to ask whether other characterised interaction sites in MeCP2 that lie outside the MBD and NID are dispensable for its function. The ExAC database shows many single nucleotide polymorphisms (SNPs) elsewhere in the protein indicating that these regions can be altered without pathogenic consequences [107]. Since mapped interaction sites tend to be large, however, it is not obvious that the neutral SNPs could disrupt binding. Exceptions are the two AT hooks, which have short core consensus sequences. Indeed, the presence of destructive SNPs in healthy hemizygous males and heterozygous females suggests that the AT hooks are not required for MeCP2 protein function [121]. Deletion of AT hook 1 in hemizygous male knock-in mice has no effect on lifespan but leads to mild symptoms that may be caused by increased MeCP2 levels in some brain regions [122]. Taken together, these findings contradict the proposal that MeCP2 functions to compact chromatin via the MBD and both AT hooks [123]. The notion that MeCP2 compacts chromatin was first raised by in vitro studies [124], but recent analysis found that chromatin in neurons lacking MeCP2 is in fact more compact than wild-type [125], consistent with smaller nuclear size [126]. It is possible that absence of MeCP2 allows access for another compacting protein, as suggested by increased levels of histone H1 [34], although this hypothesis has been questioned [127]. Another characterised short functional domain is the bipartite nuclear localisation signal (NLS) [61] which binds to importins, KPNA3 and KPNA4 [128,129]. Absence of RTT mutations that inactivate the NLS suggests that this too is nonessential [129]. In fact, disruption of the NLS does not interfere with nuclear localisation and has no phenotypic consequences in mice (Table 1) [128,129]. Interestingly, the small size of MeCP2 allows it to enter the nucleus where it can be retained by binding to methylated DNA [129].

To critically assess the functional dispensability of large regions of MeCP2 protein outside the MBD and NID, a series of knock-in mice expressing truncated versions of MeCP2 were produced: ΔN, ΔNC and ΔNIC (where N = the region N-terminal to the MBD, C = the region C-terminal to the MBD and I = the intervening region between the MBD and the NID; Fig. 3) [130]. Both ΔN and ΔNC were expressed at wild-type levels and hemizygous males expressing the mutant proteins were indistinguishable from wild-type littermates (Table 1). The dispensable N and C termini, which together make up half of the length of MeCP2, have been reported to contain interaction sites for multiple binding partners including the heterochromatin protein HP1 [131], the splicing factors FBP11, HYPC and YB-1 [132,133], and the miRNA processing factor DGC8 [134] (Fig. 3). ΔNIC mice were also viable, surviving for at least a year, but displayed mild neurological defects (Table 1) that are at least partially attributable to the reduced protein levels in these mutants (~50% of wild-type) [130]. This illustrates that evolutionarily conserved regions of MeCP2 are not needed to prevent Rett syndrome, even though they are involved in molecular interactions. Taken together, the evidence strongly suggests that MeCP2 exerts a sole key function, which is to bridge methylated DNA and the TBL1/TBLR1 subunits of the NCoR1/2
complex. An important corollary of this model is that Rett syndrome is likely to be one disease with a single root cause.

The relative dispensability of two-thirds of MeCP2 leads us to ponder why these regions are so well-conserved throughout evolution. One theory is that the other regions are required for higher functions that were not detected in the phenotypic characterisation of the mice expressing the truncated alleles [130]. Looking at variants that fall in the grey area between no phenotypic consequence (male hemizygotes in the ExAC database) and complete loss of function (Rett syndrome in females and neonatal encephalopathy in males) may shed light on this issue. MeCP2 mutations have been implicated in intellectual disability in males, mild intellectual disability in females and in cognitive disorders such as autism and schizophrenia in both males and females. Missense mutations occur throughout the protein [http://mecp2.chw.edu.au/] and could hypothetically affect MeCP2 by causing subtle differences in bridge activity or protein levels or by impacting another role of MeCP2 that is required for higher cognitive function. There is evidence that regions outside the MBD contribute subtly to its interaction with DNA, as deletions or mutations in these regions affect binding and diffusion dynamics [63,64,121,135,136]. There is also evidence that these regions are involved in regulating protein stability, for example the intervening domain contains Ser216 whose phosphorylation is reported to modulate MeCP2 levels [53]. To date, no mutations outside the MBD or NID specifically disrupt the interaction with any other binding partner. This has, however, been suggested for the MBD mutation A140V, which initially did not appear to compromise DNA binding [137]. A140V is the most common MeCP2 mutation causing intellectual disability in boys and female abnormalities in carriers [138] and its milder phenotype compared with RTT-causing mutations is recapitulated in knock-in mice (Table 1) [139,140]. The mutation weakens the interaction between MeCP2 and ATRX [137], a chromatin remodeler also mutated in cases of severe intellectual disability [141]. However, recent reports of minor changes in DNA binding affinity of the A140V protein [90,142,143] indicate that more work is needed to determine the molecular causes of pathogenicity.

