2013

Decoding 5HMC as an Active Chromatin Mark in the Brain and its Link to Rett Syndrome

Pinar Ayata

DECODING 5HMC AS AN ACTIVE CHROMATIN MARK IN THE BRAIN AND ITS LINK TO RETT SYNDROME

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Pınar Ayata

June 2013

DECODING 5HMC AS AN ACTIVE CHROMATIN MARK IN THE BRAIN AND ITS LINK TO RETT SYNDROME

Pınar Ayata, Ph.D.

The Rockefeller University 2013

Functions of the nervous system are accompanied at the cellular level by changes in gene expression, regulated by transcription factors and epigenetic mechanisms, such as histone modifications and DNA methylation, that are frequently altered in neurological disorders. 5-hydroxymethylcytosine (5hmC), a recently identified DNA base derived from 5-methylcytosine, accounts for ~40% of modified cytosines in the neuronal genomes, suggesting that 5hmC is a stable epigenetic mark and its interpretation in the nervous system may differ from the other tissues. This hypothesis was supported by the recent findings showing that 5hmC is enriched over the bodies of active genes within euchromatin in a cell-specific manner. In the first part of this study, we identify the methyl-CpG binding protein 2 (MeCP2) as the major reader of this activation mark and demonstrate that MeCP2 is the only methyl-CpG binding domain (MBD) protein that binds 5hmC and, 5mC, with high affinity. We reveal strong preferential inhibition of MeCP2 affinity to 5hmC by a Rett-syndrome-causing mutation, R133C. We then show that MeCP2 recognizes 5hmC and 5mC within CpA context. Modified CpAs were recently shown to exist in embryonic stem and neuronal cells, where it

localizes to actively transcribed gene bodies. Together these data support a model where 5hmC and MeCP2 formulate a cell-specific epigenetic mechanism for establishing active chromatin states that facilitate gene expression. This is supported by our observation of reduced chromatin accessibility of 5hmC containing loci in the absence of MeCP2.

In the second part of the study, we discover a complex in the brain nuclear extract that assembles specifically in the presence of 5hmC. We purify and identify components of this complex: the purine-rich element binding protein (Pur) α and β, which are required for the proper development of neuronal cell types. We verify the increased affinity of recombinant Purα and Purβ to 5hmC and support a binding mechanism where they separate two strands of DNA and recognize specific sequences. These findings offer a previously unknown function for Pur proteins via binding to 5hmC.

This study presents new decoders of a novel epigenetic mark, 5hmC, that is effectively and differentially employed in the brain. Unlike previous studies, we introduce readers of 5hmC as a stable activation mark. In addition, by mechanistic characterization of our model we link 5hmC to an autism spectrum disorder and offer a new avenue toward investigation of its pathophysiology.

Dedicated to my best friends, my parents, Nesrin and Kemal Ayata;

to my hero, my grandfather, Ilhan Ayata;

and to my muse, Dr. Savas Dervent, R.I.P.

Acknowledgements

The day I met Nat for the first time I had no intention to become a neuroscientist. It was my first interview for graduate school. I entered his office determined to pick my favorite transcription factor and study that for my graduate work. Maybe a second protein, just to be on the safe side. A few minutes later, I was staring at a picture of a Purkinje cell for the first time on his computer screen as he was speaking of it - like a man in love. I was sold. I joined his lab immediately and have participated in many interesting projects throughout my 6.5 years. The most vivid image of Nat for me is him walking through the lab whistling, stopping by my bench to talk about a new idea he envisioned the day before inspired by some recent findings. His creativity and enthusiasm for science made every day different from the day before, and exciting. Brainstorming with him about new experiments, recent techniques or paper figures has been among the most fun and intellectually satisfactory times in the lab for me, occasionally at the expense of a minor experiment. In addition, his way of telling a scientific story truly amazes me. I can only hope that his unique brilliance has had some influence on my personality as a scientist.

When Skirmantas, or Skiry, asked me whether I would be interested in digging into a project he started, a new scientific era began for me. Skiry taught me everything I know about Biochemistry. I would find out only years later that I have been raised in the school of Adrian Bird and became a member of the “MeCP2 Club” without knowing. Most of my novel findings have been results of those experiments that followed his suggestion “Hey, Pınar, why don’t you try (xxx)?”.

To this date I am convinced that the discoverer of the sixth base has indeed a sixth sense for discovery. He also taught me to think analytically, to carefully plan experiments as “paper figures” and to keep reproducing results until “the perfect picture” is framed and ready. Now I look back in time and I see us studying the inside of the FPLC machine and fixing it. I see us blotting dozens of large membranes with radioactive DNA probes on shakers under the hood behind screens around midnight, while cracking up with laughter. I see us arguing about Guinness versus American Ales. And I smile.

Perhaps, my favorite anecdote of Nat and Skiry starts with me and Skiry walking into Nat’s office to discuss the next set of experiments. Nat had high hopes in the beginning that my phage expression, phage display and biochemical purifications would reveal a novel protein. Although Skiry and I wanted that too, the robust and reproducible results with MeCP2 made it difficult to just quit. We kept planning more and more experiments and getting beautiful results. Nat’s reaction was: “Ok, if MeCP2 is the one and you don’t find a novel protein, then I will buy you a bottle of champagne. Otherwise you can just buy me a six pack.” Another personal favorite of mine is his justification for his overuse of exclamation points in scientific writing: “but science should be exciting!!!”

I am grateful to my thesis committee members, Dr Shai Shaham and Dr Robert Darnell, for making time to read and think about several projects i presented throughout the years and their valuable suggestions. I would also like to thank to Dr Michael Greenberg to travel to New York City from Boston to become a member of my thesis committee.

Miho and Prerana, the sisters I never had. Day after day, whenever I needed help, to talk or to listen; they were there. I cannot imagine being closer to someone more than I have been to them. Jie has been a central member of my lab family. She -although very young- has been like a mother to me. Her strength made me feel secure, knowing that she would be there to help during every crisis. And she was. Jie performed all dissections for me when my mental strength was not enough. She is an amazing teacher, and I still have not figured out the limits of her patience. Also Zoran and Mila, your presence was a big support throughout the years. To me, you were the two pillars that kept the 4th floor of the Rockefeller Research Building standing. I miss you very much.

Ivarine and Valerie; your laughter, songs and jokes added joy to each and every day. I am also thankful to JoeDock, Bea, Andreas, Hiro, Liz, Fekrije, Erika, Rajiv, Sarah, Judy, Claire and all past and current members of the lab, for the smiles and stimulating discussions we exchanged over the years and beers. I cannot forget Wendy, her amazing skills with cell culture and the late nights we spent together at the culture hood. A special thanks to Christina for adopting my orphan PKR project, continuing my MeCP2 project and teaching me her ways to handle mice. I hope she gets beautiful results. The Proteomics Facility at the Rockefeller University became like a second lab to me. Big thanks to Susan and Henry for their help with FPLC; to Henrik, Joe and Milica for their high quality MS analyses and for their time to sit down and discuss the results with me; and to Mouli for his help with HPLC and for his patience. Thanks to Fraser and TJ at the High Throughput Screening Facility for training me for BIACORE, ProteOn and ITC

and trusting me with these precious machines; and to Brian at GE Healthcare for his step-by-step instructions via many emails. A big smile to Pat, Steve, John, and Joe and at the Faculty Club; and to Haroun, David, Ed, Ken and the rest of the Faculty House crew.

My dearest classmates, Chan, Manu, Anne, Mattmatt, Daniel, Moritz, Sarah, Andrés, David, Joe, Laura, Disan, Claire and Tina: The times we shared together meant so much to me. Throughout the years we have been one big family; living, cooking, eating, traveling, camping, drinking together. I will take your friendship with me everywhere, because without it I am incomplete. My climbing, kayaking, music, couchsurfing and traveling buddies, especially Sarah, Aneta, Adam, Babitha, Finessa, Tsai, Peter, Paul, Marnie, Gary, Matt, Michael, Jerry, Amit, Anthony, Schalk, Daniel, Logan, Bob, Rob, Kim, Donald, Dana, Francis, Charles, Sozanne, Butch, Victoria, Ara, Geny, Noah, Lambabu Sherpa, Jaya, Ben, Arda, Deniz, Bert, Sounds of Science members and many others, thank you for adding extra dimensions to my life and cheers for every unique experience we shared together. I also want to thank my extended family, especially Mergen Family, Kuzen Yasemin, Leyla Hala, Babane, Umit and Nermin, and my former mentors, Batu Hoca, Cleva, Zehra Hoca, Alpay Hoca and Mali, for making me feel their love and support from far away. Sarah, you are a true friend: You have been there for me whenever I needed, sent me positive vibes, listened and understood me. And on top, you read and corrected my whole thesis! You are my angel.

Finally, I cannot thank enough to my parents, Nesrin and Kemal Ayata, for giving me every opportunity to express myself, letting me make my own choices, encouraging me to walk through life in my own way and for never judging me. Thank you for being my best friends, psychoanalysts and confidants. I love you both very much and I am proud of the way you enjoy life.

Table of Contents

List of Figures

Chapter I.

Introduction

I.I. Prologue: from brain to chromatin

Understanding how the brain works is one of the most exciting scientific quests of our age. The study of the nervous system started over a century ago with the identification of individual neuron as its "absolutely autonomous unit" by Ramón y Cajal (Ramon y Cajal 1899; Ramón y Cajal 1917). We now know that the mammalian central nervous system (CNS) is a highly sophisticated network made of hundreds of distinct classes of neurons and glia, nourished by blood vessels (Jones, Stone et al. 2011; Reid 2012).

The basic function of a neuron is to receive a chemical/physical signal at its dendritic extensions or its cell body; convert it into an electrical signal by manipulating ion gradient across its membrane; and transmit the so-called action potential to its destination on the axon where it is converted into a chemical signal by the release of signaling molecules for its target cell (Kandel, Schwartz et al. 2000). Neurons can be classified into hundreds of types, each with unique sets of functions and distinct anatomical, molecular and electrochemical profiles (Masland 2004; Lichtman, Livet et al. 2008). Glia, which far outnumber the neurons (Kandel, Schwartz et al. 2000), are housekeepers, insulators and nurses of the neurons. Some are also thought to participate in neurotransmission and to surveil, sculpt and modulate synaptic connections (Allen and Barres 2009; Graeber 2010).

The diversity of cell types in the CNS and their complex interconnectivity is established by their unique collection of expressed genes (Nelson, Sugino et al. 2006). Since the genomic information is identical in every cell, organisms use transcription factors aided by “epigenetic”, or “outside conventional genetic”, mechanisms that formulate these characteristic gene expression profiles (Jaenisch and Bird 2003). Epigenetic mechanisms include changes in the chemical, physical and topological characteristics of the chromatin, where the genomic information resides. These changes cause a discrepancy in the nuclear architecture that is readily visible under microscope: not only in the aggregation pattern of chromatin; but also in the structure of the nuclear envelope, in the distribution of nuclear pores and nuclear bodies. (Takizawa and Meshorer 2008).

Chromatin is a three-dimensional structure, made of DNA, histones and other architectural factors. The most obvious form of its structural organization is the compartmentalization into “euchromatin” and “heterochromatin”, as first described almost a century ago (Heitz 1928). Heterochromatin encompasses highly condensed regions where the genomic material is inaccessible and therefore “silent”; whereas euchromatin appears decondensed and contains actively expressed regions of genetic material. For instance, Purkinje cells (PCs), which are among the largest cells in the brain, have large nuclei and store only 10% of their DNA in heterochromatin in the center of their nucleus (Fig 1.1). In contrast, the nuclei of granule cells (GCs) are small, compact and display ~64% of their genome in large aggregates heterochromatin that are scattered in the nuclear periphery (Palay and Chan-Palay 1974; Lafarga, Berciano et al. 1991).

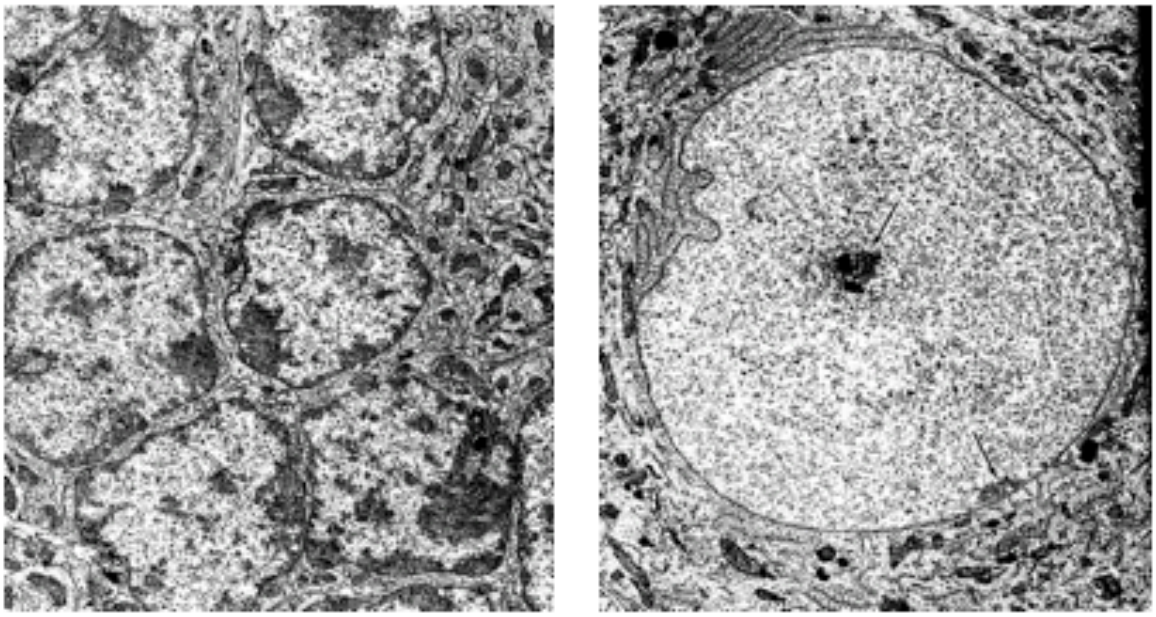


Fig 1.1. Electron microscopy image of neuronal nuclei. Granule cells (left) and Purkinje cells (right) (Palay and Chan Palay. 1974).

I.II. Chromatin: the nucleoprotein

Historical perspectives, which considered chromatin to be a static, unchanging entity, have been replaced by the current appreciation of its modular, complex and highly plastic nature. The basic building unit of chromatin is the nucleosome, which is formed by wrapping ~146 base pairs (bp) of DNA around a histone core, which consists of two copies of each histone proteins, H2A, H2B, H3 and H4 (Fig 1.2) (Hamon and Cossart 2008; Takizawa and Meshorer 2008). Each histone protein contains a globular part and an unstructured N-terminal tail that protrudes from the nucleosome (Luger, Maeder et al. 1997; Andrews and Luger 2011). These tails can accommodate a variety of chemical modifications that can influence local protein composition, arrangement of nucleosomal arrays and histone-DNA interactions. The combinatorial nature of histone modifications reveals a “histone code” that defines chromatin states (Jenuwein and Allis 2001; Andrews and Luger 2011). For example, euchromatin is normally marked with histone H3 trimethylated on lysine 4 (H3K4me3) and acetylated on lysine 9 (H3K9ac); while H3K9me3 and H4K20me3 are concentrated in constitutive heterochromatin and H3K27me3 in facultative heterochromatin. Global structure of the chromatin is also regulated by other proteins, such as chromatin remodeling enzymes, transcriptional activators and repressors, architectural proteins, and the “linker histone”, H1, that binds the “linker DNA” bridging two adjacent nucleosomes, (Misteli 2001).

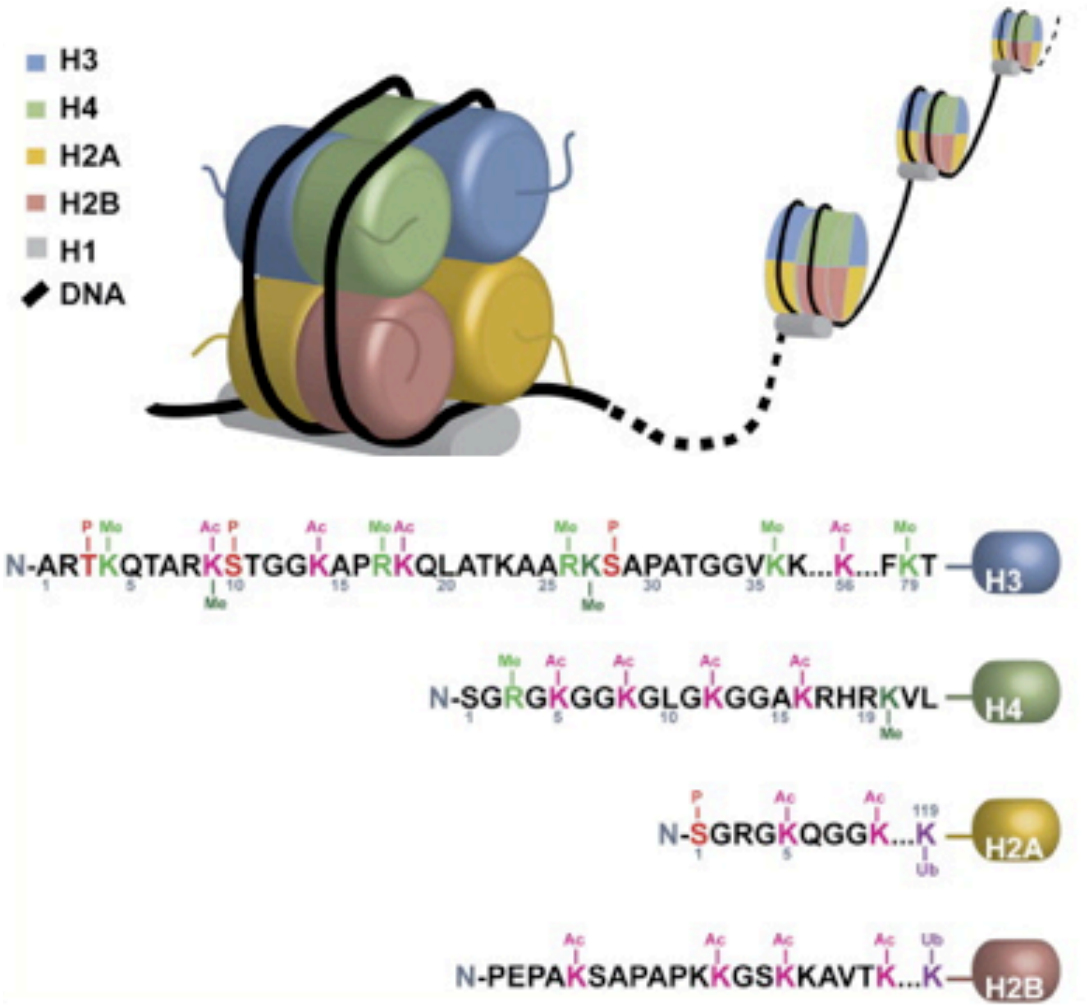


Fig 1.2. Schematic representation of nucleosome and the histone code. Nucleosome, as histone octamer made of two copies of H3, H4, H2A and H2B; the linker histone H1; and the modifications of N-terminal tails that encrypt the histone code (Hamon and Cossard, 2008)

I.III. Chromatin: nucleic acid modifications

Chromatin states are largely established by chromatin-associated factors; however they also can be influenced at base resolution by direct chemical modifications of DNA. Although nucleic acid of various organisms accommodates over 100 modified bases, in the mammalian DNA, the principle DNA modification seemed to be the methylation of cytosines (Cs) in symmetric CpG dinucleotides (5-methylcytosine, 5mC or “the fifth base”) (Bird 2002; Ratel, Ravanat et al. 2006). Today, non-CpG methylation (Ramsahoye, Biniszkiewicz et al. 2000; Pastor, Pape et al. 2011) and oxidation products of 5mC, primarily 5-hydroxymethylcytosine (5hmC or “the sixth base”) (Kriaucionis and Heintz 2009; Tahiliani, Koh et al. 2009; Ito, Shen et al. 2011), are also accepted as major mammalian DNA modifications. Other epigenetic mechanisms at the nucleic acid level include the occurrence of these modifications in the mitochondria (Pollack, Kasir et al. 1984; Dzitoyeva, Chen et al. 2012) and also a complex realm of non-coding RNAs (Mehler and Mattick 2007; Satterlee, Barbee et al. 2007).

I.IV. 5mC: discovery, genomic distribution and function

5mC entered the scientific vocabulary almost a century ago (Johnson and Coghill 1925) and its presence in the mammalian genome was discovered 25 years later (Wyatt 1950). It also is found in the genome of other vertebrates, flowering plants, invertebrates, protists, fungi and bacteria (Goll and Bestor 2005). Interestingly, the invertebrate-vertebrate boundary marks a distinctive

evolutionary shift from a fractional to a global methylation pattern, where almost all gene bodies are methylated in vertebrates (Tweedie, Charlton et al. 1997).

Eukaryotic DNA methylation is now recognized to be a chief contributor to the silent chromatin state and inhibition of gene expression (Jaenisch and Bird 2003; Ooi, O'Donnell et al. 2009). Strong evidence for the silencing role of 5mC comes from its high levels in the heterochromatin and in the repeat elements that are thus stabilized (Doskocil and Sorm 1962; Keshet, Lieman-Hurwitz et al. 1986). Another major manifestation of this role is the X chromosome inactivation, where the silent X is invariably hypermethylated (Riggs 1975; Heard, Clerc et al. 1997; Walsh, Chaillet et al. 1998). In the early studies, artificially methylated transgenes inserted into mouse cells were repressed (Stein, Razin et al. 1982). Conversely, chemical inhibition of methylation reverses the silencing of previously methylated genes (Groudine, Eisenman et al. 1981; Mohandas, Sparkes et al. 1981).

In silent chromatin, 5mC is found as a permanent mark; whereas differentially methylated regions (DMRs) are regulated in a cell-specific manner (Song, Smith et al. 2005; Hahn, Wu et al. 2011; Lister, Pelizzola et al. 2009) in concert with the repressive histone signal H3K9me3. Accordingly, activating histone marks, such as H3K27me3, H3K4me and H2A.Z variant, are excluded from methylated regions (Meissner, Mikkelsen et al. 2008; Zilberman, Coleman-Derr et al. 2008; Laurent, Wong et al. 2010; Coleman-Derr and Zilberman 2012). In many cancer cells, aberrant methylation patterns are frequently observed (Ting, McGarvey et al. 2006; Irizarry, Ladd-Acosta et al. 2009).

I.V. CpG dinucleotide and beyond

In mammals, the predominant form of DNA methylation occurs symmetrically within CpG dinucleotides and it comprises 70-80% of these dinucleotides (Bird, Taggart et al. 1985). Interestingly, CpG-rich regions (CpG islands or CGIs), which are found within promoter regions of 70% of human genes, contain very low amounts of methylation (Bird, Taggart et al. 1985; Saxonov, Berg et al. 2006). However, these may become methylated if the gene needs to be repressed in the course of development (Li, Beard et al. 1993) or differentiation (Song, Smith et al. 2005). Cytosine methylation is also a prominent cause of mutations. In humans, the mutation rate from 5mC to thymine (T) is 10-50 fold higher than other transitions (Duncan and Miller 1980), causing the CpG dinucleotide to be present only at 20% of its expected frequency (Sved and Bird 1990). The deamination of one 5mC in a CpG dinucleotide generates a 5mCpG·TpG mismatch (Wiebauer and Jiricny 1989), which can then be "corrected" to a CpA by thymine DNA glycosylase (TDG) (Neddermann, Gallinari et al. 1996).

Until recently, the methylation studies were focused on the CpG dinucleotide but newer studies revealed that non-CpG methylation, primarily CpA methylation, constitutes 20-25% of genome-wide methylation in mammalian pluripotent and differentiated cells (Ramsahoye, Biniszkiewicz et al. 2000; Lister, Pelizzola et al. 2009; Laurent, Wong et al. 2010). CpA methylation is reduced at promoters but highly enriched in the gene bodies exhibiting strong correlation with gene expression, whereas no such correlation is observed for CpG methylation (Lister,

Pelizzola et al. 2009; Laurent, Wong et al. 2010). Non-CpG methylation in gene bodies is asymmetrical almost at all times, even in CHG context, and is significantly enriched on the antisense strand. The expanded context of non-CpG context is TACA(A/T), with a periodicity of 8-10 bp between modified Cs corresponding to a single turn of the DNA helix (Lister, Pelizzola et al. 2009). Thus a pattern, where methyl groups aligned linearly on the DNA strand, becomes obvious and may be definitive in coordination of methylation-specific proteins.

I.VI. Dnmts: enzymes in the making of 5mC

5mC is generated by the covalent addition of a methyl group from S-adenosyl-L-methionine (SAM) to the 5 position of cytosine by a DNA methyltransferase, Dnmt (Fig 1.3) (Santi, Garrett et al. 1983; Bestor 2000; Lin 2011). The originally cloned Dnmt was the “maintenance” methyltransferase, Dnmt1, that methylates hemimethylated CpG dinucleotides during replication (Bestor and Ingram 1983). Later, *de novo* methyltransferases, Dnmt3a and Dnmt3b, were discovered which can methylate both unmethylated and hemimethylated target sequences (Bestor and Ingram 1983; Bestor 1988; Mund, Musch et al. 2004). The last family member, Dnmt2, has only weak activity toward DNA and is now accepted to be an RNA methylase (Okano, Xie et al. 1998; Mund, Musch et al. 2004).

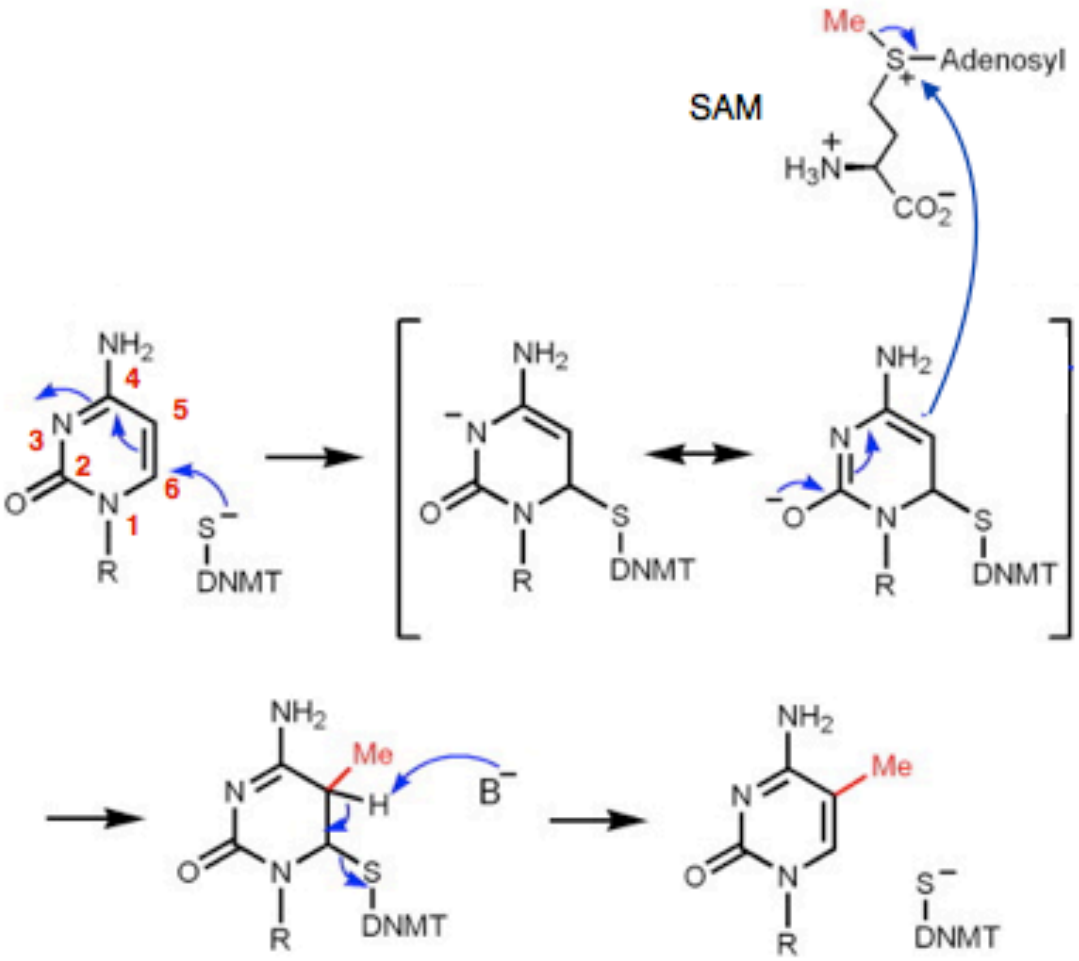


Fig 1.3. Schematic representation of methylation of cytosine. Nucleophilic attack by cysteine of DNA methyltransferase (DNMT) at the 6th position of cytosine is followed by stabilization of the resonance structure by the capture of the methyl group from S-adenosyl-L-methionine (SAM) and finally elimination of the cysteine from DNMT. (Lin 2011)

Dnmt1 is recruited to replication fork via ubiquitin-like plant homeodomain and RING finger domain 1 (Uhrf1) that recognizes hemimethylated DNA, whereas *de novo* methyltransferases bind DNA directly independent of its modification status (Bostick, Kim et al. 2007; Sharif, Muto et al. 2007). Dnmts are frequently found associated with heterochromatin regions in a cell-cycle-independent manner (Bachman, Rountree et al. 2001; Hermann, Goyal et al. 2004). Non-CpG methylation is thought to be carried out by Dnmt3a (Ramsahoye, Biniszkiewicz et al. 2000; Mund, Musch et al. 2004; Laurent, Wong et al. 2010; Arand, Spieler et al. 2012), although a role for Dnmt1 was also proposed based on the inheritance of non-CpG methylome (Grandjean, Yaman et al. 2007).

Dnmt3b is expressed within a narrow window during embryogenesis, whereas Dnmt1 and Dnmt3a are present throughout life (Okano, Bell et al. 1999). The absence of Dnmts can lead to global DNA hypomethylation, chromosomal instability and aberrant cell-cycle progression, although they are not required for embryonic stem (ES) cell integrity or self-renewal (Tsumura, Hayakawa et al. 2006). Nevertheless, the deletion of *Dnmt1* or *Dnmt3b* in mice results in embryonic lethality (Li, Bestor et al. 1992), while *Dnmt3a* $^{-/-}$ mice survive up to 4 weeks with normal methylation in heterochromatic regions (Okano, Bell et al. 1999) and *Dnmt2* $^{-/-}$ does not have a phenotype (Okano, Xie et al. 1998). Mutations in the *DNMT3B* gene has been linked to instability facial anomalies (ICF) syndrome in humans (Okano, Bell et al. 1999).

I.VII. Dnmts in the brain

The majority of DNA methylation mechanisms have been characterized in ES cells yet their influence on the CNS remains unclear. For instance, multiple isoforms of *de novo* methyltransferases are expressed in ES cells, but only full-length variant of Dnmt3a is found at high levels and in diffuse pattern in neurons (Feng, Chang et al. 2005; Nguyen, Meletis et al. 2007), suggesting that euchromatic *de novo* methylation might be specifically needed for neuronal functions. Indeed, nonpromoter methylation by Dnmt3a promotes transcription of neural genes (Wu, Coskun et al. 2010) and regulates cellular and behavioral plasticity in response to emotional stimuli in mice (LaPlant, Vialou et al. 2010). Moreover, after fear conditioning, Dnmt3a levels double whereas Dnmt1 level is unchanged, indicative of an interplay between neuronal activity and euchromatic, nonpromoter, Dnmt3a-mediated methylation (Miller and Sweatt 2007).

The lethality of the *Dnmt*-knockouts in the early stages of life necessitated alternative approaches for the study of methylation in postmitotic cells. Initially, pharmacological Dnmt inhibitors were used, such as RG108 (Brueckner, Garcia Boy et al. 2005), a competitive Dnmt1 inhibitor, or 5-azacytidine (Creusot, Acs et al. 1982; Miller and Sweatt 2007) and zebularine (Levenson, Roth et al. 2006; Miller, Gavin et al. 2010). Since the latter two chemicals are incorporated into the DNA after replication, in non-dividing cells they did not pan out to be useful.

In later studies, a brain-restricted knockout of *Dnmt1* resulted in an estimated 50% loss of methylation and produced mutant neurons and glia (Fan, Beard et al. 2001; Golshani, Hutnick et al. 2005). However when *Dnmt1* was inactivated in

the adult brain, global DNA methylation and cell survival were unaffected (Fan, Beard et al. 2001). Deletion of *Dnmt3a* in the developing CNS leads to neuromuscular defects and shortens lifespan (Nguyen, Meletis et al. 2007), whereas in its postnatal ablation in excitatory neurons long-term plasticity, learning and memory are impaired, by upregulation of immune genes (Feng, Zhou et al. 2010) that are considered breaks of synaptic plasticity (Shatz 2009).

I.VIII. MBDs: 5mC-binding proteins

The widely accepted role for 5mC is to displace some DNA binding proteins and to recruit methylation specific proteins (Watt and Molloy 1988; Klose and Bird 2004). The first “activity that could bind methylated DNA in solution” was methyl-CpG binding protein (MeCP) (Meehan, Lewis et al. 1989), that turned out to be a multiprotein complex, MeCP1 (Ng, Zhang et al. 1999; Feng and Zhang 2001). In contrast, another such activity, MeCP2, was a single polypeptide (Lewis, Meehan et al. 1992) with an 80 amino acid methyl-CpG binding domain (MBD), which is necessary and sufficient for recognition of 5mC (Nan, Meehan et al. 1993; Nan, Tate et al. 1996; Nan, Campoy et al. 1997). A database search revealed four additional proteins that contain the consensus MBD sequence: MBD1 through MBD4 (Hendrich and Bird 1998; Dhasarathy and Wade 2008).

Among the MBD family of proteins, MBD1 acts as a transcriptional repressor by recruiting histone deacetylases (HDACs) upon binding both to 5mC with its MBD and to unmodified DNA with one of its three zinc-coordinating CXXC domains (Fujita, Shimotake et al. 2000; Jorgensen, Ben-Porath et al. 2004). MBD2 and

MBD3 share 70% sequence identity and are thought to have arisen through gene duplication (Hendrich and Bird 1998; Hendrich and Tweedie 2003). MBD3 is the only MBD protein that does not recognize 5mC (Zhang, Ng et al. 1999) due to two amino acid substitutions in the MBD (Fraga, Ballestar et al. 2003). It is a core component of nucleosome remodeling and histone deacetylase repressor (NuRD) complex (Saito and Ishikawa 2002), which is recruited to DNA by MBD2 (Feng and Zhang 2001). Together they make up the MeCP1 complex. MBD4 appears to play a key role in DNA repair, by preferentially binding to and processing a 5mCpG·TpG mispair with its glycosylase domain (Hendrich, Hardeland et al. 1999).

A recent comparative analysis identified six additional proteins that contain an MBD: MBD5, MBD6, BAZ2A, BAZ2B, SETDB1 and SETDB2 (Roloff, Ropers et al. 2003). Most of these proteins do not show specificity for methylated DNA except SETDB1 (Gou, Rubalcava et al. 2010), and their functions vary (Hendrich and Tweedie 2003; Laget, Joulie et al. 2010). A novel 5mC binding protein, Kaiso, lacks MBD, but recognizes 5mC through zinc finger domains, and mediates repression by associating with the NCoR complex (Hendrich and Bird 1998; Prokhortchouk, Hendrich et al. 2001; Yoon, Chan et al. 2003).

I.IX. MBDs in the brain

The phenotypic deficits in *Mbd1* $^{-/-}$ mice are restricted to CNS, specifically in learning, memory and plasticity, due to reduced neuronal differentiation and increased genomic instability (Zhao, Ueba et al. 2003). *Mbd2* $^{-/-}$ mice show minor

behavioral deficits whereas deletion of *Mbd3* is embryonic lethal (Hendrich, Guy et al. 2001). A role for MBD3 in the neuronal development was suggested since its levels are regulated throughout developing brain, unlike MBD2 which displays minimal expression in the embryonic brain (Jung, Zhang et al. 2003). There is clinical evidence that links MBD5 to mental disorders (Williams, Mullegama et al. 2010; Noh and Graham Jr 2012), however the mechanism is still unclear. Although *Setdb1* $^{-/-}$ mice seem normal (Jiang, Matevossian et al. 2011) and the *Setdb1* gene is expressed at very low levels in the CNS, its histone methyltransferase activity might be required for proper CNS functioning (Ryu, Lee et al. 2006; Jiang, Matevossian et al. 2011). Finally, despite the fact that the X-linked Kaiso gene is highly expressed in the brain, it was not linked to CNS dysfunction (Della Ragione, Tiunova et al. 2006; Prokhortchouk, Sansom et al. 2006).

I.X. MBDs: the curious case of MeCP2

MeCP2 constitutes a unique case within the neuronal epigenetic mechanisms, because of its abundance, enigmatic functions and the clinical manifestation of its dysfunction. MeCP2 is present in all vertebrates and is in fact a “vertebrate invention” (Bird, A., personal communication). The expression of MeCP2 in mice is low at birth, increases greatly in the first three weeks and plateaus when it is expressed near nucleosome levels in the brain (Kishi and Macklis 2004; Skene, Illingworth et al. 2010). It is also expressed to some extent in the lung and spleen (Shahbazian, Young et al. 2002).

- *Structure and disorder in MeCP2*

The full-length human MeCP2 is 60% unstructured, and its secondary structure is stabilized upon binding to DNA or proteins (Adams, McBryant et al. 2007; Ghosh, Nikitina et al. 2010). The X-ray structure of MeCP2 MBD with a 5mCpG-containing DNA (Wakefield, Smith et al. 1999) revealed that 5mC and MBD make multiple contacts, several of which are maintained by immobilized water molecules in the major groove of the double helix (Ho, McNae et al. 2008). Although MeCP2 exists strictly as a monomer in solution (Ghosh, Horowitz-Scherer et al. 2010), it can form clusters upon binding to methylated DNA (Nikitina, Ghosh et al. 2007). MeCP2 affinity to DNA increases with the length of the DNA independent of methylation (Nan, Hou et al. 2007).

In addition to MBD, MeCP2 also contains a basic N-terminal domain with two consensus A/T-hook motifs which bind the minor groove of A/T-rich duplex DNA (Adams, McBryant et al. 2007), and a C-terminal TRD domain that can interact with several factors, including transcription repressors, HDACs, NCoR, mSin3a and CoREST; transcription activators, YY1, YB1, CREB and Brahma; and heterochromatin associated factors, HP1, Dnmt1 and Atrx (Nan, Ng et al. 1998; Kokura, Kaul et al. 2001; Fuks, Hurd et al. 2003; Harikrishnan, Chow et al. 2005; Young, Hong et al. 2005; Agarwal, Hardt et al. 2007; Chahrour, Jung et al. 2008; Forlani, Giarda et al. 2010). The functions of these interactions remain elusive.

- *MeCP2: a transcriptional repressor, or activator?*

Several biochemical and chromatin immunoprecipitation (ChIP) experiments showed that MeCP2 avidly binds methylated DNA (Lewis, Meehan et al. 1992;

Gregory, Randall et al. 2001; Skene, Illingworth et al. 2010), which is associated with gene silencing. The direct involvement of MeCP2 in transcriptional repression by a methylation dependent mechanism was demonstrated in cell culture studies (Boyes and Bird 1991; Nan, Campoy et al. 1997; El-Osta, Kantharidis et al. 2002) and later in reconstituted systems in *Drosophila melanogaster*, which lacks genomic methylation (Kudo 1998). In cortical cultures MeCP2 can associate with some neural genes, such as *Bdnf* (Chen, Chang et al. 2003; Kernohan, Jiang et al. 2010), and this interaction is lost upon stimulation with KCl (Harikrishnan, Bayles et al. 2010; Tian, Marini et al. 2010), leading to increased expression (Martinowich, Hattori et al. 2003).

The widely accepted role of MeCP2 as a transcriptional repressor was challenged with the availability of *Mecp2* $^{-/-}$ mice: First, a brain-restricted deletion of *Mecp2* resulted only in subtle changes in gene expression (Tudor, Akbarian et al. 2002; Jordan, Li et al. 2007) and the search for genes that are repressed by MeCP2, has so far identified only a few targets (Nuber, Kriaucionis et al. 2005; Kriaucionis, Paterson et al. 2006; Jordan, Li et al. 2007). Later analyses of gene expression in hypothalami and cerebella of *Mecp2*-deficient and *Mecp2*-overexpressing mice, revealed that the majority of genes were activated by MeCP2 (Chahrour, Jung et al. 2008; Ben-Shachar, Chahrour et al. 2009). Some groups also reported MeCP2-dependent changes in gene splicing (Young, Hong et al. 2005), microRNA profiles (Urdinguio, Fernandez et al. 2010), and transcription of repetitive elements (Skene, Illingworth et al. 2010). Thus a more

complex role for MeCP2 was proposed, where it dampens transcriptional noise as a buffer and/or regulates transcription depending on the molecular context.

- *Where is MeCP2?*

The early studies showed that MeCP2 localizes in heterochromatic foci (Hendrich and Bird 1998), that it can stably associate with nucleosomes with methylated DNA and linker DNA (Chandler, Guschin et al. 1999; Yang, van der Woerd et al. 2011), and that it can facilitate chromatin compaction (Georgel, Horowitz-Scherer et al. 2003; Nikitina, Ghosh et al. 2007; Nikitina, Shi et al. 2007). In addition, absence of *Mecp2* caused a normally silent chromatin loop to shift into an active chromatin state (Horike, Cai et al. 2005). Hence an alternative/additional role was advocated for MeCP2 in the management of the higher order chromatin landscape.

To the contrary, in two independent genome-wide analyses, MeCP2 was found broadly distributed over the genome and not localized at discrete sites, such as heterochromatic regions or promoters of repressed genes (Yasui, Peddada et al. 2007; Skene, Illingworth et al. 2010). Yet if MeCP2 covers the whole genome and its primary role is the organization of heterochromatin, then one would wonder why not all genome is condensed. Indeed, upon reexamination of nuclear staining of MeCP2 (LaSalle and Gerald 2004; Nan, Hou et al. 2007), it is evident that at heterochromatic foci MeCP2 is dense, as is DNA. The hypothesis that MeCP2 broadly covers the genome gained more support as more evidence came to light. For instance, in the absence of both *Mecp2* and *Dnmts,* the chromatin structure undergoes global changes, including altered histone

acetylation and doubling of histone H1 (Nan, Campoy et al. 1997; Shahbazian, Young et al. 2002; Skene, Illingworth et al. 2010), not only in neurons but also in glia (Ballas, Lioy et al. 2009). This is complimentary to the observation that histone H1 levels in neurons are 50% less than in other cell types (Pearson, Bates et al. 1984) and that MeCP2 can compete with H1 on nucleosomes containing methylated DNA (Ghosh, Horowitz-Scherer et al. 2010). Together these results imply a histone H1-like role for MeCP2.

In a recent study, a population of MeCP2 was found in chromatin regions that contain high levels of nucleosome and a second loosely-bound population in euchromatin. This second population was unique to the brain and absent in other tissues, suggesting a tissue-specific functional compartmentalization of MeCP2 (Thambirajah, Ng et al. 2011). An earlier chromatin fractionation study from cultured human cells showed that the majority of MeCP2 is present in the more nuclease-accessible, active regions of chromatin, while a small portion was associated with heterochromatin (Ishibashi, Thambirajah et al. 2008). Intriguingly, in the absence of *Mecp2,* the size of neuronal nuclei fails to increase at normal rates during differentiation and transcription is attenuated (Yazdani, Deogracias et al. 2012), both indicative of involvement of MeCP2 in the configuration of euchromatin state. The association of MeCP2 with active genes and active chromatin may seem to contradict its widely accepted role in repression and heterochromatin. However they can also simply be the manifestations of a dynamic, temporally dependent and activity-dependent function (Metivier, Gallais et al. 2008).

- *Rett Syndrome: A direct link to MeCP2 dysfunction*

It was not too long after discovery of MeCP2 that it was directly linked to Rett Syndrome (RTT) (Amir, Van den Veyver et al. 1999). Being an X-linked gene; mutations of *MECP2* affect males much more severely than girls, which exhibit mosaicism in heterozygocity after one X-chromosome is inactivated during dosage compensation (Adler, Quaderi et al. 1999). In line with the expression pattern of MeCP2, the symptoms of RTT arise only after 6-18 moths of age in girls: loss of speech, loss of purposeful hand use, stereotypical movements, seizures, mental retardation and hyperventilation. Only few boys are diagnosed as they die within two years of birth (Hagberg, J et al. 1983). In postmortem brain tissue from RTT patients, lower spine density in hippocampal pyramidal neurons was observed (Chapleau, Calfa et al. 2009). Neurons, derived from induced pluripotent stem cells generated from RTT patients' fibroblasts, had reduced spine density, smaller cell body and electrophysiological defects (Marchetto, Carromeu et al. 2010; Cheung, Horvath et al. 2011).

Several deletions and mutants of MeCP2 have been created in mice (Chen, Akbarian et al. 2001; Guy, Hendrich et al. 2001; Pelka, Watson et al. 2006). Although heterozygous females acquire some phenotypes only in older ages; homozygous females and hemizygous males exhibit RTT symptoms at 5 weeks and die between 6 and 12 weeks. (Guy, Hendrich et al. 2001). Brain-specific deletion of *Mecp2* during or after development results in full knockout phenotype (Chen, Akbarian et al. 2001; Nguyen, Du et al. 2012), as mice carrying MeCP2 with RTT-causing mutations (Shahbazian, Young et al. 2002; Lawson-Yuen, Liu

et al. 2007). In these mice, the gross brain architecture is normal, indicating that MeCP2 is not required for neurodevelopment. However smaller and densely packed nuclei and cell bodies were obvious. The spine number and dendritic complexity was also reduced (Na and Monteggia 2011), in agreement with a cell culture study, showing loss of *Mecp2* leads to reduction in synapse numbers, which is reversed upon gain of *Mecp2* (Chao, Zoghbi et al. 2007). In the absence of MeCP2, inhibitory activity increases and excitatory activity decreases resulting in cortical dysfunction (Dani, Chang et al. 2005).

Overexpression of *Mecp2* only by two-fold initially results in enhanced learning and plasticity (Collins, Levenson et al. 2004). At 20 weeks of age neurological symptoms start to appear and mice have shorter life span. Serine-421 (S421) of MeCP2 was initially identified as an activity induced phosphorylation site that displaces MeCP2 from DNA (Zhou, Hong et al. 2006); however a S421A mutation changed neither the genome-wide binding of MeCP2 *in vivo* nor the expression of specific genes. The mice carrying these mutations exhibited defects in synapse development in cortical pyramidal neurons and mild behavioral abnormalities (Cohen, Gabel et al. 2011). In another study, the S421A/S424A double mutants exhibited overexpression phenotypes (Li, Zhong et al. 2011). S80 is ubiquitously phosphorylated, but this is lost upon neural stimulation (Tao, Hu et al. 2009). Mice carrying S80A mutation show mild phenotypes, and DNA binding of MeCP2 is attenuated.

Deletion of *Mecp2* in inhibitory (Chao, Chen et al. 2010) and excitatory (Samaco, Hogart et al. 2005) neurons, as well as in confined regions in the brain

(Bissonnette and Knopp 2006; Gemelli, Berton et al. 2006; Fyffe, Neul et al. 2008), mimics various aspects of the disease in mice. Complementary to these results, full reactivation of *Mecp2* in its original levels and partial reactivation in neuron subpopulations reverses some or all phenotypes (Collins, Levenson et al. 2004; Luikenhuis, Giacometti et al. 2004; Giacometti, Luikenhuis et al. 2007). *Mecp2* $^{-/-}$ astrocytes cause abnormalities in neurons (Ballas, Lioy et al. 2009; Maezawa, Swanberg et al. 2009), and reactivation of *Mecp2* in astrocytes in null mice reverses majority of the disease phenotypes (Lioy, Garg et al. 2012). Interestingly, hippocampal neurons treated with conditioned medium obtained from *Mecp2*-null microglia, which release five-fold higher glutamate, develop abnormally (Maezawa and Jin 2010); and transplantation of WT myeloid cells into *Mecp2* $^{-/-}$ mice arrests major symptoms of RTT (Derecki, Cronk et al. 2012).

In summary, MeCP2 and RTT present a particular case of the epigenetics of the brain that is fundamentally different from other tissues. Here, subtle changes at the molecular level have severe consequences. Now, the addition of 5hmC to the vocabulary of brain epigenetics may provide more insight into the complexity of the mechanisms that sustain the CNS.

I.XI. 5hmC: discovery, genomic distribution and function

The existence of 5hmC in the mammalian genome was re-discovered during a comparative study of the genomic 5mC content between large and decondensed PC nuclei and compact GC nuclei (Fig 1.1), 60 years after its first discovery in T2 bacteriophages (Wyatt and Cohen 1952; Warren 1980;

Kriaucionis and Heintz 2009). Its presence in the mammalian brain was initially suggested (Penn, Suwalski et al. 1972), but this was not reproduced by others (Kothari and Shankar 1976; Gommers-Ampt and Borst 1995).

5hmC levels in the brain range from 0.4 to 0.65 % of all Cs, whereas kidney, lung and muscle tissue exhibit “medium” levels, and finally lowest levels are detected in the spleen, liver and testis. In contrast, 5mC levels are constant at 4-5 % in a variety of tissues (Globisch, Münzel et al. 2010). The change in the amount of 5mC between cell populations was complementary to the levels of 5hmC suggesting that 5mC and 5hmC may be derived from each other (Kriaucionis and Heintz 2009).