Excess NCoR1/2 Complex Recruitment is Required for Toxicity in MECP2 duplication Syndrome

At the transcriptional level, two-fold overexpression of MeCP2 has the opposite effect compared with loss-of-function mutations. Genes enriched for MeCP2 binding are upregulated in MeCP2-null mice and downregulated in mice overexpressing MeCP2 [69,70]. This suggests that neurological toxicity caused by MeCP2 overexpression is because of excess recruitment of the NCoR1/2 co-repressor complex at these genomic regions. If so, mutating the NID in the additional copy/copies should alleviate the toxic effects. This possibility has been addressed using two independent mouse models of MeCP2 overexpression: (1) a PAC transgene containing the entire human locus [16] and (2) expression of mouse Mecp2 isofrom e2 from the Mapt locus fused to the first 31 amino acids of Tau [17]. The former, which express MeCP2 at two-fold wild-type levels, displayed neurological phenotypes in behavioural tests and developed late-onset overt symptoms with 30% mortality between 5 months and 1 year of age [16]. Introduction of R306C into the transgene abolished these phenotypic consequences [95]. This result was confirmed using the second model with a more extreme phenotype. Mice with 3.8-fold levels of wild-type MeCP2 failed to survive to weaning, but this was fully rescued by the introduction of the R306C mutation into Tau-MeCP2. Not only did the mice survive to adulthood, they remained healthy with no phenotype in behavioural tests (Table 1) [94]. Thus, a functional NID is essential to confer the toxicity caused by MeCP2 overexpression.

Breaking the Bridge by Mutating TBLR1 can Occasionally Cause Rett Syndrome

The Bridge hypothesis implies that mutations in the NCoR1/2 complex might also break the DNA-MeCP2-NCoR1/2 complex bridge, resulting in RTT. The core NCoR1/2 complex is 1–2 MDa in size and contains NCoR1, NCoR2, GPS2, HDAC3, TBL1 and TBLR1 subunits. Knockout mice lacking the genes encoding core subunits (except TBL1) have been produced and show that each is essential for embryonic development [144-148]. As the NCoR1/2 complex has other molecular functions including, for example, repression via nuclear receptors in the steroid/thyroid/retinoid superfamily, it is unsurprising that loss-of-function mutations in its components have a more severe phenotype than loss of MeCP2. To specifically abolish its role in MeCP2-mediated repression, its interaction surface would need to be mutated so that MeCP2 can no longer bind without affecting other interactions. The NCoR1/2 complex consists of a central TBL1/TBLR1 tetramer, with each TBL1/TBLR1 dimer bound to one molecule of NCoR1 or NCoR2. These in turn act as scaffolds for GPS2 and HDAC3 [149]. TBL1/TBLR1 tetramerisation involves their N-terminal domains, with their C-terminal WD40 domains located on the outside of this complex, suggesting the MeCP2 interaction surface can be specifically mutated. A key difference between mutating this interface within MeCP2 and within the NCoR1/2 complex is that both TBL1 and
TBLR1 paralogs can bind MeCP2, and disrupting mutations are unlikely to occur in both genes. Another consideration is that patients with Rett syndrome are mosaics due to X chromosome inactivation, with half their cells expressing the wild-type and half the mutant copy of the X-linked MECP2 gene. The location of the gene encoding TBL1, TBL1X, on the X chromosome raised the possibility that mutations in this gene could have similar molecular consequences. However, the conditions for this scenario are not met as TBL1X is thought to be biallelically expressed [150] and TBL1 is not the major paralog in the brain—TBLR1 is in fact five times more abundant [151].

Analysis of disease-causing mutations in the TBL genes and their corresponding clinical outcome may illuminate the role of the NCoR1/2 complex in the brain. The DECIPHER database lists cases of developmental delay in both males and females caused by deletions and duplications in the genes encoding TBL1 and TBLR1, suggesting that TBL protein dosage impacts development. Furthermore, as is common for genes that escape X chromosome inactivation, dosage in males is maintained in humans and other primates by a Y-linked homolog, TBL1Y [152], and duplication of this gene also causes developmental delay [152, DECIPHER]. The DECIPHER database and case reports list numerous missense mutations in TBLR1 that cause developmental delay in heterozygotes. These are mostly located in the WD40 domain (Table 3), suggesting some may impact MeCP2 binding. Intriguingly, two patient mutations, D369E and P444R, were recently shown to abolish the bridge between TBLR1 and MeCP2, whilst retaining the ability of TBLR1 to pulldown HDAC3 (indicating that the complex remains intact) [119]. The challenge now is to compare the clinical characteristics of patients with TBLR1 WD40 domain mutations with RTT patients. Indeed, one patient with the D370N mutation was diagnosed with classical RTT, meeting all main and six out of 11 supporting criteria [153].