5hmC is invariably enriched on euchromatin regardless of the cell type or developmental stage tested (Ficz, Branco et al. 2011; Szulwach, Li et al. 2011). For instance, on metaphase chromosomes, 5hmC is located on chromosome arms but depleted on centromers that contain high levels of 5mC (Szulwach, Li et al. 2011). 5hmC levels are highest within globally decondensed nuclei enriched in active chromatin, such as those of ES cells or PCs (Meshorer and Misteli 2006; Kriaucionis and Heintz 2009; Tahiliani, Koh et al. 2009). The localization of 5hmC within the “on” state of chromatin implies that it may be involved in gene activation. Indeed gene bodies, promoters, transcription start sites (TSSs) and enhancer elements of active genes in ES cells have elevated levels of 5hmC, that strongly tracks with active enhancer marks, p300, H3K4me1, H3K4me2, H3K18ac and H3K27ac, and is mostly excluded from heterochromatin marks H3K27me3 (Ficz, Branco et al. 2011; Pastor, Pape et al. 2011; Stroud, Feng et al.

2011; Szulwach, Li et al. 2011; Wu, D'Alessio et al. 2011; Booth, Branco et al. 2012; Yu, Hon et al. 2012). Intriguingly, extensive strand-bias was detected in methylomes and hydroxymethylomes of ES cells largely in non-CpG context (Ficz, Branco et al. 2011), which increased in the absence of *Uhrf1*,*Tet1* and *Tet2;* suggesting that Dnmt3a or Dnmt3b with Tet3, or other putative enzymes might be required for asymmetric hydroxymethylation.

In tissue samples 5hmC levels correlate with the differentiation state of cells, increasing toward terminally differentiated layers in hierarchically organized tissues and mainly enriched in gene bodies (Münzel, Globisch et al. 2010; Haffner, Chaux et al. 2011; Song, Szulwach et al. 2011; Orr, Haffner et al. 2012). Aberrant hydroxymethylation patterns are observed in both imprinting disorders and cancer (Haffner, Chaux et al. 2011).

The function of 5hmC still remains unclear: an intermediate in active or passive demethylation pathways, or a bona fide epigenetic mark. It is now widely accepted that 5hmC is passively demethylated in dividing cells, since Dnmt1 does not recognize hemihydroxymethylated DNA as substrate (Valinluck and Sowers 2007; Hashimoto, Liu et al. 2012). However Uhrf1, which recruits Dnmt1 onto hemimethylated DNA, can bind hemi- and fully hydroxymethylated CpG sites with similar affinity (Sharif, Muto et al. 2007; Frauer, Hoffmann et al. 2011). Therefore, it should be investigated whether a Uhrf1-mediated mechanism maintains hemihydroxymethylated sites over cell divisions *in vivo.*

A second proposed role for 5hmC is an intermediate in active demethylation events, such as the reprogramming of methylome during embryogenesis, where

global methylation patterns are erased and a wave of *de novo* methylation follows (Mayer, Niveleau et al. 2000; Hajkova, Erhardt et al. 2002; Chen, Ueda et al. 2003; Hackett, Zylicz et al. 2012). This way a pluripotent state can be established in the zygote and then again in developing primordial germ cells (Reik, Dean et al. 2001). Within hours of fertilization, the male genome is stripped of methylation at a faster rate than cell division (Oswald, Engemann et al. 2000). Several groups have proposed that this process involves active demethylation mechanisms via DNA glycosylases, other putative DNA decarboxylases (Wu and Zhang 2010; Cortellino, Xu et al. 2011; He, Li et al. 2011; Inoue and Zhang 2011; Maiti and Drohat 2011; Hashimoto, Hong et al. 2012), or deaminases that can convert 5hmC into 5-hyroxymethyluracil (5hmU) in combination with base excision repair (BER) mechanisms (Zhu 2009; Branco, Ficz et al. 2012; Hackett, Zylicz et al. 2012; Morera, Grin et al. 2012). These hypotheses have drawn special attention after the discoveries that the demethylation of paternal genome follows a wave of hydroxymethylation (Iqbal, Jin et al. 2011; Wossidlo, Nakamura et al. 2011); and that 5hmC can be further oxidized to 5-formylcytosine (5fC) and then to 5-carboxycytosine (5caC) *in vitro* (Ito, Shen et al. 2011). However, 5fC and 5caC was not detected during zygote development (Inoue, Shen et al. 2011). Despite the recent accumulation of data, it is still not fully understood whether an active demethylation event, that involves oxidization of 5hmC into 5fC and 5caC followed by TDG-mediated base excision, takes place *in vivo*. (Cortellino, Xu et al. 2011; He, Li et al. 2011; Maiti and Drohat 2011; Hashimoto, Hong et al. 2012; Morera, Grin et al. 2012).

Finally, the accumulation of 5hmC in the brain and the differential distribution of 5hmC on different loci between brain areas strongly support a stable epigenetic role for 5hmC (Globisch, Münzel et al. 2010; Nestor, Ottaviano et al. 2011; Song, Szulwach et al. 2011; Szulwach, Li et al. 2011). In this scenario 5hmC can (1) evict proteins that bind methylated sequences; (2) recruit novel factors that avidly bind hydroxymethylated sequences; and/or (3) change DNA structure within chromatin.

I.XII. Intermission: a retrospective methodological evaluation

The addition of a third dimension to the cytosine modification status necessitated the re-evaluation of the traditional methods to analyze genomic methylation, since they operate in binary mode, where 5hmC, 5fC and 5caC have been detected in either C or 5mC population.

- *Quantification methods*

The traditional method to quantify global levels of methylation is nearest-neighbor analysis (Ramsahoye and Mills 2002), where a restriction enzyme creates a “sticky end” on DNA where it is labeled. The labeled nucleotides on these ends are separated by chromatography. A second method is use of methylation-sensitive restriction enzymes that can cleave DNA only if the target site is unmodifed (Pells, Moore et al. 2006). However both methods limit the analysis to a 4-6 bp site covering small part of the genome; they are biased for the dinucleotide context, the resistance of 5hmC, 5fC and 5caC is unknown and the former is also strongly dependent on the chromatography conditions

(Ramsahoye and Mills 2002; Ito, Shen et al. 2011). Today, the most sensitive and robust techniques to quantify modified nucleosides are high-performance liquid chromatography, mass spectrometry and nanopore amperometry (Pomerantz, McCloskey et al. 1990; Quinlivan and Gregory 2008; Clarke, Wu et al. 2009).

- *Mapping methods*

To analyze methylation status of specific loci scientists originally used MBDs to enrich methylated DNA (Cross, Chariton et al. 1994; Rauch and Pfeifer 2005; Rauch, Li et al. 2006). However each MBD exhibits a different binding pattern, making this method inherently biased. For instance, in one study the genomic localization of MeCP2 was shown to track with DNA methylation pattern (Skene, Illingworth et al. 2010), which was evaluated using the MBD of the same protein (Illingworth, Kerr et al. 2008). Furthermore, initially only a few selected loci were analyzed (Meissner, Mikkelsen et al. 2008; Irizarry, Ladd-Acosta et al. 2009) until genome-wide sequencing methods were developed. Currently methylated DNA immunoprecipitation (MeDIP-Seq) is widely used in combination with high-throughput sequencing (Weber, Davies et al. 2005), while antibodies developed for hydroxymethylated DNA immunoprecipitation (hMeDIP-Seq) show high background, especially with repetitive sequences. (Jin, Kadam et al. 2010; Ficz, Branco et al. 2011; Williams, Christensen et al. 2011; Xu, Wu et al. 2011). Therefore, antibodies were developed against cytosine-5-methylenesulphonate, product of bisulfite treatment of 5hmC (Ko, Huang et al. 2010; Pastor, Pape et al. 2011).

In a recent chemical labeling technique, T4 bacteriophage β-glucosyltransferase (βGT) transfers an azide-containing glucose moiety onto the hydroxyl group of 5hmC. The azide group then chemically incorporates a biotin group that can be used for detection, affinity enrichment or sequencing (Song, Szulwach et al. 2011; Robertson, Dahl et al. 2012). Glucosylation reaction alone can also be used for modification-resistant restriction enzymes that are blocked by glucosylation. Finally, a restriction enzyme that has been cloned almost two decades ago, cleaves DNA at hydroxymethylcytosine independent of sequence, and its activity is also blocked by glucosylation (Janosi, Yonemitsu et al. 1994).

- *Base-resolution sequencing methods*

In BS-Seq, DNA is treated with bisulfite, which converts all unmodified Cs into uracils, and the comparison of the treated sequence to the original sequence reveals 5mCs (Frommer, McDonald et al. 1992). Soon after the finding that 5hmC behaves like 5mC in BS-Seq whereas 5fC and 5caC behave like unmodified C; new techniques were invented to incorporate 5hmC into BS-Seq via an Tet-mediated or chemical oxidation step (Huang, Pastor et al. 2010; Booth, Branco et al. 2012; Yu, Hon et al. 2012). In a single-molecule real-time sequencing, kinetic signatures of a mutated DNA polymerase are monitored during incorporation of fluorescent nucleotides (Flusberg, Webster et al. 2010).

I.XIII. Tets: enzymes in the making of 5hmC

The methyl group of 5mC can be oxidized to 5hmC *in vitro* and in cell cultures by any of the three recently identified members of the Ten-eleven translocation (Tet) family of proteins, which belong to the superfamily of 2-oxoglutarate- and iron-dependent dioxygenases (Fig 1.4) (Iyer, Tahiliani et al. 2009; Tahiliani, Koh et al. 2009; Ito, D'Alessio et al. 2010). *TET1* and *TET2* are both implicated in cancer (Ono, Taki et al. 2002; Lorsbach, Moore et al. 2003; Burmeister, Meyer et al. 2009; Abdel-Wahab, Mullally et al. 2009; Ko, Huang et al. 2010); and so are enzymes that cause inhibition of Tets (Xu, Yang et al. 2011). They are also the key enzymes responsible for the presence of 5hmC in ES cells, as opposed to Tet3, which is highly expressed in terminally differentiated tissues (Ito, D'Alessio et al. 2010; Koh, Yabuuchi et al. 2011). Although these data suggest a role for Tet1 and Tet 2 in ES cell pluripotency and oncogenic transformation; their roles in ES cell self-renewal and maintenance of pluripotent state has been contradictory (Ito, D'Alessio et al. 2010; Dawlaty, Ganz et al. 2011; Koh, Yabuuchi et al. 2011; Williams, Christensen et al. 2011; Freudenberg, Ghosh et al. 2012).

All Tet enzymes can oxidize 5hmC further into 5fC and 5caC *in vitro* (Inoue, Shen et al. 2011). Evidence suggests that Tet3 mediates the demethylation of paternal genome (Iqbal, Jin et al. 2011; Wossidlo, Nakamura et al. 2011), whereas a recent report proposes a regulatory role for Tet1 in the expression of a subset of meiotic genes during generation of oocytes (Yamaguchi, Hong et al. 2012). The increased methylation at many CGIs caused by depletion of Tet1 (Xu,

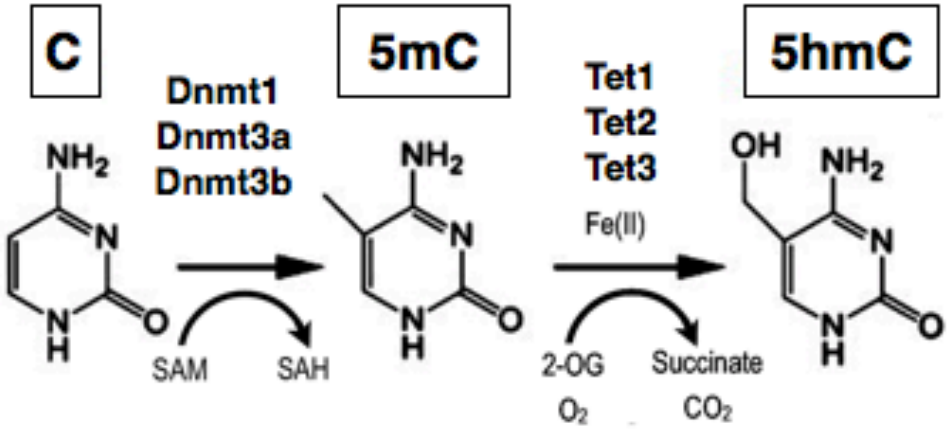


Fig 1.4. Schematic representation of methylation and hydroxylation. A methyl group from SAM is added to the 5 position on cytosine (C) by one of the Dnmts to produce 5-methylcytosine (5mC). One of the Tets add a hydroxyl group from 2-oxoglutarate onto the methyl group of 5mC to generate 5-hydroxymethylcytosine (5hmC)

Wu et al. 2011) and hypomethylation in patients with *TET2* mutations (Ko, Huang et al. 2010) also support a role for Tets in demethylation pathways, yet more research needs to be done to establish a coherent model. Although active demethylation has also been advocated in the brain, this hypothesis was based on cultured neurons that over-expressed Tet1 (Guo, Su et al. 2011). In addition, intermediate products, 5fC, 5caC and 5hmU were not detected in terminally differentiated tissues (Globisch, Münzel et al. 2010; Pfaffeneder, Hackner et al. 2011).

In ES cells Tet1 also colocalizes with both active (Pastor, Pape et al. 2011) and repressive histone marks, and also bivalent sites, that carry with both repressive and activating marks (Wu, D'Alessio et al. 2011; Xu, Wu et al. 2011), with a strong preference for unmodified CGIs within promoter regions and less within gene bodies (Williams, Christensen et al. 2011; Wu, D'Alessio et al. 2011). This suggests that it may also be functioning as a transcriptional modulator via its CXXC domain (Zhang, Zhang et al. 2010), which is indeed dispensable for its catalytic activity *in vivo* (Frauer, Rottach et al. 2011).

I.IV. 5hmC in the brain

The existing data can partly describe active demethylation events during development, yet they are still not sufficient to explain the accumulation of 5hmC in terminally differentiated cells. Although an active demethylation is advocated in the brain, this hypothesis was based on cultured neurons that overexpressed Tet1 (Guo, Su et al. 2011). In addition, intermediate products, 5fC, 5caC and

5hmU were not detected in terminally differentiated tissues (Globisch, Münzel et al. 2010; Pfaffeneder, Hackner et al. 2011). On the other hand, the location of 5hmC in the genome is significantly different in the brain, and most differentially hydroxymethylated regions (DhMRs) between different neuroanatomical regions are stable throughout life (Szulwach, Li et al. 2011; Chen, Dzitoyeva et al. 2012; Wang, Pan et al. 2012). To gain more insight, our group recently mapped 5mC and 5hmC in specific cell types in the brain, taking advantage of bacTRAP mice (Heintz 2000; Heiman, Schaefer et al. 2008), that express a ribosomal protein, L10A, tagged with enhanced green fluorescent protein (eGFP) in cell bodies (Fig 1.5.A) and nucleoli (Fig 1.5.B) of specific neuronal subpopulations. Having eGFP-tagged ribosomes allowed both the enrichment of mRNA and the isolation of genomes from specific cell types (Mellen, Ayata et al. 2012). In general, the distribution of 5hmC across the genome in specific types of neurons and glia was in agreement with previous studies of brain tissue (Song, Szulwach et al. 2011; Szulwach, Li et al. 2011): 5hmC was preferentially enriched over the entire transcription unit of expressed genes, and depleted from both the TSS and intergenic regions. In highly expressed genes 5mC was depleted over the gene bodies, whereas the enrichment of 5hmC varied between cell types. The patterns of 5hmC and 5mC were inversely correlated. These findings by our group, a recent study where 5hmC was found enriched in gene bodies of synaptic genes (Khare, Pai et al. 2012), and the detection of DhMRs, support that 5hmC is a stable epigenetic mark in the brain that is utilized in a clearly different manner for neural functions.

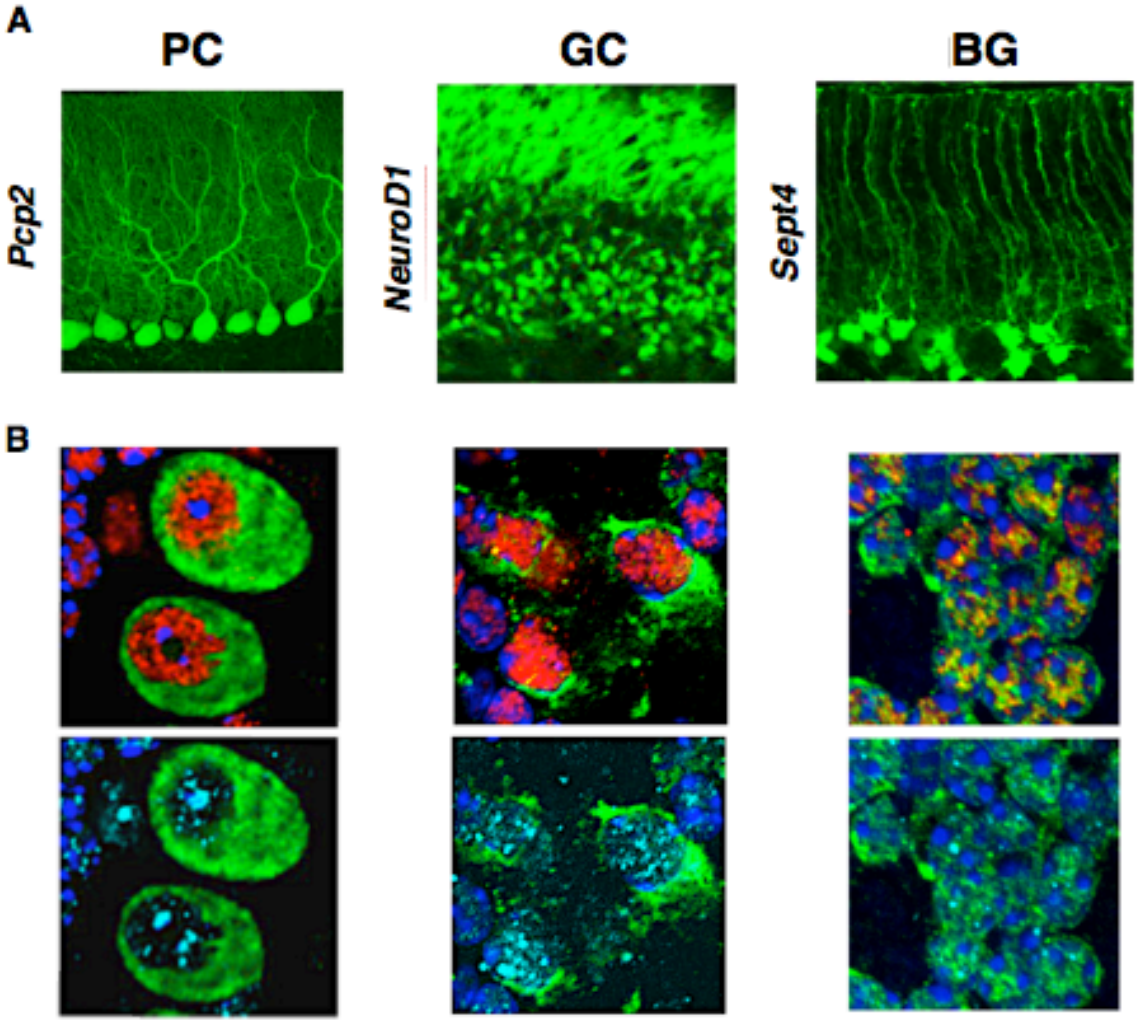


Fig 1.5. Brain sections from BAC transgenic mice. **A.** *Pcp2, NeuroD1* and *Sept4* BAC transgenic mice, expressing eGFP (green) in Purkinje Cells (PCs), granule cells (GCs) and Bergman glia (BG), respectively (www.gensat.org). **B.** Brain sections from bacTRAP lines as indicated in A but expressing eGFP/L10A (green) and visualized in higher magnification. DNA stained with DAPI (blue), and 5hmC (red) and 5mC (cyan) were stained by antibodies (Mellen, Ayata et al. 2012).

I.V. Epilogue: the search for 5hmC-binding proteins

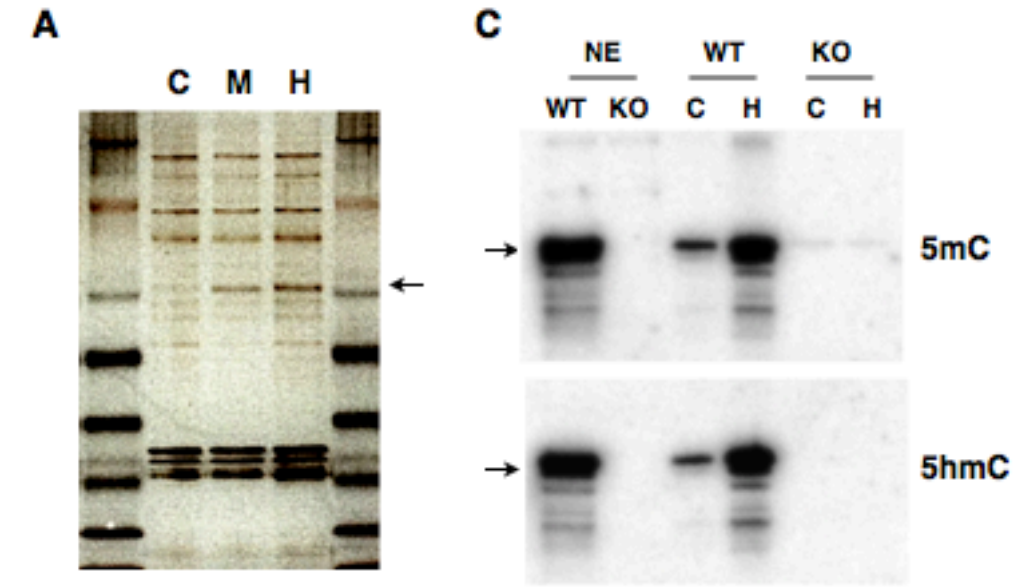
5hmC revealing itself as a stable and activating mark in the brain, raises immediate questions: (1) by which upstream events is this epigenetic code written at specific locations in the brain that differ from pluripotent cell states; and (2) by which downstream implications is it interpreted as a stable and activating mark. The aim of this study is to address the second question by trying to unveil nuclear factors that recognize the 5hmC mark in the brain. Factors that bind to methylated DNA, have elucidated the main mechanisms that convert 5mC into a repressive mark. Two 5hmC-binding proteins have been previously reported: First, Frauer et al. showed that Uhrf1 recognizes hydroxymethylated CpGs (5hmCpGs) (Frauer, Hoffmann et al. 2011). Although this finding was informative for the maintenance of 5hmC, it did not provide any insight of its interpretation as an epigenetic mark. Later, Yildirim et al. demonstrated that Tet1 and MBD3 colocalize in ES cells (Yildirim, Li et al. 2011). This study proposed an alternative control of gene expression of bivalent genes in pluripotent cells, by the repressor MBD3 and activator Brg1. Yet this group also could not establish a direct functional link between 5hmC and active chromatin states in postmitotic cells. The study we present here, attempts to solve this conundrum by identification and functional characterization of brain-specific proteins that recognize 5hmC. Thus, we hope to elucidate mechanisms that can read the 5hmC code and translate it into an activation mark in the brain.

Chapter II.

MeCP2 binds to 5hmC at CpA dinucleotides enriched in the bodies of the active genes in euchromatin

II.I. MeCP2 is an abundant 5hmC binding protein

To identify the factors that may be responsible for decoding 5hmC in the brain, nuclear extracts prepared from rodent brain (Klose and Bird 2004) were incubated with magnetic beads coated with DNA containing unmodified C, 5mC or 5hmC in the presence of excess non-specific DNA competitor. After the beads were isolated, proteins captured on the beads were eluted and separated on SDS PAGE. Silver staining of this gel revealed a band of ~70 kDa enriched with both 5mC and 5hmC, but not with C (Fig 2.1.A). This band was excised from a preparative gel of this type, analyzed by mass spectroscopy and identified as MeCP2 in a peptide database search with ~50% sequence coverage (Fig 2.1.B). Since it is possible that novel low-abundance 5hmC binding proteins might be obscured by the abundant MeCP2, we repeated this experiment in the absence of MeCP2. Nuclear extracts were prepared from *Mecp2*-KO animals using beads coated with DNA containing C or 5hmC. Upon visualization by Coomassie or more sensitive Silver stain, we did not detect any 5hmC-specific bands in the eluates from KO animals (data not shown). As an alternative detection, we transferred the electrophoresed eluates from wild type (WT) and KO onto a charged membrane and analyzed it by Southwestern blotting method



B

```
  1 MAAAAAAAAA AAAAAAAAAA AAAAAPSGGG GGEEERLEEK SEDQDLQGLK
 51 EKPLKFKKVK KDKKEDKEGK HEPLQPSAHH SAEPAEAGKA ETSESSGSAP
101 AVPEASASPK QRRSIIRDRG PMYDDPTLPE GWTRKLKQRK SGRSAGKYDV
151 YLINPQGKAF RSKVELIAYF EKVGDTSLDP NDFDFTVTGR GSPSRREQKP
201 PKKPKSPKAP GTGRGRGRPK GSGTGRPKAA ASEGVQVKRV LEKSPGKLLV
251 KMPFQASPGG KGEGGGATTS AQVMVIKRPG RKRKAEADPQ AIPKKRGRKP
301 GSVVAAAAAE AKKKAVKESS IRSVQETVLP IKKRKTRETV SIEVKEVVKP
351 LLVSTLGEKS GKGLKTCKSP GRKSKESSPK GRSSSASSPP KKEHHHHHHH
401 AESPKAPMPL LPPPPPPEPQ SSEDPISPPE PQDLSSSICK EEKMPRAGSL
451 ESDGCPKEPA KTQPMVAAAA TTTTTTTTTV AEKYKHRGEG ERKDIVSSSM
501 PRPNREEPVD SRTPVTERVS
```

Fig 2.1. DNA containing 5mC and 5hmC binds MeCP2. **A.** SDS PAGE gel with Silver stain. The lanes contain proteins pulled down with DNA containing C, 5mC (M) or 5hmC (H). Arrow points at the excised band **B.** The sequence of rat MeCP2 as identified by mass spectrometry of the excised band followed by database search. Red sequences are peptide hits from M and H; blue was only found in M. **C.** Membranes analyzed by Southwestern blotting with radioactive probes containing 5mC (top) or 5hmC (bottom). First two lanes contain nuclear extract (NE) from wild type (WT) or *Mecp2*-knockout (KO) mice as input; lanes 3 and 5 are pulldowns with DNA containing C, lanes 4 and 6 with DNA containing 5hmC (H); from WT and KO, respectively. Arrows mark MeCP2.

(Campoy, Meehan et al. 1995). Thus, membrane bound re-natured proteins were probed with DNA end-labeled with ^{32}P isotope containing either 5mC or 5hmC, revealing a protein of the correct molecular weight that can bind both 5mC and 5hmC containing DNA probes from WT animals, and that is not present in samples prepared from KO animals (Fig 2.1.C). To our surprise, no other abundant protein with high specificity for 5hmC DNA was revealed in these studies, even in the absence of MeCP2.

II.II. Recombinant MeCP2 specifically binds to 5mC and 5hmC

To test the binding of MeCP2 to 5hmC directly, a His-tagged N-terminal fragment (NT) of human MeCP2 containing the MBD (residues 1-205) was produced in *E. coli* and purified using Nickel-nitrilotriacetic acid (Ni-NTA) resin (Fig 2.2.A). Probes with C, 5mC or 5hmC nucleotides were prepared (Fig 2.2.B) and used in electrophoretic mobility shift assays (EMSA) to measure binding. At all concentrations tested, the MeCP2 NT failed to bind the C-containing probe, while avidly binding both the 5mC and 5hmC probes (Fig 2.2.C).We also have purified the minimal MBD (residues 77-167) of MeCP2 and confirmed that MBD was sufficient for this binding (Fig 2.2.D). As an additional control, EMSAs also were performed using probes reacted with T-4 phage β-glucosyltransferase (βGT), which adds a glucose from uridine diphosphoglucose (UDP-glucose) specifically to the -OH group of 5hmC without affecting 5mC and C nucleotides (Szwagierczak, Bultmann et al. 2010). The binding of MeCP2 NT to the glucosylated 5hmC probe was blocked due to glucosylation of 5hmC residues

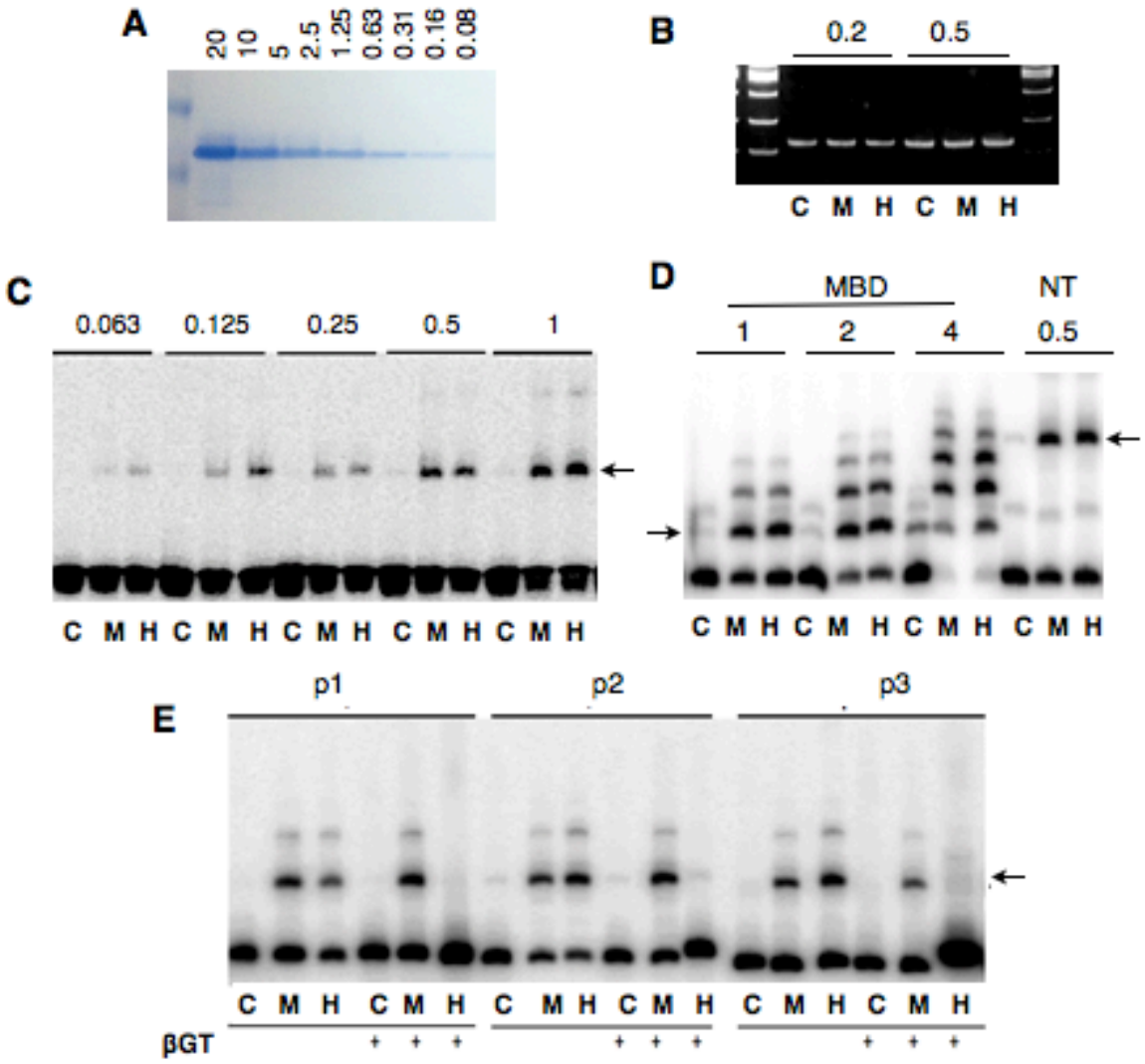


Fig 2.2. Recombinant human MeCP2 binds 5mC and 5hmC via MBD. **A.** SDS PAGE gel containing purified MeCP2 NT (1-205 aa) stained with Coomassie. Concentrations in ng. **B.** EtBr-stained agarose gel with DNA probes containing C, 5mC (M) and 5hmC (H) prior radioactive labeling. Concentrations in µg. **C.** EMSA gel with 5 fmoles of C-, M- and H-containing DNA probes and increasing concentrations of MeCP2 NT. Concentrations in pmols. **D.** EMSA gel as in C. increasing concentrations of MeCP2 MBD (77-167 aa). concentrations in pmols. **E.** EMSA gel as in C each lane with 0.5 pmoles of MeCP2 NT. Probes, that are marked with +, are reacted with β-glucosyltransferase (βGT) prior to labeling. p1-p3 are different sequences selected from the mouse genome. Arrows point at protein-dependent DNA complexes.

and retained as 5mC, which is refractive to glucosylation (Fig 2.2.E). MeCP2 binding to 5hmC was not sequence specific since the binding properties of MeCP2 to a variety of probes selected from the mouse genome did not vary significantly.

We set up to examine the modification-specific interaction of other MBD family proteins. A schematic view of the recombinant fragments of MBDs is shown in Fig 2.3.A. We purified His-tagged MBDs 1 through 4 using a column prepared with Ni-NTA (Fig 2.3.B) (Janknecht, de Martynoff et al. 1991). In contrast MBD1, 2 and 4 all bound strongly to 5mC containing DNA, and did show specificity to 5hmC containing probes (Fig 2.3.C). As previously reported (Yildirim, Li et al. 2011), binding of MBD3 was observed to both 5mC and 5hmC DNAs, and the mobility of the MBD3/5hmC complex was slightly retarded relative to the MBD3/5mC complex. MBD3 binding to both 5mC and 5hmC was much weaker, requiring amounts that are two orders of magnitude higher than the other MBD proteins, and its binding to 5hmC DNA was sensitive to glucosylation. A weak, 5hmC-specific and βGT-sensitive was also observed with MBD4.

We verified these observations by competition assays: in the absence of a nonspecific competitor, unlabeled, or “cold” probes containing 5mC and 5hmC can compete out the binding of MeCP2 to radioactively labeled, or “hot”, 5mC-containing DNA probe with the same efficiency (Fig 2.4.A). This effect is much less pronounced for C. On the contrary, “cold” 5mC-containing probe could compete out the binding of MBD1 and 2 to radioactive probe containing 5mC, much more efficiently than “cold” probes containing C or 5hmC (Fig 2.4.B and C).

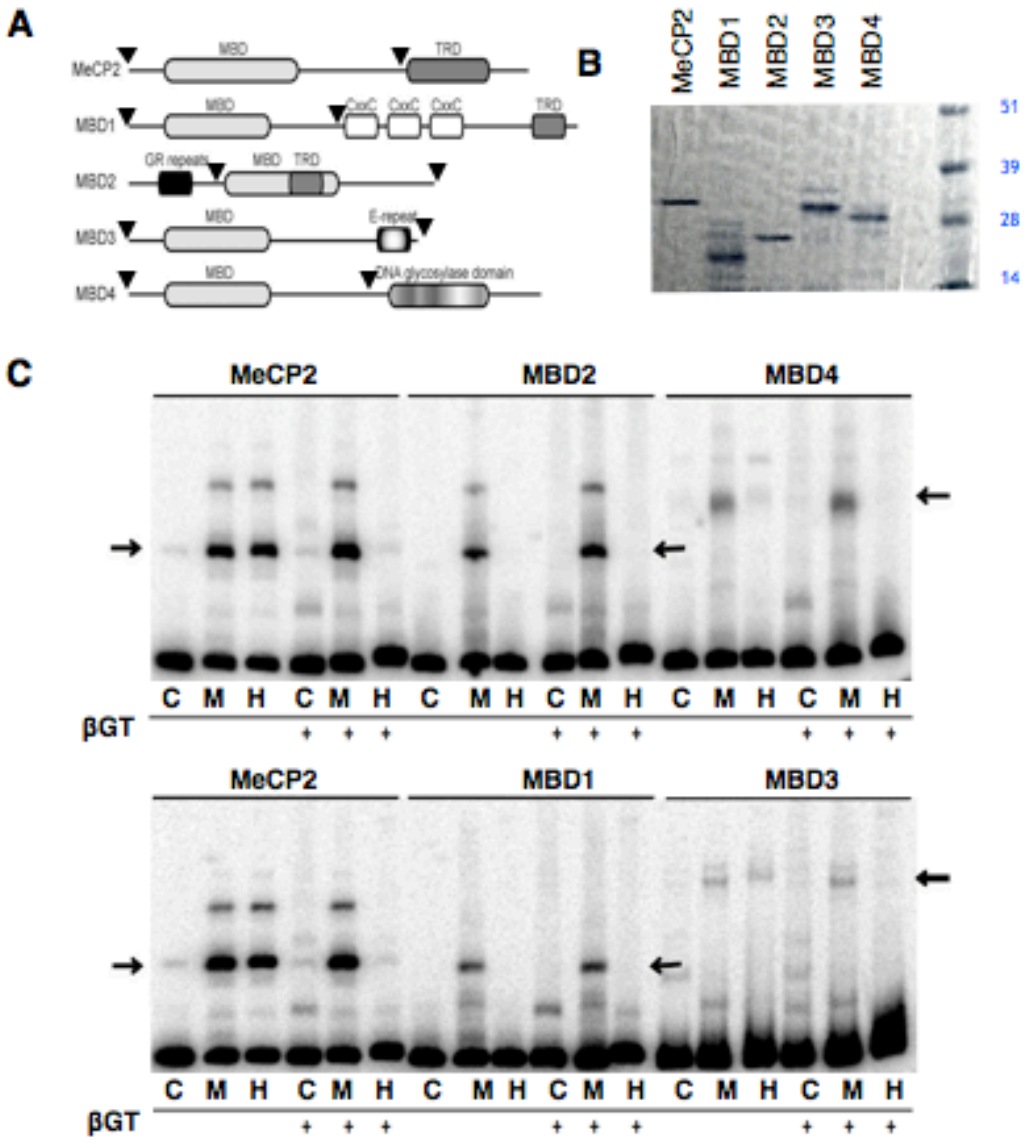


Fig 2.3. Binding of major MBD proteins to C, 5mC and 5hmC. **A.** Schematic view of recombinant MBD used in this study (Dhasarathy and Wade, 2008). Triangles frame cloned fragments **B.** SDS PAGE stained with Coomassie, containing purified recombinant MBD proteins as shown in A. **C.** EMSA gel with 5 fmoles of C-, M- and H-containing DNA probes in the presence of 1 pmol of MeCP2 (aa 1-205), 1.2 pmol of MBD1 (aa 1-144), 0.5 pmol of MBD2 (aa 141-319), 50 pmol of MBD3 (full-length aa 1-259) or 2.5 pmol of MBD4 (aa 1-197). The probes were reacted with βGT prior to labeling if marked with +. Arrows point at protein-dependent DNA complexes.

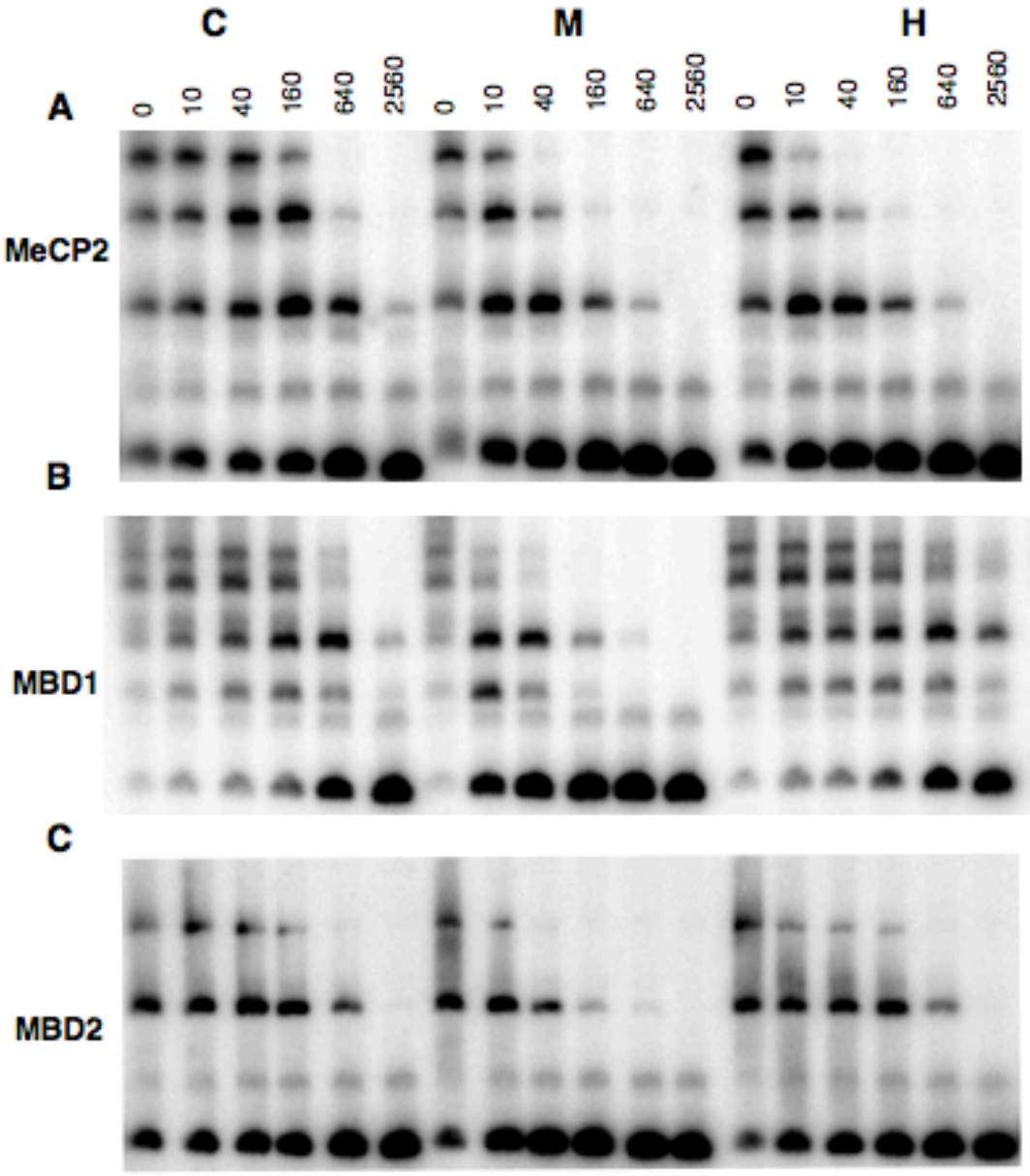


Fig 2.4. Competition of 5mC binding of MeCP2 by 5hmC and 5mC. **A.** EMSA gel of 1 pmol of MeCP2 with 5 fmoles of radioactive DNA probe containing 5mC in absence of nonspecific competitor but in presence of increasing concentrations of unlabeled probe containing C, 5mC (M) or 5hmC (H). Concentrations represented as "fold of the amount of radioactive probe".
B. Same as A, but with 1 pmol of MBD1 instead of MeCP2 **C.** Same as A, but with 0.4 pmol of MBD2 instead of MeCP2.

II.III. RTT mutation R133C preferentially inhibits binding to 5hmC

If binding of MeCP2 to 5hmC is critical for its role in the regulation of neuronal nuclear function and gene expression, then it is possible that a subset of the MeCP2 mutations that cause RTT disrupt 5hmC binding without strongly impacting 5mC interaction. To determine if this is the case, binding of MeCP2 MBDs (aa 1-205) carrying a variety of previously characterized RTT mutations were assayed (Kudo, Nomura et al. 2003). Here we focused on residues that (1) do not alter the MeCP2 binding and nuclear localization significantly, (2) located in the DNA binding pocket, and (3) show atypical or mild phenotypes in RTT database (http://mecp2.chw.edu.au/mecp2/). We also included a newly identified mutation S134F (Lima, Brunoni et al. 2009). To represent two extreme cases of DNA binding activity three mutations were selected: D121G, that abolishes 5mC binding, and L100V and A140V, that don't directly interact with DNA, maintain wild type affinity to 5mC and show milder phenotypes (Orrico, Lam et al. 2000; Couvert, Bienvenu et al. 2001; Jentarra, Olfers et al. 2010). The rest of the disease-causing mutations in the MBD were chosen because they showed no or little disruption of nuclear localization or 5mC-binding. We produced these in *E. coli* and purified as previously described (Fig 2.5.A). Although the general effect of mutations in this series was to inhibit binding to both 5mC and 5hmC DNAs to a similar degree or to remain unchanged, we observed a pronounced decrease in the interaction with 5hmC relative to 5mC DNA with the MeCP2 MBD carrying the R133C substitution (Fig 2.5.B).

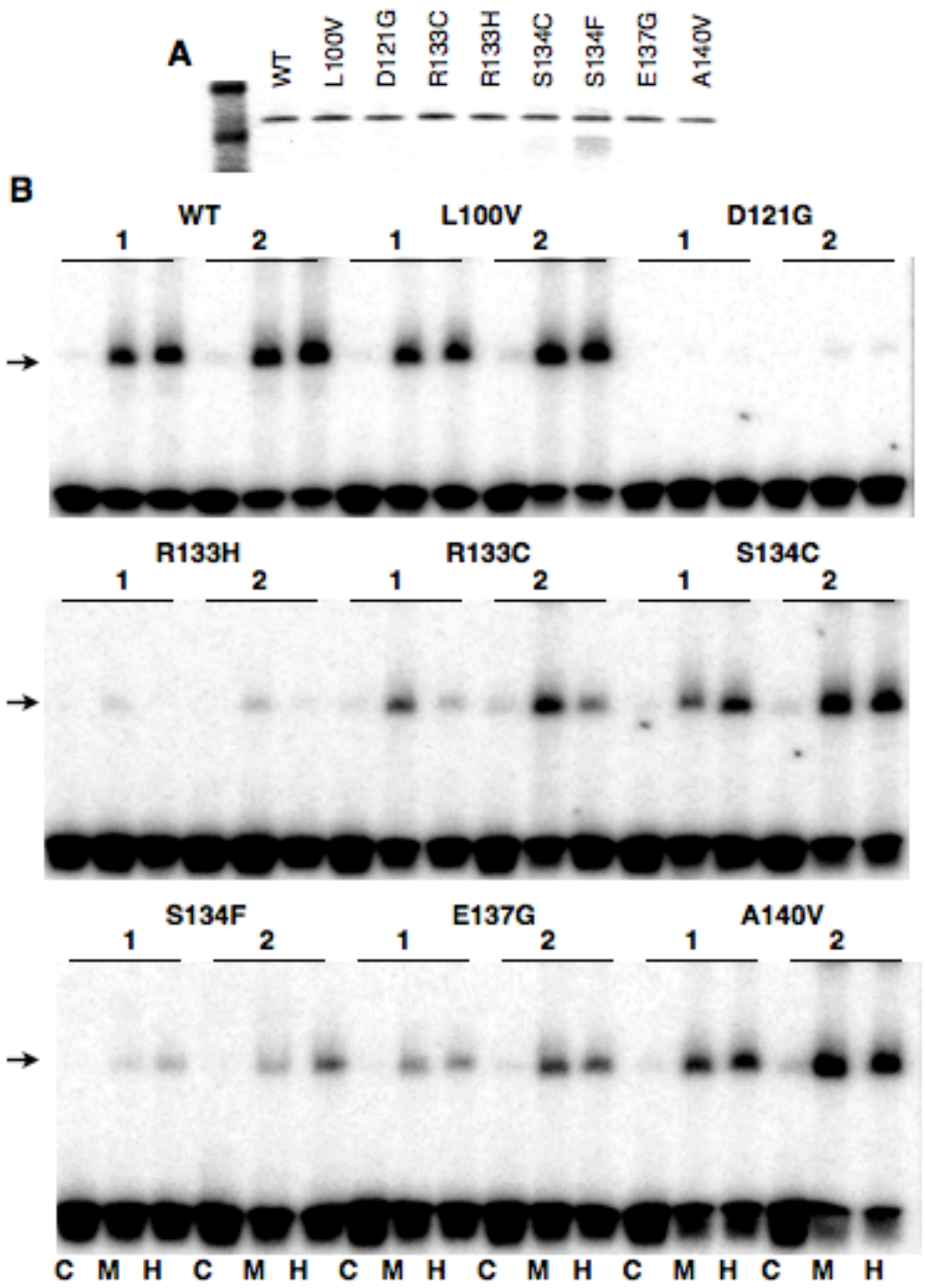


Fig 2.5. Binding of MeCP2 with RTT missense mutations to 5mC and 5hmC. **A.** Coomassie stained SDS PAGE showing purified recombinant MeCP2 with mutations indicated. **B.** EMSA gel with 5 fmoles of C-, M- and H-containing DNA probes and 1 and 2 pmols of MeCP2 NT, WT or with indicated point mutations. Arrows point at protein-dependent DNA complexes.

Since R133C and R133H are mutations at the same location, yet give different affinities for modified probes, next we wondered if it would be possible to “improve” this differential binding by designing artificial mutants that bound 5mC as efficiently as WT and showed nonspecific binding to both 5hmC and C. Hence we designed R133 mutants by replacing the positively charged R residue with another positively charged residue of smaller size (R133K), a negatively charged residue (R133E), a hydrophobic residue (R133M) and a hydrophilic residue (R133S). The binding of most of these mutants to 5mC was not affected yet their binding to 5hmC was reduced (Fig 2.6.A). This discrepancy was pronounced with residues that contain a negatively charged atom such as sulfur (in R133C and R133M) and oxygen (in R133E and R133S). We hypothesize that this is due to their vicinity to the oxygen of the hydroxyl group of 5hmC. The binding of mutants containing R133H and R133K to both 5mC and 5hmC was strongly reduced (Fig 2.6.B). To test whether this effect is also conserved in the full-length (FL) proteins, we purified FL WT MeCP2, the RTT mutation R133C and two of the artificial mutants R133E and R133S. Since the FL protein cannot be separated in EMSA gel, we separated these proteins on SDS page (Fig 2.6.C), transferred them onto a charged membrane, and analyzed it by Southwestern method as before. The membrane-bound re-natured proteins were probed with radioactive DNA probes containing either 5mC or 5hmC (Fig 2.6.D). We normalized the signal of both membranes to the signal from WT NT of MeCP2 and detected no binding activity of R133C and R133E mutants to 5hmC containing DNA, whereas R133S showed weak binding to 5hmC.

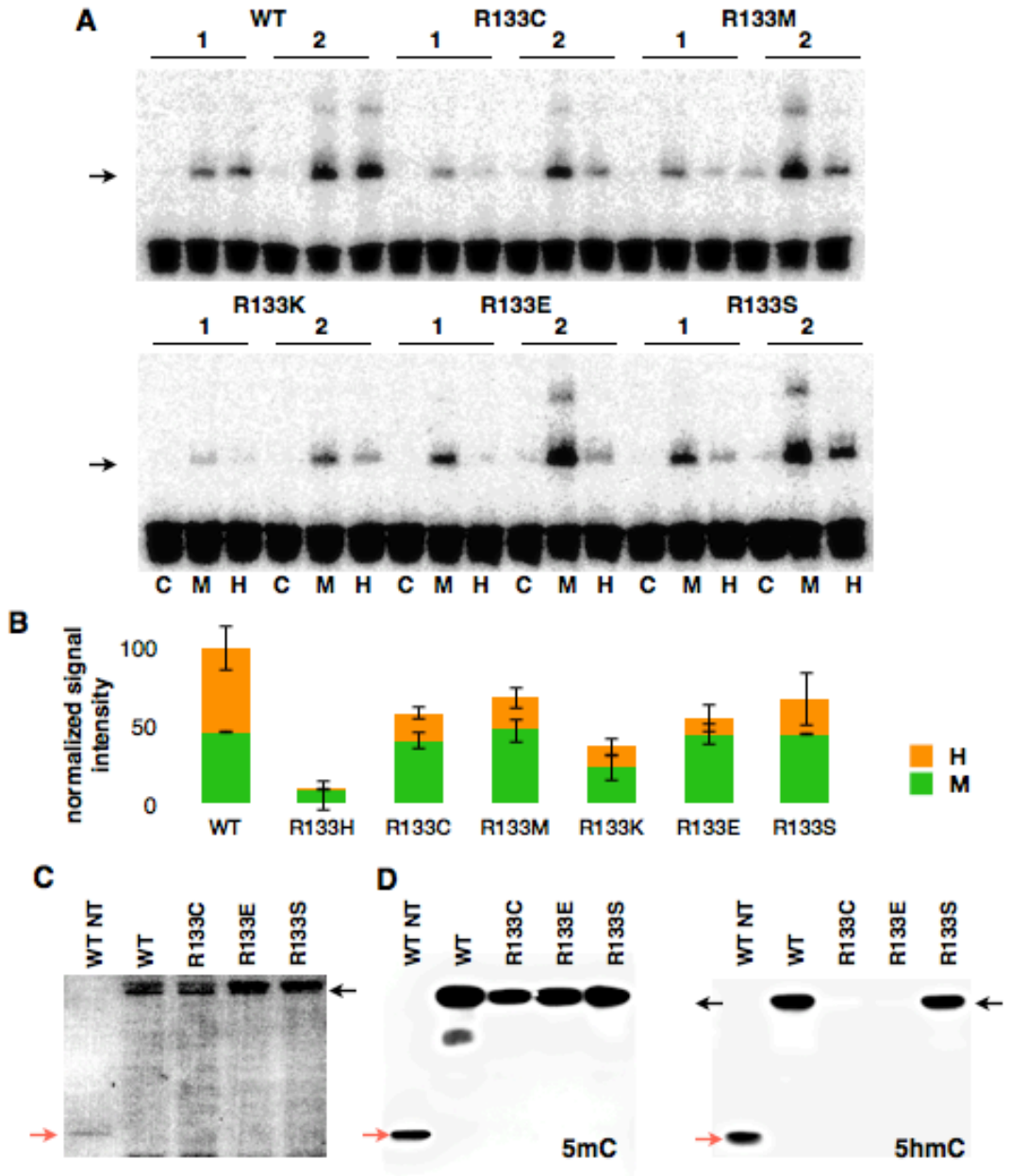


Fig 2.6. Affect of mutations at R133 to its binding to 5hmC. **A.** EMSA gel with 5 fmoles of C-, M- and H-containing DNA probe and 1 and 2 pmols of MeCP2 NT, WT or with indicated point mutations. Arrows point at protein-dependent DNA complexes. **B.** Quantification of EMSA signal intensities from 5mC and 5hmC in A and Fig 2.5.B. Standard error of mean is shown (n=3) **C.** SDS PAGE of 8 pmoles of full-length MeCP2 proteins WT or with indicated point mutations. Stained with Coomassie. **D.** Membranes incubated with radioactive DNA probes containing 5mC (left) or 5hmC (right) with 20 pmoles of full-length MeCP2, WT or mutant. In B and C, red arrows mark MeCP2 NT and black arrows mark MeCP2 FL.

II.IV. Binding of MBD proteins to C, 5mC and 5hmC

To investigate these findings in more detail, and to provide independent data supporting the conclusions of the EMSA assays presented above, surface plasmon resonance (SPR) assays were used to measure the binding of full length MeCP2, the MeCP2 MBD, other MBD proteins, and the MeCP2 carrying R133C mutation to 5hmC, 5mC and C containing DNA (Malmqvist 1999). 5'-biotinylated DNA probes prepared using C, 5mC or 5hmC nucleotides were immobilized on parallel flow cells (Fc) of a streptavidin-coated sensor chip to a level of 500 (±25) resonance units (RU). Several dilutions of purified recombinant proteins were then introduced to each Fc in parallel and the change in SPR response over time was recorded. By visual inspection of overlay plots of such response by MeCP2 NT and MBD2, the specific binding of MeCP2 to both 5mC and 5hmC and the specific binding of MBD2 to 5mC were readily observed (Fig. 2.7.A). Next, we plotted the SPR response of these proteins at saturation against the corresponding protein concentration, to represent the steady-state binding (Fig 2.7.B). As predicted, MeCP2 NT showed specific binding to both 5mC and 5hmC containing DNA that was strongly dependent on protein concentration, whereas binding to C-containing DNA plateaued at very low protein concentrations, consistent with nonspecific binding. In contrast, MBD2 bound strongly to 5mC-containing DNA and did not bind to DNA containing C or 5hmC. We then expanded these assays to other MBDs, FL MeCP2 and FL MeCP2 containing the R133C point mutation. In SPR experiments MBD3 did not show specific binding to any probe (data not shown). As expected, MBD1 and 4

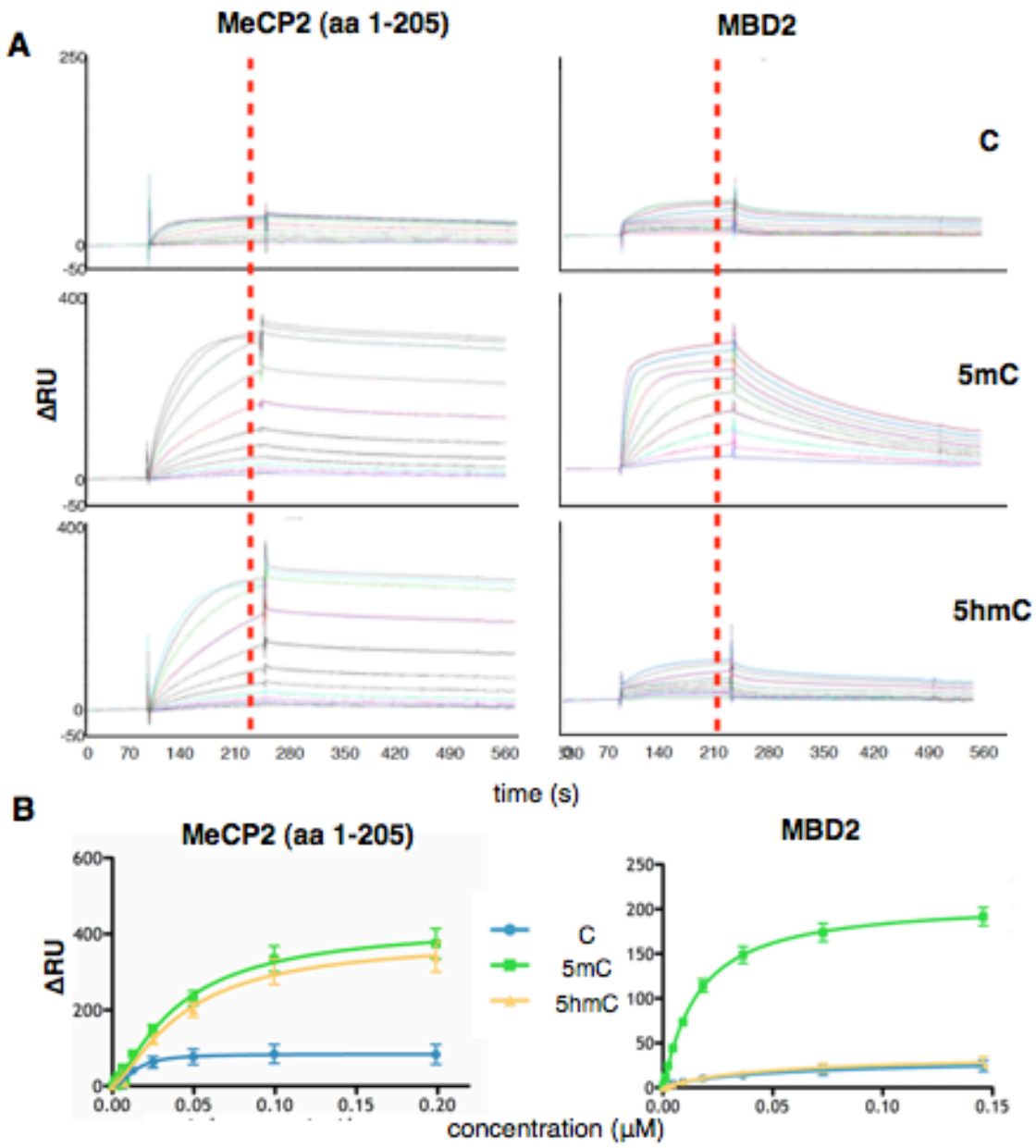


Fig 2.7. The SPR analysis of MeCP2 NT and MBD2. **A.** The dynamic SPR response of flow cell surfaces coupled with biotinylated probe containing C, 5mC or 5hmC. Overlay plots showing the change in SPR response over time by application of several dilutions of MeCP2 (left) or MBD2 (right). The dotted red line indicates the point of saturation binding. **B.** The steady-state binding curves. The mean SPR response at saturation binding from A. is plotted against protein concentration, averaged over 8 independent experiments with standard error of mean is shown.

showed specific binding only to 5mC DNA (Fig 2.8.A). As predicted, the binding characteristics of FL MeCP2 was as NT (Fig 2.8.A), consistent with the pull down experiments, the Southwestern results and the EMSA data presented above. Interestingly, binding of the MeCP2 R133C mutant to 5hmC was very strongly depressed relative to binding to 5mC DNA, although a small effect on overall binding to 5mC was evident. To further assess these results, the Bmax of proteins binding to each probe were calculated from steady-state binding curves to generate quantitative binding data for each protein (Fig 2.8.B). No significant difference was observed in the measured Bmax of MeCP2 binding to 5mC and 5hmC. The most interesting and unexpected outcome of these calculations is that the R133C MeCP2 mutant retained most of its 5mC binding capability (Bmax = 76% of WT, p=0.77) despite loss of specific binding to 5hmC (Bmax = 25% of WT, p = 0.0029). The fact that this single substitution in the MeCP2 MBD can strongly and preferentially impact the substrate binding properties of MeCP2 is important because identification of MeCP2 mutations that retain WT 5mC binding, in the R133C variant yet retain severely diminished 5hmC binding can provide an important avenue for assessing the role of MeCP2 binding to 5hmC in the pathophysiology of RTT. Furthermore, these data demonstrate that small changes in the structure of MeCP2 may influence its relative binding properties to 5mC and 5hmC, raising the interesting possibility that the posttranslational modifications to MeCP2 that have been shown to occur in response to a variety of stimuli (Chen, Chang et al. 2003; Tao, Hu et al. 2009; Rutlin and Nelson 2011) could alter its substrate specificity and downstream functions.

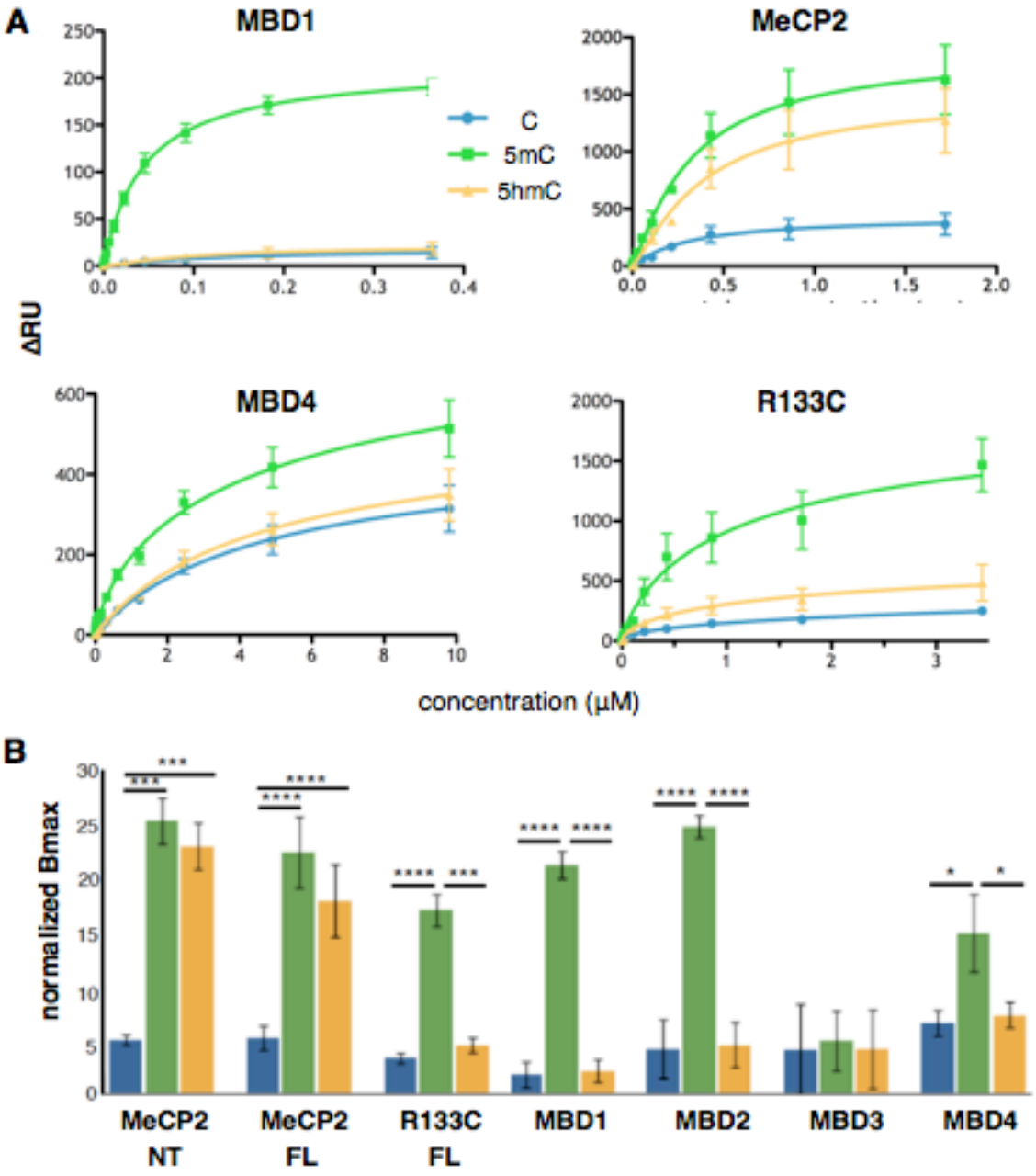


Fig 2.8. The steady state analysis of MBD proteins. **A.** The steady-state SPR response of indicated proteins, as Fig 2.7.B. **B.** The mean maximum binding (Bmax) of each MBD to DNA containing C, 5mC or 5hmC is extracted. Standard error of mean is shown over 4-8 independent experiments, **** p-val< 0.0001; *** p-val<0.001; ** p-val<0.01; * p-val< 0.05 in F test to compare variances.

II.V. MeCP2 binds to 5hmC and 5mC in CpA context

The identification of MeCP2 as a major 5hmC binding protein in rodent brain is surprising given previous in vitro studies reporting that it binds 5mC-containing DNA much more avidly than 5hmC-containing DNA (Valinluck, Tsai et al. 2004; Bostick, Kim et al. 2007). The most notable difference between previous studies and ours was the preparation of the DNA: Other studies utilized probes prepared by dimerization of chemically synthesized short oligonucleotides with a single modification in CpG context, where we amplified probes from a native genomic sequence using dCTP, d5mTP or d5hmTP. The resulting probe is a 120 bp probe that is densely modified (as in some regions of the genome) where the modified residues can exist in every possible dinucleotide context. To address this issue directly, we prepared probes that were modified on a single C in all dinucleotide and trinucleotide contexts. In order not to compromise the DNA binding of MeCP2 we used long, 75 bp, probes rich in A/T bases near the modification. To our surprise, MeCP2 failed to bind to the 5hmCpG dinucleotide, in agreement with the earlier studies, yet its binding was nearly as strong for the 5mCAC trinucleotide as well as 5hmCAC (Fig 2.9.A). It bound with similar affinity to both 5mCAT and 5hmCAT trinucleotides. We observed considerably low binding with 5mCGA and 5mCAA trinucleotides, but binding was abolished in 5hmCGA and 5hmCAA. MeCP2 also did not bind any CpC or CpT nucleotides or any of the hybrid CpGs, where the modification status of the two strands differed (data not shown). In addition, R133C mutant conserved the binding to 5mCpG and 5mCpA but lost its affinity to hydroxymethylated CpAs (5hmCpAs), indicating

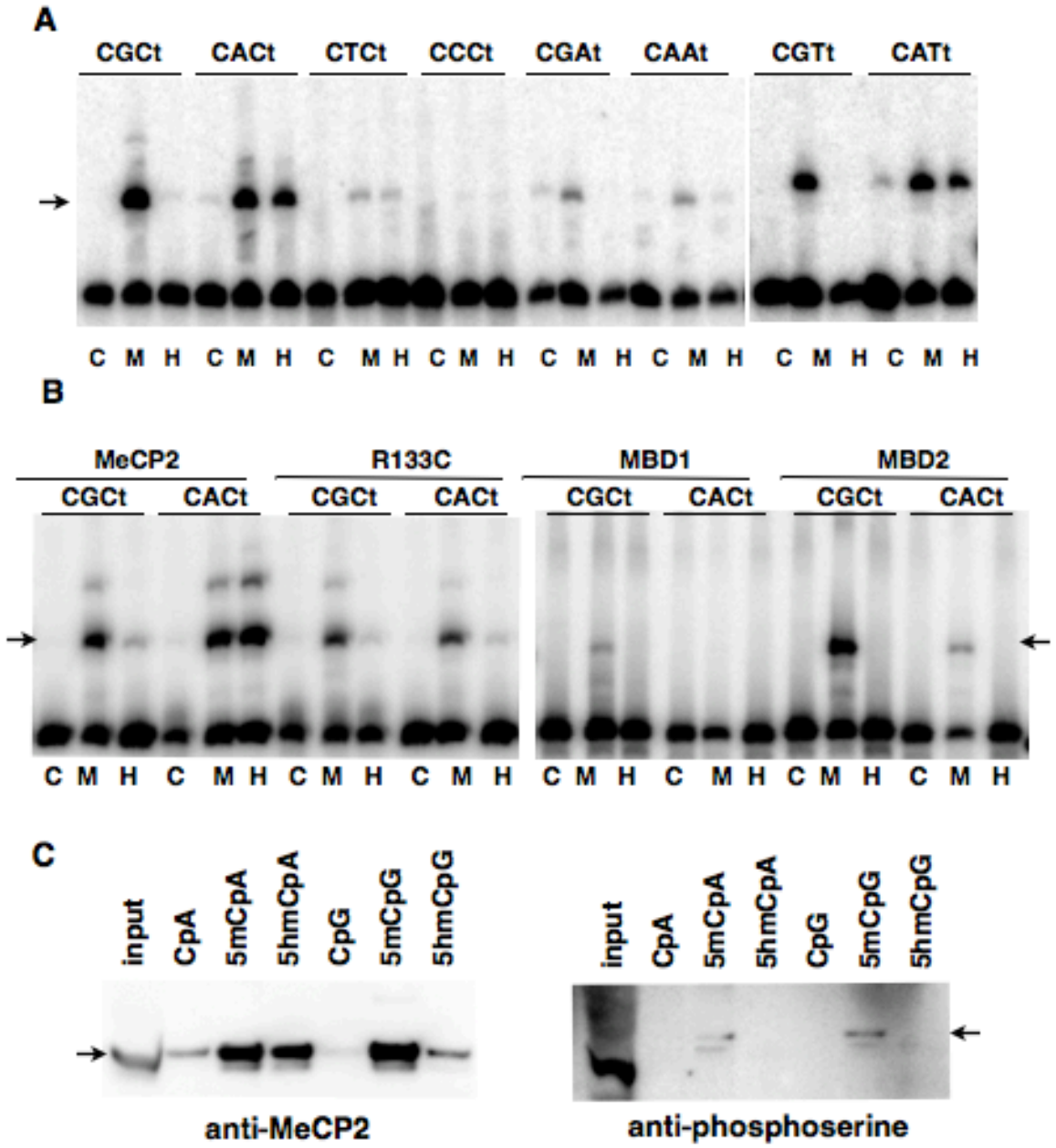
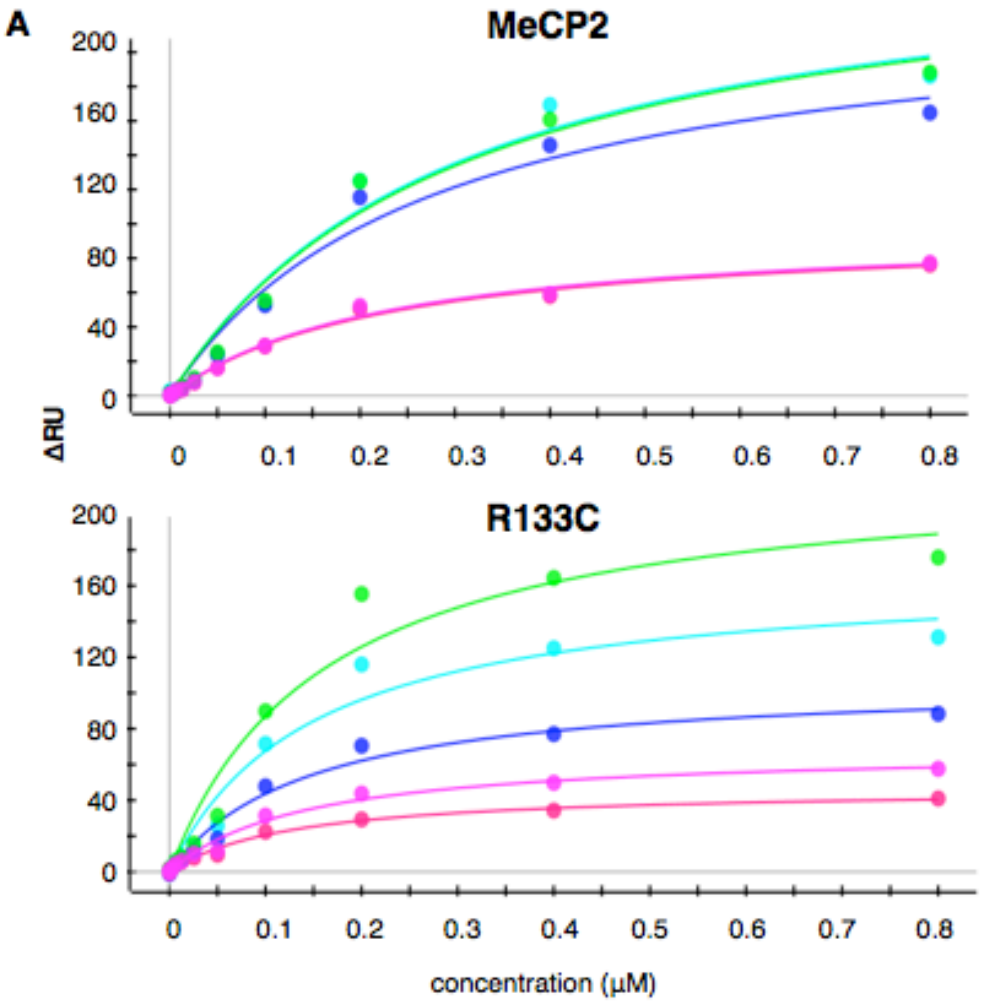


Fig 2.9. MeCP2 binds 5mC and 5hmC within CpA context. **A.** EMSA gel in presence of 1 pmol of MeCP2 NT with 75 bp probes containing single C, M and H within indicated trinucleotide context. Arrows point at protein-dependent DNA complexes. **B.** EMSA gel as in A. with MeCP2 NT WT and R133C mutant, 1 pmol of MBD1 and 0.5 pmol of MBD2. **C.** Western blot with anti-MeCP2 (left) and anti-phosphoserine antibody on membrane containing input and eluates from pulldowns from nuclear extract from rodent brain using probes containing indicated modified dinucleotides. Arrows point at specific bands.

that R133 residue is crucial for 5hmC recognition but not for the recognition of the subsequent pyrimidine (Fig 2.9.B). As expected, both MBD1 and MBD2 did not bind either DNA containing C or 5hmC, or 5mC in CpA context. MBD2 bound DNA with a single 5mCpG strongly; however, MBD1 showed low binding. The identity of the fourth nucleotide did not affect binding (data not shown).

We next performed pulldowns in nuclear extract from rodent brain using DNA baits carrying a single modification in CpA or CpG context, enriched MeCP2 in DNA-bound eluates by immunoprecipitation and confirmed our binding data by Western blot of the euates using an antibody against MeCP2 (Fig 2.9.D, left). Interestingly, when we used an anti-phosphoserine antibody, we saw a band specific for proteins eluted from the methylated DNA bait (Fig 2.9.D, right). However, we cannot currently conclude that the source of this band is MeCP2.

To analyze this binding more in detail, we conducted a new generation SPR analysis (Abdiche, Lindquist et al. 2011), where we immobilized biotinylated DNA probes containing a single CGCT, 5mCGCT, 5hmCGCT, 5mCACT and 5hmCACT on a neutravidin-coated array to a level of 550 (±40) RU and measured the response by several dilutions of WT MeCP2 NT and R133C in parallel. The equilibrium analysis was performed using ProteOn Software. As expected, MeCP2 showed specific binding to 5mCpG, 5mCpA and 5hmCpA (Fig 2.10.A). Interestingly, the binding of R133C mutant to 5mCpG was at WT levels, yet its binding to 5mCpA was reduced. We observed more significant effect on its 5hmCpA binding. To verify our observations, we extracted Bmax values using ProteOn Software for each protein and DNA modification (Fig 2.10.B).



B

	MeCP2	R133C
CpG	97.27	47.00
5mCpA	273.67	167.66
5hmCpA	233.34	107.60
5mCpG	272.62	226.94
5hmCpG	98.36	68.76

Fig 2.10. The binding of WT MeCP2 and R133C to 5(h)mCpG and 5(h) mCpA. **A.** The equilibrium SPR response of indicated proteins to surface-immobilized DNA containing single 5mCpG, 5hmCpG, 5mCpA or 5hmCpA, as measured by ProteOn protein interaction array system. **B.** Bmax of WT MeCP2 and R133C for each DNA modification as extracted by ProteOn Software.

These results showed that MeCP2 preferentially binds a methylated site within the context Cp(A/G)p(T/C) > Cp(A/G)p(A/G), whereas hydroxymethylated target sites are within the context CpAp(T/C). Since non-CpG modification, primarily CpA(C), makes up 20-25% of the modified cytosines (Laurent, Wong et al. 2010), it is likely that a substantial population of MeCP2 is bound on modified CpAs *in vivo*. Additionally, modified non-CpG sites are are preceded by a TA dinucleotide upstream and followed by an A or T (Lister, Pelizzola et al. 2009). Such pattern may be stabilizing the binding by MeCP2 which possesses two consensus A/T-hook motifs of MeCP2 (Ho, McNae et al. 2008). Given that modified CpAs are concentrated in the bodies of active genes and strongly correlate with gene expression (Lister, Pelizzola et al. 2009), and that genes that are highly expressed lack methylation but are largely hydroxymethylated (Mellen, Ayata et al. 2012), one can expect that such genes have high levels of hydroxymethylated CpAs. Although single nucleotide data is not yet available, asymmetrical hydroxymethylation is observed at high levels throughout the transcription unit of active genes, and such regions contain extensive amounts of modified non-CpG dinucleotides (Ficz, Branco et al. 2011). If a considerable fraction of MeCP2 target sites are hydroxymethylated CpAs enriched in the bodies of active genes within euchromatin, we hypothesize that these sites may be bound *in vivo* by the previously identified subpopulation of MeCP2, that is loosely bound and associated with active states of genes and chromatin (Thambirajah, Ng et al. 2011).

II.VI. MeCP2 facilitates chromatin accessibility around 5hmC

Given that MeCP2 can not only tightly bind densely packed regions of silent chromatin, enriched in 5mC, but also is loosely associated with accessible regions of chromatin, which are where active genes, 5hmC and modified CpAs are extensively found, we next were interested in assessing its potential role in global regulation of chromatin accessibility. To do so, cerebellar nuclei were isolated from five-week-old WT and KO male mice (Guy, Hendrich et al. 2001). For each sample, a time course of MNase digestion was performed, and the release of 5hmC- and 5mC-enriched DNA fragments prepared by Southern blotting and assayed with antibodies against 5mC and 5hmC (Fig 2.11.A). To quantitate and normalize the data from different digestions, the signal from quadrant 1 (Q1) to Q4 in each lane was measured in four independent cohorts of WT and KO mice, and the data quantified as the percentage of total signal in each time of digestion (Fig 2.11.B). We denoted the signal present in the Q1 fraction as nuclease-resistant fraction and plotted that against the early digestion times (Fig 2.11.C). Two interesting results were obtained: First, we observed that 5hmC-enriched DNA is released readily from chromatin by MNase digestion, whereas 5mC-containing chromatin is significantly more resistant to digestion. This is consistent with the analysis of individual genes, and confirms previous studies demonstrating the 5mC enriched DNA is present in MNase-resistant compact structures (Karymov, Tomschik et al. 2001). Second, in *Mecp2*-null mice a significant, small delay in digestion of 5hmC-containing DNA was observed, whereas no reproducible difference in the sensitivity of 5mC-containing DNA to

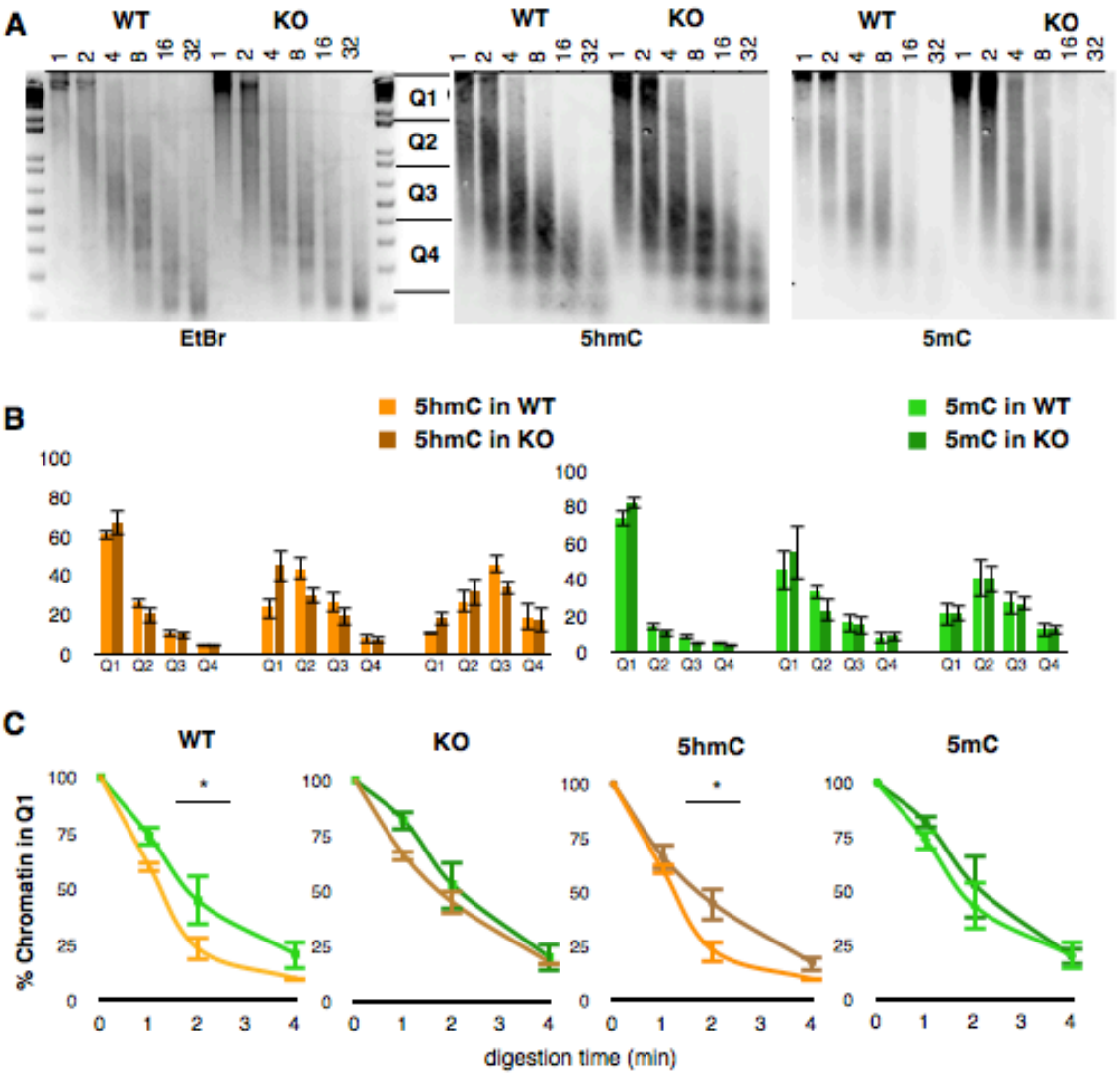


Fig 2.11. MeCP2 facilitates chromatin accessibility around 5hmC. **A.** EtBr stained agarose gel containing DNA of MNase digested nuclei from WT or KO animals (left). Same sample was assayed with anti-5hmC antibody (middle), and then anti 5-mC antibody (right) after Southern transfer. Digestion time shown in minutes. **B.** The signal of each lane in A. measured in quadrants Q1-4 and normalized to total signal in lane. Mean over 4 independent experiments was calculated with standard error of mean. **C.** The Q1 values from B plotted. with standard error of mean and * for p-val< 0.05 in unpaired t-test.

MNase was evident. These data demonstrate that MeCP2 regulates the accessibility of 5hmC-containing DNA to MNase, supporting a model in which MeCP2 binding to 5hmC within highly expressed genes may facilitate transcription through its effects on chromatin organization.

II.VII. Discussion

The data presented here identify a novel role for MeCP2 in the regulation of chromatin structure in support of a new model for the organization of accessible chromatin states around expressed genes that is specific to the vertebrate nervous system, in addition to the traditional repressive and silencing functions it elicits upon its binding to 5mCpG dinucleotides (Guy, Hendrich et al. 2001). Based on the previous literature and our findings, we propose that binding of 5hmCpA by MeCP2 plays a central role in the fine tuning of chromatin states that facilitate expression of neural genes. The mechanism by which MeCP2 binding to 5hmCpA regulates chromatin accessibility is evidently different from its repressive role within heterochromatin regions enriched in 5mC and remains to be deciphered. However this model, based on our binding data, is substantiated by the fact that both 5hmC and MeCP2 are at least an order of magnitude more abundant in the nervous system than in the periphery (Kriaucionis and Heintz 2009; Skene, Illingworth et al. 2010); that both modified CpA dinucleotides and 5hmC are abundant on bodies of active neural genes and euchromatin; that these regions are depleted of 5mC (Lister, Pelizzola et al. 2009; Mellen, Ayata et

al. 2012); and that MeCP2 has a yet uncharacterized association with accessible chromatin states (Thambirajah, Ng et al. 2011).

- *5hmCpA, a new epigenetic code in the brain*

The accumulation of modified CpA throughout the transcription unit of active genes (Lister, Pelizzola et al. 2009; Ficz, Branco et al. 2011), combined with the signature depletion of 5mC and accumulation of 5hmC throughout bodies of highly expressed genes (Mellen, Ayata et al. 2012), imply a new depiction of epigenetic information in the form of 5hmCpA in correlation with gene expression. This differs significantly from the traditional epigenetic language, where 5mCpGs compact chromatin into repressive states via MBD and other 5mC-binding proteins (Guy, Hendrich et al. 2001; Yildirim, Li et al. 2011). Since CpA modification is inherently “asymmetric”, it is destined to be lost over cell divisions. Hence in post-mitotic cells, *de novo* mechanisms are necessary to establish such code that potentially define cell identity and function. This is crucial for a complex network, like the brain, where each cell fulfills a slightly or fundamentally different function from its sister cell. Indeed, non-CpG modification of transcription units of highly expressed neuronal genes has been previously reported (Backdahl, Herberth et al. 2009; Cortese, Lewin et al. 2011). It is likely that neuronal genes can be methylated at CpA dinucleotides by Dnmt3a in an activity-dependent manner; given its neuron-specific euchromatic localization (Wu, Coskun et al. 2010); its ability to methylate CpAs *in vivo* (Mund, Musch et al. 2004); and its activity-dependent regulation in neurons (Feng, Chang et al. 2005; Miller and Sweatt 2007; Nguyen, Meletis et al. 2007; LaPlant, Vialou et al. 2010). Moreover,

the *de novo* DNA methyltransferase, Dnmt3b, is not detected in postmitotic tissue (Okano, Bell et al. 1999). Although not much is known about Tet3, we suspect that its functions may encompass the hydroxylation of 5mCpAs, since all Tet enzymes can hydroxylate methylated CpAs *in vitro* (Tahiliani, M, personal communication); Tet3 is highly expressed in the brain (Szwagierczak, Bultmann et al. 2010); and in the absence of Tet1 and Tet2, asymmetric non-CpG hydroxymethylation increases (Ficz, Branco et al. 2011). Together these data suggest that Dnmt3 and Tet3, or other putative enzymes, may be generating high amounts of 5hmCpA. In the brain, both 5hmC and modified CpAs share a similar pattern: both are depleted in TSS, but enriched the bodies of highly expressed genes within euchromatin (Lister, Pelizzola et al. 2009; Mellen, Ayata et al. 2012), therefore we expect that a large proportion of these modified CpAs to be hydroxymethylated. This is strongly supported by the finding that these genes are invariably depleted from 5mC.

- *MeCP2, the reader of the new language*

In this study, we have shown that MeCP2 binds 5hmCpA *in vitro*. To appreciate this binding at the molecular level, we have modified the existing structural information (Wakefield, Smith et al. 1999) using Pymol software. The original structure (Fig 2.12.A) conveys the hydrophobic stabilization of the methyl group of 5mC by the arginine chain, as well as the ionic interaction between the negative oxygen of guanine (G) and the amino group of arginine (R). When the methyl group is oxidized in 5hmC (Fig 2.12.B), then two highly negative oxygen molecules in 5hmC and G make the interaction unfavorable, whereas the

replacement of G by adenine (A) restores the ionic balance. The molecular representation of a cysteine residue replacing the arginine (R133C) also provides mechanistic insight to the preferential inhibition of 5hmC binding of MeCP2. In a hypothetical model, where R133C MeCP2 is bound to 5hmCpA (Fig 2.12.D), the nucleophilic sulfur of cysteine would be adjacent to the oxygen of 5hmC. That would create an energetically unfavorable state and thus binding would not occur. On the other hand, in the 5mCpG-bound R133C MeCP2, the oxygen of G and the sulfur atom of cysteine are separated where water molecules may be accommodated and stabilize 5mCpG in the binding pocket (Fig 2.12.C). Same effect can be observed in some of the other R133 mutants we have created, where the distance between the oxygen of 5hmC and an other nucleophilic atom, such as sulfur in R133M (Fig 2.13.B) or oxygen in R133E or R133S (Fig 2.13.D and E), is too small to be energetically favored. In the case of R133K and R133H, such strong repulsion is not the case and accordingly, the discrepancy between 5mC and 5hmC binding is smaller. However, the 5mC binding is more strongly reduced. This may be due to steric hinderance and/or displacement of water molecules that stabilize the binding pocket (Ho, McNae et al. 2008).

Given our binding data; and that both 5hmC and MeCP2 are at least an order of magnitude more abundant in the nervous system than in the periphery (Kriaucionis and Heintz 2009; Skene, Illingworth et al. 2010); we hypothesize that MeCP2 binding to 5hmCpA is crucial in the decryption of the new neuronal epigenetic code, which we proposed earlier. This code is not accessible to other MBD proteins, as they neither recognize 5hmC nor modified CpAs. On the other

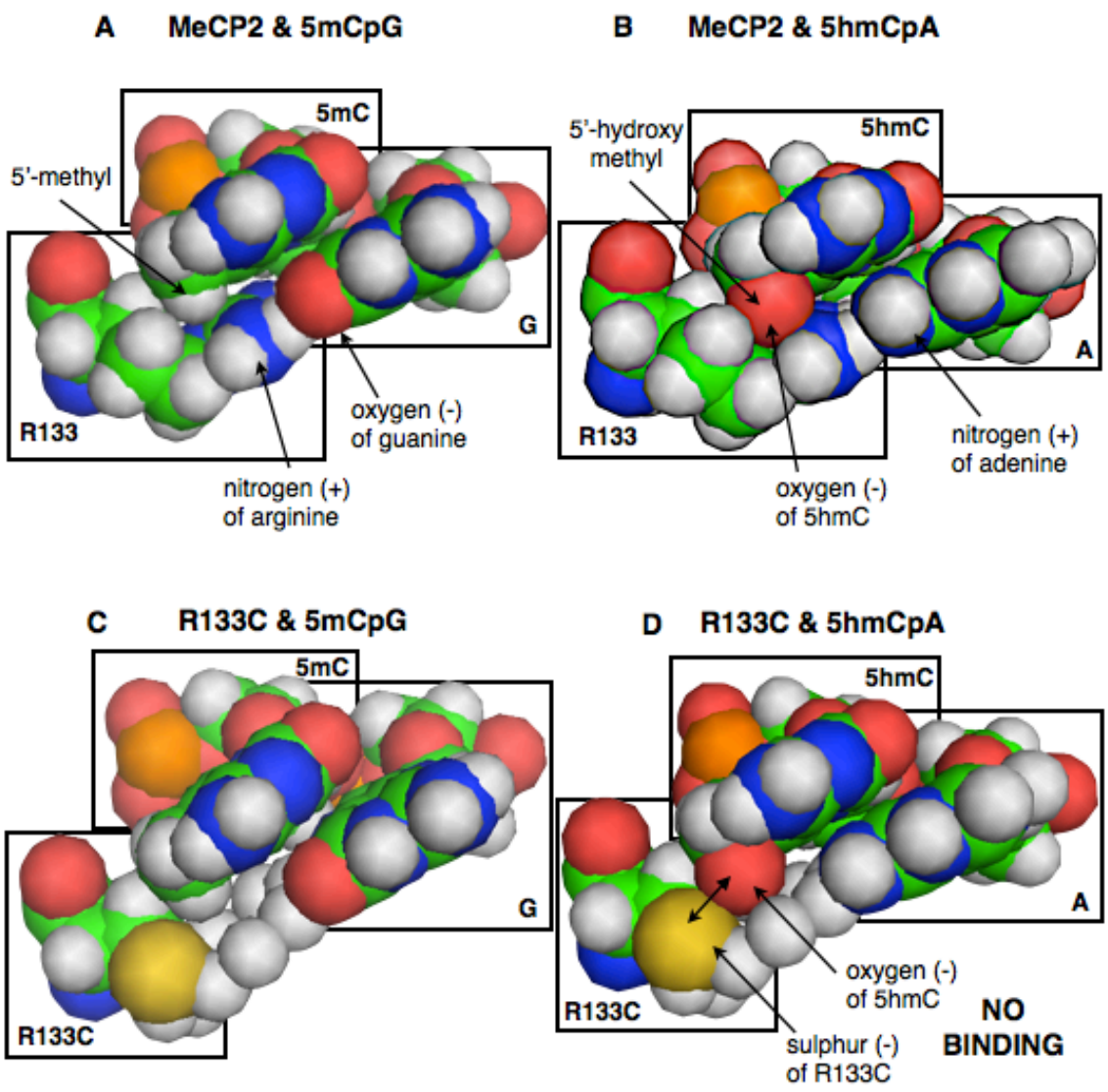


Fig 2.12. Molecular representation of R133 of WT MeCP2 and R133C with 5mCpG and 5hmCpA. **A.** Sphere representation of R133 from WT MeCP2 MBD with 5mCpG as previously reported (Wakefield, Smith et al. 1999). **B.** WT MeCP2 with 5hmCpA, as modified from A. **C.** R133C MeCP2 with 5mCpG, as modified from A. **D.** R133C MeCP2 with 5hmCpA, as modified from B. **From A-D**: Carbon (green), Hydrogen (grey), Nitrogen (blue), Oxygen (red), Phosphorous (orange), Sulphur (yellow), Adenine (A), Guanine (G), Cytosine (C), Arginine133 to Cysteine (R133C)

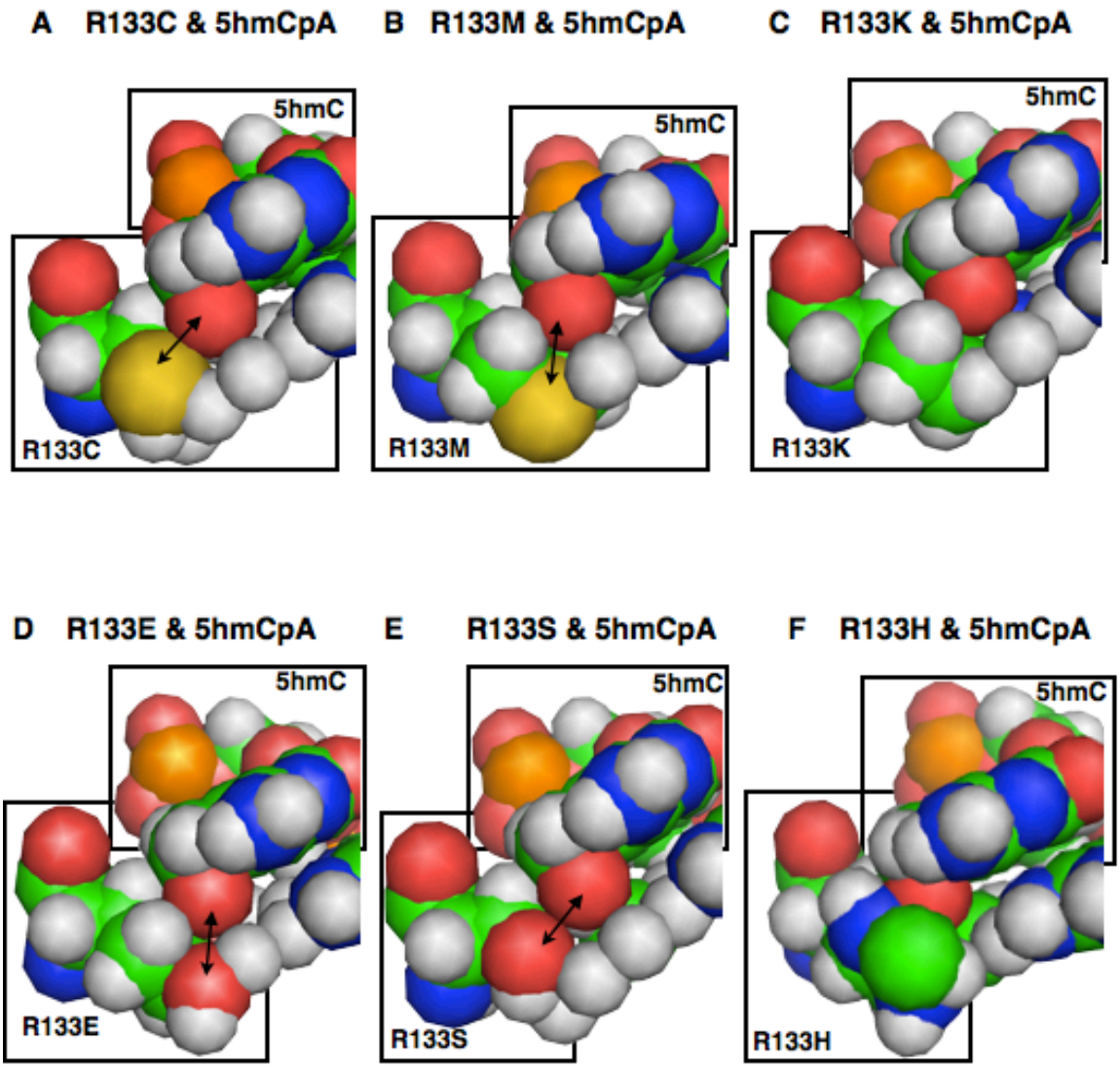


Fig 2.13. Molecular representation of R133 mutants with 5hmC. **From A-D**: Structures modified from the original X-ray crystal structure previously reported (Wakefield, Smith et al. 1999) using Pymol (www.pymol.org). Carbon (green), Hydrogen (grey), Nitrogen (blue), Oxygen (red), Phosphorous (orange), Sulphur (yellow), Adenine (A), Guanine (G), Cytosine (C), Arginine133 to Cysteine (R133C). Arrows show the repulsion between adjacent negatively charged atoms.

hand, 5hmCpA is likely to be occupied by MeCP2 *in vivo*, not only because MeCP2 avidly binds 5hmCpA *in vitro*, but also because MeCP2 contains two consensus A/T-hook motifs (Ho, McNae et al. 2008), that may be stabilized by the TA dinucleotide upstream of modified non-CpG and the A/T downstream (Lister, Pelizzola et al. 2009). This suggests that the interaction interphase of MeCP2 with 5hmCpA may be different than stable complexes established around 5mCpG by binding of MeCP2 or other less abundant MBD family proteins (Lopez-Serra, Ballestar et al. 2006), since the target site is proven to be an activating mark within accessible chromatin.