### Concluding Remarks

There is now strong evidence that MeCP2 is an important reader of the DNA methylome in neurons. Loss-of-function mutations in MECP2 that cause Rett syndrome in patients pinpoint the MBD (responsible for binding to methylated DNA) and the NID (responsible for binding to the NCoR1/2 co-repressor complex) as the key domains essential for

### Table 3. Missense mutations in the WD40 domain of TBLR1 cause intellectual disability.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>MeCP2 binding</th>
<th>NCoR1/2 complex binding</th>
<th>Clinical description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>H213Q</td>
<td>Abolished</td>
<td>Yes, HDAC3</td>
<td>Global DD, delayed speech and language, microcephaly, facial dysmorphism</td>
<td>[153], DECIPHER 280701</td>
</tr>
<tr>
<td>L282P</td>
<td>Yes, HDAC3</td>
<td>Yes, HDAC3</td>
<td>ASD</td>
<td>[195,196]</td>
</tr>
<tr>
<td>A311P</td>
<td>Abolished</td>
<td>Yes, HDAC3</td>
<td>Abnormality of the nervous system</td>
<td>DECIPHER 322627</td>
</tr>
<tr>
<td>D369E</td>
<td>Abolished</td>
<td>Yes, HDAC3</td>
<td>Severe ID, ASD, ADHD, facial dysmorphism</td>
<td>DECIPHER 260528D370N A: RTT-like criteria, including regression (loss of learned speech and following by general hand movements with stereotyped); B: Not described</td>
</tr>
<tr>
<td>Y237D</td>
<td>Abolished</td>
<td>Yes, HDAC3</td>
<td>ID, delayed speech and language, hyperactivity, facial dysmorphism, sleep disturbance</td>
<td>DECIPHER 251463</td>
</tr>
<tr>
<td>D370N</td>
<td>Abolished</td>
<td>Yes, HDAC3</td>
<td>RTT</td>
<td>DECIPHER 285549</td>
</tr>
<tr>
<td>Y446C</td>
<td>Abolished</td>
<td>Yes, HDAC3</td>
<td>PIerpont syndrome: developmental delay, facial dysmorphism, abnormal fat distribution in distal limbs</td>
<td>DECIPHER 285757</td>
</tr>
<tr>
<td>G460D</td>
<td>Abolished</td>
<td>Yes, HDAC3</td>
<td>ID, developmental delay, ID = intellectual disability, ASD = autism spectrum disorder, ADHD = attention deficient hyperactivity disorder, RTT = Rett syndrome</td>
<td>DECIPHER 261213</td>
</tr>
</tbody>
</table>
| DD = Developmental delay, ID = intellectual disability, ASD = autism spectrum disorder, ADHD = attention deficient hyperactivity disorder, RTT = Rett syndrome.

**Table 3.** Missense mutations in the WD40 domain of TBLR1 cause intellectual disability.
MeCP2 protein function. Analysis of the co-crystal structures of these interactions shows in atomic detail how the mutated residues contribute to binding to these macromolecules, and most mutations have been demonstrated experimentally to impair binding. The bridge model, which proposes that MeCP2 recruits the NCoR1/2 complex to methylated DNA, is supported by the observation that all RTT-causing mutations throughout the length of MeCP2 affect this function, and deletion of the regions outside these two domains has only mild phenotypic consequences. These findings significantly enhance our understanding of the underlying biology of MECP2-related disorders and link these molecular mechanisms to disease-causing mutations in TBLR1. After many years of research, recent findings now convey a coherent model for the role of MeCP2 in globally modulating gene expression, particularly in neurons. While there is evidently more to learn about the downstream metabolic consequences of MeCP2 dysfunction, the new knowledge promises to underpin efforts to devise therapies for RTT and other MECP2-related disorders.

Acknowledgements

The authors thank Jacky Guy, Jim Selfridge, Matthew Lyst and John Connelly for comments on the manuscript. R.T. is a Sir Henry Wellcome Postdoctoral Fellow (210913/Z/18/Z). A.B. is supported by a Wellcome Trust Centre grant (091580/Z/10/Z), a Wellcome Investigator Award (107930/Z/15/Z), a Rett Syndrome Research Trust Consortium grant, a European Research Council Advanced grant (EC 694295 Gen-Epix) and is a member of the Simons Initiative for the Developing Brain. The University of Edinburgh is a charitable body, registered in Scotland, with registration number SC005336.

Conflicts of Interest

None.

References


MeCP2 function in Brain


MeCP2 function in Brain


MeCP2 function in Brain


F.M. Piccolo, Z. Liu, P. Dong, C.-L. Hsu, E.I. Stoyanova, A. Rao, R. Tijan, N. Heintz, MeCP2 nuclear dynamics in live neurons results from low and high affinity chromatin