In individual cell types, the level and genomic location of 5hmC and 5mC are tightly regulated (Szulwach, Li et al. 2011; Mellen, Ayata et al. 2012) and the local protein composition associated with them varies (Lopez-Serra, Ballestar et al. 2006; Clouaire and Stancheva, 2010). Then a delicate balance between 5mC, 5hmC, MeCP2 and other MBD proteins may explain the fine tuning chromatin states that enable elaborate adjustments of gene expression patterns. As a result, changes in the function of MeCP2 (Amir, Van den Veyver et al. 1999; Tao, Hu et al. 2009; Adkins and Georgel 2011) will disrupt such balance at varying severity in each cell type, and the phenotypic consequences will be cell-type and circuit specific. In mice, after the neurodevelopment is completed in the cerebellum and MeCP2 expression reaches a plateau (Skene, Illingworth et al. 2010), ~85% of DhMRs within specific tissues are stably maintained (Szulwach, Li et al. 2011). Interestingly, these stable DhMRs are not altered in the absence of *Mecp2*, supporting the downstream role of MeCP2 in these regions. Although

the total 5hmC signal increased by <20% in the absence of *Mecp2*, this change was restricted to gene body DhMRs, leading the scientists to propose a gene-body specific role for MeCP2 (Szulwach, Li et al. 2011).

- *Bilingual MeCP2*

The exact mechanisms by which 5hmCpA is interpreted by MeCP2 into open chromatin states to facilitate gene transcription is still unclear. Binding to 5hmCpA within euchromatin and active genes, in addition to 5mCpG within heterochromatin and silent genes, juxtaposes two contradicting roles for MeCP2. Intriguingly, a similar contradiction has been presented in the earlier studies: the apparent action of MeCP2, that is more akin to a linker histone coating the whole genome (Skene, Illingworth et al. 2010) and the detection of two populations of MeCP2 in the brain, one loosely bound in highly accessible chromatin domains and the other tightly bound in heterochromatin regions that are rich in nucleosomes (Thambirajah, Ng et al. 2011). It is likely that this latter population of MeCP2 can stably associate with nucleosomes participating in methylated regions in heterochromatin (Chandler, Guschin et al. 1999). The euchromatic MeCP2 population, however, may be occupying expressed genes through its binding to 5hmCpA (Fig 2.14.A). Since the 50% of the demand for histone H1 is supplied by MeCP2 in the brain (Skene, Illingworth et al. 2010; Ghosh, Horowitz-Scherer et al. 2010), the regulation of global chromatin state in neurons might as well be largely mediated by the distribution of different populations of MeCP2 over the neuronal genome. We propose that a population of MeCP2 is loosely bound to 5hmCpA within active transcription units and facilitates chromatin

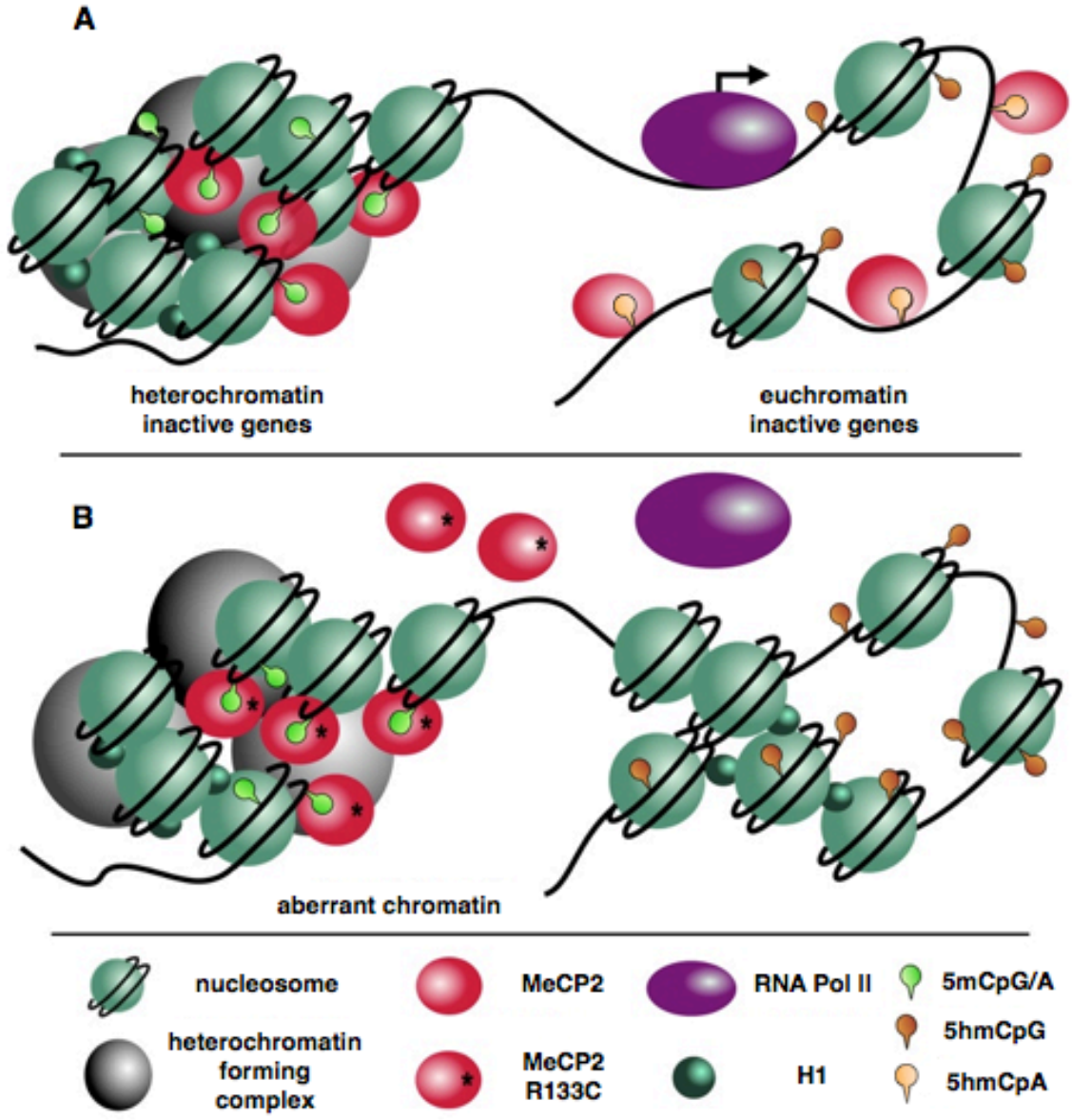


Fig 2.14. Model for MeCP2 function in neural chromatin. **A.** WT MeCP2 populations distributed in euchromatin and heterochromatin. **B.** R133C mutation causes shift in the balance of genomic localization of MeCP2

accessibility by unknown mechanisms, with other populations still being tightly associated with silent chromatin. This model provides a possible explanation for the recent demonstration that *Mecp2* gene dosage positively correlates with the expression of the majority of the genes in the brain (Chahrour, Jung et al. 2008; Ben-Shachar, Chahrour et al. 2009). If the distribution of MeCP2 over the whole genome fine tunes accessible and silent chromatin states, then a dose-dependency should be crucial. Indeed, changes in the level or activity of MeCP2 that disturbs this balance results in alterations of chromatin structure and, consequently, gene expression. This has been manifested by the disease phenotypes that arise when MeCP2 is overexpressed by only two-fold (Collins, Levenson et al. 2004) and when the activity-dependent posttranslational modifications of MeCP2 were impaired (Cohen, Gabel et al. 2011; Li, Zhong et al. 2011). An interesting strategy to observe the consequences of such balance shifts would be to manipulate 5hmC levels *in vivo* and observe how MeCP populations dislocate and how this is reflected in the phenotype.

- *R133C, lost in translation*

Our understanding of the pathophysiology of RTT must now encompass both the role of MeCP2 binding to 5mC in the silent chromatin states (Chahrour, Jung et al. 2008), and present results supporting a model in which MeCP2 binds to 5hmC within active transcription units (Fig 2.14.A). Our finding suggests that disease-causing mutation R133C displaces euchromatic MeCP2 and shifts the balance towards heterochromatin-associated MeCP2 (Fig 2.14.B). Because of the observations that the distribution of 5hmC, 5mC and their relationship to gene

expression vary depending on the cell type, the change in the distribution of MeCP2 population upon R133C mutation may present an important avenue toward understanding the biochemical mechanisms causing qualitative and quantitative aspects of RTT phenotype. It is well documented that patients carrying the R133C mutation have a milder form of RTT that is characterized by delayed-onset regression, with improved speech and motor skills (Bebbington, Anderson et al. 2008). However, for many other characteristics, including breathing abnormalities, sleep problems, mood disturbances, and epilepsy prevalence, no significant differences are evident between patients bearing R133C or other mutations (Bebbington, Anderson et al. 2008).

Although some studies presented data supporting that R133C mutation impairs the binding of 5mC by MeCP2 (Ballestar, Yusufzai et al. 2000; Ballestar and Wolffe 2001), our results and other studies contradicted this idea, showing that this mutation shows binding characteristics that are closer to WT than other mutations in the MBD (Kudo, Nomura et al. 2003; Ghosh, Horowitz-Scherer et al. 2008; Kumar, Kamboj et al. 2008; Mund, Musch, et al. 2004; Fan, Nikitina, et al. 2005). Although it is possible that a mild impairment in 5mC binding might be important for the RTT phenotypes, another attractive scenario is that the complete loss of 5hmC binding (Fig 2.12.D) is the primary cause of these latter clinical features of RTT. In this scenario, the R133C mutant of MeCP2 retains its 5mC binding ability to maintain repression of silent genes, whereas the fine tuning of accessible chromatin states via 5hmC-binding are disrupted.

- *Where is 5hmCpA?*

Although this study identifies a new role for MeCP2 bound to 5hmCpA, there are still many unanswered questions: First, we still don't know the mechanism of how the CpA is hydroxymethylated in resting or activated states in the CNS. With the recent advancement in the sequencing techniques, it is now possible to sequence hydroxymethylomes genome-wide and in base-resolution (Flusberg, Webster et al. 2010; Huang, Pastor et al. 2010; Booth, Branco et al. 2012; Yu, Hon et al. 2012). Comparison of hydroxymethylomes and methylomes within cell populations or individual cells, and between active and resting states, will introduce the epigenetic information in the brain. It is now widely accepted that a quarter of global methylation occurs within non-CpG context (Ramsahoye, Biniszkiewicz et al. 2000; Lister, Pelizzola et al. 2009), yet an increased ratio of non-CpA dinucleotides within active gene bodies is highly probable, since an enrichment of 5mCpA is observed in such regions.

- *What are multiple identities of MeCP2?*

We don't know the identity of "populations" of MeCP2. Based on recent studies we postulate that one or more posttranslational modifications might be a switch for distribution of MeCP2 in such populations (Cohen, Gabel et al. 2011). It is also probable that the binding of MeCP2 to 5hmCpA is mechanistically different from its binding to 5mC (a hypothetical rearrangement of R133 residue of MeCP2 is shown in Fig 2.15.B-C), causing a structural change in MeCP2 and its readout in downstream protein-protein interactions. On the other hand, due to the intrinsic

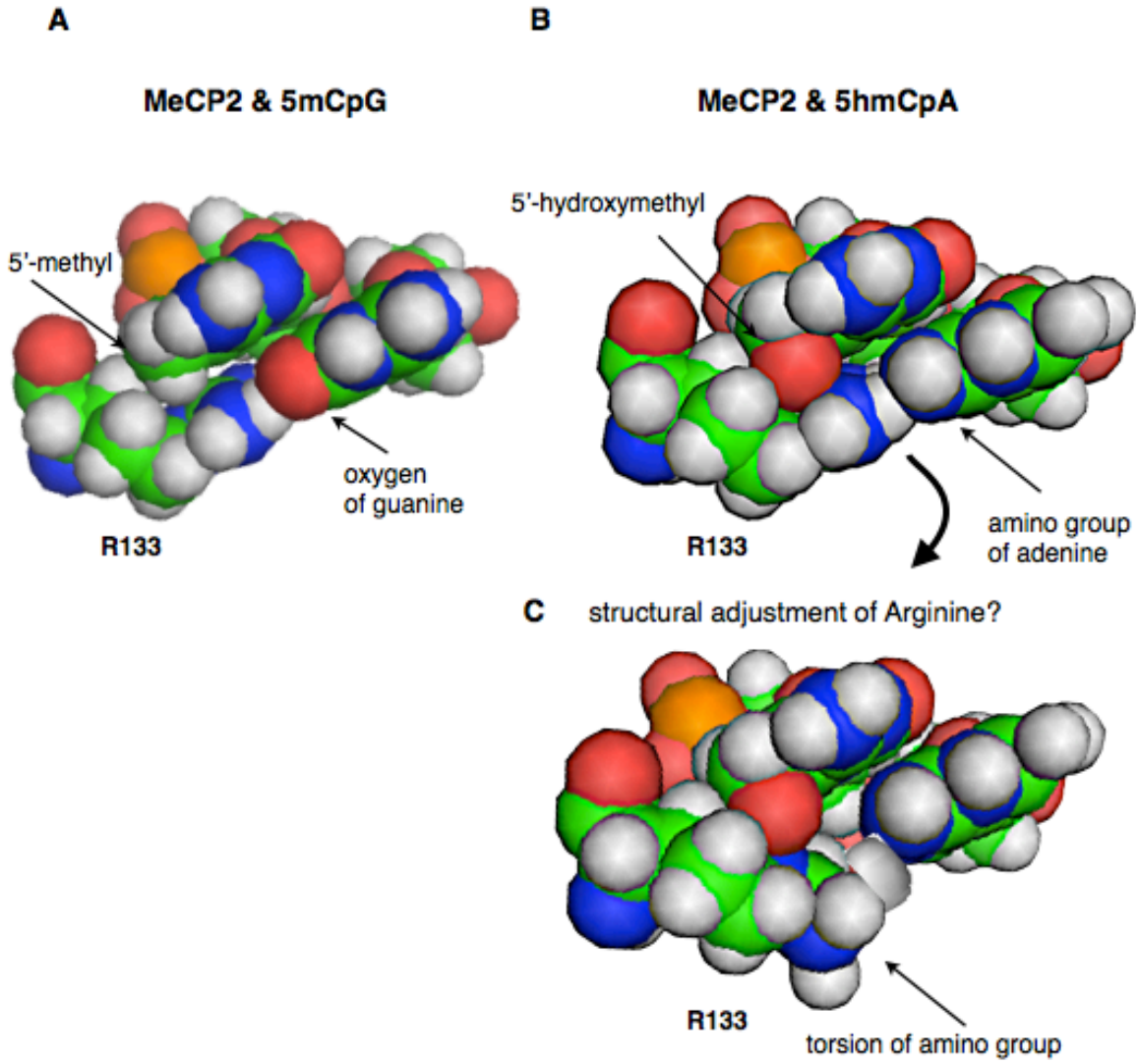


Fig 2.15. Molecular representation of a hypothetical structural change in MeCP2 by rearrangement or R133 upon binding 5hmCpA (www.pymol.org) **A.** Sphere representation of R133 from WT MeCP2 MBD with 5mCpG as reported (Wakefield, Smith et al. 1999). **B.** WT MeCP2 with 5hmCpA, modified from A. **C.** MeCP2 with 5hmCpA, modified from B, by torsion of one amino group. Carbon (green), Hydrogen (grey), Nitrogen (blue), Oxygen (red), Phosphorous (orange), Sulphur (yellow)

disorder of MeCP2, even if this kind of structural change was taking place, it may not have long range effects on the protein structure.

We do not have *in vivo* binding information of MeCP2. Since MeCP2 is a dynamically bound protein that exists in different populations that occupy different genomic regions at different times, a simple ChIP assay will bring down mixed populations of MeCP2 and, thus, will not be informative. For this kind of analysis, the identity of different MeCP2 populations is a prerequisite. Complimentary to the identification of *in vivo* binding sites of MeCP2, comparison of changes in the distribution of 5hmC within euchromatin and heterochromatin in the absence of *Mecp2* also would shed light into the contribution of MeCP2 species to the arrangement of chromatin states. This could be achieved by immunostaining 5hmC and 5mC; and comparison of the layout of the chromatin around hydroxymethylated regions in different cell types between the KO and WT mice. In previous studies, smaller and more compact nuclei were observed in *Mecp2* deficiency (Chen, Akbarian et al. 2001; Cheung, Horvath et al. 2011). We have also observed a subtle decrease in the overall nuclease sensitivity of chromatin that is more pronounced in 5hmC-rich regions.

We expect a great deal of insight into MeCP2 function to be revealed by *in vivo* studies with mice carrying R133C mutation, where the balance between MeCP2 populations is disrupted. If 5hmC-bound MeCP2 is involved in the fine regulation of the expression of certain neural genes, depending on the neuronal activity and cell type, then it is possible that 5hmC plays a role in the phenotypes that result in categorization of RTT as an Autism Spectrum Disorder.

- *What else is out there?*

Currently we do not have single-base resolution hydroxymethylome of the brain tissue. However we can expect that there still is a substantial amount of 5hmC within CpG context, which is not recognized by MeCP2. Therefore, alternative mechanisms may contribute to the decoding of 5hmC. We do not know whether 5hmC-mediated demethylation events play a role in the activity-dependent changes in the neuronal epigenome (Guo, Su et al. 2011) or if demethylation pathways are fundamentally different epigenetic mechanisms, carried out by a different subset of Tet enzymes during development. This finds support in the high level of DhMRs that are lost during neurodevelopment (Szulwach, Li et al. 2011). We also do not understand where MBD3 binding to 5hmC in ES cells (Yildirim, Li et al. 2011) fits in this model. We cannot presently answer these questions, yet in this study we have identified a novel MeCP2-mediated decoding mechanism of the epigenetic cryptogram that is unique to the brain and found an unexpected link to a critical new function for the RTT-causing protein MeCP2.

Chapter III.

Identification of a novel 5hmC-specific complex containing Purα

III.I. A novel 5-hmC-specific complex in brain nuclear extract

Although MeCP2 binding to 5hmC offers a new way of understanding the regulation of chromatin in the nervous system, it is possible that there are other novel 5hmC-binding proteins in the brain that were not detected in our studies previously due to methodological limitations. To further investigate this possibility, crude nuclear extract from rodent brain (Klose and Bird 2004) was directly assayed in EMSAs with radioactively labeled DNA probes containing C, 5mC or 5hmC in the presence of an excess of non-specific DNA competitor at a range of dilutions. To our surprise, a 5hmC-dependent low-mobility complex was apparent at extract concentrations ~0.8 µg (Fig 3.1.A). This complex was much less obvious with probes containing C or 5mC nucleotides. The quantification of the signal at the observed mobility revealed 4- to 10-fold enhancement of 5hmC-dependent signal at the observed location compared to the other probes (Fig 3.1.B). The appearance of this complex was not tightly dependent on the amount of nonspecific competitor; however, when we competed the “hot” 5hmC probe in this complex with increasing concentrations of “cold” probe containing either C, 5mC or 5hmC, the latter probe competed out the hot 5hmC probe slightly better than the others (Fig 3.1.C). This effect was especially clear when the competitor

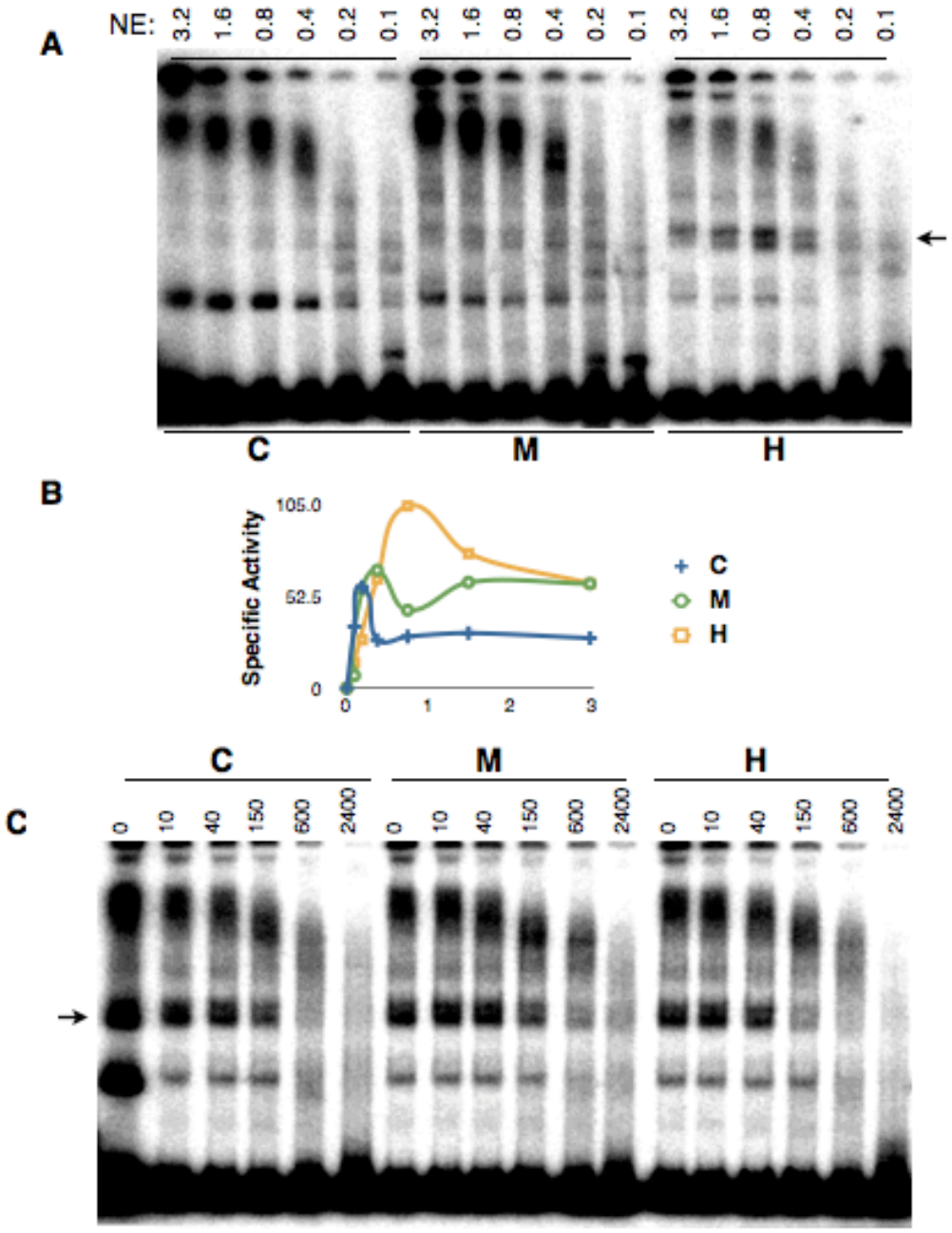


Fig 3.1. Detection and verification of a 5hmC-specific complex in nuclear extracts from rodent brain. **A.** EMSA gel with 1 µg nonspecific competitor and 1 fmol probe with increasing amounts of nuclear extract (NE), as indicated in µgs. **B.** Quantification of signal intensity at location marked with arrow in A, background subtracted. **C.** EMSA gel as in B but with 1 µg nonspecific competitor and increasing amounts of unlabeled competitor, as indicated in "fold of labeled probe". Arrow points at the 5hmC-specific complex in A and C.

concentration was 150-fold of labeled probe. Although this specificity seemed subtle under our assay conditions, it was robust and reproducible. Hence we attempted to biochemically purify and identify the proteins that cause this specific activity, with the expectation of finding another decoding mechanism for 5hmC in the brain.

III.II. Biochemical purification of 5-hmC-specific complex

First we prepared nuclear extract from 80 rat cerebella and fractionated it over a column coated with heparin, which has been used as a biomimetic polymer of DNA because it is a highly sulfated glycosaminoglycan and a linear negatively charged polymer with a size ranging from 5-30 kDa (Kiss, Kakkar et al. 1976; Farooqui 1980). Therefore heparin chromatography is a suitable way to separate DNA-binding proteins from the rest on nuclear proteins. Thus, highly concentrated nuclear extracts were applied on a heparin column at physiological conditions and extensively washed. The bound proteins were eluted by gradually increasing the salt concentration from 0.15 to 1 molar (1M), collected in sequential fractions and assayed for their DNA binding.

In our initial gradient elution (Δ[KCl] = 43 mM/min), we observed a 5hmC-specific complex that eluted at salt concentrations 0.4-0.55M. The complex resulted EMSA shift of similar distance from the free probe as the shift observed in NE (Fig 3.2.A). The signal intensity of this 5hmC-specific complex in fraction 28 was two-fold of 5mC-containing complex (Fig 3.2.B). When we used a less steep gradient (Δ[KCl] = 20 mM/min), we observed a second, higher mobility complex

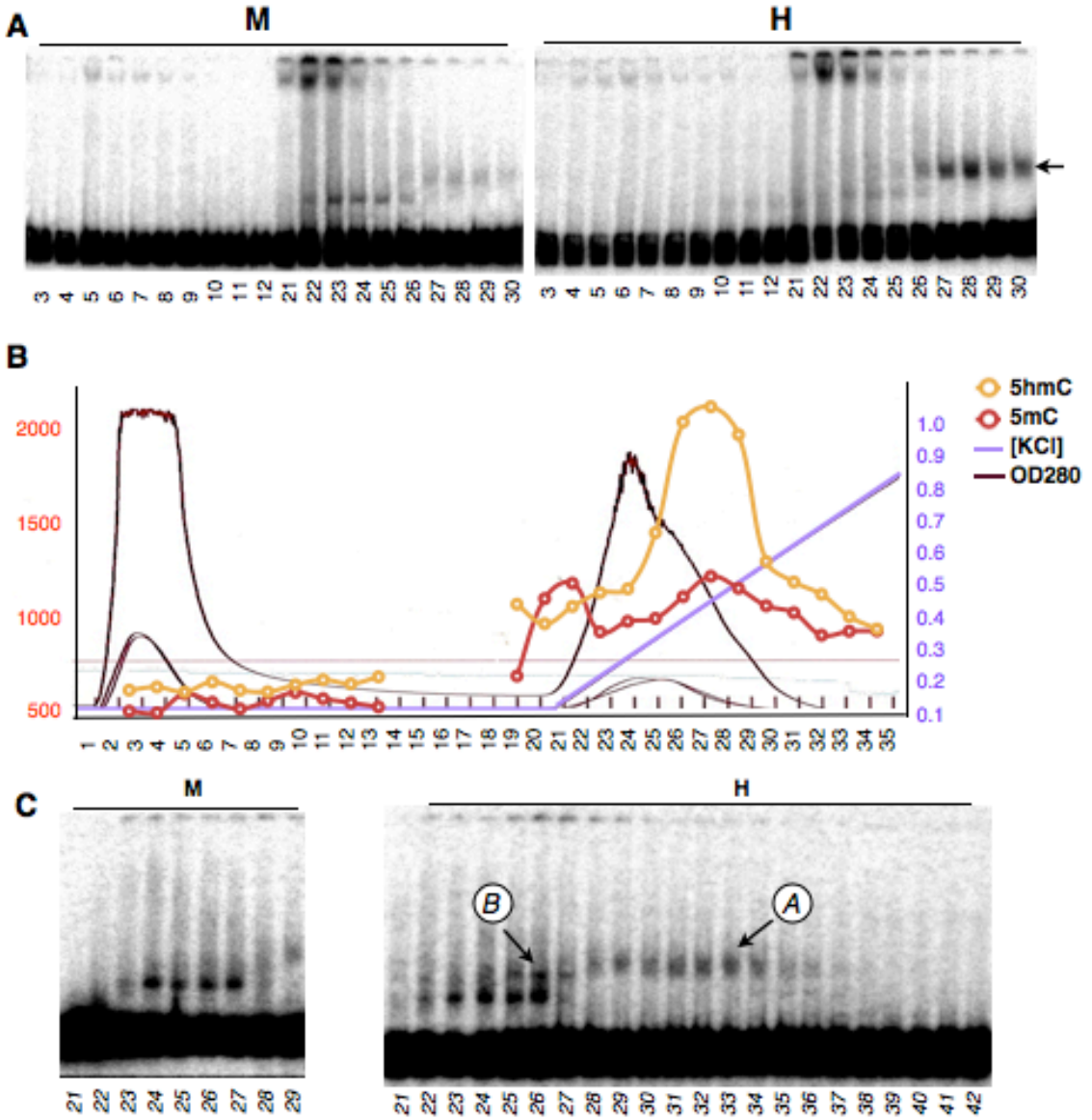


Fig 3.2. 5hmC-specific complexes are detected in the heparin bound fractions. **A.** EMSA gel with 4µl of unbound heparin fractions (3-12) and eluted fractions (21-30) with 1 µg nonspecific competitor and 1 fmol probe. Arrow points at the 5hmC-specific complex. **B.** Flow of heparin chromatography. Fractions are numbered from the time of sample injection. KCl (purple line) concentration starts at 0.15M and increases at Δ[KCl] = 43 mM/min after fraction 21. Protein concentration is detected by OD280 (brown). Signal intensity in A. at position marked with arrow is shown for 5mC (red) and 5hmC (orange). **C.** EMSA gel as in A but fractions are eluted at Δ[KCl] = 20 mM/min. Arrows point complexes A and B.

that eluted at lower salt concentrations, at 0.27-0.39 M (Fig 3.2.C), that was not present when the probe was made with 5mC. Since the mobility of both complexes were very close, it could be possible that the initial 5hmC activity we observed was made up of more than one species. Since we could biochemically separate these two species, we decided to purify them in parallel. A schematic flow of our biochemical purification is shown in Fig 3.3.A. Basically, we pooled fractions containing the later eluting complex (fractions 28 to 34 in 3.2.C, referred as complex A) and fractions containing the earlier eluting complex (fractions 24-27, referred as complex B) separately. We applied both mixtures sequentially on a size exclusion column. The fractions that still showed the specific activity were once more pooled for each complex and applied on affinity columns that were coated with DNA containing 5hmC; and finally the bound proteins were eluted with a steep salt gradient. The fractions with the specific activity were pooled, concentrated and separated in SDS PAGE. Finally the bands that were visible after Coomassie stain were excised and identified by MS. The pooled fractions at each purification step were analyzed by SDS PAGE followed by Coomassie staining (Fig 3.3.B) as a confirmation after each purification step. Finally, by mass pectrometry based database search, the proteins that co-purified with complex A were identified as Purine-rich element binding proteins, Purα and Purβ, in addition to Parp1 and Snrnp200 (Fig 3.3.C); while complex B co-eluted with a family of heterogenous nuclear ribonucleoprotein AB (hnrnpAB), hnrnpD, hnrnpR and hnrnpU (Fig 3.3.D).

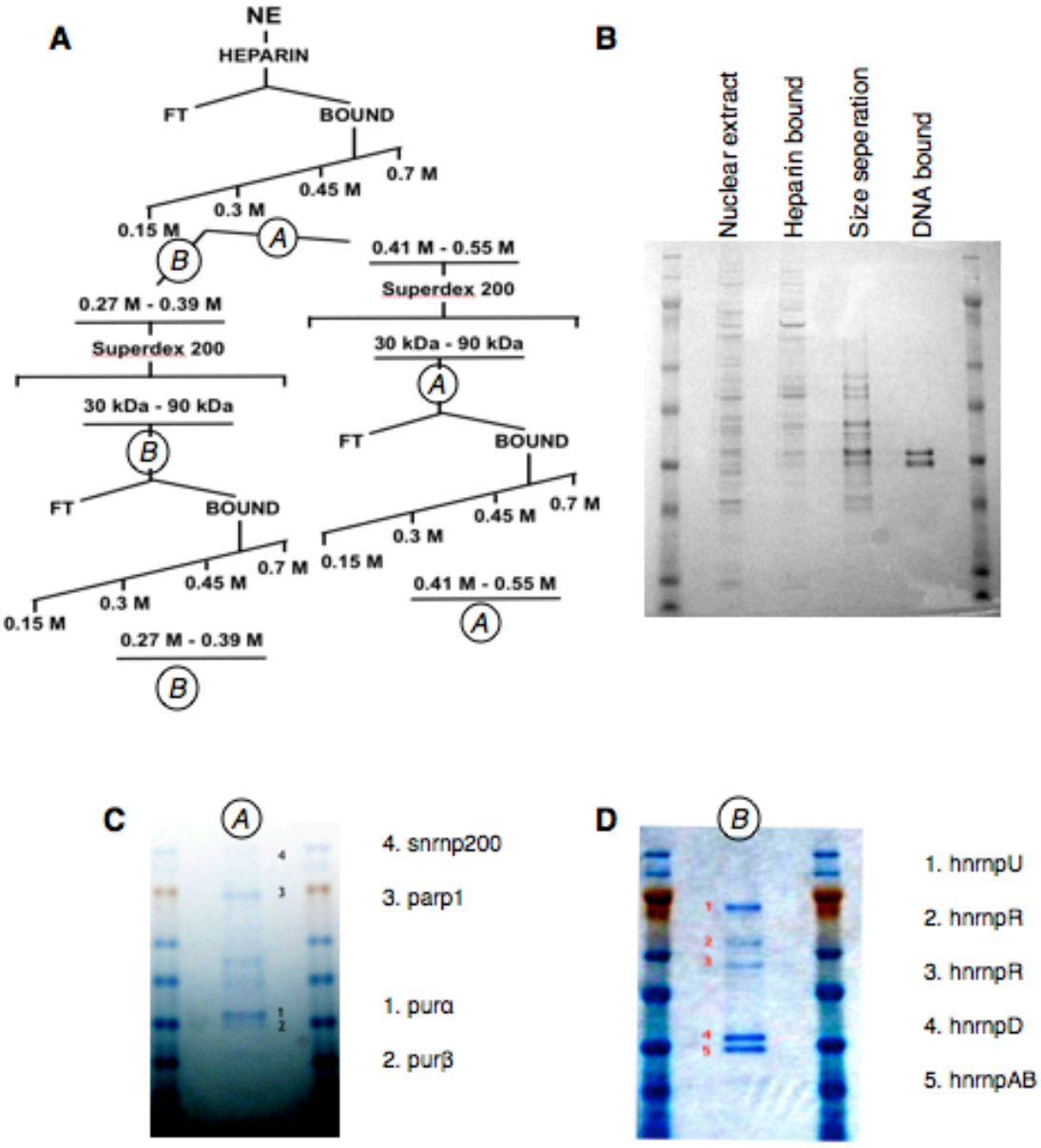


Fig 3.3. Purification and identification of 5hmC-complexes. **A.** Flow chart of biochemical purifications performed in parallel. **B.** SDS PAGE stained with Coomassie, showing the protein mixture at each purification step. **C.** SDS PAGE stained with Coomassie, with concentrated proteins of complex A and the identities of each portein band. **D.** Same as C. but complex B is shown.

III.III. 5-hmC-specific complex contains Purα

Previously, in our experiments, unlike MeCP2 NT (23 kDa), full-length MeCP2 (70 kDa) failed to enter the native gel from the wells. It is likely that full-length MeCP2 is too large to penetrate through the pores of the acrylamide matrix, though its intrinsic disorder may be a factor as well. Based on this criterion, we focused on Purα, Purβ, hnrnpAB and hnrnpD, all of which are less than 50kDa. To check the presence of these proteins in the 5hmC-specific complex that we originally observed, we first incubated the extract with antibodies against these proteins, then the competitor and the radioactive probe. The complex was clearly depleted by anti-Purα antibody without affecting other complexes within the lane, whereas the other antibodies did not show reactivity (Fig 3.4.A). Although this result gave us confidence that Purα was the strongest candidate to be the 5hmC-specific protein, it was not enough to exclude the possibility that other proteins might be involved, too. This could be due simply to low affinity of antibodies. Indeed, in our *in vitro* experiments, anti-Purβ antibody failed to deplete the recombinant Purβ/5hmC complexes (data not shown); we interpret the supershift assays using this antibody as inconclusive. When we reacted the anti-Purα antibody with the purified complex A, we observed the supershift, as expected (Fig 3.4.B); however, the specificity of the complex was compromised during purification.

As an additional control, we applied nuclear extracts directly onto DNA affinity column containing 5hmC and eluted at Δ[KCl] = 20 mM/min. After visual inspection fractions in SDS PAGE, we pooled fractions that contain similar

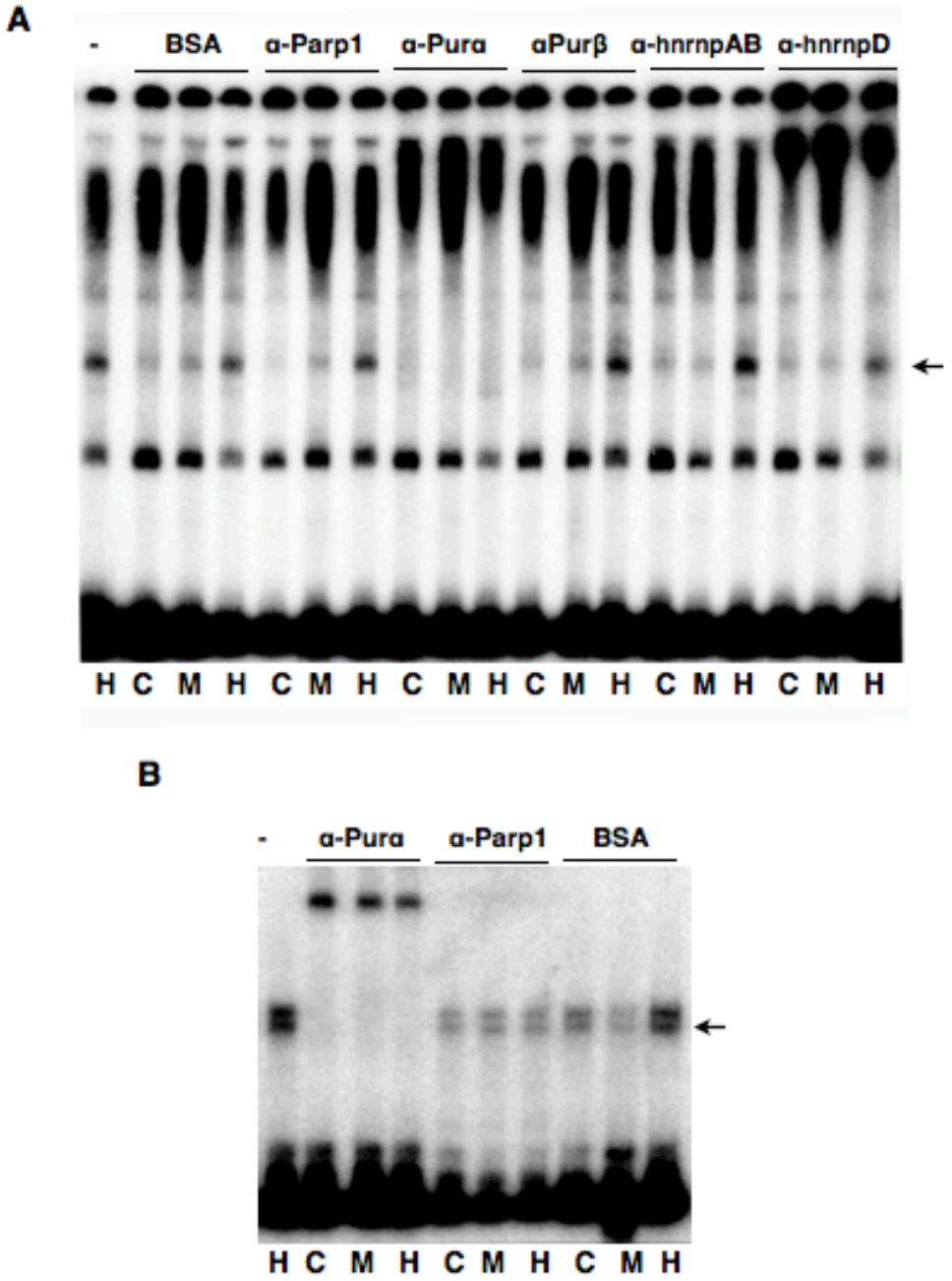


Fig 3.4. The 5hmC-specific complex contains Purα. **A.** EMSA gel with 1 µg nonspecific competitor, 1 fmol probe and 1 µg nuclear extract incubated with 0.1 µg antibody prior to reaction. **B.** EMSA gel with 0.1 µg nonspecific competitor, 1 fmol probe and 10 λ purified complex A, incubated with 1 ng antibody prior to reaction.. Arrow points at the 5hmC-specific complex in A and B.

protein composition (Fig 3.5.A) and analyzed the binding of these pooled fractions to DNA as previously by Southern blot using radioactive probes containing C, 5mC and 5hmC (Fig 3.5.B) and EMSA (Fig 3.5.C). The Southern blots did not reveal any 5hmC-specific bands; however, we observed a protein in fraction pools 41-44 and 45-49 with specific binding to 5mC and 5hmC containing DNA at the size of MeCP2. Indeed, by MS we identified the protein band (marked with red dot adjacent to number 1) to be MeCP2, confirming our previous results. Additionally, the band of the correct size but in early fractions did not contain any MeCP2. The bands adjacent to numbers 2 and 3 were identified as Purα and Purβ, respectively. Intriguingly, the bands of this size at pool 37-40 contained several Hnrnps and low amounts of Purα or Purβ. Fraction 41-44 contained similar amounts of Hnrnps and Purα or Purβ, whereas fraction 45-49 contained exclusively Purα or Purβ. The increase in the amounts of Pur proteins in the later fractions as the elution conditions get more stringent also tracks with the specific activity observed in the EMSA, confirming that Pur proteins are most likely the major components of the 5hmC-specific complex. The proteins marked with 4 and 5 are identified as several subclasses of Histone H1.

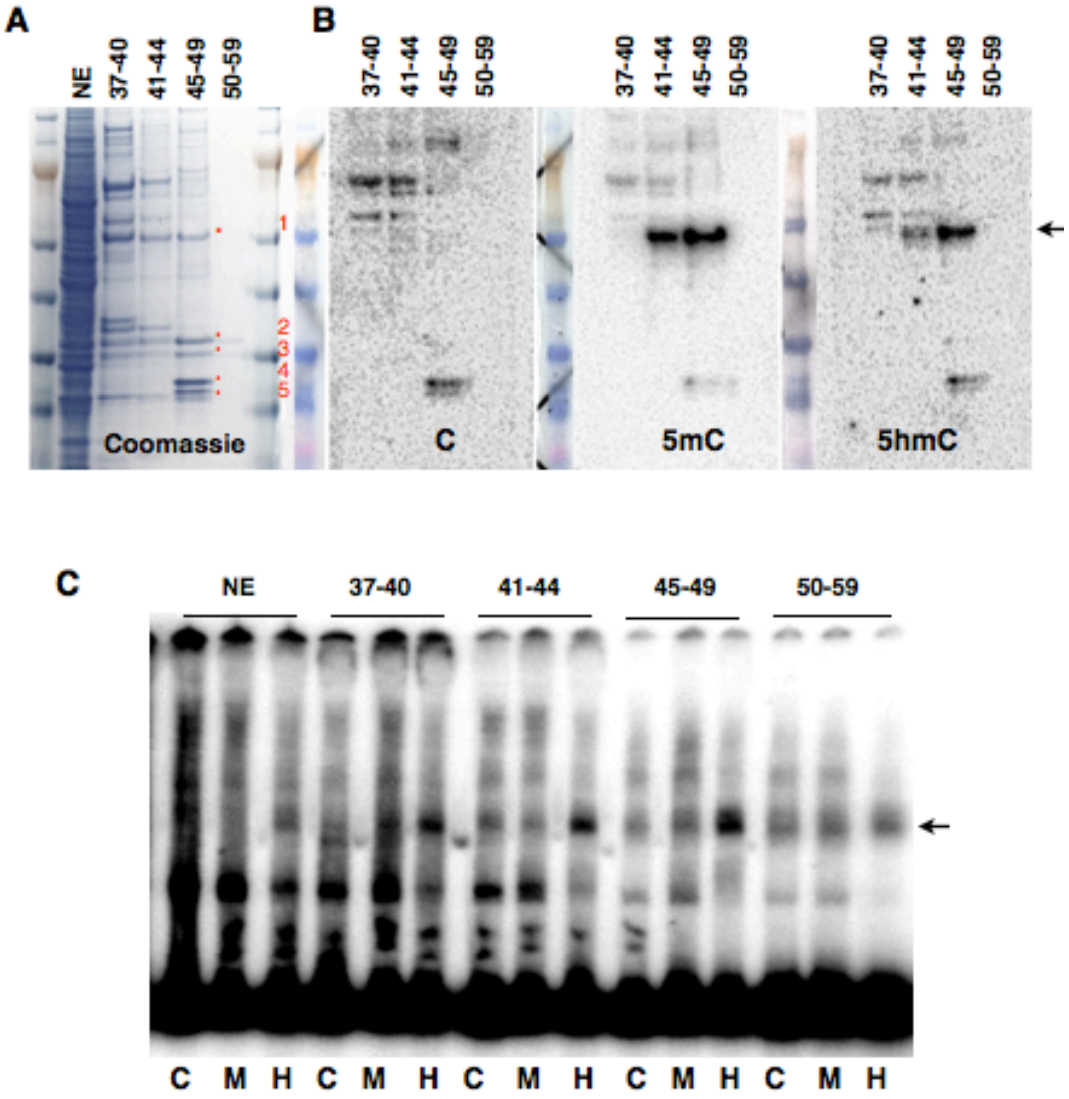


Fig 3.5. Direct application of nuclear extract to DNA-affinity column. **A.** SDS PAGE gel stained with Coomassie, containing eluates as indicated, where 37-59 for KCl concentration 0.15M-1M. Bands identified by MS are marked by red dots, adjacent to MeCP2 (1), Purα (2), Purβ (3) and Histone H1 subtypes (4 and 5). **B.** Membranes, containing same samples as in A. except NE; incubated with radioactive DNA containing C, 5mC or 5hmC. **C.** EMSA gel with 1λ pooled fraction, 1 fmol probe and 1 μg nonspecific competitor. Arrow points at the 5hmC-specific complex.

III.IV. Binding of recombinant hnrnps to 5hmC

Although hnrnps are first described as a family of proteins, which bound nascent RNA transcripts (Choi and Dreyfuss 1984), their functions vary greatly (Han, Tang et al. 2010). HnrnpAB and D have also been implicated in DNA-linked processes as well, such as telomere maintenance (Ford, Wright et al. 2002; Chai, Zheng et al. 2003), transcription (Chai, Zheng et al. 2003) and replication (Campillos, Lamas et al. 2003). We expected the hnrnps to be a false positive; given that they are predominantly involved in RNA processes, combined with our findings that (1) their amounts of hnrnps negatively correlated with the specific activity in fractions eluted from the DNA column, (2) the specific complex reacts with anti-Purα antibody and not with anit-hnrnp antibodies, and (3) in the competition assays the complex A (containing Pur proteins) behaves like the 5hmC-specific complex in the crude extract. To confirm that hnrnps were indeed nonspecific, hnrnpAB and hnrnpD were produced in *E.coli* and purified using a His tag (Fig 3.6.A, left panel). To our surprise, when we assayed the DNA binding characteristics of these proteins at very low concentrations, both proteins formed a weak 5hmC-specific low mobility complex in presence of competitor DNA (Fig 3.6.B). The 5hmC-dependent signal of hnrnpAB was 4-fold of both C- and 5mC-dependent signals (Fig 3.6.C, left), whereas in the case of hnrnpD this effect was around 2-fold (Fig 3.6.C, right). Interestingly, as the concentration of the hnrnpAB increased, the low mobility complex disappeared (Fig 3.6.C, black arrow) and a higher mobility complex started to appear (Fig 3.6.C, red arrow), the intensity of which was independent of the DNA modification. We postulated that the basis of

these two different complexes with different binding characteristics might be explained by the structural properties of these proteins: hnrnpAB and D contain two RNA recognition (RRM) domains (Hoffman, Query et al. 1991). RRM1 is highly conserved between hnrnpAB and hnrnpD, and so is RRM2; whereas the homology between RRM1 and RRM2 is much lower. It is then possible that RRM1 and 2 might have different characteristics for nucleic acid binding.

It is possible that the low mobility complex is formed when the probe is in excess and thus both RRMs are occupied by one radioactive probe; whereas when only on RRM is bound by a probe, then a higher mobility complex can be observed. If the remaining RRM is bound by the nonspecific competitor, the mobility of the complex will vary, causing a background smear. To directly address this issue, we produced recombinant RRM1 and RRM2 of both hnrnps (Fig 3.6.A, right panel) and assayed their DNA binding. Their DNA binding efficiency was significantly reduced, but preliminary results indicated that RRM1 of hnrpAB is marginally more specific to 5hmC-containing DNA (Fig 3.6.E).

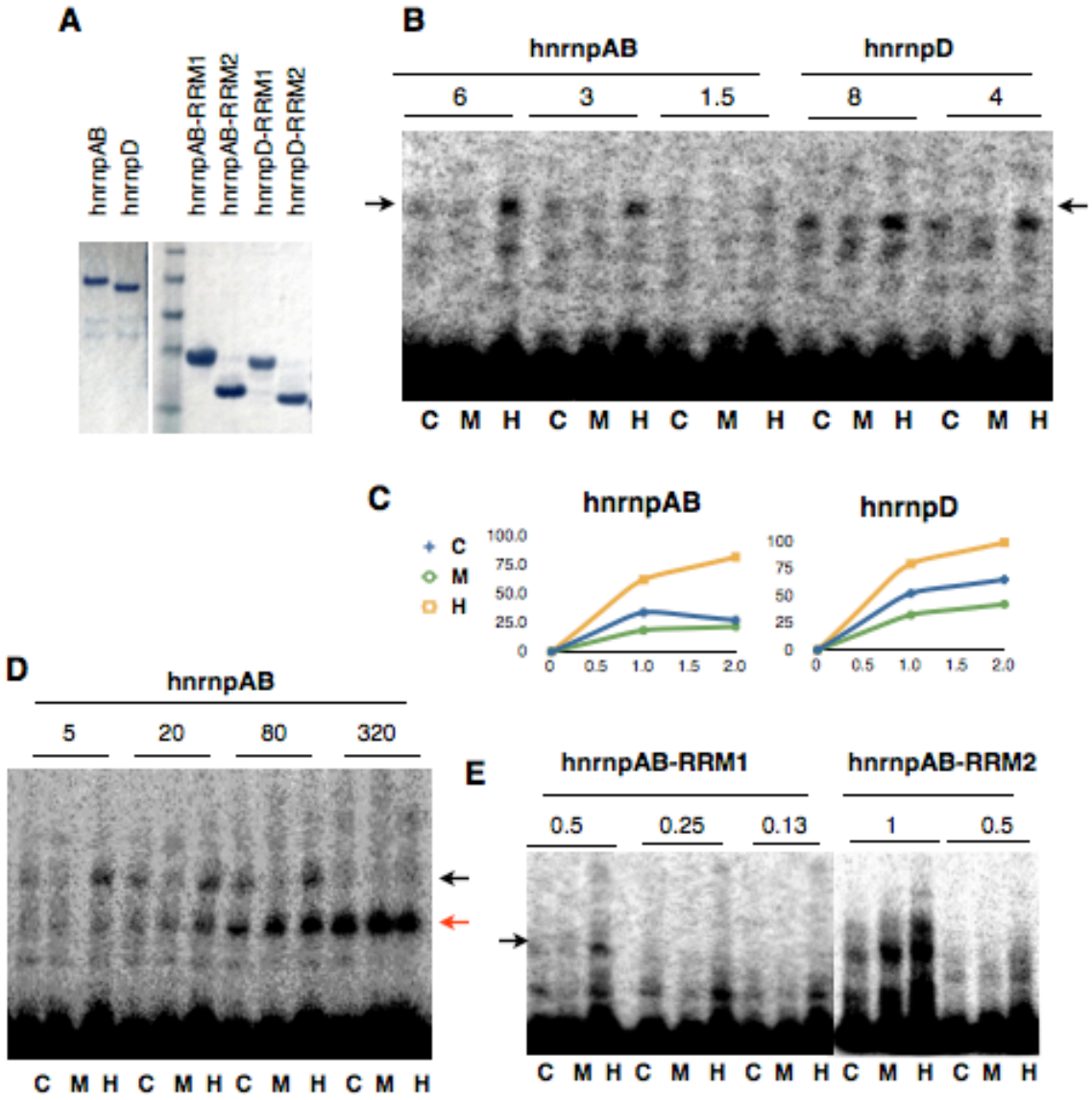


Fig 3.6. hnrnps have low affinity but specific binding to 5hmC through RRM1. **A.** SDS PAGE stained with Coomassie, containing purified recombinant hnrnps, full-length (left) or domains only (right). **B.** EMSA gel with dilutions of hnrnpAB and D proteins (in pmols) with 0.1 µg nonspecific competitor, 0.5 µg BSA and 1 fmol probe. Black arrow in A and B mark lower mobility complexes. Red arrow points at higher mobility complex. **C.** Quantification of signal intensity at location marked with arrow in B. **D.** EMSA gel as in A. **E.** EMSA as in A. hnrnpAB-RRM1 and hnrnpAB-RRM2 concentrations in nmoles. Arrow points at the 5hmC-specific hnrnpAB-RRM1-dependent complex.

III.V. Binding of recombinant Purs to 5hmC

Purα together with Purβ and two isoforms of Purγ constitute a distinct and highly conserved class of PUR family that has been implicated in multiple roles in cellular and viral regulation of nucleic acids, including some neuronal genes in higher eukaryotes (Gallia, Johnson et al. 2000; White, Johnson et al. 2009). *Purα* expression increases during development with highest levels observed in bodies and dendrites in PCs, where the highest 5hmC levels also are observed (Kriaucionis and Heintz 2009). *Purα*-deficient mice die after birth with severe neurological pathologies (Khalili, Del Valle et al. 2003). They also are linked to leukemia (Lezon-Geyda, Najfeld et al. 2001), just as *Tet* genes (Kosmider, Delabesse et al. 2011; Lorsbach, Moore et al. 2003). Therefore Purα might be an important player in neural epigenetic regulatory mechanisms.

Although not much is known about *in vivo* functions of Purβ, *in situ* hybridization data show that it is highly enriched in PCs as well (http://mouse.brain-map.org/). To have insight into their binding characteristics, we produced recombinant proteins in bacteria (Fig 3.7.A). We observed some DNA binding activity by each of these proteins with marginal preference for 5hmC (Fig 3.7.B). Recombinant Purα, when bound to unmodified or methylated probe formed two complexes with different mobilities. When unmodified, these complexed were of equal intensity. Upon methylation, the lower mobility complex seemed to be conserved whereas the level of the higher mobility complex decreased, and this complex also migrated a little faster than the corresponding complex with C-containing probe.

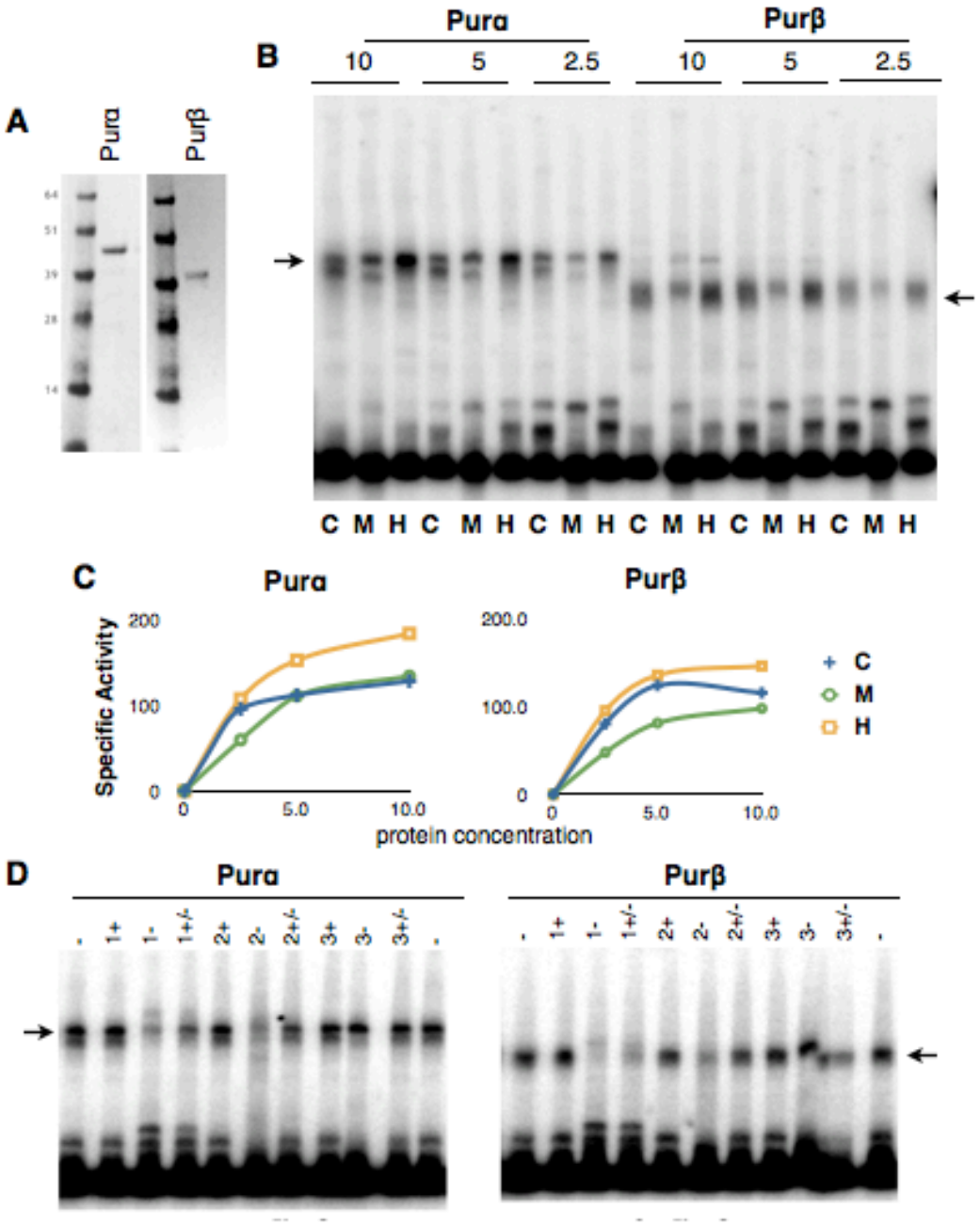


Fig 3.6. Purα and Purβ binding to 5hmC. **A.** SDS PAGE stained with Coomassie, containing purified recombinant full-length Purα and Purβ **B.** EMSA gel with dilutions of Purα and Purβ (in pmols) with 0.1 µg nonspecific competitor and 1 fmol probe. Arrows mark protein-dependent complexes. **C.** Quantification of signal intensity at location marked with arrow in B. **D.** EMSA gel as in B but with 10 fmols of short, unlabeled, single-stranded (+ or - strand or both), unmodified competitors, with sequence (1-49, 38-86, 80-128 of radioactive probe 1-128 bp). Arrows point protein-dependent complexes.

In contrast, Purα formed a single lower mobility complex upon binding to hydroxymethylated DNA, which seemed to have higher levels compared to the other lanes, although this might simply be due to co-migration of two distinct complexes. Purβ forms a smear of low-mobility complexes with unmodified DNA, and the binding is much weaker when DNA is methylated. The intensity of Purβ/ 5hmC smear was slightly higher compared to Purβ/C smear by visual inspection. We then quantified the signal intensities of the lowest mobility complex of each sample and derived a binding curve (Fig 3.7.C). The preference for 5hmC of both proteins was < 2-fold compared to other probes. Even at highest concentrations, these proteins failed to produce an SPR response on the DNA-coated surfaces (data not shown).

We suspected that the loss in specificity might be due to a co-factor that is not present in the *in vitro* binding reaction. Therefore we included increasing amounts of crude extract; however, it did not make a difference in the binding characteristics of these proteins (data not shown). Since Purα and Purβ can strongly interact with each other in a DNA-dependent as well as DNA-independent way (Kelm, Cogan et al. 1999), we have also combined both proteins in the expectation that they might show cooperativity; unfortunately, incubation of these proteins prior to introduction of nucleic acid did not make a difference (data not shown). Endogenous Purs might be post-translationally modified in the 5hmC-specific complex, whereas the bacterially expressed recombinant proteins are not. Another alternative is that the crude extract contains a variety of proteins that bind unmodified DNA and others that recognize

5mC-containing DNA and do not give access to Pur proteins; however, 5hmC is not preferred by many proteins so they occupy 5hmC sites. Under artificial conditions they are not selective for DNA modification. It is possible that these complexes also form with DNA containing C or 5mC in the nuclear extracts; however, they are bound by additional factors and migrate at much lower rates and therefore are not visible at the location we detect the 5hmC-dependent complex.

III.VI. Support for strand separation by Purs

Pur proteins consist of a glycine-rich flexible N terminus, a central nucleic acid binding core, with highest sequence conservation, and a C-terminal region, that is thought to be involved in protein interactions. The two nearly-identical PUR repeats interact with each other to form a PUR domain, which form two independent DNA binding surfaces, whereas a third PUR repeat facilitates the dimerization of Purα (Graebsch, Roche et al. 2009). The Pur proteins have been characterized by their binding to single-stranded DNA (ssDNA), double-stranded DNA (dsDNA), and ssRNA with a preference for GGN-repeats (Knapp, Ramsey et al. 2006). Purα is able to unwind duplex DNA, displacing the pyrimidine-rich strand while maintaining contact with the purine-rich strand (Wortman, Johnson et al. 2005). Since 5hmC might destabilize the duplex structure (Thalhammer, Hansen et al. 2011) we next asked whether the preference for 5hmC could be simply because this modification expedites the strand separation for the PUR proteins. To test this idea we competed the “hot” 128 bp probe with short,

unlabeled and unmodifed ssDNA sequences (1-49, 38-86, 80-128 of radioactive probe). Indeed, only couple strands, but not their complimentary strand, could abolish binding to 5hmC even at very low concentrations (Fig 3.7.D). Their effect was weaker when the complimentary strand was added to the reaction, indicating that sequence-specific ssDNA binding is an essential mechanism for the DNA recognition of Pur proteins, although one strand did not contain a GGN repeat. Together this study proposes several candidates that might recognize 5hmC *in vivo*; however, their biological relevance still remains unclear.

III.VI. Discussion

The study presented here identifies a 5hmC-specific complex in the nuclear extracts from cerebellum and introduces new candidates for the decoding of 5hmC code in the brain. Although there is a lot to be done to characterize the biological function of these proteins upon binding to 5hmC, we present some convincing evidence that Purα is an integral component of this complex. The significant role and abundance of both 5hmC and Purα in PCs (Kriaucionis and Heintz 2009), combined with our results showing that in crude nuclear extracts there is a Purα-complex that associates with hydroxymethylated DNA up to 10-fold more avidly than unmodified or methylated DNA, suggest that this interaction might be crucial for neuronal function. In the brain Purα is found at high levels in Purkinje cells (Khalili, Del Valle et al. 2003; Hokkanen, Feldmann et al. 2011) with a peak of expression during critical periods of brain development, and in *Purα* $^{-/-}$ mice showing neurodevelopmental defects (Khalili,

Del Valle et al. 2003). Also high levels of 5hmC are present in both developing and adult brains (Ruzov, Tsenkina et al. 2011; Szulwach, Li et al. 2011). Together these data suggest a complex yet significant role for Purα protein in the brain.

- *Mysterious Purα*

Although we were able to discover a Purα-complex with preference for 5hmC-containing DNA, we still don't know if hydroxymethylated sites are occupied by this complex *in vivo* and more importantly what the downstream events are. We also don't know the other components of this complex, if there are any. To answer the first question, a cell-specific genome-wide ChIP-Seq experiment for Purα can be performed. Yet we still don't know how this interaction is functionally interpreted by Purα, since the functional data of Purα is controversial. It has been shown to function as a transcriptional activator (Haas, Gordon et al. 1993; Haas, Thatikunta et al. 1995; Zambrano, De Renzis et al. 1997) and repressor (Penberthy, Zhao et al. 2004; Subramanian, Polikandriotis et al. 2004), in addition to its various roles in both RNA- and DNA-dependent processes (Li, Koike et al. 2001; White, Johnson et al. 2009), including DNA repair (Kaminski, Cheeseboro et al. 2010). This data makes the interpretation of 5hmC-dependent role of Purα difficult to judge. It is likely that Purβ is a component to the Purα-complex, since Purα and Purβ can strongly interact with each other (Kelm, Cogan et al. 1999). Hnrnps may also be included, since their interaction with Pur proteins has been previously suggested (Li, Koike et al. 2001).

Moreover the binding of recombinant or purified Purα, as well as Purβ and hnrnps, to 5hmC was compromised. This may be due to limitations of *in vitro*

reconstitution assays. First, the recombinant proteins expressed in bacteria lack posttranslational modifications. This problem can be solved by switching to a mammalian expression system. Secondly, both recombinant and purified proteins were stripped from any co-factors or other binding partners. To overcome this we have supplemented the reaction with nuclear extract but did not observe any recovery. On the other hand, if there is a dose dependence then this strategy can be improved by first depleting the endogenous Pur proteins using an antibody, and then adding this to the recombinant Purα. In the assay conditions these proteins are very labile. In addition, with each purification step, the addition of albumin was increasingly necessary, indicative of high background non-specific binding.

- *Specific, not specific*

Although this complex appears specific in the crude extract, we could not detect a strong preference for recombinant Purα to 5hmC *in vitro.* Although it is highly probable that the *in vitro* conditions failed to reproduce the endogenous environment, we also have to consider the possibility that Purα-complex does not bind 5hmC preferentially. We propose several explanations for this assumption: First, the fact that mammalian proteome contains thousands of DNA binding proteins (Kadonaga 2004), whereas only a few proteins are known to specifically recognize methylated sequences (Ballestar and Wolffe 2001; Fournier, Sasai et al. 2012). Although some DNA-binding proteins still may recognize their cognate DNA sequences when they are modified, it is reasonable to assume that a substantial fraction of these are displaced upon DNA modifications within their

binding sites. Therefore, even if *in vitro* binding of a protein to its target shows only subtle changes at different modification states, these may be translated to *in vivo* occupation of the majority of hydroxymethylated sites by such factors, since there is much bigger competition for an unmodified site. Since the nuclear extract represents a system that is close to the *in vivo* environment, it is very likely that this interaction has biological relevance.

A second probable explanation is that the complex forms with each modification of DNA. However when it contains unmodified or methylated DNA, it is bound by additional factors and therefore it does not have the same mobility. If 5hmC causes formation of a complex that has different protein composition, this is still significant since it can potentially explain the downstream events following the formation of this complex, the composition of which differs depending on the DNA modification. Finally it is also possible that one probe (128 bp) can accommodate more than one complex independently.

- *The strand-separation hypothesis*

A noteworthy but simple model for the preferential interaction between 5hmC and Pur proteins is that the hydroxymethylated DNA is energetically “cheaper” for strand-separation by Pur proteins and therefore the binding kinetics are enhanced. If Purα is able to unwind duplex DNA to maintain contact with the purine-rich strand (Wortman, Johnson et al. 2005), then the pyrimidine-rich strand might be stabilized by another protein in the nuclear extract. When this protein is not present after purification of the reconstituted system, the unbound strand might be energetically too “costly” to maintain the preferential binding.

Indeed, an example for such interaction has been previously demonstrated, where Pur proteins strongly interacted with a pyrimidine-rich DNA-binding protein, MSY1 (Kelm, Cogan et al. 1999). Although this interaction was reproduced in the absence of DNA as well, it does not exclude the possibility that this kind of communication between proteins might be a significant biological process. However this implies that as an ssDNA competitor abolishes binding as in Fig 3.6.D, the complimentary strand then should then have the opposite effect, since it would stabilize the displaced strand of the original probe. This has not been observed in our experiments. Strand separation model also could explain the different mobilities of Pur/C, Pur/5mC and Pur/5hmC complexes. These could be simply two populations where the Pur protein is bound to one strand or the complimentary. Indeed, the molecular weight between strands is 0.27 kDa. Since the mobility in the native gel might depend on factors other than molecular size, this small difference might explain this phenomenon.

- *What else is out there and how do we search?*

In this study we identified a Purα-complex that binds 5hmC specifically. Yet our biochemical understanding of this complex is still limited and we do not have information on the functional meaning of this interaction. In addition, we still don't know which proteins occupy the rest of the hydroxymethylated sites, or whether they are depleted of proteins. We can improve and extend our strategies to find new candidates: First there is a large population of proteins that remain tightly bound to the chromatin after salt extraction and thus not subject to our analysis, whereas high salt concentration may be disruptive for certain proteins. Secondly,

the ionic conditions in gel shift assays might cause interactions to be lost. By enzymatic degradation of nucleic acids to get nuclear proteins, combined with further extending the detection methods more candidates may be screened. Finally, using different sequences as DNA baits that are modified in various contexts can increase the screening efficiency.

Identifying the readers of the 5hmC code will contribute greatly to the understanding of neuronal epigenetics, although it is also possible that 5hmC simply modifies the physical properties of the duplex DNA, changing the structure of chromatin, re-arranging nucleosomes, displacing methylation-specific factors or increasing accessibility to proteins that bind ssDNA, which includes RNA Polymerase II. To decipher the 5hmC code and understand its true meaning, our studies need to encompass many possible scenarios, some of which may be occurring in concert. However in the words of Albert Einstein “[A]s our circle of knowledge expands, so does the circumference of darkness surrounding it.”

Materials and Methods

Preparation of Probes

C, 5mC and 5hmC probes were prepared using either unmodified dCTP (New England Biolabs), 5'-methylated dCTP (Trilink Biotechnologies, San Diego, CA, USA) or 5'-hydroxymethylated dCTP (Bioline, Taunton, MA, USA), along with dATP, dTTP and dGTP (New England Biolabs). A 120 bp region of mouse BDNF promoter was amplified with ChoiceTaq (Denville, South Plainfield, NJ, USA) according to manufacturer's recommendations using 5'-biotinylated or 5'-unmodified forward primers and 5'-unmodified reverse primers (Genelink, Hawthorne, NY, USA) listed below. For CG-rich probe we amplified a 200 bp region of the pUC plasmid (Invitrogen). Probes were purified via Qiaquick PCR Purification Columns (Qiagen). 1 pmol of probe was end-labeled with 1 or 10 uCi of γ-P32 dATP (Perkin Elmer, Waltham, MA, USA) using T4 Polynucleotide Kinase (New England Biolabs) for 45 min at 37°C and purified using Illustra ProbeQuant G-50 Micro Columns (GE Healthcare, Little Chalfont, UK). ^{32}P was counted in scintillation fluid Ready Safe (Beckman Coulter) in 1209 Rackbeta scintillation counter (LKB-Wallac Pharmacia, Turku, Finland).

primer	sequence
BDNF F	GCGTGAATTTGCTAGGACTGG
BDNF R	GAATTACCAGAATCAGAATTCCG
CG-rich F	GTGAGCTAACTCACATTAATTG
CG-rich R	TGCTCACATGTTCTTTCCTG
probe 1 (p1) F	GCCCATTCTTTCTTGATAGATTTTAGTTGTTCAAC
probe 1 (p1) R	TGAGCATCCTGGCGAGCATAG

probe 2 (p2) F	GCCCACAATAAACCAGCCACTTACCAG
probe 2 (p2) R	GGGAGATGTTCTGGCCTCTCAG
probe 3 (p3) F	GTAAGCTGATGGAAAACTGCTGTTG
probe 3 (p3) R	AGCCAACTCTCTAAGAGATGGACCTCAC
single-site F	CATGAATAAAACATCTACAAAATCAACCCTAAAAC
single-site R	TCGACCATTTAAAATATACATCTTTCCTATTAAAC
CGCt	AATAAAACATCTACAAAATCAACCCTAAAACcgcTTTAATAGGAAAGATGTATATTTTAAATGG
CACt	AATAAAACATCTACAAAATCAACCCTAAAACcacTTTAATAGGAAAGATGTATATTTTAAATGG
CCCt	AATAAAACATCTACAAAATCAACCCTAAAACcccTTTAATAGGAAAGATGTATATTTTAAATGG
CTCt	AATAAAACATCTACAAAATCAACCCTAAAACctcTTTAATAGGAAAGATGTATATTTTAAATGG
CGA	AATAAAACATCTACAAAATCAACCCTAAAACcgaTTTAATAGGAAAGATGTATATTTTAAATGG
CAA	AATAAAACATCTACAAAATCAACCCTAAAACcaaTTTAATAGGAAAGATGTATATTTTAAATGG
CGT	AATAAAACATCTACAAAATCAACCCTAAAACcgTTTAATAGGAAAGATGTATATTTTAAATGG
CAT	AATAAAACATCTACAAAATCAACCCTAAAACcaTTTAATAGGAAAGATGTATATTTTAAATGG
1+	GCGTGAATTTGCTAGGACTGGAAGTGAAAACATCTACAAAGCATGCAAT
1-	ATTGCATGCTTTGTAGATGTTTTCACTTCCAGTCCTAGCAAATTC ACGC
2+	CAAAGCATGCAATGCCCTGGAACGGAATTCTTCTAATAAAAGATGTATC
2-	GATACATCTTTTATTAGAAGAATTCCGTTCCAGGGCATTGCATGCTTTG
3+	ATAAAAGATGTATCATTTTAAATGCGCGGAATTCTGATTCTGGTAATTC
3-	GAATTACCAGAATCAGAATTCCGCGCATTTAAAATGATACATCTTTTAT

Nuclear protein extracts from cerebella

Nuclear extracts were prepared from 80 frozen rat cerebella (Pel-Freez, Rogers, AR, USA) as previously described (Klose and Bird 2004). Briefly, rat brains were diluted 5 volumes to 1 in ice-cold Buffer A containing 10 mM Hepes pH 7.5, 25 mM KCl, 0.15 mM spermine, 0.5 mM spermidine, 1 mM EDTA, 2 M sucrose, 10% glycerol, and Halt Protease and Phosphatase Inhibitor Cocktail (Pierce Biotechnology, Rockford, IL, USA) followed by homogenization in a 60

mL Dounce (Bronwill Scientific Inc., Rochester, NY) on Glas-Col Tissue Homogenizer (Cole-Parmer, Vernon Hills, IL, USA) (5 strokes at 4,000 rpm). The homogenate was layered onto a 10 mL cushion of the Buffer A and centrifuged in pre-chilled SW2 (Beckman Coulter) rotor at 24,000 revolutions per minute (rpm) in Beckman XL-70 ultracentrifuge for 1 hour at 2 °C. Recovered nuclei were incubated in 5 volumes of Buffer B containing 20 mM Hepes (pH 7.9), 150 mM KCl, 1 mM EDTA, 0.5 mM DTT, Halt Protease and Phosphatase Inhibitor Cocktail on ice 2 min and then resuspended. The nuclei were counted by Neubauer haemocytometer (Hausser Scientific, Horsham, PA, USA). The nuclear proteins were extracted by stepwise addition of 3M KCl until a final concentration of 400 mM. The extraction was allowed to proceed for 20 min on ice, and then the nuclei were pelleted at 13,000 rpm for 10 min at 4°C. The supernatant was dialyzed in Slide-A-Lyzer cassettes with 10 kDa molecular weight cutoff (Pierce) overnight at 4°C against Buffer B. The protein concentration was measured by Quick Start Bradford Assay (BioRad) according to manufacturer's instructions. For fast protein liquid chromatography, 10 mL of the dilate was collected and centrifuged 5 min at 14,000 rpm. The cleared lysate then was concentrated in Amicon15 10K MWCO (Millipore) 20 min at 4,000 rpm at 4°C or until the extract is down to 4 mL. Alternatively, nuclear extracts were prepared from 10 wild type (WT) or ECP2 KO mice using SW41 rotor with a 2 mL cushion of Buffer A.

5hmC affinity pull-down

10 ug of 5'-biotinylated C, 5mC or 5hmC BDNF probe was immobilized on

Dynabeads M-280 Streptavidin (Invitrogen) following manufacturer's recommendations and incubated with 2 mg of nuclear extract in presence of 1 mg of poly-dIdC competitor (Sigma Aldrich) in Buffer B supplemented with 1% Triton X-100 (Sigma Aldrich) 1 hour at room temperature. The isolated proteins were washed extensively and eluted in 1X LDS Sample Buffer (Invitrogen) under reducing conditions, separated by gel electrophoresis in 4-12% gradient BisTris Gels (Invitrogen) and stained by GelCode Blue (Pierce) or Silver Stain Kit (Pierce). Protein bands were excised from a GelCode stained gel and analyzed by mass spectrometry (MS). Alternatively, nuclear extracts prepared from WT or KO mice were incubated with magnetic beads coated with 5'-Biotinylated CG-rich probes. Nuclei from *Mecp2*$^{tm1.1Bird}$ hemizygous males were a gift from Dr Adrian Bird.

MeCP2 immunoprecipitation

40 ug of 5'-biotinylated probe containing CpG, 5mCpG, 5hmCpG, CpA, 5mCpA and 5hmCpA was immobilized on Dynabeads M-280 Streptavidin (Invitrogen) following manufacturer's recommendations. 10 mg of nuclear extract supplemented with 0.5 mg of poly-dIdC competitor (Sigma Aldrich) in Buffer B was incubated with the DNA-coated beads for 1 hour at room temperature. The isolated proteins were washed extensively and eluted for 30 min in Buffer N containing 400 mM NaCl, 50 mM Hepes pH 7.9, 1 mM EDTA, 10 mM DTT, 0.5% SDS and 10 units/ml of Benzonase (Sigma). 50 ug of anti-MeCP2 antibody (Cell Signaling) was coupled to Dynabeads Protein A (Invitrogen) in Buffer B

supplemented with 0.5% Igepal (Sigma) according to manufacturer's recommendations. The eluates in Buffer N were diluted 1:5 in Buffer B and incubated with MeCP2-coated beads for 2 hours at room temperature. The proteins bound to beads were eluted in 1X LDS Sample Buffer (Invitrogen) under reducing conditions, separated by gel electrophoresis in 4-12% gradient BisTris Gels (Invitrogen) and assayed by Western blotting.

Western blotting

For Western blotting, the extracts denatured in sample buffer (Invitrogen) under reducing conditions, and eluates prepared as described above, were separated on a pre-cast Bis-Tris SDS-gel (Invitrogen) via electrophoresis and transferred on a PVDF-membrane (Biorad). The membrane was blotted with primary antibodies, anti-MeCP2 (Upstate) antibody in combination with horse-radish-peroxidase-conjugated secondary antibody (Pierce) or anti-phosphoserine-HRP (Abcam). The chemiluminescence was detected by the application of the ECL substrate (Pierce).

Southwestern Blotting

Nuclear extract (isolated as described above) was separated on duplicate gels and transferred to PVDF membranes (BioRad) in Mini Trans-Blot Electrophoretic Transfer Cell (BioRad) at 100 V for 2 hours at room temperature and assayed for DNA binding as previously described (Campoy, Meehan et al. 1995). The blotted proteins were re-natured by immersing the blot in Buffer B

with 6 M guanidine HCl, which was then serially diluted to 3, 1.5, 0.75, 0.375, 0.188, and 0.094 M using binding Buffer B with incubation at 4°C for 5 min each time. The blot was blocked at room temperature for 30 min in Buffer B containing 5% non-fat milk, then incubated with 10^5 cpm/mL CG-rich 5mC or 5hmC probes with 10 ug/mL poly-dIdC for 2 hours at room temperature, extensively washed with Buffer B and air-dried. The autoradiography was measured by exposing a storage phosphor screen (Molecular Dynamics, Sunnyvale, CA) that was scanned by Typhoon Imager (GE Healthcare) and analyzed by ImageJ software.

Expression and purification of recombinant proteins

cDNA clones of human MeCP2, and MBD family of proteins were purchased from Open Biosystems (Lafayette, CO) and cloned into pet28a vector (Novagen, Madison, WI, USA) using primers listed below with NdeI and XhoI (NEB) restriction sites.

primer	sequence
MeCP2 NT (aa 1-205) F	ATGGTAGCTGGGATGTTAGGGCTCAGGGAAG
MeCP2 NT (aa 1-205) R	AGACCCAAGGCGGCCACGTCA
MeCP2 MBD (aa 77-167) F	GCTACCGGACTCAGATCTCGAGCTCAACGCCACCATGGCTTCTGCCT CCCCCAAACAGCG
MeCP2 MBD (aa 77-167) R	CCGCGGTACCGTCGACTGCAGAATTCGACCGGGAGGGGCTCCCTCT CCC
MeCP2 (aa 1-486) F	ATGGTAGCTGGGATGTTAGGGCTCAGGGAAG
MeCP2 (aa 1-486) R	GCTAACTCTCTCGGTCACGGGCGTCCGGCTGTCCAC

MBD1 (aa 1-144) F	ATGGCTGAGGACTGGCTGGACTGCCCG
MBD1 (aa 1-144) R	GCCATCCCCTGAGAAGCTGATTCCACAGTTCTCACAGCAC
MBD2 (aa 141-319) F	ATGGCCACGGAGAGCGGGAAGAGGATGGATTGCCCGGCCCTCCCCCCCG
MBD2 (aa 141-319) R	CTCATCATTGCTACCTGGACCAACTCCTTGAAGACCTTTGGGTAG
MBD3 (aa 1-259) F	ATGGAGCGGAAGAGCCCGAGCGGGAAGAAGTTC
MBD3 (aa 1-259) R	GTGCTCCATCTCCGGGTCCGGGTCG
MBD4 (aa 1-197) F	ATGGGCACGACTGGGCTGGAGAGTCTGAGTCTG
MBD4 (aa 1-197) R	CTCTGAACTACTACTTGGCGGCATAAACACATCCTTTTTGC
L100V Rin	GTGTCCAGCCTTCAGGCACGGTGGG
L100V Fin	GTGCCTGAAGGCTGGACACGGAAGC
D121G Rin	CTTTTCCCTGGGGATTGATCAAATACACACCATACTTCCCAG
D121G Fin	GTGTGTATTTGATCAATCCCCAGGGAAAAGCCTTTC
R133H Rin	ACTCTAAAGTGGAGTTGATTGCGTACTTCGAAAAGGTAGGCGAC
R133H Fin	CGAAGTACGCAATCAACTCCACTTTAGAGTGAAAGGCTTTTC
R133C Rin	TGCTCTAAAGTGGAGTTGATTGCGTACTTCGAAAAGGTAG
R133C Fin	GTACGCAATCAACTCCACTTTAGAGCAAAAGGCTTTTC
S134C Rin	GTAAAGTGGAGTTGATTGCGTACTTCGAAAAGGTAGGC
S134C Fin	CGAAGTACGCAATCAACTCCACTTTACAGCGAAAGGC
S134F Rin	TTAAAGTGGAGTTGATTGCGTACTTCGAAAAGGTAGGCGAC
S134F Fin	CTTTTCGAAGTACGCAATCAACTCCACTTTAAAGCGAAAGGCTTTTC
E137G Rin	CTACCTTTTCGAAGTACGCAATCAACCCCACTTTAGAG
E137G Fin	GGTTGATTGCGTACTTCGAAAAGGTAGGCGACAC
A140V Rin	GTGTCGCCTACCTTTTCGAAGTACACAATCAACTCCAC
A140V Fin	TGTACTTCGAAAAGGTAGGCGACACATCCCTGGAC
R133M Rin	GAAGTACGCAATCAACTCCACTTTAGACATAAAGGCTTTTC
R133M Fin	ATGTCTAAAGTGGAGTTGATTGCGTACTTCGAAAAGGTAGGCGAC
R133K Rin	GAAGTACGCAATCAACTCCACTTTAGACTTAAAGGCTTTTC
R133K Fin	AAGTCTAAAGTGGAGTTGATTGCGTACTTCGAAAAGGTAGGCGAC
R133E Rin	GAAGTACGCAATCAACTCCACTTTAGACTCAAAGGCTTTTC
R133E Fin	GAGTCTAAAGTGGAGTTGATTGCGTACTTCGAAAAGGTAGGCGAC

R133S Rin	GAAGTACGCAATCAACTCCACTTTAGAGCTAAAGGCTTTTC
R133S Fin	AGCTCTAAAGTGGAGTTGATTGCGTACTTCGAAAAGGTAGGCGAC
pura F	ATGGCGGACCGAGACAGCGGC
pura R	ATCTTCTTCCCCTTCTTCCTCACCCTG
purβ F	GCGGACGGCGACAGCGGC
purβ R	ATCCTCATCCACCTCCTCACCCTC
hnrnp AB F	ATGTCGGACGCGGCTGAGGAGC
hnrnp AB R	GTATGGCTTGTAGTTATTCTGATGAC
hnrnp D F	ATGTCGGAGGAGCAGTTCGGAG
hnrnp D R	GTATGGTTTGTAGCTATTTTGATGTCCACCTCG
hnrnp AB R1 F	ATGTCGGACGCGGCTGAGGAGC
hnrnp AB R1 R	CATAGCCATAGCCTTTTTAGGGTC
hnrnp AB R2 F	ATGGCTATGAAGAAGGACCCTG
hnrnp AB R2 R	GTATGGCTTGTAGTTATTCTGATGAC
hnrnp D R1 F	ATGTCGGAGGAGCAGTTCGGAG
hnrnp D R1 R	TCTTTTGTTTTCATGGCTTTG
hnrnp D R2 F	GAGCCTGTCAAAAAAATTTTTG
hnrnp D R2 R	GTATGGTTTGTAGCTATTTTGATGTCCACCTCG

The C-terminally 6-His tagged recombinant proteins were expressed in BL21 strain of E. coli (Invitrogen) in Overnight Express Autoinduction System (EMD Millipore, Billerica, MA, USA), purified using Ni-NTA Spin columns (Qiagen) according to manufacturer's recommendations and dialyzed against Buffer B supplemented with 10% glycerol. The protein concentration was measured by Quick Start Bradford Assay (BioRad).

Electrophoretic mobility shift assay (EMSA)

Unless stated otherwise, 1 pmol of MeCP2 (aa 1-205), 1.2 pmol of MBD1, 0.5 pmol of MBD2, 50 pmol of MBD3 or 2.5 pmol of MBD4 was incubated with 10 fmol of ^{32}P-end-labeled C, 5mC or 5hmC BDNF probe in presence of 1 μg of pdIdC in Buffer B supplemented with 5% Ficoll for 30 min at room temperature (RT). 10 µL of the binding reaction was electrophoresed in native 6% 29:1 acrylamide:bis-acrylamide gel in 0.5 % TBE buffer at 4°C for 2-4 h. The gels were then dried in vacuum (Fisher Scientific, Hampton, NH, USA) for 1 h at 70°C, exposed to a storage phosphor screen (Molecular Dynamics) and scanned by Typhoon Imager (GE Healthcare). For assays with crude nuclear extracts, concentrations ranging from 60 ng to 2 µg were first incubated with 5-10 µgs of polydIdC for 10 min at RT and then with 2.5-10 fmoles of radioactive probe. For heparin fractions, a total of 4 µL of the pool, 2.5 µg polydIdC and 2.5 fmoles of probe was used. For size-exclusion and affinity column fractions, a total of 20 µL of the pool was used with 0.5 µg polydIdC and 5 fmoles of probe.

Surface Plasmon Resonance

Four flow cells (Fc) of Streptavidin (SA)-Sensor Chips (GE Healthcare) were cleaned with 3, 1-min pulses of 1M NaCl in 50mM NaOH in Biacore 3000 (GE Healthcare). 5'-biotinylated C, 5mC and 5hmC BDNF probes were injected onto Fc 2, 3 and 4, respectively, at 2 µL/min in Buffer B until SPR response increased by 500(+/-25) Resonance Units (RU). Fc 1 was immobilized with 2 mM

Biotin to serve as reference cell. 8-12 serial dilutions of protein samples within 0.01 Kd to 10 Kd range in Buffer B supplemented with 0.02% Tween and 50 ng/µL poly-dIdC were injected onto each Fc in parallel at 30 µL/min for 3 min. The surfaces were regenerated by 1 min injection of Buffer B with 1.5 M KCl (Buffer C). The change in SPR response with respect to reference cell 5 seconds before the end of injection was recorded for steady state analysis. Each run was performed in duplicates and each experiment was repeated with at least 4 samples of each protein preparation on two independently prepared surfaces. Bmax values and their standard error were automatically extracted in GraphPad Prism Software from steady state values and they were divided by the corresponding protein size. The F test to compare variances was performed using GraphPad Prism software.

The new generation SPR analysis was performed using ProteON XPR36 protein interaction array system (Bio-Rad). The NLC Sensor Chip (Bio-Rad) was preconditioned according to manufacturer's instructions. biotinylated DNA probes, prepared as described above, at 10 ng/µl were applied vertically at 30 µl/min for 5 min or until the resonance of each surface reached 550 (+/-40) RU. One channel was immobilized with 2 mM Biotin to serve as reference cell. 15 serial dilutions of protein samples within 0.01 Kd to 10 Kd range in Buffer B supplemented with 50 ng/µL poly-dIdC were injected onto each Fc in parallel at 100 µL/min for 2 min. The surfaces were regenerated by 18 sec injection of Buffer B with 1.5 M KCl (Buffer C).

Southeastern Blot

18 million nuclei from cerebella of WT or KO mice were resuspended in buffer containing 10 mM Tris pH 8.0 and 150 mM NaCl supplemented with 5mM $CaCl_2$, 10 mM Na-butyrate, 0.15 mM spermine, 0.5 mM spermidine and Halt Protease and Phosphatase Inhibitor Cocktail (Pierce). The nuclei were partially digested by the addition of 100 Kunitz units of Micrococcal Nuclease (NEB) and aliquots of the reaction were stopped by the addition of excess amount of EGTA at various time points. The nuclei were then incubated with RNase A/T1 Mix (Fermentas) lysed in 1% SDS and incubated with Proteinase K for at 37°C for 30 min and 65°C for 30 min. The DNA was extracted with phenol: chloroform: isoamyl alcohol (25:24:1), precipitated in 70% ethanol and dissolved in Tris-EDTA buffer. 0.5-1 μg of DNA was resolved in 2% TAE agarose gel and transferred to nitrocellulose membrane (Whatman plc, Maidstone, Kent, UK) overnight by capillary action in 20X SSC buffer (3M NaCl, 300mM Na-Citrate) following denaturation in 0.4 M NaOH and 0.6 M NaCl and neutralization in 0.5 M Tris (pH 7.5) and 1.5 M NaCl of the agarose gel. The membrane was equilibrated in 6X SSC buffer, air-dried, UV-crosslinked, blocked in 5% fat-free milk in TBS buffer with 0.2% Tween and incubated with rabbit anti 5hmC antibody (1:500, Active Motif) or mouse anti 5mC antibody (1:250, Diagenode, Denville NJ, USA) for 2 hours at room temperature. The membranes were washed for 3 X 5 minutes and incubated with a secondary antibody conjugated with horse radish peroxidase (1:4000, Abcam) for 30 min at room temperature. The blots were

washed for 3 X 10 minutes, rinsed with water and then reacted with ECL substrate (Pierce). The signal was detected by exposing reacted blots to chemiluminescence films (Sigma Aldrich) that were developed in film processor (Konica Minolta, Osaka, Japan). The gel images from five gels from three independent experiments were analyzed in ImageJ software. The average values, standard error of the mean and p values by unpaired t-test were calculated via GraphPad Prism Software. We refer to this method as Southeastern blot, since it combines the transfer of an electrophoretically separated DNA sample onto a membrane as in classical Southern blot technique with the assaying of chemical modifications as in Eastern blot technique

Fast protein liquid chromatography

All fast protein liquid chromatography (FPLC) experiments were performed in ÄKTAFPLC system (GE Healthcare). A 5mL HiTrap Heparin HP column was attached to the system after washing the system with MilliQ water and Buffer C , and then the upstream of column with running buffer, Buffer B. The flow started at 5 ml/min, 0% B for at least 10 min, such that the column was equilibrated with ten-fold volumes of running buffer. A 5 mL injection loop was attached and also equilibrated with 10 volumes of running buffer. The concentrated and precleared extract was injected manually into the loop, the flow was lowered to 2 mL/min the flow path was changed to “inject”. 2 mL fractions were collected throughout the run. After 6 volumes of running buffer passed through the column for 15 min, a gradient elution process started such that the concentration of Buffer C increased

from 0% to 100 % in 45 min. The run was stopped when the system was saturated with 100% buffer C and the fractions were assayed by EMSA.

For size-exclusion chromatography, Superdex 200 HR 10/30 (GE Healthcare) was used. The system was washed with MilliQ water prior to attaching column, and column was equilibrated at 0.5 mL/min with 2 volumes of MilliQ water (50 mL) and 2 volumes of Buffer B, adjusted to KCl concentration equivalent to that of the fractions. A 250 µL injection loop was also attached and washed with 2 mL of this buffer. The heparin fractions containing the specific activity were then pooled, concentrated to a total volume of less than 250 µL, precleared and injected manually. The run was let to proceed at 0.2 mL/min until one volume buffer B was allowed to flow through. 0.5 mL fractions were collected.

For DNA-affinity column, 1 mL HiTrap Streptavidin HP column was attached to the system after washing the system with MilliQ water and Buffer Bind & Wash (BW), containing 5 mM Tris·HCl pH 7.5, 0.5 mM EDTA pH 8.0 and 1 M NaCl. The column was equilibrated with ten-fold volumes of buffer BW at 1 mL/min. A 1 mL injection loop was attached and also equilibrated with 10 volumes of buffer BW. 100 µg biotinylated DNA probe containing C, 5mC or 5hmC (described above) was injected manually. The binding was allowed to proceed at 0.05 mL/min for 20 min and additional volumes of buffer BW was run through the column. Next, the column was detached from FPLC system. The system was washed with MilliQ water and Buffer C, and then the upstream of column with Buffer B. The column was re-attached and was equilibrated with ten-fold volumes of running buffer at 0.5 ml/min. The size exclusion fractions containing the specific activity were

pooled, concentrated to a total volume of less than 1mL, precleared and injected manually, after the injection loop was also equilibrated with 10 volumes of buffer B. The flow was lowered to 0.4 mL/min the flow path was changed to “inject”. 1 mL fractions were collected throughout the run.

Mass spectrometric identification of isolated proteins.

MS analysis was conducted by the Proteomics Resource Center at The Rockefeller University, New York, NY. Protein bands were reduced and alkylated and hereafter subjected to in-gel trypsinization following published protocol. (Shevchenko, Wilm et al. 1996). Post digestion, peptides were extracted and separated by reversed phase based nano flow liquid chromatography (Dionex, Boston, MA, USA) connected to an Orbitrap XL mass spectrometer (Thermo, San Jose, CA, USA). Peptides were subjected to Tandem MS in CID mode. Tandem MS data were extracted using ProteomeDiscoverer v. 1.3 (Thermo, Bremen, Germany) and queried against the Rat International Protein Index (IPI rat) v 3.87 using MASCOT 2.3 (Matrixscience, London, UK). For each identified protein a rough measure of amount was calculated based on the average area of the three most abundant peptides (Silva, Gorenstein et al. 2006).

Animal protocols

All protocols involving animals were approved by the Rockefeller University Institute Animal Care and Use Committee (protocol number #08114) in accordance with the National Institute of Health (NIH) guidelines. C57BL/6J mice an. *Mecp2*$^{tm1.1Bird}$ hemizygous males were purchased from Jackson Labs at 4 weeks of age. Animals were sacrificed between 7 to 11 weeks of age for experimentation, unless otherwise indicated. All mice were maintained on a 12 h light/dark cycle and given *ad libitum* access to food and water.

Appendix A

A new method for transient, drug-dependent and cell-specific inhibition of global translation *in vivo*

A.I. Aim

This study tries to examine the requirement for protein synthesis in specific cell populations in the brain during memory consolidation. To achieve this, we engineered an endogenous translational inhibitor to become drug-inducible. We then generated genetically modified mice that carry this engineered construct under a conditional and ubiquitous promoter, such that the inhibitor will only be activated in cells that co-express the appropriate recombinase. In such mouse lines, that express the recombinase in specific cell populations, we can block protein synthesis temporarily and in a cell-specific manner. This way, the requirement of protein synthesis in variety of neuronal subpopulations for different stages and types of memory consolidation can be systematically tested.

A.II. Background: long term memory and protein synthesis

Memory is broadly defined as the capacity of an organism to encode, store and retrieve information. Since Ebbinghaus' seminal research in 19th century (Ebbinghaus 1885), clinical (Scoville and Milner 1957; Squire, Cohen et al. 1984; Zola-Morgan, Squire et al. 1986), anatomical (Squire, Amaral et al.

1990; Smith and Jonides 1997), genetic (Jones, Errington et al. 2001; Schmitt, Deacon et al. 2003) and molecular studies continue to investigate the involvement of neuroanatomical regions in different stages and types of memory formation (Gooney, Shaw et al. 2002). The current cortical memory consolidation scheme starts with the reciprocal strengthening of hippocampal-cortical connections by the simultaneous activation of these areas (Marr 1970; Sutherland and McNaughton 2000; Wittenberg, Sullivan et al. 2002). Then, the intracortical connections strengthen until the memory trace becomes hippocampus-independent and stable (McClelland, McNaughton et al. 1995; Graham and Hodges 1997; Dudai 2004).

The first clues that *de novo* protein synthesis might be necessary for long-term memory (LTM) formation came about in the 1960s, when chemical protein synthesis inhibitors blocked LTM leaving the short-term memory (STM) intact (Flexner, Flexner et al. 1963). Upon LTM-inducing training events, enhanced translation activity (Martin, Casadio et al. 1997) and activated signaling events that trigger gene expression (Bartsch, Casadio et al. 1998; Malleret, Haditsch et al. 2001) were recorded. Although current techniques provide compelling evidence on mechanisms of memory formation; some are invasive (Alberini 2008; Rudy 2008) and may result in convoluted phenotypes, while noninvasive imaging techniques do not provide sufficient molecular details. In addition, data in cell-type resolution is not yet available. By combining a genetic targeting with a switch by a nontoxic molecule, the strategy presented here offers an alternative, temporal, noninvasive and cell-specific way to solve questions raised in this field.

A.III. Background: mammalian protein synthesis

Protein synthesis is tightly regulated in mammalian cells at the level of initiation, that being the first and the rate-limiting step (Kaufman 1994). The 40S ribosome associates with eukaryotic initiation factors (eIF) 1A and eIF3 and a ternary complex, made of eIF2, initiator methionyl-tRNA and GTP; and forms the 43S pre-initiation complex. eIF4E, eIF4A and eIF4G together form the cap-binding complex, which recruits the ribosome complex on mRNA. 40S ribosome starts scanning the mRNA until it encounters a start codon, where 60S ribosome joins in and elongation starts while eIFs are sequentially released (Mathews, Sonenberg et al. 2007). eIF2 and eIF4E are primary targets for translational control by cellular mechanisms, in addition to cleavage of eIF4G and other eIFs by caspases or exogenous proteases (Fig A.1.A) (Holcik and Sonenberg 2005). Several 4E-binding proteins (4E-BPs) regulate the cap-dependent translation while phosphorylation of eIF2α by one of the four stress-induced protein kinases prevents the formation of the ternary complex and inhibits the general protein synthesis except some mRNAs that carry regulatory sequences as escape mechanism (Fig A.1.B) (Richter and Sonenberg 2005; Wek, Jiang et al. 2006).

A.IV. Background: effect of protein synthesis inhibitors in brain

GCN2 is the major eIF2α-kinase in the brain (Berlanga, Santoyo et al. 1999; Sood, Porter et al. 2000). *Gcn2* $^{-/-}$ mice show mpaired hippocampus-dependent forms of LTM leaving amygdala-dependent LTM unaffected (Costa-Mattioli, Gobert et al. 2005). Interestingly these mice have enhanced memory

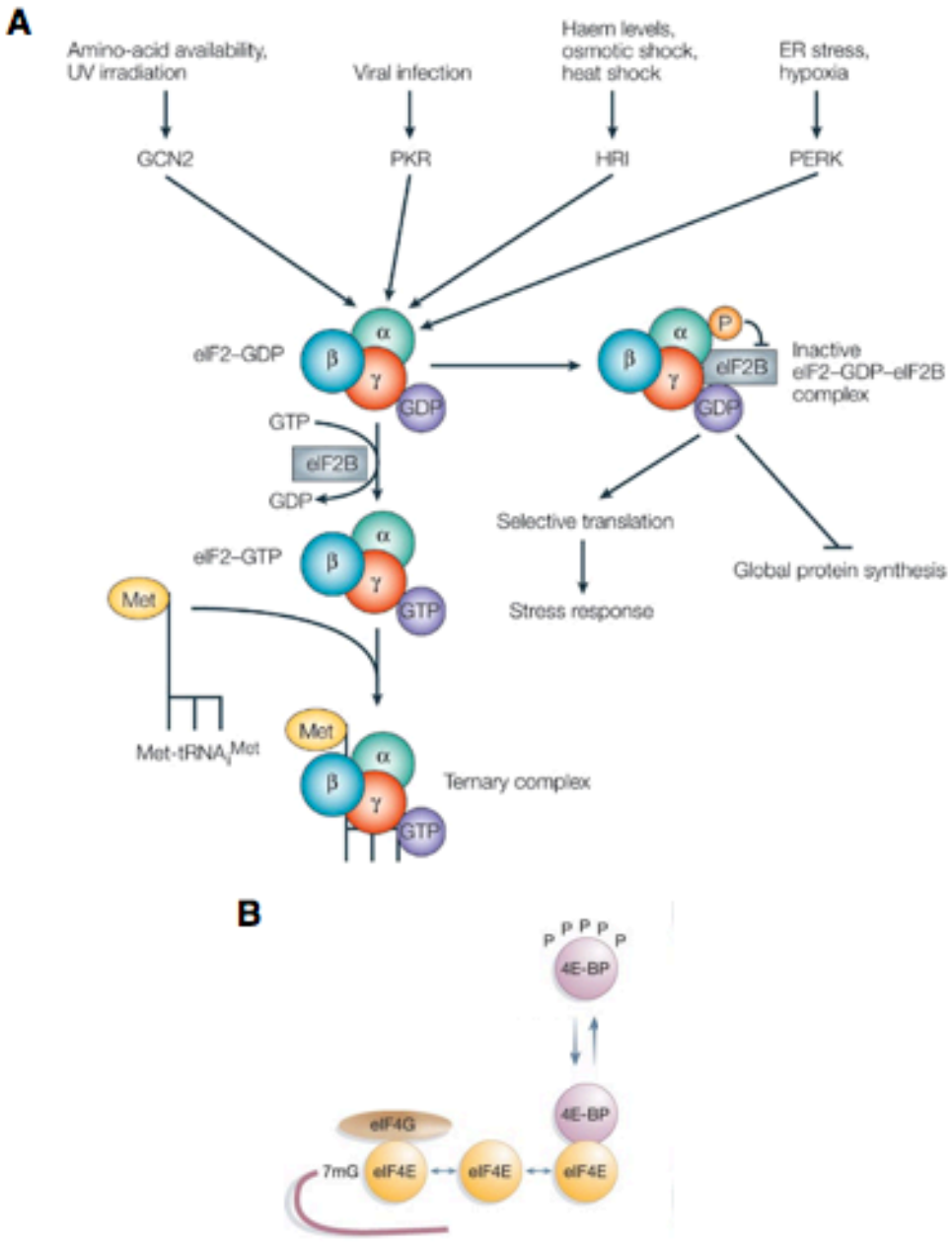


Fig A.1. Mammalian mechanisms for inhibition of translation. **A.** Schematic representation of inhibition by eIF2α-kinases (Holcik and Sonenberg, 2005) **B.**Schematic representation of cap-dependent translation (Richter and Sonenberg, 2005)

with weak training protocols. Similar results were obtained with eIF2α$^{+/S51A}$ mice that show decrease basal eIF2α phosphorylation (Costa-Mattioli, Gobert et al. 2007). Conversely, Sal003, an inhibitor of eIF2α-dephosphorylation, impairs LTM. Decreased hippocampal eIF4E phosphorylation was observed in another study shortly after training (Kelleher Iii, Govindarajan et al. 2004). Two of the eIF4E-BPs are inactivated by induction of LTM (Banko, Hou et al. 2006; Gelinas, Banko et al. 2007). Hippocampal slices from *4E-BP2* $^{-/-}$ mice had lowered threshold for LTP. Interestingly in these mice the hippocampus-dependent LTM was impaired (Banko, Poulin et al. 2005) whereas the cortex-dependent memory was enhanced (Banko, Merhav et al. 2007).

Among the four eIF2α-kinases (Fig A.1.A), RNA-dependent protein kinase or protein kinase R (PKR) inhibits global translation in response to viral infection (Holcik and Sonenberg 2005; Wek, Jiang et al. 2006). There are two ways that activate PKR: Homodimerization of N-terminal regulatory region upon binding viral dsRNA and subsequent autophosphorylation of the kinase domain; or the cleavage of the regulatory domain from the constitutively active kinase domain (aa 252-551; PKRkin) by virally induced caspases. Dimerization of PKR can be chemically induced by replacing the N-terminal domain of PKR with a well established FKBP domain, which can be dimerized via a synthetic ligand, AP20187. This method can effectively block protein synthesis in cultured cortical neurons, in parallel with spine growth and long term potentiation (Je, Lu et al. 2009). Yet this study was reproduced *in vivo* studies.

A.V. Background: design of site specific cleavage of PKRkin

In theory, the constitutively active PKRkin can be engineered to be a target of a selected protease by replacing an endogenous 13-residue protease-sensitive site (aa 338-350), that separates residues that make up the catalytic site (Dar, Dever et al. 2005), by the target sequence of that protease, if the protease does not have other targets in the host organism. The non-structural protein 3 (NS3) form Hepatitis C Virus (HCV) is a small monomeric protein with an unusual but well characterized substrate specificity (Steinkuehler, Urbani et al. 1996). It has previously been expressed in mammalian cells without noticeable toxicity and it is the primary drug target in the HCV (Huang, Murray et al. 2006).

In HVC, NS3 forms a complex with a 14-residue strand of its cofactor, NS4A protein, and together they process the viral polyprotein into individual components. A recombinant protease, NS3/4A, was made by fusing the N-terminal proteolytic domain of NS3 to NS4A with a flexible linker.This fusion stabilizes the structure and the active site, and substantially increases both the half-life and the enzymatic efficiency (Kim, Morgenstern et al. 1996; Taremi, Beyer et al. 1998). Among the cleavage sites of NS3 on HCV polyprotein, the NS5A/B junction (DTEDVVCC`SMSY) shows highest efficiency. In a recent study, NS3/4A was used as a tool to tag newly synthesized proteins (Lin, Glenn et al. 2008). The authors expressed NS3/4A fused to the C-terminal of the protein of interest (POI) and N-terminal of a tag, fused via the cleavage site. This way NS3/4A cleaved itself off from both the POI and the tag. In the presence of the inhibitor, the POI was expressed as fusion with the tag and NS3/4A. In mice,

there are two identified targets of NS3 protease, anti-viral response proteins TIR-domain-containing adapter-inducing IFN-β(Li, Sun et al. 2005) and mitochondrial antiviral signaling protein (Johnson, Owen et al. 2007), meaning that in transgenic mice with this construct may have compromised immune response.

NS3/4A is an important candidate for new generation drug development. Several drugs have been designed in the recent years and few even were approved by Food and Drug Administration as HCV drugs (Romano, Ali et al. 2012). A product-derived inhibitor of NS3/4A, BILN 2061, with sub-nanomolar inhibition constant, reduces the viral load below detection limit within 48 hours in clinical trials (Yee, Farina et al. 2006). The half-life of this complex is estimated in the order of seconds (Bartenschlager, Lohmann et al. 1995). The development of the drug was stopped in Phase II trials due to cardiotoxicity when administrated in high doses over extended periods (Reiser, Hinrichsen et al. 2005).

A.VI. Background: design of the expression system in mice

In this method, NS3/4A, PKRkin, engineered to be a target of NS3/4A, PKRkin*, and eGFP were expressed in a single mouse as a multicistronic construct (Fig A.2.A), each pair flanking a self-cleaving 14-residue proteinase, 2A, from foot-and-mouth disease virus (FMDV) (Donnelly, Gani et al. 1997). The cleavage mediated by this protein is shown to take place during translation by ribosome-skipping (Donnelly, Luke et al. 2001). Once 2A is synthesized it cleaves itself off from the growing peptide chain fused to the upstream and the translation of the downstream protein resumes following the cleavage. 2A is shown to

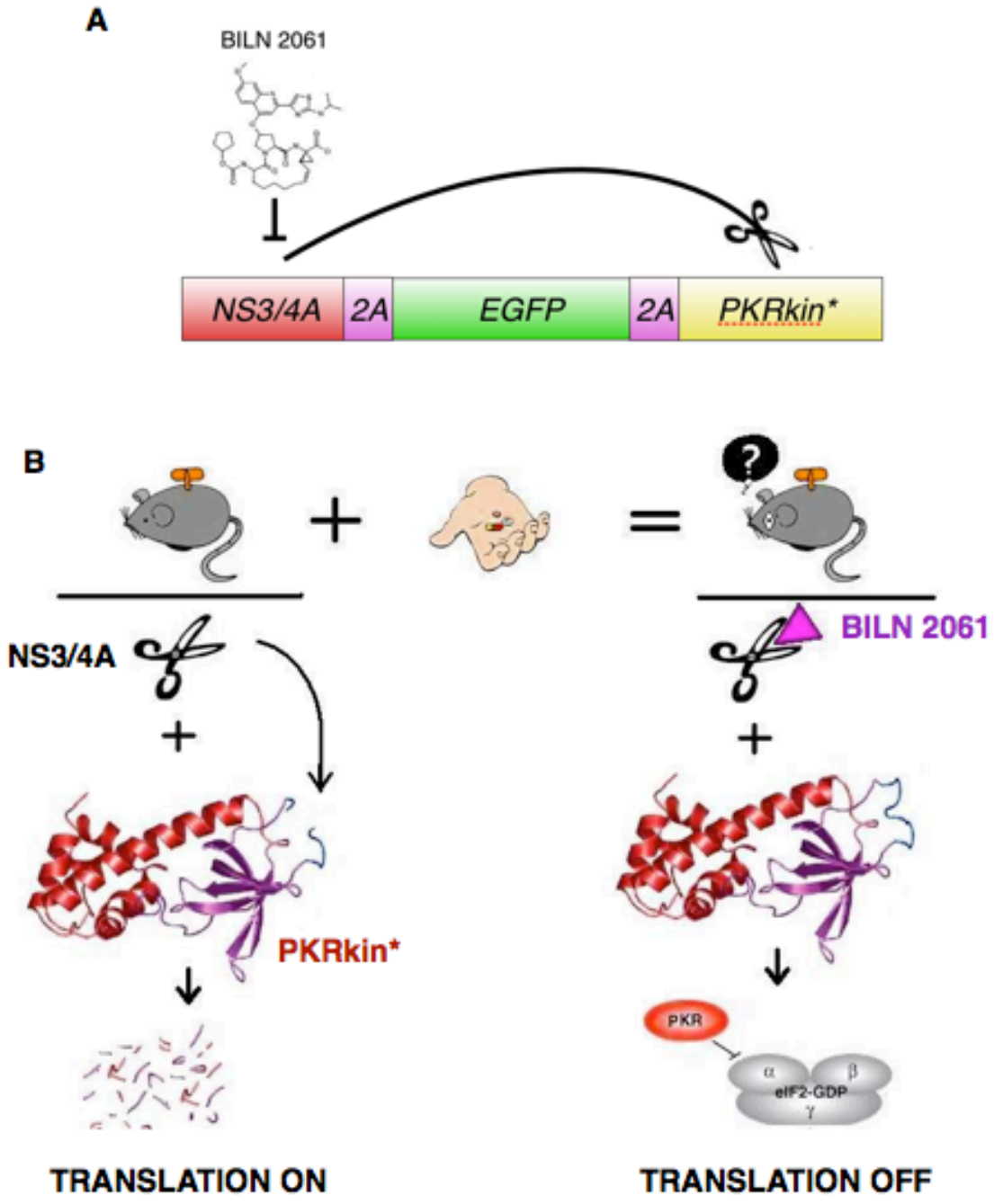


Fig A.2. Schematic representation of method. **A.** Structure of the multicistronic construct. **B.** If the construct is expressed in a cell typed that is required for the memory type tested, upon application of BILN 2061, translation will be suppressed in these cells and memory formation will be impaired

function exclusively in cis and it has no detectable toxicity. The cleavage efficiency increases from 80% to 99% when 2A includes its upstream 39 amino acids from FMDV polyprotein. This protein has been successfully used in several studies as a tool to make multicistronic constructs (de Felipe, Martin et al. 1999; de Felipe 2004). Alternatively NS3/4A and PKRkin* could be expressed as a fusion protein with a flexible linker, allowing proximity of the enzyme to its substrate and 1:1 ratio of enzyme and its substrate. NS3/4A in fusion with its substrate has been successfully used *in vivo* (Lin, Glenn et al. 2008).

In order to make this construct conditionally expressed in specific cell types, a variation of the well established genetic scheme involving Cre/loxP recombination was chosen (Sauer 1998). Briefly, in this method, the mice are double transgenic with a Cre recombinase that can excise a sequence flanked by 13 bp long Cre recognition sites, called loxP, or “floxed”; and with the gene of interest that is preceded by a floxed “STOP” cassette, made of a selectable marker followed by an SV40 early polyadenylation signal sequence, a false translational start and a splice donor signal (Lakso, Sauer et al. 1992). The *Eef1a1* locus, encoding a translation elongation factor, was chosen as a recipient site for transgenic knock-in since it is ubiquitously expressed at high levels and shows no phenotype upon deletion of one alleles (Klinakis, Szabolcs et al. 2009). The final construct was designed of 5’ and 3’ homology regions of the first intron of *Eef1a1* gene locus, a splice acceptor site, a floxed STOP cassette and and the multicistronic construct. The rational of the strategy is shown in Fig A.2.B.

A.VII. Results: NS3/4A cleaves PKRkin* but not PKRkin

To test the efficiency of proteolysis in mammalian cells, two NS3/4A target sequences, junction sequences from NS5A/B (DTEDVVCC`SMSY) or NS4A/B (DEMEEC`ASHLPY) were inserted into a fusion of eGFP and GST. These were then expressed in 293T cells in a 2A-multicistronic construct containing NS3/4A. The day following transfection, when equal amounts of cytosolic extracts were assayed with anti-GFP antibody the cleavage of the NS5A/B junction sequence was observed to be the most efficient (Fig A.3.A). This also shows that the 2A had very high cleavage efficiency, such that the polyprotein was absent in the blots. We next asked whether the expression of genes depends on their position in the multicistronic construct. In shuffled constructs, we noticed that the gene at the first position is always expressed at higher levels than at other positions (Fig A.3.B. left panel), while other positions did not affect the expression levels (Fig A. 3.B. right panel).

The cells were then transfected with multicistronic construct containing PKRkin* with NS5A/B junction site. The extracts blotted with an anti-PKR antibody show that significantly lower PKRkin* levels in the presence of NS3/4A, whereas PKRkin was not affected (Fig A.3.C). At any concentration tested, BILN 2061 was able to induce the expression of PKRkin*, but did not affect the levels of PKRkin* in the absence of NS3 or wild type PKRkin. 1 uM BILN 2061 was applied to cells at various times and cells were lysed simultaneously. Western blot analysis

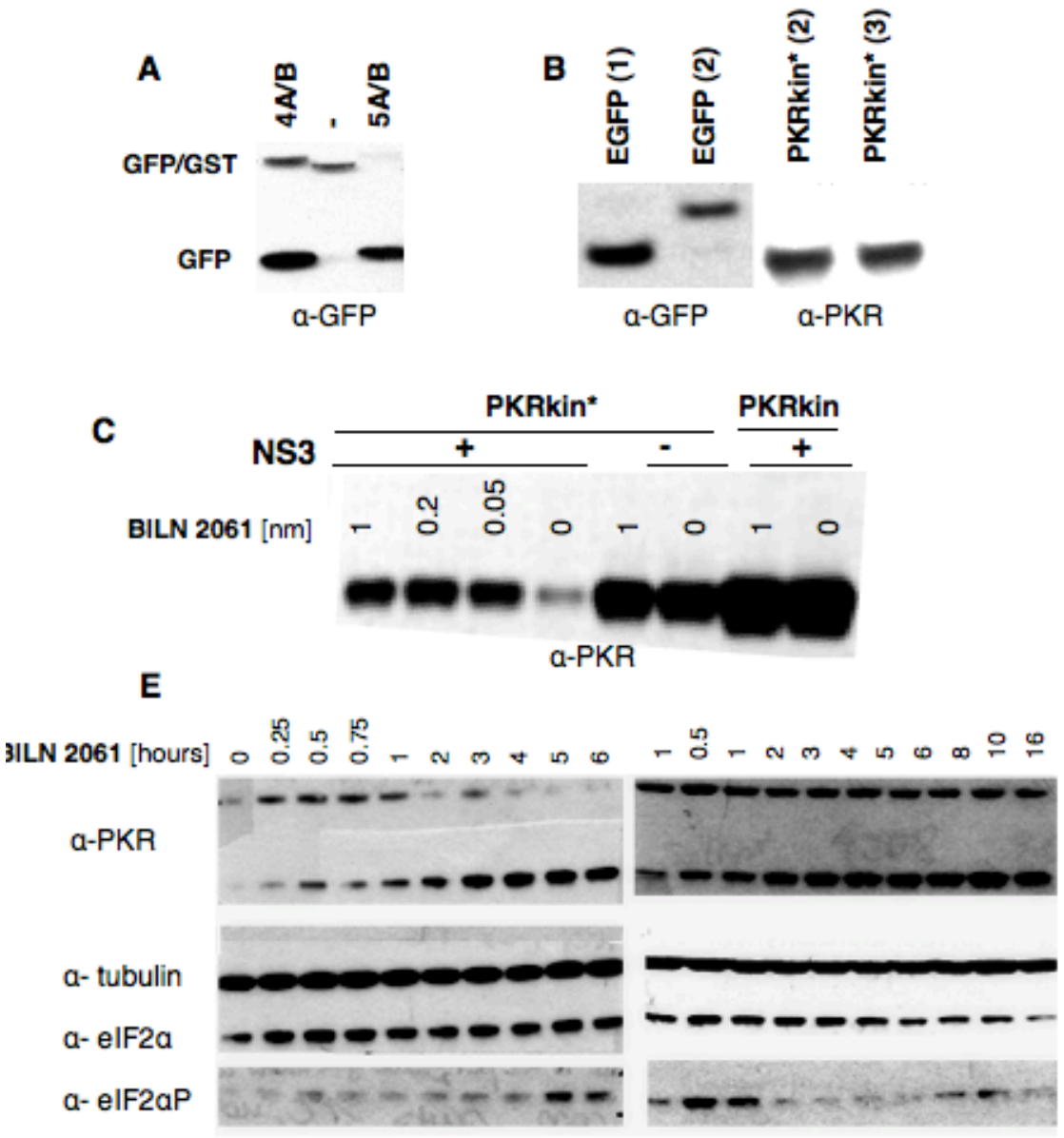


Fig A.3. PKRkin* is cleaved by NS3 and induced by BILN 2061. **A.** Western blot showing GFP/GST fused via indicated cleavage sites on top. Detection antibody is indicated in A-E with α. **B.** Western blot showing the recombinant proteins and their position in the multicistronic construct is shown in parenthesis. **C.** Western blot with PKRkin or PKRkin*; in presence or absence of NS3 after overnight application of different concentrations of BILN 2061. **D.** Western blot indicated hours after application of 0.1 nm BILN 2061. α- eIF2αP is phospho-specific antibody.

showed that by 3-4 hours the PKR levels plateau (Fig A.3.D) which phosphorylates its target eIF2α.Cells expressing PKRkin were observed to remain healthy for at least 3 days.

A.VIII Results: PKRkin* inhibits translation in absence of NS3/4A

To test the translation inhibition upon drug administration, the multicistronic construct with PKRkin*, NS3/4A and eGFP was co-transfected with a *Luciferase* gene from firefly, *Luc2p*. At all concentrations tested, 16 hours after application of BILN 2061, PKRkin* was able to suppress Luciferase activity by more than 10-fold compared to a control (Fig A.4.A). PKRkin* had increased activity with respect to wild-type PKRkin, possibly due to some adaptation or resistance mechanisms taking place in the cell, including inherent susceptibility of the wild-type PKRkin to endogenous proteases.

To test the inhibition of global translation, newly synthesized proteins were metabolically labeled in a medium containing ^{35}S after starvation. In cells transfected with PKRkin* and NS3/4A, translation rate was comparable to that of wild-type cells, whereas in the absence of protease translation rate decreased substantially (Fig A.4.B,C and D).

Although metabolic labeling is a robust way of assessing translation rates in cultured cells, when this method is applied *in vivo* in subpopulations of cells, it is no longer applicable. This necessitated a cell-specific assay of translation rates. Therefore we used eGFP/L10A reporter-ribosomal fusion. This way not only the inhibition of translation can be directly assessed by eGFP levels, but also the

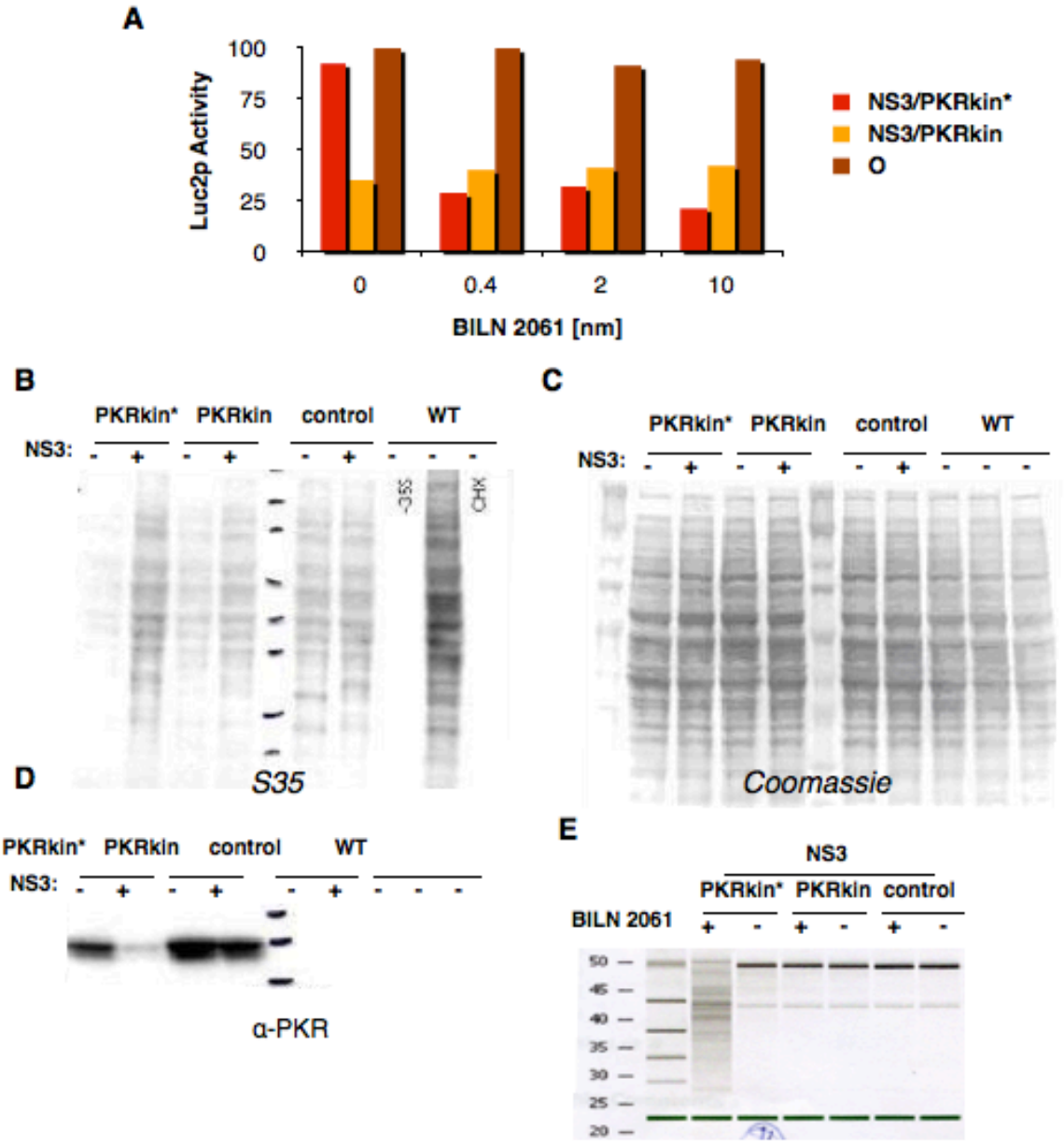


Fig A.4. PKRkin* inhibits translation in response to BILN 2061. **A.** Luc2p activity normalized to the cell with *Luc2p* vector co-transfected with control vector (0) or others. **B.** Membrane containing cell lysate expressing indicated proteins. Last three lanes are untransfected cells, where last lane is treated with Cycloheximide (CHX), third from the last does not contain S35. **C.** Coomassie stained gel as in B. **D.** Western blot of membrane from B with α-PKR antibody. **E.** RNA quality of BILN 2061 treated cells expressing indicated proteins.

gene expression profile at different time points and between different mouse lines can be compared, complementary to the information obtained by inhibition of translation. We purified RNA from transfected cells with eGFP/L10A as reporter, using antibodies against eGFP. High quality of RNA was observed in every sample, except when BILN2061, NS3/4A and PKRkin* were present (Fig A.4.E)., indicating lack of eGFP/L10A protein.

Since *in vivo* the effect will vary depending on the cell type this is expressed and the level it is expressed, we sought to design other constructs with varying induction levels. To improve kinetics we produced a fusion of NS3/4A/PKRkin* and added more target sites to other loops of PKR (a-c in Fig A.5.A). To make a self-inducing system we introduced regulatory sequences of genes that are upregulated by PKR, such as *Atf4*, upstream of PKR (Blais, Filipenko et al. 2004). In parallel we also tested the less efficient NS4A/B cleavage site and used the less active forms of the protease, NS3 alone and a T54A mutant NS3 that has compromised proteolytic activity (Lin, Glenn et al. 2008). The drug-induced inhibition of translation by some of these constructs is shown in Fig A.5.A. In the cells, transfected with the NS3/4A/PKRkin* fusions carrying NS4A/B and NS5A/B cleavage sites, PKRkin* was efficiently cleaved (Fig A.5.B). These and the original multicistronic NS3/4A & PKRkin* with NS5A/B site were most suitable for genetic studies, since the drug-induction was the most efficient.

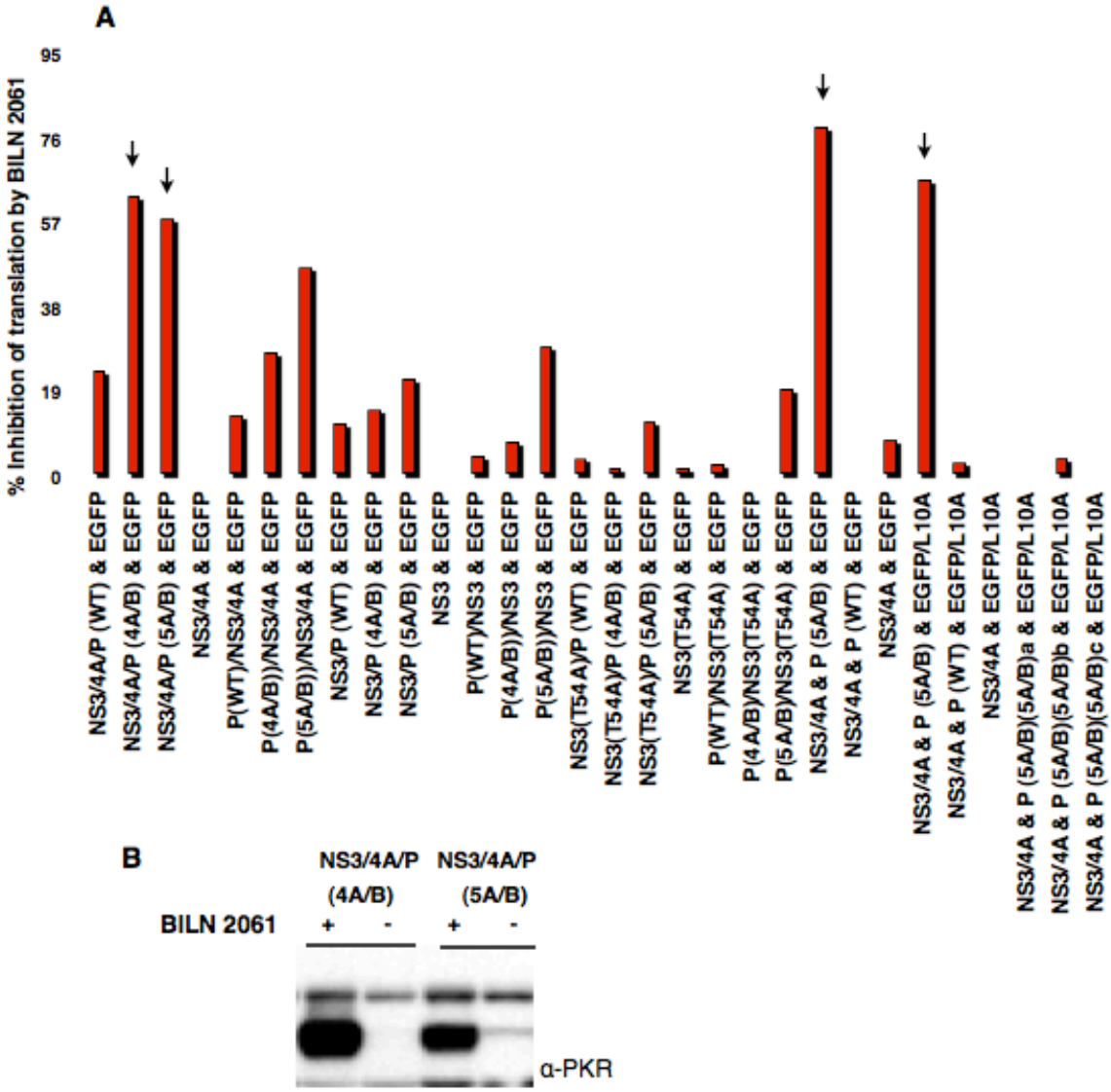


Fig A.5. Comparison of constructs. **A.** Percent inhibition of Luc2p activity normalized to the cells transfected with control vector to those with *Luc2p* vector. P (WT) is PKRkin with endogenous loop, P(4A/B) and P(5A/B) are PKRkin* with indicated cleavage sites. (5A/B) a-c are PKRkin* with 5A/B carrying extra cleavage sites. Arrows show constructs selected for knockin **B.** Western blot showing the fusion proteins in presence or absence of drug.

A.IX. Results: generation and verification of knockin mice

This construct was cloned downstream of a floxed STOP cassette and targeted to a ubiquitous and high expressing *Eef1a1* locus (Fig A.6.A) (Klinakis, Szabolcs et al. 2009). Knockin mice were generated, clones that were positive in genotyping were confirmed by Southern blotting (Fig A.6.B). These mice can now be bred to a mouse line with a specific Cre recombinase expression pattern. When these mice were bred to *Emx1* BAC-Cre mice (Schmidt, Warner-Schmidt et al. 2012), expected eGFP expression patterns was observed in brain sections without immunostaining (Fig A.7).

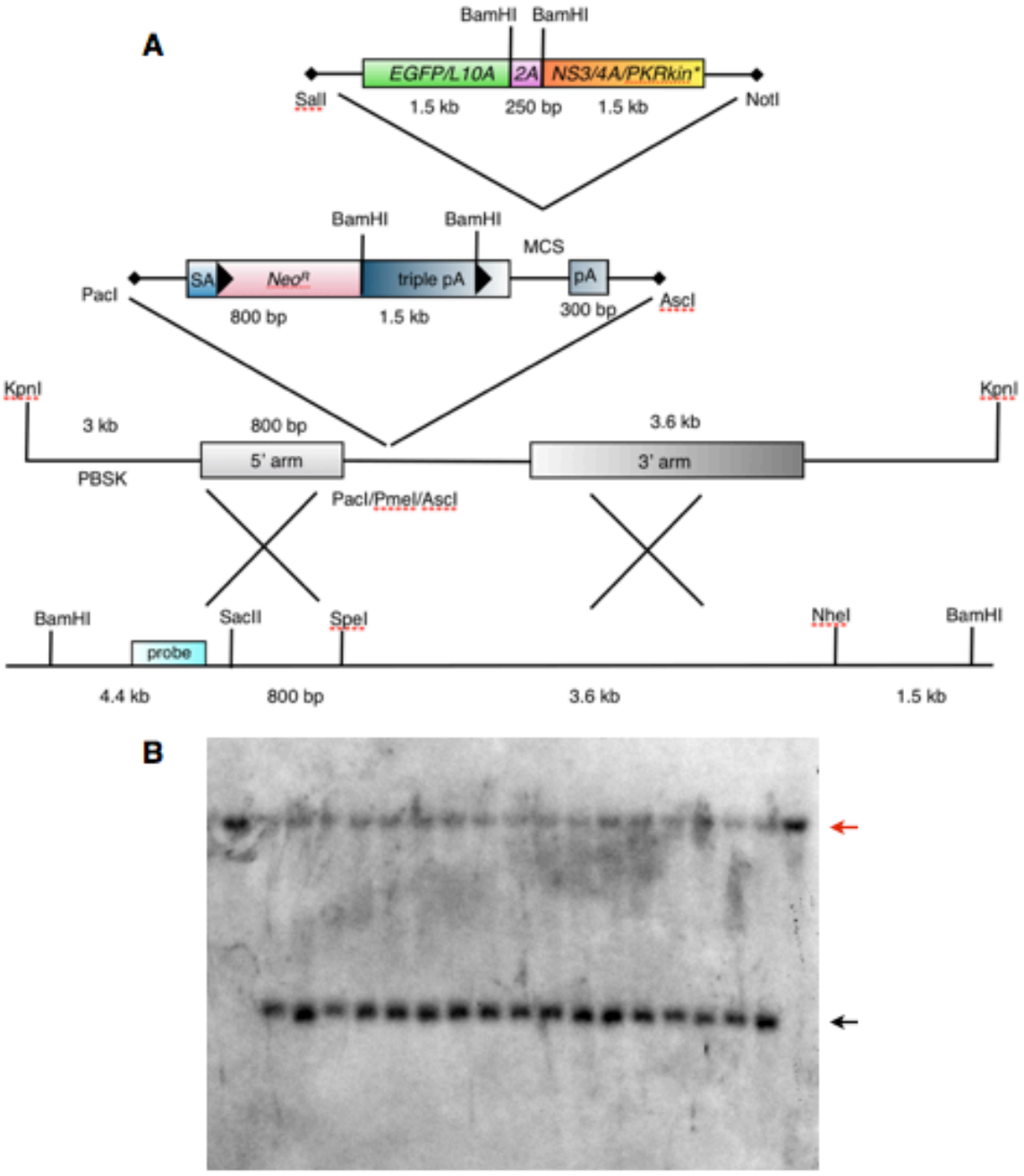


Fig A.6. Knockin of PKRkin*/NS3. **A.** Schematic representation of the targeting strategy (Klinakis, Szabolcs et al. 2009) **B.** Southern blot using the probe sequence shown in A. Genomic DNA of knockin mice digested with KpnI. First and last lanes are WT. WT bands shown with red arrow. Positive bands of 400 bp are shown with black arrow.

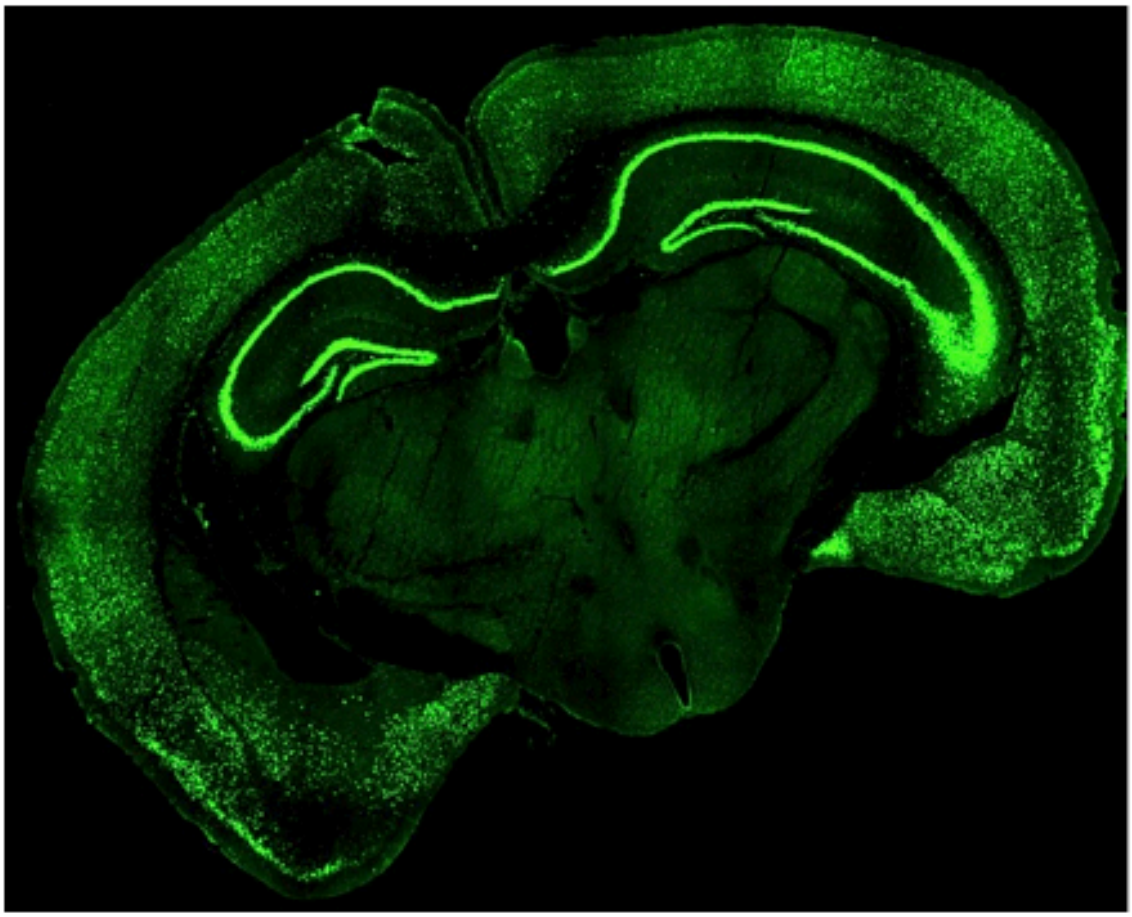


Fig A.7. Brain section from double transgenic mice. *Emx1* bacCre mice with NS3/4A/PKRkin* and eGFP/L10A/2A (green). Courtesy of Dr Christina Pyrgaki.

A.X. Discussion

This new way of modifying translation rates *in vivo* offers many advantages compared to the pre-existing techniques. First, by using mouse lines that express Cre recombinase in specific subpopulation, the requirement of protein translation can be tested cell specifically. Second, applying the drug at various time points, at various dosages and in specific brain regions the duration, level and neuroanatomical extent of inhibition can also be regulated. Conditional mice that carry constructs with different induction characteristics can also offer additional dosage regulation. Third, since none of the elements that are used in the system have invasive side effects, the observed effects are likely to be the direct result of blocking the protein synthesis in the target cell population. Finally, since the method is reversible and not invasive, a single animal can be tested over various stages in a behavioral paradigm. In addition, the eGFP /L10A tag allows direct assessment of gene expression in these cells in comparison to other tissues or over different time points. This method is devised to test whether protein synthesis is required in selected cell types for different types of memory consolidation. However they can be of multipurpose; for example in the case of tinnitus, ringing in the ears, which might be related to dysfunction in plasticity (Engineer, Riley et al. 2011).

A.XI. Materials and Methods:

Chemicals and construction of recombinant plasmids

The gene for HCV polyprotein was kindly provided by Dr. Charles Rice. The 2A gene was kindly provided by Dr. Martin Ryan. The PKR gene was a gift from Dr. Prerana Shresta and the eGFP/L10A was a gift from Dr Joseph Dougherty. eGFP and GST proteins were fused via control linker, or the junction sequences from NS5A/B or NS4A/B (DEMEEC`ASHLPY) by chimeric primers and overlap extension PCR. PKRkin* was produced the same way. The elements of the constructs were amplified using Phusion polymerase (NEB) and cloned into peGFP-N1 vector (Clontech) using appropriate restriction enzymes (NEB).

primer	sequence
PKRkin F	GAAACAAAGTATACTGTGGACAAGAG
PKRkin R	ACATGTGTGTCGTTCATTTTTCTCTG
PKRkin Rin*	GTCCAGGTGTAGCTCATGCTGCAGCACACCACGTCCTCGGTGTC
PKRkin Fin*	GCTGCAGCATGAGCTACACCTGGACCGGCAAGAGCAAAAATAG
2A F	GTCACCGAGTTGCTTTACCGGATGAAG
2A R	CGAGTCCAACCCTGGGCCCTTTTTTTTTAGT
NS3/4A F	GGCAGCGTGGTCATTGTGGGCAGGATC
NS3/4A R	CTATGGAAACCACTATGCGGGGGGC
eGFP-L10A F	CAGCTCCTCGCCCTTGCTCACCAT
eGFP-L10A R	ATACAGACGCTGGGGCTTGCCCATG
PKRkin mF-in*	CGAGATGGAGGAGTGCGCCAGCCACCTGCCCTACAGCAAAAATAGTTCAAG GTCAAAGAC
PKRkin mRin*	GGTGGCTGGCGCACTCCTCCATCTCGTCACTGGTCTCAGGATCATAATC

PKRkin 1-Rin*	CTCATGCTACAGCACACCACGTCCTCGGTGTCATATTTAACACGTTTAATAAC GTAAGTCTTTCCGTCAATTC
PKRkin 1-Fin*	GACGTGGTGTGCTGTAGCATGAGCTACACCTGGACCGGCAACGAGAAGGCG GAGCGTGAAGTAAAAGCATTG
PKRkin 2-Rin*	GTAGCTCATGCTACAGCACACCACGTCCTCGGTGTCCTCGCCTCTTCTTTTT TCAATCCATTGTTC
PKRkin 2-Fin*	GTGTGCTGTAGCATGAGCTACACCTGGACCGGCCTAGACAAAGTTTTGGCTT TGGAACTCTTTG
PKRkin 3-Rin*	GTAGCTCATGCTACAGCACACCACGTCCTCGGTGTCATTTTTCAGAGATGTTA CAAGTCCAAAGTC
PKRkin 3-Fin*	GTGTGCTGTAGCATGAGCTACACCTGGACCGGCAAGCGAACAAGGAGTAAG GGAAC
Ns3/4A B-F	ATTACGGCCTACTCCCAACAGACGC
Ns3/4A B-R	CTATGGAAACCACTATGCGGGGGGC
Ns3/4A C-Fin	TCCTGGCGACCTGCGTCAATGGCGTGTGTTGGGC
Ns3/4A C-Rin	GCACCATGATAGACAGCCCAACACACGCCATTG
eGFP F	CAGCTCCTCGCCCTTGCTCACCAT
eGFP R	CTTGTACAGCTCGTCCATGCCGAGAGTGATC
GST F	TCCCCTATACTAGGTTATTGGAAAATTAAGGGC
GST R	ACGCGGAACCAGATCCGATTTTGGAGGATG
Rluc F	GCTTCCAAGGTGTACGACCCCG
Rluc R	CTGCTCGTTCTTCAGCACGCG

Purifications were done using PCR extraction kit (Qiagen). The ligations were performed with Quick Ligase Kit (NEB) and the products were transformed into chemically competent TOP10 cells (Invitrogen) that were grown in LB medium (Sigma) in presence of 35 mg/ml Kanamycin (Sigma). The plasmids were purified using a Miniprep kit (Qiagen). Every construct was verified by sequencing.

The plasmid containing the STOP sequence and the Eef1a1 targeting plasmid were kindly provided by Dr. Apostolos Klinakis. The mammalian expression constructs were sub-cloned into the gene targeting plasmids as previously reported (Klinakis, Szabolcs et al. 2009). The final construct was extracted with phenol: chloroform: isoamyl alcohol (25:24:1), precipitated in 70% ethanol and dissolved in Tris-EDTA buffer.

Cell culture and transfection

Unless otherwise stated the cell culture experiments were done the following way: 293T cells were grown in a 24-well plate in growth medium (Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen) to have 90% confluency the day of transfection. They were transfected with 1 ug of a reference plasmid using 2 uL 293Tfectin (Invitrogen) reagent. The amounts of the remaining plasmids were adjusted to have the same molarity per well. Cells were lysed on ice in phosphate-buffered saline supplied with 1% TritonX100 (Sigma) and inhibitors against proteases (78437, Pierce) phosphatases (78420, Pierce) and translation machinery (100 ug/ml cycloheximide, Sigma); and the cytosolic extracts were isolated after precipitating the insoluble fraction. The protein concentrations were measured using BCA protein assay (Pierce). BILN 2061 was kindly provided by Boehringer-Ingelheim Inc (Yee, Farina et al. 2006).

The Luciferase assay was performed according to instruction manual after cells were lysed Dual Luciferase assay kit lysis buffer (Promega).

Western blotting

For Western blotting the extracts were denatured in sample buffer (Invitrogen) under reducing conditions, separated on a pre-cast Bis-Tris SDS-gel (Invitrogen) via electrophoresis and transferred on a PVDF-membrane (Biorad). The membrane was blotted with primary antibodies, anti-GFP (gift of Dr. Brian Chait), anti-PKR (Abcam), anti-eIF2α (Cell Signaling), anti-phospho-eIF2α (Cell Signaling) and anti-tubulin (Abcam). The chemiluminescence was detected by the application of the ECL substrate (Pierce) on the appropriate horse-radish-peroxidase-conjugated secondary antibodies.

Metabolic Labeling

To visualize the inhibition of protein synthesis metabolic labeling experiments were conducted. 20 hours post transfection with 5 ug PKRkin* construct (with or without NS3/4A) and 10 uL reagent cells in 6-well plate were labeled with 10 uCi TRAN^{35}S label (MP Biomedicals) for 30 minutes following 30 minutes of starvation before the addition of the label into L-methionine and L-cysteine-free DMEM media (Invitrogen). One well was incubated with cycloheximide for 10 minutes prior labeling as a control for translation inhibition. Equal amounts of protein from extracts were separated by electrophoresis and transferred on a membrane. The membrane was dried in 100% methanol and analyzed for autoradiography, later it was blotted with antibodies and finally stained with Coommassie Plus stain (Pierce). Extracts containing equal amounts of protein were added into scintillation liquid (PerkinElmer) and the radioactivity

was measured via scintillation counter. The radioactivity was calculated taking that of the cells in presence of cycloheximide as zero and that of the wild-type cells as 100%.

Quality control of ribosomal RNA

Two custom made mouse anti-GFP antibodies, clones 19C8 and 19F7 (Heiman, Schaefer et al. 2008), were captured on Dynal magnetic beads (Invitrogen Corporation) coupled with protein L (Pierce). The precleared lysate from cells was incubated with antibody coupled beads at 4°C with end-over-end rotation for approximately 16 hours. Beads were subsequently collected on a magnetic rack, washed three times with high-salt wash buffer (10 mM HEPES [pH 7.4], 350 mM KCl, 5 mM $MgCl_2$, 1% NP-40, 0.5mM DTT, 100 µg/ml cycloheximide) and RNA was released and purified using Rneasy Micro Kit (Qiagen) with in-column DNase digestion. RNA quantity and quality were determined with a Nanodrop 1000 spectrophotometer (Wilmington) and Agilent 2100 Bioanalyzer using RNA 6000 Pico Chip.

Gene Targeting, Southern Blotting and Genotyping

Gene targeting, generation of knockin mice and Southern blotting was conducted by the Gene Targeting Facility at The Rockefeller University, New York, NY as previously described (Klinakis, Szabolcs et al. 2009). Following primers were used for genotyping. The PCRs were performed with ChoiceTaq (Denville) following manufacturer's instructions.

primer	sequence
β-actin F	GGCTGTATTCCCCTCCATCG
β-actin R	CCAGTTGGTAACAATGCCATGT
eGFP F	GCAGAAGAACGGCATCAAGGT
eGFP R	ACGAACTCCAGCAGGACCATG
GA5 F	GGACCTTTACTTGGTCACGAGGCATGCCGATG
GA5 R	TGCCGGTCCAGGTGTAGCTCATGCTGCAGC
GA4 F	GGACCTTTACTTGGTCACGAGGCATGCCGATG
GA4 R	CAGGTGGCTGGCGCACTCCTCCATCTCGTC
GA F	GGACCTTTACTTGGTCACGAGGCATGCCGATG
GA R	CTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCG
A2G5 F	CACCATGGGCAAGCCCCAGCGTCTGTATG
A2G5 R	TGCCGGTCCAGGTGTAGCTCATGCTGCAGC

Transgenic mice and visualization of eGFP expression

All protocols involving animals were approved by the Rockefeller University Institute Animal Care and Use Committee (protocol number #09024) in accordance with the National Institute of Health (NIH) guidelines. All mice were maintained on a 12 h light/dark cycle and given *ad libitum* access to food and water. The *Emx1*-Cre and *Ntsr1*-Cre mice were a gift from Dr Eric Schmidt. The eGFP/L10A expression was detected as previously described (Schmidt, Warner-Schmidt et al. 2012). Mice were deeply anesthetized and transcardially perfused with PBS followed by 4% paraformaldehyde. Brains were dissected and postfixed for 1-12 hr at 4°C, cryopreserved in 30% sucrose solution, sectioned on a freezing microtome (35 µm sections). All sections were imaged on a Zeiss LSM700 confocal microscope.

Appendix B

A new drug-inducible Cre recombinase

B.I. Aim

Cre recombinase strategy expanded biomedical research using genetically modified animal models, by introducing spatial control of gene modifications. However precise temporal control is still not ideal by the currently available strategies. Here we developed a drug-inducible Cre, that can be used in a plethora of *in vivo* gene modifications.

B.II. Background: Cre recombinase and transgenic strategies

Our ability to design genetically modified mice made previously impossible *in vivo* studies in mammalian organisms possible (Wasserman and DePamphilis 1993). This technology was advanced by the use of recombinase switches that are utilized to study lethal genes or to isolate phenotypes in adulthood or in specific cell populations (Sauer 1998)such as those of the brain (Gaveriaux-Ruff and Kieffer 2007; Gong, Doughty et al. 2007).

Cre site-specific DNA recombinase is a powerful tool for the design of such switches simply by inclusion of short recombination sites at the site of genetic modification. Cre is a 38 kDa product of bacteriophage P1 where is catalyzes the cyclization of P1 plasmid after injection (Sternberg, Sauer et al. 1986), by recognizing a 34 bp site, called *loxP* locus. Cre can recognize two *loxP* sites in opposite orientation and invert the sequence in between; however when it

encounters two parallel *loxP* sites, it excises the sequence flanked by these (floxed) and ligates them together.

Two strategies were developed that enable Cre to turn on or off genes of interest (GOI). To turn on genes by Cre, either GOI is reversed between two anti-parallel *loxP* sites, or a floxed sequence is cloned upstream of GOI such that the expression of GOI is disrupted (Gu, Marth et al. 1994). For this purpose, a floxed "STOP" cassette is used, that is made of a selectable marker followed by an SV40 early polyadenylation signal sequence, a false translational start and a splice donor signal (Lakso, Sauer et al. 1992). To turn genes off by Cre, an exon of GOI can simply be floxed in a way that it does not compromise the expression or function of the gene in absence of Cre.

Although the use of native gene expression profiles to express Cre, provides great tissue and cell types specificity, temporal control may not always be possible. To achieve this, drug responsive promoters have been used (Gossen and Bujard 1992). Elegant strategies were generated by fusing a mutated human estrogen receptor (ER^T) or progestrone receptor (PR^T) to Cre. In absence of inducer (tamoxifen or mifepristone, respectively) Cre-ER^T or PR^T is excluded from the nucleus, however when inducer binds ER^T or PR^T and then Cre can penetrate into nucleus (Metzger, Clifford et al. 1995; Brocard, Warot et al. 1997). Unfortunately these methods showed interference with endogenous signaling pathways, very high background activities or low efficiencies; or they required drugs at toxic levels (Brocard, Warot et al. 1997; Kellendonk, Tronche et al. 1999; Rossant and McMahon 1999; Garcia and Mills 2002).

An ideal inducible Cre recombinase system has three basic features: (1) no basal activity in absence of the inducer, (2) nontoxic inducer and (3) inducer that is effective in low doses. With that in mind we used the rational of the drug-inducible protein function, as explained in Appendix A (A.I.IV) for making a drug-inducible Cre. The locations to insert cleavage sites on the Cre recombinase were selected based on the disordered loops observed on the X-ray structure of Cre recombinase (Guo, Gopaul et al. 1997); by non-inactivating sites that were revealed by a previous pentapeptide insertional mutagenesis screen (Petyuk, McDermott et al. 2004); and by regions where non-homologous loops were found only in homologs of Cre. The selected sites are shown in Fig B.1.A.

To test the recombination efficiency using one plasmid, we cloned a floxed *Luc2p* after start codon and a Renilla *Luciferase* gene*, RLuc,* that is not expressed due to frameshift, such that upon excision of floxed region, *RLuc* is in-frame and therefore expressed (Fig B.1.B). We could easily measure the activity of each Luciferase by using a dual-Luciferase reporter assay (Sherf, Navarro et al. 1996) and quantify the recombination efficiency by the ratio of two Luciferases.

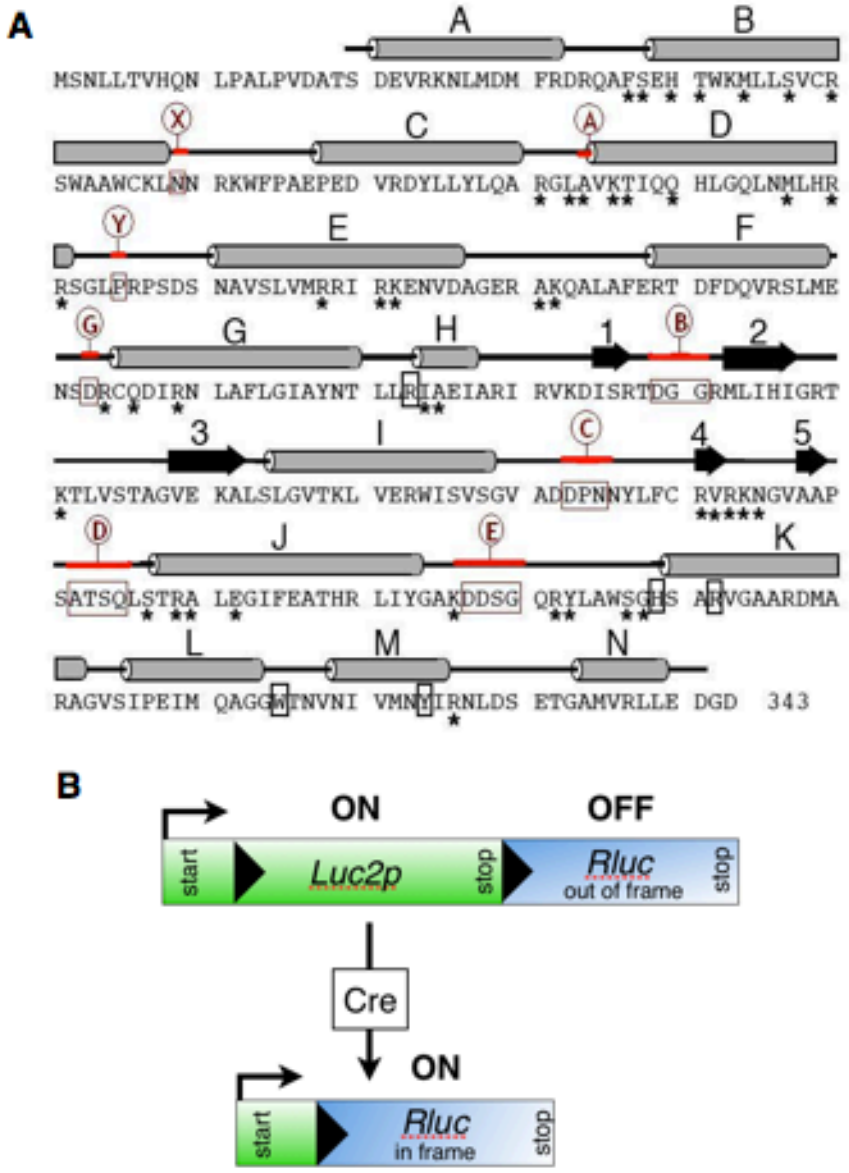


Fig B.1. Strategies for inducible Cre. **A.** Amino acid sequence of Cre recombinase. Marked with red are the sequences that were replaced by NS4A/B cleavage site **B.** Schematic representation of strategy used to quantify detection. In absence of Cre, *Luc2p* is on and *Rluc* gene is off, since it is not in frame. Upon Cre activity, *loxP* sites (black triangles) are recombined and *Luc2P* is excised, allowing Rluc to be expressed

B.III. Results: drug-inducible Cre recombinase

We have located 8 target sites to introduce the cleavage site as an insertion. After the initial screen by Luciferase recombination assay, we observed complete abolishment of Cre activity on some sites (A & D in Fig B.2.A) and no sign of cleavage on others . By making various combinations of cleavage sites we managed to generate several double and triple mutants that showed high induction levels (XY, XG and YGE in Fig B2A). We confirmed our observations in Western blots, as well (Fig B.2.B).

B.IV. Discussion

The drug-inducible Cre described his strategy has a great potential to improve the currently available inducible recombination strategies. With the availability of Cre mutants with a variety of background and induction levels, this technique offers a possibility of making transgenic with desired inducibility of Cre. This strategy still needs large scale *in vivo* screens of several transgenic lines, however the reward will be very useful for temporally controlled *in vivo* genetic manipulations.

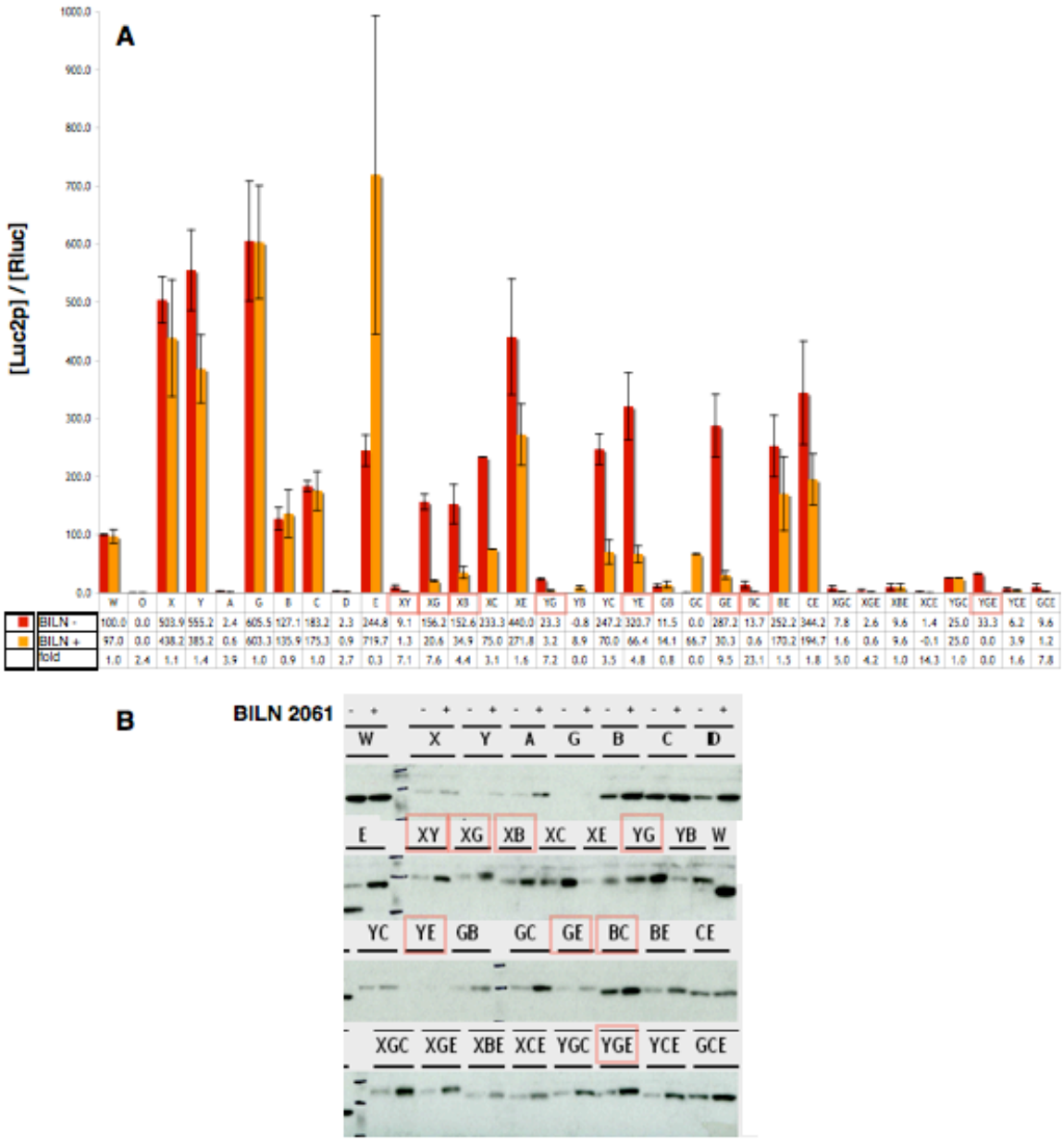


Fig B.2. Cleavage of Cre mutants by NS3A/B and induction of recombination by Cre. **A.** Bars represent ratio of Luc2p activity versus Rluc activity, in cells co-transfected with the *Cre* mutants and reporter plasmid, in absence (red) or presence (orange) of BILN 2061. Standard variation is shown. **B.** Western blot with α-Cre antibody showing cells expressing Cre mutants in presence or absence of BILN 2061. Mutants selected for *in vivo* testing are shown in red rectangles.

B.V. Materials and Methods

The gene for Cre recombinase was a gift of Dr. Shiaoching Gong. the plasmid containing the STOP cassette was from Dr Ana Domingos. For all protocols, see Materials and Methods from Appendix A.

primer	sequence
Cre F	ATGTCCAATTTACTTACCGTACACC
Cre R	CTAATCGCCATCTTCCAGCAGGCG
Cre X Rin	GTGTAGCTCATGCTGCAGCACACCACGTCCTCGGTGTCATTCAACTTG CACCATGCCG
Cre X Fin	GCTGCAGCATGAGCTACACCTGGACCGGCAAGAACCGGAAATGGTTTC CCG
Cre Y Rin	GTGTAGCTCATGCTGCAGCACACCACGTCCTCGGTGTCCAGCCCGGA CCGACGATG
Cre Y Fin	GCTGCAGCATGAGCTACACCTGGACCGGCAAGCGACCAAGTGACAGC AATG
Cre A Rin	GCTCATGCTGCAGCACACCACGTCCTCGGTGTCTACTGCCAGACCGC GCGCCTG
Cre A Fin	GGTGTGCTGCAGCATGAGCTACACCTGGACCGGCGTAAAAACTATCCA GCAACATTTGGGC
Cre B Rin	GCTCATGCTGCAGCACACCACGTCCTCGGTGTCAGTACGTGAGATATC TTTAACCCTGATCC
Cre B Fin	GGTGTGCTGCAGCATGAGCTACACCTGGACCGGCAGAATGTTAATCCA TATTGGCAGAACG
Cre C Rin	GCTCATGCTGCAGCACACCACGTCCTCGGTGTCAGCTACACCAGAGAC GGAAATCC
Cre C Fin	GGTGTGCTGCAGCATGAGCTACACCTGGACCGGCAACTACCTGTTTTG CCGGGTCAG
Cre D Rin	GCTCATGCTGCAGCACACCACGTCCTCGGTGTCAGATGGCGCGGCAA CACC

Cre D Fin	GGTGTGCTGCAGCATGAGCTACACCTGGACCGGCTCAACTCGCGCCCTGGAAG
Cre E Rin	GCTCATGCTGCAGCACACCACGTCCTCGGTGTCCTTAGCGCCGTAAATCAATCGATG
Cre E Fin	GGTGTGCTGCAGCATGAGCTACACCTGGACCGGCCAGAGATACCTGGCCTGGTCTG
Cre G Rin	GCTCATGCTGCAGCACACCACGTCCTCGGTGTCGCTATTTTCCATGAGTGAACGAACCTG
Cre G Fin	GGTGTGCTGCAGCATGAGCTACACCTGGACCGGCCGCTGCCAGGATATACGTAATCTG
Luc2p F	ATGGAAGATGCCAAAAACATTAAGAAG
Luc2p R	TTAGACGTTGATCCTGGCGCTGG

Bibliography

Abdel-Wahab, O., A. Mullally, C. Hedvat, G. Garcia-Manero, J. Patel, M. Wadleigh, S. Malinge, J. Yao, O. Kilpivaara, R. Bhat, K. Huberman, S. Thomas, I. Dolgalev, A. Heguy, E. Paietta, M. M. Le Beau, M. Beran, M. S. Tallman, B. L. Ebert, H. M. Kantarjian, R. M. Stone, D. G. Gilliland, J. D. Crispino and R. L. Levine (2009). "Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies." Blood **114**(1): 144-147.

Abdiche, Y. N., K. C. Lindquist, A. Pinkerton, J. Pons and A. Rajpal (2011). "Expanding the ProteOn XPR36 biosensor into a 36-ligand array expedites protein interaction analysis." Analytical Biochemistry **411**(1): 139-151.

Adams, V. H., S. J. McBryant, P. A. Wade, C. L. Woodcock and J. C. Hansen (2007). "Intrinsic Disorder and Autonomous Domain Function in the Multifunctional Nuclear Protein, MeCP2." Journal of Biological Chemistry **282**(20): 15057-15064.

Adkins, N. L. and P. T. Georgel (2011). "MeCP2: structure and function." Biochemistry and Cell Biology **89**(1): 1-11.

Adler, D., N. Quaderi, S. Brown, V. Chapman, M. J, P. Tate and C. Disteche (1999). "The X-linked methylated DNA binding protein, Mecp2, is subject to X inactivation in the mouse." Mamm Genome **6**(8): 491-492.

Agarwal, N., T. Hardt, A. Brero, D. Nowak, U. Rothbauer, A. Becker, H. Leonhardt and M. C. Cardoso (2007). "MeCP2 interacts with HP1 and modulates its

heterochromatin association during myogenic differentiation." Nucleic Acids Research **35**(16): 5402-5408.

Alberini, C. M. (2008). "The role of protein synthesis during the labile phases of memory: Revisiting the skepticism." Neurobiology of Learning and Memory **89**(3): 234-246.

Allen, N. and B. Barres (2009). "Neuroscience: Glia - more than just brain glue." Nature **457**(7230): 675-677.

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

Andrews, A. J. and K. Luger (2011). "Nucleosome Structure(s) and Stability: Variations on a Theme." Annual Review of Biophysics **40**(1): 99-117.

Arand, J., D. Spieler, T. Karius, M. R. Branco, D. Meilinger, A. Meissner, T. Jenuwein, G. Xu, H. Leonhardt, V. Wolf and J. Walter (2012). "*In Vivo* Control of CpG and Non-CpG DNA Methylation by DNA Methyltransferases." PLoS Genet **8**(6): e1002750.

Bachman, K. E., M. R. Rountree and S. B. Baylin (2001). "Dnmt3a and Dnmt3b Are Transcriptional Repressors That Exhibit Unique Localization Properties to Heterochromatin." Journal of Biological Chemistry **276**(34): 32282-32287.

Backdahl, L., M. Herberth, G. Wilson, P. Tate, L. S. Campos, R. Cortese, F. Eckhardt and S. Beck (2009). "Gene body methylation of the

dimethylarginine dimethylamino-hydrolase 2 (*Ddah2*) gene is an epigenetic biomarker for neural stem cell differentiation." Epigenetics **4**(4): 248-254.

Ballas, N., D. Lioy, C. Grunseich and M. G (2009). "Non-cell autonomous influence of MeCP2-deficient glia on neuronal dendritic morphology." Nat Neurosci **12**(3): 311-317.

Ballestar, E. and A. P. Wolffe (2001). "Methyl-CpG-binding proteins." European Journal of Biochemistry **268**(1): 1-6.

Ballestar, E., T. M. Yusufzai and A. P. Wolffe (2000). "Effects of Rett Syndrome Mutations of the Methyl-CpG Binding Domain of the Transcriptional Repressor MeCP2 on Selectivity for Association with Methylated DNA." Biochemistry **39**(24): 7100-7106.

Banko, J. L., L. Hou, F. Poulin, N. Sonenberg and E. Klann (2006). "Regulation of Eukaryotic Initiation Factor 4E by Converging Signaling Pathways during Metabotropic Glutamate Receptor-Dependent Long-Term Depression." The Journal of Neuroscience **26**(8): 2167-2173.

Banko, J. L., M. Merhav, E. Stern, N. Sonenberg, K. Rosenblum and E. Klann (2007). "Behavioral alterations in mice lacking the translation repressor 4E-BP2." Neurobiology of Learning and Memory **87**(2): 248-256.

Banko, J. L., F. Poulin, L. Hou, C. T. DeMaria, N. Sonenberg and E. Klann (2005). "The Translation Repressor 4E-BP2 Is Critical for eIF4F Complex Formation, Synaptic Plasticity, and Memory in the Hippocampus." The Journal of Neuroscience **25**(42): 9581-9590.

Bartenschlager, R., V. Lohmann, T. Wilkinson and J. Koch (1995). "Complex formation between the NS3 serine-type proteinase of the hepatitis C virus and NS4A and its importance for polyprotein maturation." Journal of Virology **69**(12): 7519-7528.

Bartsch, D., A. Casadio, K. A. Karl, P. Serodio and E. R. Kandel (1998). "CREB1 Encodes a Nuclear Activator, a Repressor, and a Cytoplasmic Modulator that Form a Regulatory Unit Critical for Long-Term Facilitation." Cell **95**(2): 211-223.

Bebbington, A., A. Anderson, D. Ravine, S. Fyfe, M. Pineda, N. de Klerk, B. Ben-Zeev, N. Yatawara, A. Percy, W. E. Kaufmann and H. Leonard (2008). "Investigating genotype-phenotype relationships in Rett syndrome using an international data set." Neurology **70**(11): 868-875.

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

Berlanga, J. J., J. Santoyo and C. de Haro (1999). "Characterization of a mammalian homolog of the GCN2 eukaryotic initiation factor 2α kinase." European Journal of Biochemistry **265**(2): 754-762.

Bestor, T. (1988). "Cloning of mammalian DNA methyltransferase." Gene **74**(1): 9-12.

Bestor, T. H. (2000). "The DNA methyltransferases of mammals." Human Molecular Genetics **9**(16): 2395-2402.

Bestor, T. H. and V. M. Ingram (1983). "Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA." Proceedings of the National Academy of Sciences **80**(18): 5559-5563.

Bird, A. (2002). "DNA methylation patterns and epigenetic memory." Genes & Development **16**(1): 6-21.

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

Bissonnette, J. and S. Knopp (2006). "Separate respiratory phenotypes in methyl-CpG-binding protein 2 (Mecp2) deficient mice." Pediatr Res **59**(4): 513-518.

Blais, J. D., V. Filipenko, M. Bi, H. P. Harding, D. Ron, C. Koumenis, B. G. Wouters and J. C. Bell (2004). "Activating Transcription Factor 4 Is Translationally Regulated by Hypoxic Stress." Molecular and Cellular Biology **24**(17): 7469-7482.

Booth, M. J., M. R. Branco, G. Ficz, D. Oxley, F. Krueger, W. Reik and S. Balasubramanian (2012). "Quantitative Sequencing of 5-Methylcytosine and 5-Hydroxymethylcytosine at Single-Base Resolution." Science **336** (6083): 934-937.

Bostick, M., J. K. Kim, P.-O. Esteve, A. Clark, S. Pradhan and S. E. Jacobsen (2007). "UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells." Science **317**(5845): 1760-1764.

Boyes, J. and A. Bird (1991). "DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein." Cell **64**: 1123-1134.

Branco, M. R., G. Ficz and W. Reik (2012). "Uncovering the role of 5-hydroxymethylcytosine in the epigenome." Nat Rev Genet **13**: 7-13.

Brocard, J., X. Warot, O. Wendling, N. Messaddeq, J.-L. Vonesch, P. Chambon and D. Metzger (1997). "Spatio-temporally controlled site-specific somatic mutagenesis in the mouse." Proceedings of the National Academy of Sciences **94**(26): 14559-14563.

Brueckner, B., R. Garcia Boy, P. Siedlecki, T. Musch, H. C. Kliem, P. Zielenkiewicz, S. Suhai, M. Wiessler and F. Lyko (2005). "Epigenetic Reactivation of Tumor Suppressor Genes by a Novel Small-Molecule Inhibitor of Human DNA Methyltransferases." Cancer Research **65**(14): 6305-6311.

Burmeister, T., C. Meyer, S. Schwartz, J. Hofmann, M. Molkentin, E. Kowarz, B. r. Schneider, T. Raff, R. Reinhardt, N. Gokbuget, D. Hoelzer, E. Thiel and R. Marschalek (2009). "The MLL recombinome of adult CD10-negative B-cell precursor acute lymphoblastic leukemia: results from the GMALL study group." Blood **113**(17): 4011-4015.

Campillos, M., J. Lamas, M. A. Garcia, M. J. Bullido, F. Valdivieso and J. Vazquez (2003). "Specific interaction of heterogeneous nuclear ribonucleoprotein A1 with the -219T allelic form modulates APOE promoter activity." Nucleic Acids Research **31**(12): 3063-3070.

Campoy, F. J., R. R. Meehan, S. McKay, J. Nixon and A. Bird (1995). "Binding of Histone H1 to DNA Is Indifferent to Methylation at CpG Sequences." Journal of Biological Chemistry **270**(44): 26473-26481.

Chahrour, M., S. Y. Jung, C. Shaw, X. Zhou, S. T. C. Wong, J. Qin and H. Y. Zoghbi (2008). "MeCP2, a Key Contributor to Neurological Disease, Activates and Represses Transcription." Science **320**(5880): 1224-1229.

Chai, Q., L. Zheng, M. Zhou, J. J. Turchi and B. Shen (2003). "Interaction and Stimulation of Human FEN-1 Nuclease Activities by Heterogeneous Nuclear Ribonucleoprotein A1 in α-Segment Processing during Okazaki Fragment Maturation." Biochemistry **42**(51): 15045-15052.

Chandler, S. P., D. Guschin, N. Landsberger and A. P. Wolffe (1999). "The Methyl-CpG Binding Transcriptional Repressor MeCP2 Stably Associates with Nucleosomal DNA." Biochemistry **38**(22): 7008-7018.

Chandler, S. P., D. Guschin, N. Landsberger and A. P. Wolffe (1999). "The Methyl-CpG Binding Transcriptional Repressor MeCP2 Stably Associates with Nucleosomal DNA." Biochemistry **38**(22): 7008-7018.

Chao, H.-T., H. Chen, R. C. Samaco, M. Xue, M. Chahrour, J. Yoo, J. L. Neul, S. Gong, H.-C. Lu, N. Heintz, M. Ekker, J. L. R. Rubenstein, J. L. Noebels, C. Rosenmund and H. Y. Zoghbi (2010). "Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes." Nature **468**(7321): 263-269.

Chao, H.-T., H. Y. Zoghbi and C. Rosenmund (2007). "MeCP2 Controls Excitatory Synaptic Strength by Regulating Glutamatergic Synapse Number." Neuron **56**(1): 58-65.

Chapleau, C. A., G. D. Calfa, M. C. Lane, A. J. Albertson, J. L. Larimore, S. Kudo, D. L. Armstrong, A. K. Percy and L. Pozzo-Miller (2009). "Dendritic spine pathologies in hippocampal pyramidal neurons from Rett syndrome brain and after expression of Rett-associated MECP2 mutations." Neurobiology of Disease **35**(2): 219-233.

Chen, H., S. Dzitoyeva and H. Manev (2012). "Effect of aging on 5-hydroxymethylcytosine in the mouse hippocampus." Restorative Neurology and Neuroscience **30**(3): 237-245.

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

Chen, T., Y. Ueda, J. E. Dodge, Z. Wang and E. Li (2003). "Establishment and Maintenance of Genomic Methylation Patterns in Mouse Embryonic Stem Cells by Dnmt3a and Dnmt3b." Molecular and Cellular Biology **23**(16): 5594-5605.

Chen, W. G., Q. Chang, Y. Lin, A. Meissner, A. E. West, E. C. Griffith, R. Jaenisch and M. E. Greenberg (2003). "Derepression of BDNF Transcription Involves Calcium-Dependent Phosphorylation of MeCP2." Science **302** (5646): 885-889.

Cheung, A. Y. L., L. M. Horvath, D. Grafodatskaya, P. Pasceri, R. Weksberg, A. Hotta, L. Carrel and J. Ellis (2011). "Isolation of MECP2-null Rett Syndrome patient hiPS cells and isogenic controls through X-chromosome inactivation." Human Molecular Genetics **20**(11): 2103-2115.

Choi, Y. D. and G. Dreyfuss (1984). "Isolation of the heterogeneous nuclear RNA-ribonucleoprotein complex (hnRNP): a unique supramolecular assembly." Proceedings of the National Academy of Sciences **81**(23): 7471-7475.

Clarke, J., H.-C. Wu, L. Jayasinghe, A. Patel, S. Reid and H. Bayley (2009). "Continuous base identification for single-molecule nanopore DNA sequencing." Nat Nano **4**(4): 265-270.

Clouaire, T. and I. Stancheva (2008). "Methyl-CpG binding proteins: specialized transcriptional repressors or structural components of chromatin?" Cellular and Molecular Life Sciences **65**(10): 1509-1522.

Cohen, S., H. W. Gabel, M. Hemberg, A. N. Hutchinson, L. A. Sadacca, D. H. Ebert, D. A. Harmin, R. S. Greenberg, V. K. Verdine, Z. Zhou, W. C. Wetsel, A. E. West and M. E. Greenberg (2011). "Genome-Wide Activity-Dependent MeCP2 Phosphorylation Regulates Nervous System Development and Function." Neuron **72**(1): 72-85.

Coleman-Derr, D. and D. Zilberman (2012). "Deposition of Histone Variant H2A.Z within Gene Bodies Regulates Responsive Genes." PLoS Genet **8**(10): e1002988.

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

Cortellino, S., J. Xu, M. Sannai, R. Moore, E. Caretti, A. Cigliano, M. Le Coz, K. Devarajan, A. Wessels, D. Soprano, L. K. Abramowitz, M. S. Bartolomei, F. Rambow, M. R. Bassi, T. Bruno, M. Fanciulli, C. Renner, A. J. Klein-Szanto, Y. Matsumoto, D. Kobi, I. Davidson, C. Alberti, L. Larue and A. Bellacosa (2011). "Thymine DNA Glycosylase Is Essential for Active DNA Demethylation by Linked Deamination-Base Excision Repair." Cell **146**(1): 67-79.

Cortese, R., J. Lewin, L. Backdahl, M. Krispin, R. Wasserkort, F. Eckhardt and S. Beck (2011). "Genome-Wide Screen for Differential DNA Methylation Associated with Neural Cell Differentiation in Mouse." PLoS ONE **6**(10): e26002.

Costa-Mattioli, M., D. Gobert, H. Harding, B. Herdy, M. Azzi, M. Bruno, M. Bidinosti, C. Ben Mamou, E. Marcinkiewicz, M. Yoshida, H. Imataka, A. C. Cuello, N. Seidah, W. Sossin, J.-C. Lacaille, D. Ron, K. Nader and N. Sonenberg (2005). "Translational control of hippocampal synaptic plasticity and memory by the eIF2α kinase GCN2." Nature **436**(7054): 1166-1173.

Costa-Mattioli, M., D. Gobert, E. Stern, K. Gamache, R. Colina, C. Cuello, W. Sossin, R. Kaufman, J. Pelletier, K. Rosenblum, K. i. Krnjevic, J.-C.

Lacaille, K. Nader and N. Sonenberg (2007). "eIF2α Phosphorylation Bidirectionally Regulates the Switch from Short- to Long-Term Synaptic Plasticity and Memory." Cell **129**(1): 195-206.

Couvert, P., T. Bienvenu, C. Aquaviva, K. Poirier, C. Moraine, C. Gendrot, A. Verloes, C. Andres, A. C. Le Fevre, I. Souville, J. Steffann, V. des Portes, H.-H. Ropers, H. G. Yntema, J.-P. Fryns, S. Briault, J. Chelly and B. Cherif (2001). "MECP2 is highly mutated in X-linked mental retardation." Human Molecular Genetics **10**(9): 941-946.

Creusot, F., G. Acs and J. K. Christman (1982). "Inhibition of DNA methyltransferase and induction of Friend erythroleukemia cell differentiation by 5-azacytidine and 5-aza-2'-deoxycytidine." Journal of Biological Chemistry **257**(4): 2041-2048.

Cross, S., J. Chariton, X. Nan and A. Bird (1994). "Purification of CpG islands using a methylated DNA binding column." Nat Genet **6**(3): 236-240.

Dani, V. S., Q. Chang, A. Maffei, G. G. Turrigiano, R. Jaenisch and S. B. Nelson (2005). "Reduced cortical activity due to a shift in the balance between excitation and inhibition in a mouse model of Rett Syndrome." Proc of the Nat Acad of Sci of the USA **102**(35): 12560-12565.

Dar, A. C., T. E. Dever and F. Sicheri (2005). "Higher-Order Substrate Recognition of eIF2 by the RNA-Dependent Protein Kinase PKR." Cell **122** (6): 887-900.

Dawlaty, M. M., K. Ganz, B. E. Powell, Y.-C. Hu, S. Markoulaki, A. W. Cheng, Q. Gao, J. Kim, S.-W. Choi, D. C. Page and R. Jaenisch (2011). "Tet1 Is

Dispensable for Maintaining Pluripotency and Its Loss Is Compatible with Embryonic and Postnatal Development." Cell stem cell **9**(2): 166-175.

de Felipe, P. (2004). "Skipping the co-expression problem: the new 2A "CHYSEL" technology." Genet Vaccines Ther **2**(13).

de Felipe, P., V. Martin, M. Cortes, M. Ryan and M. Izquierdo (1999). "Use of the 2A sequence from foot-and-mouth disease virus in the generation of retroviral vectors for gene therapy." Gene Therapy **6**(2): 198-208.

Della Ragione, F., A. Tiunova, M. Vacca, M. Strazzullo, E. Gonz√°lez, J. Armstrong, R. Valero, C. Campanile, M. Pineda, M. Hulten, E. Monros, M. D'Esposito and E. Prokhortchouk (2006). "The X-linked methyl binding protein gene Kaiso is highly expressed in brain but is not mutated in Rett syndrome patients." Gene **373**(0): 83-89.

Derecki, N., J. Cronk, Z. Lu, E. Xu, S. Abbott, P. Guyenet and J. Kipnis (2012). "Wild-type microglia arrest pathology in a mouse model of Rett syndrome. ." Nature **484**: 105-109.

Dhasarathy, A. and P. A. Wade (2008). "The MBD protein family-Reading an epigenetic mark?" Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis **647**(1,Äì2): 39-43.

Donnelly, M. L., D. Gani, M. Flint, S. Monaghan and M. D. Ryan (1997). "The cleavage activities of aphthovirus and cardiovirus 2A proteins." Journal of General Virology **78**(1): 13-21.

Donnelly, M. L. L., G. Luke, A. Mehrotra, X. Li, L. E. Hughes, D. Gani and M. D. Ryan (2001). "Analysis of the aphthovirus 2A/2B polyprotein ‘cleavage’

mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal ‘skip’." Journal of General Virology **82**(5): 1013-1025.

Doskocil, J. and F. Sorm (1962). "Distribution of 5-methylcytosine in pyrimidine sequences of deoxyribonucleic acids." Biochimica et Biophysica Acta **55** (6): 953-959.

Dudai, Y. (2004). "The Neurobiology of Consolidations, Or, How Stable is the Engram?" Annual Review of Psychology **55**(1): 51-86.

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

Dzitoyeva, S., H. Chen and H. Manev (2012). "Effect of aging on 5-hydroxymethylcytosine in brain mitochondria." Neurobiology of Aging **33** (12): 2881-91.

Ebbinghaus, H. (1885). Uber das Gedaechtnis Leipzig, Duncker & Humbolt.

El-Osta, A., P. Kantharidis, J. R. Zalcberg and A. P. Wolffe (2002). "Precipitous Release of Methyl-CpG Binding Protein 2 and Histone Deacetylase 1 from the Methylated Human Multidrug Resistance Gene (MDR1) on Activation." Molecular and Cellular Biology **22**(6): 1844-1857.

Engineer, N., J. Riley, J. Seale, W. Vrana, J. Shetake, S. Sudanagunta, M. Borland and M. Kilgard (2011). "Reversing pathological neural activity using targeted plasticity." Nature **470**(101-104).

Fan, G., C. Beard, R. Z. Chen, G. r. Csankovszki, Y. Sun, M. Siniaia, D. Biniszkiewicz, B. Bates, P. P. Lee, R. Kuhn, A. Trumpp, C.-S. Poon, C. B.

Wilson and R. Jaenisch (2001). "DNA Hypomethylation Perturbs the Function and Survival of CNS Neurons in Postnatal Animals." The Journal of Neuroscience **21**(3): 788-797.

Fan, Y., T. Nikitina, J. Zhao, T. J. Fleury, R. Bhattacharyya, E. E. Bouhassira, A. Stein, C. L. Woodcock and A. I. Skoultchi (2005). "Histone H1 Depletion in Mammals Alters Global Chromatin Structure but Causes Specific Changes in Gene Regulation." Cell **123**(7): 1199-1212.

Farooqui, A. A. (1980). "Purification of enzymes by heparin-sepharose affinity chromatography." Journal of Chromatography A **184**(3): 335-345.

Feng, J., H. Chang, E. Li and G. Fan (2005). "Dynamic expression of de novo DNA methyltransferases Dnmt3a and Dnmt3b in the central nervous system." Journal of Neuroscience Research **79**(6): 734-746.

Feng, J., Y. Zhou, S. L. Campbell, T. Le, E. Li, J. D. Sweatt, A. J. Silva and G. Fan (2010). "Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons." Nat Neurosci **13**(4): 423-430.

Feng, Q. and Y. Zhang (2001). "The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes." Genes Dev **15**: 827-832.

Ficz, G., M. Branco, S. Seisenberger, F. Santos, F. Krueger, T. Hore, C. J. Marques and W. Reik (2011). "Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation." Nature **473**: 398-402.

Flexner, J., L. Flexner and E. Stellar (1963). "Memory in mice as affected by inracerebral puromycin. ." Science **141**(3575): 57-59.

Flusberg, B. A., D. R. Webster, J. H. Lee, K. J. Travers, E. C. Olivares, T. A. Clark, J. Korlach and S. W. Turner (2010). "Direct detection of DNA methylation during single-molecule, real-time sequencing." Nat Meth **7**(6): 461-465.

Ford, L., W. Wright and J. Shay (2002). "A model for heterogeneous nuclear ribonucleoproteins in telomere and telomerase regulation." Oncogene **21** (4): 580-583.

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

Fournier, A., N. Sasai, M. Nakao and P.-A. Defossez (2012). "The role of methyl-binding proteins in chromatin organization and epigenome maintenance." Briefings in Functional Genomics **11**(3): 251-264.

Fraga, M. F., E. Ballestar, G. Montoya, T. P, P. A. Wade and M. Esteller (2003). "The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic binding properties." Nucleic Acids Research **31**: 1765-1774.

Frauer, C., T. Hoffmann, S. Bultmann, V. Casa, M. C. Cardoso, I. Antes and H. Leonhardt (2011). "Recognition of 5-Hydroxymethylcytosine by the Uhrf1 SRA Domain." PLoS ONE **6**(6): e21306.

Frauer, C., A. Rottach, D. Meilinger, S. Bultmann, K. Fellinger, S. Hasenoder, M. Wang, W. Qin, J. Soing, F. Spada and H. Leonhardt (2011). "Different Binding Properties and Function of CXXC Zinc Finger Domains in Dnmt1 and Tet1." PLoS ONE **6**(2): e16627.

Freudenberg, J. M., S. Ghosh, B. L. Lackford, S. Yellaboina, X. Zheng, R. Li, S. Cuddapah, P. A. Wade, G. Hu and R. Jothi (2012). "Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity." Nucleic Acids Research **40**(8): 3364-3377.

Frommer, M., L. E. McDonald, D. S. Millar, C. M. Collis, F. Watt, G. W. Grigg, P. L. Molloy and C. L. Paul (1992). "A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands." Proc of the Nat Acad of Sci **89**(5): 1827-1831.

Fujita, N., N. Shimotake, I. Ohki, T. Chiba, H. Saya, M. Shirakawa and M. Nakao (2000). "Mechanism of Transcriptional Regulation by Methyl-CpG Binding Protein MBD1." Molecular and Cellular Biology **20**(14): 5107-5118.

Fuks, F., P. J. Hurd, D. Wolf, X. Nan, A. P. Bird and T. Kouzarides (2003). "The Methyl-CpG-binding Protein MeCP2 Links DNA Methylation to Histone Methylation." Journal of Biological Chemistry **278**(6): 4035-4040.

Fyffe, S. L., J. L. Neul, R. C. Samaco, H.-T. Chao, S. Ben-Shachar, P. Moretti, B. E. McGill, E. H. Goulding, E. Sullivan, L. H. Tecott and H. Y. Zoghbi (2008). "Deletion of *Mecp2* in Sim1-Expressing Neurons Reveals a Critical

Role for MeCP2 in Feeding Behavior, Aggression, and the Response to Stress." Neuron **59**(6): 947-958.

Gallia, G. L., E. M. Johnson and K. Khalili (2000). "Survey and Summary: Purα: a multifunctional single-stranded DNA- and RNA-binding protein." Nucleic Acids Research **28**(17): 3197-3205.

Garcia, E. L. and A. A. Mills (2002). "Getting around lethality with inducible Cre-mediated excision." Seminars in Cell & Developmental Biology **13**(2): 151-158.

Gaveriaux-Ruff, C. and B. L. Kieffer (2007). "Conditional gene targeting in the mouse nervous system: Insights into brain function and diseases." Pharmacology & Therapeutics **113**(3): 619-634.

Gelinas, J. N., J. L. Banko, L. Hou, N. Sonenberg, E. J. Weeber, E. Klann and P. V. Nguyen (2007). "ERK and mTOR Signaling Couple Œ≤-Adrenergic Receptors to Translation Initiation Machinery to Gate Induction of Protein Synthesis-dependent Long-term Potentiation." Journal of Biological Chemistry **282**(37): 27527-27535.

Gemelli, T., O. Berton, E. D. Nelson, L. I. Perrotti, R. Jaenisch and L. M. Monteggia (2006). "Postnatal Loss of Methyl-CpG Binding Protein 2 in the Forebrain is Sufficient to Mediate Behavioral Aspects of Rett Syndrome in Mice." Biological Psychiatry **59**(5): 468-476.

Georgel, P. T., R. A. Horowitz-Scherer, N. Adkins, C. L. Woodcock, P. A. Wade and J. C. Hansen (2003). "Chromatin Compaction by Human MeCP2." Journal of Biological Chemistry **278**(34): 32181-32188.

Ghosh, R. P., R. A. Horowitz-Scherer, T. Nikitina, L. M. Gierasch and C. L. Woodcock (2008). "Rett Syndrome-causing Mutations in Human MeCP2 Result in Diverse Structural Changes That Impact Folding and DNA Interactions." Journal of Biological Chemistry **283**(29): 20523-20534.

Ghosh, R. P., R. A. Horowitz-Scherer, T. Nikitina, L. S. Shlyakhtenko and C. L. Woodcock (2010). "MeCP2 Binds Cooperatively to Its Substrate and Competes with Histone H1 for Chromatin Binding Sites." Molecular and Cellular Biology **30**(19): 4656-4670.

Ghosh, R. P., T. Nikitina, R. A. Horowitz-Scherer, L. M. Gierasch, V. N. Uversky, K. Hite, J. C. Hansen and C. L. Woodcock (2010). "Unique Physical Properties and Interactions of the Domains of Methylated DNA Binding Protein 2." Biochemistry **49**(20): 4395-4410.

Giacometti, E., S. Luikenhuis, C. Beard and R. Jaenisch (2007). "Partial rescue of MeCP2 deficiency by postnatal activation of MeCP2." Proceedings of the National Academy of Sciences **104**(6): 1931-1936.

Globisch, D., M. Münzel, M. Müller, S. Michalakis, M. Wagner, S. Koch, T. Brückl, M. Biel and T. Carell (2010). "Tissue Distribution of 5-Hydroxymethylcytosine and Search for Active Demethylation Intermediates." PLoS ONE **5**(12): e15367.

Goll, M. G. and T. H. Bestor (2005). "Eukaryotic Cytosine Methyltransferases." Annual Review of Biochemistry **74**(1): 481-514.

Golshani, P., L. Hutnick, F. Schweitzer and G. Fan (2005). "Conditional Dnmt1 deletion in dorsal forebrain disrupts development of somatosensory barrel

cortex and thalamocortical long-term potentiation." Thalamus Relat Syst **3** (3): 227-233.

Gommers-Ampt, J. H. and P. Borst (1995). "Hypermodified bases in DNA." The FASEB Journal **9**(11): 1034-42.

Gong, S., M. Doughty, C. R. Harbaugh, A. Cummins, M. E. Hatten, N. Heintz and C. R. Gerfen (2007). "Targeting Cre Recombinase to Specific Neuron Populations with Bacterial Artificial Chromosome Constructs." The Journal of Neuroscience **27**(37): 9817-9823.

Gooney, M., K. Shaw, A. Kelly, S. O'Mara and M. Lynch (2002). "Long-term potentiation and spatial learning are associated with increased phosphorylation of TrkB and extracellular signal regulated kinase (ERK) in dentate gyrus: evidence for a role for brain-derived neurotrophic factor." Behavioral Neuroscience **116**(3): 455-463.

Gossen, M. and H. Bujard (1992). "Tight control of gene expression in mammalian cells by tetracycline-responsive promoters." Proceedings of the National Academy of Sciences **89**(12): 5547-5551.

Gou, D., M. Rubalcava, S. Sauer, F. Mora-Bermudez, H. Erdjument-Bromage, P. Tempst, E. Kremmer and F. Sauer (2010). "SETDB1 Is Involved in Postembryonic DNA Methylation and Gene Silencing in *Drosophila*." PLoS ONE **5**(5): e10581.

Graeber, M. B. (2010). "Changing Face of Microglia." Science **330**(6005): 783-788.

Graebsch, A., S. p. Roche and D. Niessing (2009). "X-ray structure of Purα reveals a Whirly-like fold and an unusual nucleic-acid binding surface." Proceedings of the National Academy of Sciences **106**(44): 18521-18526.

Graham, K. S. and J. R. Hodges (1997). "Differentiating the roles of the hippocampus complex and the neocortex in long-term memory storage: Evidence from the study of semantic dementia and Alzheimer's disease." Neuropsychology **11**(1): 77-89.

Grandjean, V., R. Yaman, F. Cuzin and M. Rassoulzadegan (2007). "Inheritance of an Epigenetic Mark: The CpG DNA Methyltransferase 1 Is Required for De Novo Establishment of a Complex Pattern of Non-CpG Methylation." PLoS ONE **2**(11): e1136.

Gregory, R. I., T. E. Randall, C. A. Johnson, S. Khosla, I. Hatada, L. P. O'Neill, B. M. Turner and R. Feil (2001). "DNA Methylation Is Linked to Deacetylation of Histone H3, but Not H4, on the Imprinted Genes Snrpnand U2af1-rs1." Molecular and Cellular Biology **21**(16): 5426-5436.

Groudine, M., R. Eisenman and H. Weintraub (1981). "Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation." Nature **292**: 311-317.

Gu, H., J. Marth, P. Orban, H. Mossmann and K. Rajewsky (1994). "Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting." Science **265**(5168): 103-106.

Guo, F., D. N. Gopaul and G. D. Van Duyne (1997). "Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse." Nature **389**(6646): 40-46.

Guo, J. U., Y. Su, C. Zhong, G.-l. Ming and H. Song (2011). "Hydroxylation of 5-Methylcytosine by TET1 Promotes Active DNA Demethylation in the Adult Brain." Cell **145**(3): 423-434.

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

Haas, S., J. Gordon and K. Khalili (1993). "A developmentally regulated DNA-binding protein from mouse brain stimulates myelin basic protein gene expression." Molecular and Cellular Biology **13**(5): 3103-3112.

Haas, S., P. Thatikunta, A. Steplewski, E. M. Johnson, K. Khalili and S. Amini (1995). "A 39-kD DNA-binding protein from mouse brain stimulates transcription of myelin basic protein gene in oligodendrocytic cells." The Journal of Cell Biology **130**(5): 1171-1179.

Hackett, J. A., J. J. Zylicz and M. A. Surani (2012). "Parallel mechanisms of epigenetic reprogramming in the germline." Trends in Genetics **28**(4): 164-174.

Haffner, M. C., A. Chaux, A. K. Meeker, D. Esopi, J. Gerber, L. G. Pellakuru, A. Toubaji, P. Argani, C. Iacobuzio-Donahue, W. G. Nelson, G. J. Netto, A. DeMarzo and S. Yegnasubramanian (2011). “Global 5-hydroxymethylcytosine content is significantly reduced in tissue stem/

progenitor cell compartments and in human cancers.” Oncotarget **2**(8): 627-37.

Hagberg, B., A. J, D. K and R. O (1983). "A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases." Ann Neurol **14**(4): 471-479.

Hahn, M. A., X. Wu, A. X. Li, T. Hahn and G. P. Pfeifer (2011). "Relationship between Gene Body DNA Methylation and Intragenic H3K9me3 and H3K36me3 Chromatin Marks." PLoS ONE **6**(4): e18844.

Hajkova, P., S. Erhardt, N. Lane, T. Haaf, O. El-Maarri, W. Reik, J. Walter and M. A. Surani (2002). "Epigenetic reprogramming in mouse primordial germ cells." Mechanisms of Development **117**(1-2): 15-23.

Hamon, M. A. and P. Cossart (2008). "Histone Modifications and Chromatin Remodeling during Bacterial Infections." Cell Host & Microbe **4**(2): 100-109.

Han, S., Y. Tang and R. Smith (2010). "Functional diversity of the hnRNPs: past, present and perspectives." Biochem. J. **430**: 379-392.

Harikrishnan, K., R. Bayles, G. D. Ciccotosto, S. Maxwell, R. Cappai, G. J. Pelka, P. P. L. Tam, J. Christodoulou and A. El-Osta (2010). "Alleviating Transcriptional Inhibition of the Norepinephrine Slc6a2 Transporter Gene in Depolarized Neurons." The Journal of Neuroscience **30**(4): 1494-1501.

Harikrishnan, K., M. Chow, E. Baker, S. Pal, S. Bassal, D. Brasacchio, L. Wang, J. Craig, P. Jones, S. Sif and A. El-Ostra (2005). "Brahma links the SWI/

SNF chromatin remodeling complex with MeCP2-dependent transcriptional silencing." Nat Genet **37**(3): 254-264.

Hashimoto, H., S. Hong, A. S. Bhagwat, X. Zhang and X. Cheng (2012). "Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation." Nucleic Acids Research **40**(20): 10203-10214.

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

He, Y.-F., B.-Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C.-X. Song, K. Zhang, C. He and G.-L. Xu (2011). "Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA." Science **333**(6047): 1303-1307.

Heard, E., P. Clerc and P. Avner (1997). "X-chromosome Inactivation in Mammals." Annual Review of Genetics **31**(1): 571-610.

Heiman, M., A. Schaefer, S. Gong, J. D. Peterson, M. Day, K. E. Ramsey, M. Suarez-Farinas, C. Schwarz, D. A. Stephan, D. J. Surmeier, P. Greengard and N. Heintz (2008). "A Translational Profiling Approach for the Molecular Characterization of CNS Cell Types." Cell **135**(4): 738-748.

Heintz, N. (2000). "Analysis of mammalian central nervous system gene expression and function using bacterial artificial chromosome-mediated transgenesis." Human Molecular Genetics **9**(6): 937-943.

Heitz, E. (1928). "Das Heterochromatin der Moose." Jahrb Wiss Botanik **69**: 762-818.

Hendrich, B. and A. Bird (1998). "Identification and Characterization of a Family of Mammalian Methyl-CpG Binding Proteins." Molecular and Cellular Biology **18**(11): 6538-6547.

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

Hendrich, B., U. Hardeland, H. Ng, J. Jiricny and A. Bird (1999). "The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites." Nature **401**: 301-304.

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

Hermann, A., R. Goyal and A. Jeltsch (2004). "The Dnmt1 DNA-(cytosine-C5)-methyltransferase Methylates DNA Processively with High Preference for Hemimethylated Target Sites." Journal of Biological Chemistry **279**(46): 48350-48359.

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

Hoffman, D. W., C. C. Query, B. L. Golden, S. W. White and J. D. Keene (1991). "RNA-binding domain of the A protein component of the U1 small nuclear

ribonucleoprotein analyzed by NMR spectroscopy is structurally similar to ribosomal proteins." Proceedings of the National Academy of Sciences **88** (6): 2495-2499.

Hokkanen, S., H. M. Feldmann, H. Ding, C. Jung, L. Bojarski, I. Renner-Muller, U. Schuller, H. Kretzschmar, E. Wolf and J. Herms (2011). "Lack of Purα alters postnatal brain development and causes megalencephaly." Human Molecular Genetics.

Holcik, M. and N. Sonenberg (2005). "Translational control in stress and apoptosis." Nat Rev Mol Cell Biol **6**(4): 318-327.

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

Huang, Y., W. A. Pastor, Y. Shen, M. Tahiliani, D. R. Liu and A. Rao (2010). "The Behaviour of 5-Hydroxymethylcytosine in Bisulfite Sequencing." PLoS ONE **5**(1): e8888.

Huang, Z., M. G. Murray and J. A. Secrist Iii (2006). "Recent development of therapeutics for chronic HCV infection." Antiviral Research **71**(2,Äì3): 351-362.

Illingworth, R., A. Kerr, D. DeSousa, H. Jorgensen, P. Ellis, J. Stalker, D. Jackson, C. Clee, R. Plumb, J. Rogers, S. Humphray, T. Cox, C. Langford and A. Bird (2008). "A Novel CpG Island Set Identifies Tissue-Specific Methylation at Developmental Gene Loci." PLoS Biol **6**(1): e22.

Inoue, A., L. Shen, Q. Dai, C. He and Y. Zhang (2011). "Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development." Cell Res **21**(12): 1670-1676.

Inoue, A. and Y. Zhang (2011). "Replication-Dependent Loss of 5-Hydroxymethylcytosine in Mouse Preimplantation Embryos." Science **334** (6053): 194.

Iqbal, K., S.-G. Jin, G. P. Pfeifer and P. E. Szabo (2011). "Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine." Proceedings of the National Academy of Sciences **108** (9): 3642-3647.

Irizarry, R. A., C. Ladd-Acosta, B. Wen, Z. Wu, C. Montano, P. Onyango, H. Cui, K. Gabo, M. Rongione, M. Webster, H. Ji, J. B. Potash, S. Sabunciyan and A. P. Feinberg (2009). "The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores." Nat Genet **41**(2): 178-186.

Ishibashi, T., A. A. Thambirajah and J. Ausio (2008). "MeCP2 preferentially binds to methylated linker DNA in the absence of the terminal tail of histone H3 and independently of histone acetylation." FEBS Letters **582**(7): 1157-1162.

Ito, S., A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers and Y. Zhang (2010). "Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification." Nature **466**(7310): 1129-1133.

Ito, S., L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He and Y. Zhang (2011). "Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine." Science **333**(6047): 1300-1303.

Iyer, L. M., M. Tahiliani, A. Rao and L. Aravind (2009). "Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids." Cell Cycle **8**(11): 1698-1710.

Jaenisch, R. and A. Bird (2003). "Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals." Nat Genet.

Janknecht, R., G. de Martynoff, J. Lou, R. A. Hipskind, A. Nordheim and H. G. Stunnenberg (1991). "Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus." Proceedings of the National Academy of Sciences **88**(20): 8972-8976.

Janosi, L., H. Yonemitsu, H. Hong and A. Kaji (1994). "Molecular Cloning and Expression of a Novel Hydroxymethylcytosine-specific Restriction Enzyme (PvuRts1I) Modulated by Glucosylation of DNA." Journal of Molecular Biology **242**(1): 45-61.

Je, H.-S., Y. Lu, F. Yang, G. Nagappan, J. Zhou, Z. Jiang, K. Nakazawa and B. Lu (2009). "Chemically Inducible Inactivation of Protein Synthesis in Genetically Targeted Neurons." The Journal of Neuroscience **29**(21): 6761-6766.

Jentarra, G., S. Olfers, S. Rice, N. Srivastava, G. Homanics, M. E. Blue, S. Naidu and V. Narayanan (2010). "Abnormalities of cell packing density and

dendritic complexity in the MeCP2 A140V mouse model of Rett syndrome/ X-linked mental retardation." BMC Neuroscience **11**(19).

Jenuwein, T. and C. D. Allis (2001). "Translating the Histone Code." Science **293** (5532): 1074-1080.

Jiang, Y., A. Matevossian, Y. Guo and S. Akbarian (2011). "Setdb1-mediated histone H3K9 hypermethylation in neurons worsens the neurological phenotype of Mecp2-deficient mice." Neuropharmacology **60**(7,Äì8): 1088-1097.

Jin, S.-G., S. Kadam and G. P. Pfeifer (2010). "Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine." Nucleic Acids Research **38**(11): e125.

Johnson, C. L., D. M. Owen and M. Gale (2007). "Functional and Therapeutic Analysis of Hepatitis C Virus NS3/4A Protease Control of Antiviral Immune Defense." Journal of Biological Chemistry **282**(14): 10792-10803.

Johnson, T. B. and R. D. Coghill (1925). "Researches on Pyrimidines. The Discovery of 5-methyl-cytosine in tuberculunic acid, the nucleic acid of the *Tubercle bacillus*." Journal of the American Chemical Society **47**(11): 2838-2844.

Jones, E. G., J. M. Stone and H. J. Karten (2011). "High-resolution digital brain atlases: a Hubble telescope for the brain." Annals of the New York Academy of Sciences **1225**(S1): E147-E159.

Jones, M. W., M. L. Errington, P. J. French, A. Fine, T. V. P. Bliss, S. Garel, P. Charnay, B. Bozon, S. Laroche and S. Davis (2001). "A requirement for

the immediate early gene Zif268 in the expression of late LTP and long-term memories." Nat Neurosci **4**(3): 289-296.

Jordan, C., H. Li, H. Kwan and U. Francke (2007). "Cerebellar gene expression profiles of mouse models for Rett syndrome reveal novel MeCP2 targets." BMC Medical Genetics **8**(36).

Jorgensen, H. F., I. Ben-Porath and A. P. Bird (2004). "Mbd1 Is Recruited to both Methylated and Nonmethylated CpGs via Distinct DNA Binding Domains." Molecular and Cellular Biology **24**(8): 3387-3395.

Jung, B. P., G. Zhang, R. Nitsch, J. Trogadis, S. Nag and J. H. Eubanks (2003). "Differential expression of methyl CpG-binding domain containing factor MBD3 in the developing and adult rat brain." Journal of Neurobiology **55** (2): 220-232.

Kadonaga, J. T. (2004). "Regulation of RNA Polymerase II Transcription by Sequence-Specific DNA Binding Factors." Cell **116**(2): 247-257.

Kaminski, R., L. Cheeseboro, S. Amini, E. M. Johnson, M. K. White, K. Khalili and A. Darbinyan (2010). "Role of Purα in the cellular response to ultraviolet-C radiation." Cell Cycle **9**(20): 4164-4173.

Kandel, E., J. Schwartz and T. Jessell (2000). Principles of Neural Science. New York, McGraw Hill.

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

Kaufman, R. (1994). "Control of gene expression at the level of translation initiation." Curr Opin Biotechnol **5**(5): 550-557.

Kelleher Iii, R. J., A. Govindarajan, H.-Y. Jung, H. Kang and S. Tonegawa (2004). "Translational Control by MAPK Signaling in Long-Term Synaptic Plasticity and Memory." Cell **116**(3): 467-479.

Kellendonk, C., F. O. Tronche, E. Casanova, K. Anlag, C. Opherk and G. N. Schutz (1999). "Inducible site-specific recombination in the brain." Journal of Molecular Biology **285**(1): 175-182.

Kelm, R. J., J. G. Cogan, P. K. Elder, A. R. Strauch and M. J. Getz (1999). "Molecular interactions between single-stranded DNA-binding proteins associated with an essential MCAT element in the mouse smooth muscle α-actin promoter." The Journal of biological chemistry **274**(20): 14238-14245.

Kernohan, K. D., Y. Jiang, D. C. Tremblay, A. C. Bonvissuto, J. H. Eubanks, M. R. W. Mann and N. G. Berube (2010). "ATRX Partners with Cohesin and MeCP2 and Contributes to Developmental Silencing of Imprinted Genes in the Brain." Developmental Cell **18**(2): 191-202.

Keshet, I., J. Lieman-Hurwitz and H. Cedar (1986). "DNA methylation affects the formation of active chromatin." Cell **44**(4): 535-543.

Khalili, K., L. Del Valle, V. Muralidharan, W. J. Gault, N. Darbinian, J. Otte, E. Meier, E. M. Johnson, D. C. Daniel, Y. Kinoshita, S. Amini and J. Gordon (2003). "Purα Is Essential for Postnatal Brain Development and Developmentally Coupled Cellular Proliferation As Revealed by Genetic

Inactivation in the Mouse." Molecular and Cellular Biology **23**(19): 6857-6875.

Khare, T., S. Pai, K. Koncevicius, M. Pal, E. Kriukiene, Z. Liutkeviciute, M. Irimia, P. Jia, C. Ptak, M. Xia, R. Tice, M. Tochigi, S. Morera, A. Nazarians, D. Belsham, A. H. C. Wong, B. J. Blencowe, S. C. Wang, P. Kapranov, R. Kustra, V. Labrie, S. Klimasauskas and A. Petronis (2012). "5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary." Nat Struct Mol Biol **19**(10): 1037-1043.

Kim, J. L., K. A. Morgenstern, C. Lin, T. Fox, M. D. Dwyer, J. A. Landro, S. P. Chambers, W. Markland, C. A. Lepre, E. T. O'Malley, S. L. Harbeson, C. M. Rice, M. A. Murcko, P. R. Caron and J. A. Thomson (1996). "Crystal Structure of the Hepatitis C Virus NS3 Protease Domain Complexed with a Synthetic NS4A Cofactor Peptide." Cell **87**(2): 343-355.

Kishi, N. and J. Macklis (2004). "MECP2 is progressively expressed in post-migratory neurons and is involved in neuronal maturation rather than cell fate decisions. ." Cell Neurosci **27**: 306-321.

Kiss, J., V. Kakkar and D. Thomas (1976). Heparin Chemistry and Clinical Usage. New York, Academic Press.

Klinakis, A., M. Szabolcs, G. Chen, S. Xuan, H. Hibshoosh and A. Efstratiadis (2009). "Igf1r as a therapeutic target in a mouse model of basal-like breast cancer." Proceedings of the National Academy of Sciences **106**(7): 2359-2364.

Klose, R. J. and A. P. Bird (2004). "MeCP2 Behaves as an Elongated Monomer That Does Not Stably Associate with the Sin3a Chromatin Remodeling Complex." Journal of Biological Chemistry **279**(45): 46490-46496.

Knapp, A. M., J. E. Ramsey, S.-X. Wang, K. E. Godburn, A. R. Strauch and R. J. Kelm (2006). "Nucleoprotein Interactions Governing Cell Type-dependent Repression of the Mouse Smooth Muscle α-Actin Promoter by Single-stranded DNA-binding Proteins Purα and Purβ." Journal of Biological Chemistry **281**(12): 7907-7918.

Ko, M., Y. Huang, A. M. Jankowska, U. J. Pape, M. Tahiliani, H. S. Bandukwala, J. An, E. D. Lamperti, K. P. Koh, R. Ganetzky, X. S. Liu, L. Aravind, S. Agarwal, J. P. Maciejewski and A. Rao (2010). "Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2." Nature **468** (7325): 839-843.

Koh, K. P., A. Yabuuchi, S. Rao, Y. Huang, K. Cunniff, J. Nardone, A. Laiho, M. Tahiliani, C. A. Sommer, G. Mostoslavsky, R. Lahesmaa, S. H. Orkin, S. J. Rodig, G. Q. Daley and A. Rao (2011). "Tet1 and Tet2 Regulate 5-Hydroxymethylcytosine Production and Cell Lineage Specification in Mouse Embryonic Stem Cells." Cell stem cell **8**(2): 200-213.

Kokura, K., S. C. Kaul, R. Wadhwa, T. Nomura, M. M. Khan, T. Shinagawa, T. Yasukawa, C. Colmenares and S. Ishii (2001). "The Ski Protein Family Is Required for MeCP2-mediated Transcriptional Repression." Journal of Biological Chemistry **276**(36): 34115-34121.

Kosmider, O., E. Delabesse, V. r. M.-D. Mas, P. Cornillet-Lefebvre, O. Blanchet, A. Delmer, C. Recher, S. Raynaud, D. Bouscary, F. Viguie, C. Lacombe, O. A. Bernard, N. Ifrah, F. o. Dreyfus and M. l. Fontenay (2011). "TET2 mutations in secondary acute myeloid leukemias: a French retrospective study." Haematologica **96**(7): 1059-1063.

Kothari, R. M. and V. Shankar (1976). "5-methylcytosine content in the vertebrate deoxyribonucleic acids: Species specificity." Journal of Molecular Evolution **7**(4): 325-329.

Kriaucionis, S. and N. Heintz (2009). "The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain." Science **324**(5929): 929-930.

Kriaucionis, S., A. Paterson, J. Curtis, J. Guy, N. MacLeod and A. Bird (2006). "Gene Expression Analysis Exposes Mitochondrial Abnormalities in a Mouse Model of Rett Syndrome." Molecular and Cellular Biology **26**(13): 5033-5042.

Kudo, S. (1998). "Methyl-CpG-Binding Protein MeCP2 Represses Sp1-Activated Transcription of the Human Leukosialin Gene When the Promoter Is Methylated." Molecular and Cellular Biology **18**(9): 5492-5499.

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

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

Lafarga, M., M. T. Berciano and L. M. Garcia-Segura (1991). "Freeze-fracture organization of chromatin and cytoplasm in neurons and astroglia of rat cerebellar cortex." Journal of Neurocytology **20**(7): 533-551.

Laget, S., M. Joulie, F. Le Masson, N. Sasai, E. Christians, S. Pradhan, R. J. Roberts and P.-A. Defossez (2010). "The Human Proteins MBD5 and MBD6 Associate with Heterochromatin but They Do Not Bind Methylated DNA." PLoS ONE **5**(8): e11982.

Lakso, M., B. Sauer, B. Mosinger, E. J. Lee, R. W. Manning, S. H. Yu, K. L. Mulder and H. Westphal (1992). "Targeted oncogene activation by site-specific recombination in transgenic mice." Proceedings of the National Academy of Sciences **89**(14): 6232-6236.

LaPlant, Q., V. Vialou, H. E. Covington, D. Dumitriu, J. Feng, B. L. Warren, I. Maze, D. M. Dietz, E. L. Watts, S. D. Iniguez, J. W. Koo, E. Mouzon, W. Renthal, F. Hollis, H. Wang, M. A. Noonan, Y. Ren, A. J. Eisch, C. A. Bolanos, M. Kabbaj, G. Xiao, R. L. Neve, Y. L. Hurd, R. S. Oosting, G. Fan, J. H. Morrison and E. J. Nestler (2010). "Dnmt3a regulates emotional behavior and spine plasticity in the nucleus accumbens." Nat Neurosci **13**(9): 1137-1143.

LaSalle, J. M. and P. S. Gerald (2004). Paradoxical Role of Methyl-CpG-Binding Protein 2 in Rett Syndrome. Current Topics in Developmental Biology, Academic Press. **Volume 59:** 61-86.

Laurent, L., E. Wong, G. Li, T. Huynh, A. Tsirigos, C. T. Ong, H. M. Low, K. W. Kin Sung, I. Rigoutsos, J. Loring and C.-L. Wei (2010). "Dynamic changes in the human methylome during differentiation." Genome Research **20**(3): 320-331.

Lawson-Yuen, A., D. Liu, L. Han, Z. I. Jiang, G. E. Tsai, A. C. Basu, J. Picker, J. Feng and J. T. Coyle (2007). "Ube3a mRNA and protein expression are not decreased in Mecp2R168X mutant mice." Brain Research **1180**(0): 1-6.

Levenson, J. M., T. L. Roth, F. D. Lubin, C. A. Miller, I.-C. Huang, P. Desai, L. M. Malone and J. D. Sweatt (2006). "Evidence That DNA (Cytosine-5) Methyltransferase Regulates Synaptic Plasticity in the Hippocampus." Journal of Biological Chemistry **281**(23): 15763-15773.

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

Lezon-Geyda, K., V. Najfeld and E. M. Johnson (2001). "Deletions of PURA, at 5q31, and PURB, at 7p13, in myelodysplastic syndrome and progression to acute myelogenous leukemia." Leukemia **15**(6): 954-962.

Li, E., C. Beard and R. Jaenisch (1993). "Role for DNA methylation in genomic imprinting." Nature **366**: 362-365.

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

Li, H., X. Zhong, K. F. Chau, E. C. Williams and Q. Chang (2011). "Loss of activity-induced phosphorylation of MeCP2 enhances synaptogenesis, LTP and spatial memory." Nat Neurosci **14**(8): 1001-1008.

Li, X.-D., L. Sun, R. B. Seth, G. Pineda and Z. J. Chen (2005). "Hepatitis C virus protease NS3/4A cleaves mitochondrial antiviral signaling protein off the mitochondria to evade innate immunity." Proceedings of the National Academy of Sciences of the United States of America **102**(49): 17717-17722.

Li, Y., K. Koike, S. Ohashi, T. Funakoshi, M. Tadano, S. Kobayashi, K. Anzai, N. Shibata and M. Kobayashi (2001). "Purα; Protein Implicated in Dendritic RNA Transport Interacts with Ribosomes in Neuronal Cytoplasm." Biological and Pharmaceutical Bulletin **24**(3): 231-235.

Lichtman, J. W., J. Livet and J. R. Sanes (2008). "A technicolour approach to the connectome." Nat Rev Neurosci **9**(6): 417-422.

Lima, F. T. d., D. Brunoni, J. S. o. Schwartzman, M. C. Pozzi, F. Kok, Y. Juliano and L. d. V. Pereira (2009). "Genotype-phenotype correlation in Brazillian Rett syndrome patients." Arquivos de Neuro-Psiquiatria **67**: 577-584.

Lin, H. (2011). "S-Adenosylmethionine-dependent alkylation reactions: When are radical reactions used?" Bioorganic Chemistry **39**(5,Äì6): 161-170.

Lin, M. Z., J. S. Glenn and R. Y. Tsien (2008). "A drug-controllable tag for visualizing newly synthesized proteins in cells and whole animals." Proceedings of the National Academy of Sciences **105**(22): 7744-7749.

Lioy, D., S. Garg, C. Monaghan, J. Raber, K. Foust, B. Kaspar, P. Hirrlinger, F. Kirchhoff, J. Bissonnette, N. Ballas and G. Mandel (2012). "A role for glia in the progression of Rett's syndrome. ." Nature **475**: 497-500.

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

Lopez-Serra, L., E. Ballestar, M. F. Fraga, M. Alaminos, F. Setien and M. Esteller (2006). "A Profile of Methyl-CpG Binding Domain Protein Occupancy of Hypermethylated Promoter CpG Islands of Tumor Suppressor Genes in Human Cancer." Cancer Research **66**(17): 8342-8346.

Lorsbach, R. B., J. Moore, S. Mathew, S. C. Raimondi, S. T. Mukatira and J. R. Downing (2003). "TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23)." Leukemia **17**(3): 637-641.

Luger, K., A. Maeder, R. Richmond, D. Sergent and T. Richmond (1997). "Crystal structure of the nucleosome core particle at 2.8 Å resolution." Nature **389**: 251-260.

Luikenhuis, S., E. Giacometti, C. F. Beard and R. Jaenisch (2004). "Expression of MeCP2 in postmitotic neurons rescues Rett syndrome in mice." Proceedings of the National Academy of Sciences of the United States of America **101**(16): 6033-6038.

Maezawa, I. and L.-W. Jin (2010). "Rett Syndrome Microglia Damage Dendrites and Synapses by the Elevated Release of Glutamate." The Journal of Neuroscience **30**(15): 5346-5356.

Maezawa, I., S. Swanberg, D. Harvey, J. M. LaSalle and L.-W. Jin (2009). "Rett Syndrome Astrocytes Are Abnormal and Spread MeCP2 Deficiency through Gap Junctions." The Journal of Neuroscience **29**(16): 5051-5061.

Maiti, A. and A. C. Drohat (2011). "Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxylcytosine: Potential Implications for Active Demethylation of CpG Sites." Journal of Biological Chemistry **286**(41): 35334-35338.

Malleret, G. I., U. Haditsch, D. Genoux, M. W. Jones, T. V. P. Bliss, A. M. Vanhoose, C. Weitlauf, E. R. Kandel, D. G. Winder and I. M. Mansuy (2001). "Inducible and Reversible Enhancement of Learning, Memory, and Long-Term Potentiation by Genetic Inhibition of Calcineurin." Cell **104**(5): 675-686.

Malmqvist, M. (1999). "BIACORE: an affinity biosensor system for characterization of biomolecular interactions." Biochem Soc Trans **27**(2): 335-340.

Marchetto, M. C. N., C. Carromeu, A. Acab, D. Yu, G. W. Yeo, Y. Mu, G. Chen, F. H. Gage and A. R. Muotri (2010). "A Model for Neural Development and Treatment of Rett Syndrome Using Human Induced Pluripotent Stem Cells." Cell **143**(4): 527-539.

Marr, D. (1970). "A Theory for Cerebral Neocortex." Proceedings of the Royal Society of London. Series B. Biological Sciences **176**(1043): 161-234.

Martin, K. C., A. Casadio, H. Zhu, Y. E, J. C. Rose, M. Chen, C. H. Bailey and E. R. Kandel (1997). "Synapse-Specific, Long-Term Facilitation of Aplysia Sensory to Motor Synapses: A Function for Local Protein Synthesis in Memory Storage." Cell **91**(7): 927-938.

Martinowich, K., D. Hattori, H. Wu, S. Fouse, F. He, Y. Hu, G. Fan and Y. E. Sun (2003). "DNA Methylation-Related Chromatin Remodeling in Activity-Dependent *Bdnf* Gene Regulation." Science **302**(5646): 890-893.

Masland, R. H. (2004). "Neuronal cell types." Current Biology **14**(13): R497-R500.

Mathews, M., N. Sonenberg and J. Hershey (2007). Translational control in biology and medicine. New York, Cold Spring Harbor Laboratory Press.

Mayer, W., A. Niveleau, J. Walter, R. Fundele and T. Haaf (2000). "Demethylation of the zygotic paternal genome." Nature **403**: 501-502.

McClelland, J. L., B. L. McNaughton and R. C. O'Reilly (1995). "Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory." Psychological Review **102**(3): 419-457.

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

Mehler, M. F. and J. S. Mattick (2007). "Noncoding RNAs and RNA Editing in Brain Development, Functional Diversification, and Neurological Disease." Physiological Reviews **87**(3): 799-823.

Meissner, A., T. Mikkelsen, H. Gu, M. Wernig, J. Hanna, A. Sivachenko, X. Zhang, B. BE, C. Nusbaum, J. DB, A. Gnirke, R. Jaenisch and E. Lander (2008). "Genome-scale DNA methylation maps of pluripotent and differentiated cells." Nature **454**: 766-770.

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

Meshorer, E. and T. Misteli (2006). "Chromatin in pluripotent embryonic stem cells and differentiation." Nat Rev Mol Cell Biol **7**(7): 540-546.

Metivier, R., R. Gallais, C. Tiffoche, C. Le Peron, R. Z. Jurkowska, R. P. Carmouche, D. Ibberson, P. Barath, F. Demay, G. Reid, V. Benes, A. Jeltsch, F. Gannon and G. Salbert (2008). "Cyclical DNA methylation of a transcriptionally active promoter." Nature **452**(7183): 45-50.

Metzger, D., J. Clifford, H. Chiba and P. Chambon (1995). "Conditional site-specific recombination in mammalian cells using a ligand-dependent chimeric Cre recombinase." Proceedings of the National Academy of Sciences **92**(15): 6991-6995.

Miller, C. A., C. F. Gavin, J. A. White, R. R. Parrish, A. Honasoge, C. R. Yancey, I. M. Rivera, M. D. Rubio, G. Rumbaugh and J. D. Sweatt (2010). "Cortical DNA methylation maintains remote memory." Nat Neurosci **13**(6): 664-666.

Miller, C. A. and J. D. Sweatt (2007). "Covalent Modification of DNA Regulates Memory Formation." Neuron **53**(6): 857-869.

Misteli, T. (2001). "Protein Dynamics: Implications for Nuclear Architecture and Gene Expression." Science **291**(5505): 843-847.

Mohandas, T., R. S. Sparkes and L. J. Shapiro (1981). "Reactivation of an inactive human X-chromosome: evidence for X inactivation by DNA methylation." Science **211**: 393-396.

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

Mund, C., T. Musch, M. Strodicke, B. Assmann, E. Li and F. Lyko (2004). "Comparative analysis of DNA methylation patterns in transgenic Drosophila overexpressing mouse DNA methyltransferases." Biochem. J. **378**(3): 763-768.

Münzel, M., D. Globisch, T. Brückl, M. Wagner, V. Welzmiller, S. Michalakis, M. Müller, M. Biel and T. Carell (2010). "Quantification of the Sixth DNA Base Hydroxymethylcytosine in the Brain." Angewandte Chemie International Edition **49**(31): 5375-5377.

Na, E. S. and L. M. Monteggia (2011). "The role of MeCP2 in CNS development and function." Hormones and Behavior **59**(3): 364-368.

Nan, X., F. J. Campoy and A. Bird (1997). "MeCP2 Is a Transcriptional Repressor with Abundant Binding Sites in Genomic Chromatin." Cell **88**(4): 471-481.

Nan, X., J. Hou, A. Maclean, J. Nasir, M. J. Lafuente, X. Shu, S. Kriaucionis and A. Bird (2007). "Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation." Proceedings of the National Academy of Sciences **104**(8): 2709-2714.

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

Nan, X., H. Ng, C. Johnson, C. Laherty, B. Turner, R. Eisenman and A. Bird (1998). "Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex." Nature **393**: 886-889.

Nan, X., P. Tate, E. Li and A. Bird (1996). "DNA methylation specifies chromosomal localization of MeCP2." Molecular and Cellular Biology **16** (1): 414-21.

Neddermann, P., P. Gallinari, T. Lettieri, D. Schmid, O. Truong, J. J. Hsuan, K. Wiebauer and J. Jiricny (1996). "Cloning and Expression of Human G/T Mismatch-specific Thymine-DNA Glycosylase." Journal of Biological Chemistry **271**(22): 12767-12774.

Nelson, S. B., K. Sugino and C. M. Hempel (2006). "The problem of neuronal cell types: a physiological genomics approach." Trends in Neurosciences **29** (6): 339-345.

Nestor, C. E., R. Ottaviano, J. Reddington, D. Sproul, D. Reinhardt, D. Dunican, E. Katz, J. M. Dixon, D. J. Harrison and R. R. Meehan (2011). "Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes." Genome Research **22**(3): 467-477.

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

Nguyen, M. V. C., F. Du, C. A. Felice, X. Shan, A. Nigam, G. Mandel, J. K. Robinson and N. Ballas (2012). "MeCP2 Is Critical for Maintaining Mature Neuronal Networks and Global Brain Anatomy during Late Stages of Postnatal Brain Development and in the Mature Adult Brain." The Journal of Neuroscience **32**(29): 10021-10034.

Nguyen, S., K. Meletis, D. Fu, S. Jhaveri and R. Jaenisch (2007). "Ablation of de novo DNA methyltransferase Dnmt3a in the nervous system leads to neuromuscular defects and shortened lifespan." Developmental Dynamics **236**(6): 1663-1676.

Nikitina, T., R. P. Ghosh, R. A. Horowitz-Scherer, J. C. Hansen, S. A. Grigoryev and C. L. Woodcock (2007). "MeCP2-Chromatin Interactions Include the Formation of Chromatosome-like Structures and Are Altered in Mutations

Causing Rett Syndrome." Journal of Biological Chemistry **282**(38): 28237-28245.

Nikitina, T., X. Shi, R. P. Ghosh, R. A. Horowitz-Scherer, J. C. Hansen and C. L. Woodcock (2007). "Multiple Modes of Interaction between the Methylated DNA Binding Protein MeCP2 and Chromatin." Molecular and Cellular Biology **27**(3): 864-877.

Noh, G. J. and J. M. Graham Jr (2012). "2q23.1 microdeletion of the MBD5 gene in a female with seizures, developmental delay and distinct dysmorphic features." European Journal of Medical Genetics **55**(1): 59-62.

Nuber, U. A., S. Kriaucionis, T. C. Roloff, J. Guy, J. Selfridge, C. Steinhoff, R. Schulz, B. Lipkowitz, H. H. Ropers, M. C. Holmes and A. Bird (2005). "Up-regulation of glucocorticoid-regulated genes in a mouse model of Rett syndrome." Human Molecular Genetics **14**(15): 2247-2256.

Okano, M., D. W. Bell, D. A. Haber and E. Li (1999). "DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development." Cell **99**(3): 247-257.

Okano, M., S. Xie and E. Li (1998). "Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells." Nucleic Acids Research **26**(11): 2536-2540.

Ono, R., T. Taki, T. Taketani, M. Taniwaki, H. Kobayashi and Y. Hayashi (2002). "LCX, Leukemia-associated Protein with a CXXC Domain, Is Fused to MLL in Acute Myeloid Leukemia with Trilineage Dysplasia Having t(10;11)(q22;q23)." Cancer Research **62**(14): 4075-4080.

Ooi, S. K. T., A. H. O'Donnell and T. H. Bestor (2009). "Mammalian cytosine methylation at a glance." Journal of Cell Science **122**(16): 2787-2791.

Orr, B. A., M. C. Haffner, W. G. Nelson, S. Yegnasubramanian and C. G. Eberhart (2012). "Decreased 5-Hydroxymethylcytosine Is Associated with Neural Progenitor Phenotype in Normal Brain and Shorter Survival in Malignant Glioma." PLoS ONE **7**(7): e41036.

Orrico, A., C.-W. Lam, L. Galli, M. T. Dotti, G. Hayek, S.-F. Tong, P. M. K. Poon, M. Zappella, A. Federico and V. Sorrentino (2000). "MECP2 mutation in male patients with non-specific X-linked mental retardation." FEBS Letters **481**(3): 285-288.

Oswald, J., S. Engemann, N. Lane, W. Mayer, A. Olek, R. Fundele, W. Dean, W. Reik and J. Walter (2000). "Active demethylation of the paternal genome in the mouse zygote." Current Biology **10**(8): 475-478.

Palay, S. and V. Chan-Palay (1974). Cerebellar cortex: cytology and organization. . New York, Springer.

Pastor, W. A., U. J. Pape, Y. Huang, H. R. Henderson, R. Lister, M. Ko, E. M. McLoughlin, Y. Brudno, S. Mahapatra, P. Kapranov, M. Tahiliani, G. Q. Daley, X. S. Liu, J. R. Ecker, P. M. Milos, S. Agarwal and A. Rao (2011). "Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells." Nature **473**(7347): 394-397.

Pearson, E. C., D. L. Bates, T. D. Prospero and J. O. Thomas (1984). "Neuronal nuclei and glial nuclei from mammalian cerebral cortex." European Journal of Biochemistry **144**(2): 353-360.

Pelka, G. J., C. M. Watson, T. Radziewic, M. Hayward, H. Lahooti, J. Christodoulou and P. P. L. Tam (2006). "Mecp2 deficiency is associated with learning and cognitive deficits and altered gene activity in the hippocampal region of mice." Brain **129**(4): 887-898.

Pells, S., H. Moore, R. Meehan and L. Young (2006). Methylation-Sensitive Polymerase Chain Reaction. Nuclear Reprogramming, Humana Press. **325:** 239-249.

Penberthy, W. T., C. Zhao, Y. Zhang, J. R. Jessen, Z. Yang, O. Bricaud, A. Collazo, A. Meng and S. Lin (2004). "Purα and Sp8 as opposing regulators of neural gata2 expression." Developmental Biology **275**(1): 225-234.

Penn, N. W., R. Suwalski, C. O'Riley, K. Bojanowski and R. Yura (1972). "The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid." Biochem. J. **126**(4): 781-790.

Petyuk, V., J. McDermott, M. Cook and B. Sauer (2004). "Functional Mapping of Cre Recombinase by Pentapeptide Insertional Mutagenesis." Journal of Biological Chemistry **279**(35): 37040-37048.

Pfaffeneder, T., B. Hackner, M. Truß, M. Münzel, M. Müller, C. A. Deiml, C. Hagemeier and T. Carell (2011). "The Discovery of 5-Formylcytosine in Embryonic Stem Cell DNA." Angewandte Chemie International Edition **50** (31): 7008-7012.

Pollack, Y., J. Kasir, R. Shemer, S. Metzger and M. Szyf (1984). "Methylation pattern of mouse mitochondrial DNA." Nucleic Acids Research **12**(12): 4811-4824.

Pomerantz, S. C., J. A. McCloskey and A. M. James (1990). Analysis of RNA hydrolyzates by liquid chromatography-mass spectrometry. Methods in Enzymology, Academic Press. **Volume 193:** 796-824.

Prokhortchouk, A., B. Hendrich, H. Jorgensen, A. Ruzov, M. Wilm, G. Georgiev, A. Bird and E. Prokhortchouk (2001). "The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor." Genes & Development **15**(13): 1613-1618.

Prokhortchouk, A., O. Sansom, J. Selfridge, I. M. Caballero, S. Salozhin, D. Aithozhina, L. Cerchietti, F. G. Meng, L. H. Augenlicht, J. M. Mariadason, B. Hendrich, A. Melnick, E. Prokhortchouk, A. Clarke and A. Bird (2006). "Kaiso-Deficient Mice Show Resistance to Intestinal Cancer." Molecular and Cellular Biology **26**(1): 199-208.

Quinlivan, E. P. and J. F. Gregory (2008). "DNA methylation determination by liquid chromatography-tandem mass spectrometry using novel biosynthetic [U-15N]deoxycytidine and [U-15N]methyldeoxycytidine internal standards." Nucleic Acids Research **36**(18): e119.

Ramon y Cajal, S. (1899). Textura del sistema nervioso del hombre y de los vertebrados. . Madrid, Imprenta y Libreria de Nicolas Moya.

Ramón y Cajal, S. (1917). Recuerdos de mi vida. Madrid, Moya.

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

Ramsahoye, B. H. and K. I. Mills (2002). Nearest-Neighbor Analysis: DNA Methylation Protocols. J. M. Walker, Springer New York. **200:** 9-16.

Ratel, D., J.-L. Ravanat, F. Berger and D. Wion (2006). "N6-methyladenine: the other methylated base of DNA." BioEssays **28**(3): 309-315.

Rauch, T., H. Li, X. Wu and G. P. Pfeifer (2006). "MIRA-Assisted Microarray Analysis, a New Technology for the Determination of DNA Methylation Patterns, Identifies Frequent Methylation of Homeodomain-Containing Genes in Lung Cancer Cells." Cancer Research **66**(16): 7939-7947.

Rauch, T. and G. P. Pfeifer (2005). "Methylated-CpG island recovery assay: a new technique for the rapid detection of methylated-CpG islands in cancer." Lab Invest **85**(9): 1172-1180.

Reid, R. (2012). "From Functional Architecture to Functional Connectomics." Neuron **75**(2): 209-217.

Reik, W., W. Dean and J. Walter (2001). "Epigenetic Reprogramming in Mammalian Development." Science **293**(5532): 1089-1093.

Reiser, M., H. Hinrichsen, Y. Benhamou, H. W. Reesink, H. Wedemeyer, C. Avendano, N. Riba, C.-L. Yong, G. Nehmiz and G. G. Steinmann (2005). "Antiviral efficacy of NS3-serine protease inhibitor BILN-2061 in patients with chronic genotype 2 and 3 hepatitis C." Hepatology **41**(4): 832-835.

Richter, J. D. and N. Sonenberg (2005). "Regulation of cap-dependent translation by eIF4E inhibitory proteins." Nature **433**(7025): 477-480.

Riggs, A. (1975). "X inactivation, differentiation, and DNA methylation." Cytogenet Cell Genet **14**(1): 9-25.

Robertson, A. B., J. A. Dahl, R. Ougland and A. Klungland (2012). "Pull-down of 5-hydroxymethylcytosine DNA using JBP1-coated magnetic beads." Nat. Protocols **7**(2): 340-350.

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

Romano, K. P., A. Ali, C. Aydin, D. Soumana, A. ü. l. Ozen, L. M. Deveau, C. Silver, H. Cao, A. Newton, C. J. Petropoulos, W. Huang and C. A. Schiffer (2012). "The Molecular Basis of Drug Resistance against Hepatitis C Virus NS3/4A Protease Inhibitors." PLoS Pathog **8**(7): e1002832.

Rossant, J. and A. McMahon (1999). "Cre-ating mouse mutants-a meeting review on conditional mouse genetics." Genes & Development **13**(2): 142-145.

Rudy, J. W. (2008). "Is there a baby in the bathwater? Maybe: Some methodological issues for the de novo protein synthesis hypothesis." Neurobiology of Learning and Memory **89**(3): 219-224.

Rutlin, M. and S. B. Nelson (2011). "MeCP2: Phosphorylated Locally, Acting Globally." Neuron **72**(1): 3-5.

Ruzov, A., Y. Tsenkina, A. Serio, T. Dudnakova, J. Fletcher, Y. Bai, T. Chebotareva, S. Pells, Z. Hannoun, G. Sullivan, S. Chandran, D. C. Hay, M. Bradley, I. Wilmut and P. De Sousa (2011). "Lineage-specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development." Cell Res **21**(9): 1332-1342.

Ryu, H., J. Lee, S. W. Hagerty, B. Y. Soh, S. E. McAlpin, K. A. Cormier, K. M. Smith and R. J. Ferrante (2006). "ESET/SETDB1 gene expression and

histone H3 (K9) trimethylation in Huntington's disease." Proceedings of the National Academy of Sciences **103**(50): 19176-19181.

Saito, M. and F. Ishikawa (2002). "The mCpG-binding domain of human MBD3 does not bind to mCpG but interacts with NuRD/Mi2 components HDAC1 and MTA2." J. Biol. Chem. **277**: 35434-35439.

Samaco, R. C., A. Hogart and J. M. LaSalle (2005). "Epigenetic overlap in autism-spectrum neurodevelopmental disorders: MECP2 deficiency causes reduced expression of UBE3A and GABRB3." Human Molecular Genetics **14**(4): 483-492.

Santi, D. V., C. E. Garrett and P. J. Barr (1983). "On the mechanism of inhibition of DNA-cytosine methyltransferases by cytosine analogs." Cell **33**(1): 9-10.

Satterlee, J. S., S. Barbee, P. Jin, A. Krichevsky, S. Salama, G. Schratt and D.-Y. Wu (2007). "Noncoding RNAs in the Brain." The Journal of Neuroscience **27**(44): 11856-11859.

Sauer, B. (1998). "Inducible Gene Targeting in Mice Using the Cre/loxSystem." Methods **14**(4): 381-392.

Saxonov, S., P. Berg and D. L. Brutlag (2006). "A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters." Proceedings of the National Academy of Sciences of the United States of America **103**(5): 1412-1417.

Schmidt, E. F., J. L. Warner-Schmidt, B. G. Otopalik, S. B. Pickett, P. Greengard and N. Heintz (2012). "Identification of the Cortical Neurons that Mediate Antidepressant Responses." Cell **149**(5): 1152-1163.

Schmitt, W. B., R. M. J. Deacon, P. H. Seeburg, J. N. P. Rawlins and D. M. Bannerman (2003). "A Within-Subjects, Within-Task Demonstration of Intact Spatial Reference Memory and Impaired Spatial Working Memory in Glutamate Receptor-A-Deficient Mice." The Journal of Neuroscience **23** (9): 3953-3959.

Scoville, W. B. and B. Milner (1957). "Loss of Recent Memory After Bilateral Hippocampal Lesions." Journal of Neurology, Neurosurgery & Psychiatry **20**(1): 11-21.

Shahbazian, M., J. Young, L. Yuva-Paylor, C. Spencer, B. Antalffy, J. Noebels, D. Armstrong and H. Zoghbi (2002). "Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3. ." Neuron **35**: 243-254.

Sharif, J., M. Muto, S.-i. Takebayashi, I. Suetake, A. Iwamatsu, T. A. Endo, J. Shinga, Y. Mizutani-Koseki, T. Toyoda, K. Okamura, S. Tajima, K. Mitsuya, M. Okano and H. Koseki (2007). "The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA." Nature **450**(7171): 908-912.

Shatz, C. J. (2009). "MHC Class I: An Unexpected Role in Neuronal Plasticity." Neuron **64**(1): 40-45.

Sherf, B., S. Navarro, R. Hannah and K. Wood (1996). "Dual-Luciferase® reporter assay: An advanced coreporter technology integrating firefly and Renilla luciferase assays. ." Promega Notes MAg **57**: 2-8.

Shevchenko, A., M. Wilm, O. Vorm and M. Mann (1996). "Mass spectrometric sequencing of proteins silverstained polyacrylamide gels." Anal Chem **68**: 850-858.

Silva, J. C., M. V. Gorenstein, G. Z. Li, J. P. Vissers and S. J. Geromanos (2006). "Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition." Mol Cell Proteomics **5**: 144-156.

Skene, P. J., R. S. Illingworth, S. Webb, A. R. W. Kerr, K. D. James, D. J. Turner, R. Andrews and A. P. Bird (2010). "Neuronal MeCP2 Is Expressed at Near Histone-Octamer Levels and Globally Alters the Chromatin State." Molecular cell **37**(4): 457-468.

Smith, E. E. and J. Jonides (1997). "Working Memory: A View from Neuroimaging." Cognitive Psychology **33**(1): 5-42.

Song, C.-X., K. E. Szulwach, Y. Fu, Q. Dai, C. Yi, X. Li, Y. Li, C.-H. Chen, W. Zhang, X. Jian, J. Wang, L. Zhang, T. J. Looney, B. Zhang, L. A. Godley, L. M. Hicks, B. T. Lahn, P. Jin and C. He (2011). "Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine." Nat Biotech **29**(1): 68-72.

Song, F., J. F. Smith, M. T. Kimura, A. D. Morrow, T. Matsuyama, H. Nagase and W. A. Held (2005). "Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression." Proceedings of the

National Academy of Sciences of the United States of America **102**(9): 3336-3341.

Sood, R., A. C. Porter, D. Olsen, D. R. Cavener and R. C. Wek (2000). "A Mammalian Homologue of GCN2 Protein Kinase Important for Translational Control by Phosphorylation of Eukaryotic Initiation Factor-α." Genetics **154**(2): 787-801.

Squire, L., D. Amaral and G. Press (1990). "Magnetic resonance imaging of the hippocampal formation and mammillary nuclei distinguish medial temporal lobe and diencephalic amnesia." The Journal of Neuroscience **10**(9): 3106-3117.

Squire, L. R., N. J. Cohen and J. A. Zouzounis (1984). "Preserved memory in retrograde amnesia: Sparing of a recently acquired skill." Neuropsychologia **22**(2): 145-152.

Stein, R., R. A and H. Cedar (1982). "In vitro methylation of the hamster adenine phosphoribosyltransferase gene inhibits its expression in mouse L cells." Proceedings of the National Academy of Sciences **79**: 3418-3422.

Steinkuehler, C., A. Urbani, L. Tomei, G. Biasiol, M. Sardana, E. Bianchi, A. Pessi and R. De Francesco (1996). "Activity of purified hepatitis C virus protease NS3 on peptide substrates." Journal of Virology **70**(10): 6694-700.

Sternberg, N., B. Sauer, R. Hoess and K. Abremski (1986). "Bacteriophage P1 cre gene and its regulatory region: Evidence for multiple promoters and for regulation by DNA methylation." Journal of Molecular Biology **187**(2): 197-212.

Stroud, H., S. Feng, S. Kinney, S. Pradhan and S. E. Jacobsen (2011). "5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells." Genome Biol **12**(6).

Subramanian, S. V., J. A. Polikandriotis, R. J. Kelm, J. J. David, C. G. Orosz and A. R. Strauch (2004). "Induction of Vascular Smooth Muscle α-Actin Gene Transcription in Transforming Growth Factor α-Activated Myofibroblasts Mediated by Dynamic Interplay between the Pur Repressor Proteins and Sp1/Smad Coactivators." Molecular Biology of the Cell **15**(10): 4532-4543.

Sutherland, G. R. and B. McNaughton (2000). "Memory trace reactivation in hippocampal and neocortical neuronal ensembles." Current Opinion in Neurobiology **10**(2): 180-186.

Sved, J. and A. Bird (1990). "The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model." Proceedings of the National Academy of Sciences **87**(12): 4692-4696.

Szulwach, K. E., X. Li, Y. Li, C.-X. Song, J. W. Han, S. Kim, S. Namburi, K. Hermetz, J. J. Kim, M. K. Rudd, Y.-S. Yoon, B. Ren, C. He and P. Jin (2011). "Integrating 5-Hydroxymethylcytosine into the Epigenomic Landscape of Human Embryonic Stem Cells." PLoS Genet **7**(6): e1002154.

Szulwach, K. E., X. Li, Y. Li, C.-X. Song, H. Wu, Q. Dai, H. Irier, A. K. Upadhyay, M. Gearing, A. I. Levey, A. Vasanthakumar, L. A. Godley, Q. Chang, X. Cheng, C. He and P. Jin (2011). "5-hmC-mediated epigenetic dynamics

during postnatal neurodevelopment and aging." Nat Neurosci **14**(12): 1607-1616.

Szwagierczak, A., S. Bultmann, C. S. Schmidt, F. Spada and H. Leonhardt (2010). "Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA." Nucleic Acids Research **38**(19): e181.

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

Takizawa, T. and E. Meshorer (2008). "Chromatin and nuclear architecture in the nervous system." Trends in Neurosciences **31**(7): 343-352.

Tao, J., K. Hu, Q. Chang, H. Wu, N. E. Sherman, K. Martinowich, R. J. Klose, C. Schanen, R. Jaenisch, W. Wang and Y. E. Sun (2009). "Phosphorylation of MeCP2 at Serine 80 regulates its chromatin association and neurological function." Proceedings of the National Academy of Sciences **106**(12): 4882-4887.

Taremi, S. S., B. Beyer, M. Maher, N. Yao, W. Prosise, P. C. Weber and B. A. Malcolm (1998). "Construction, expression, and characterization of a novel fully activated recombinant single-chain hepatitis C virus protease." Protein Science **7**(10): 2143-2149.

Thalhammer, A., A. S. Hansen, A. H. El-Sagheer, T. Brown and C. J. Schofield (2011). "Hydroxylation of methylated CpG dinucleotides reverses

stabilisation of DNA duplexes by cytosine 5-methylation." Chemical Communications **47**(18): 5325-5327.

Thambirajah, A. A., M. K. Ng, L. J. Frehlick, A. Li, J. J. Serpa, E. V. Petrotchenko, B. Silva-Moreno, K. K. Missiaen, C. H. Borchers, J. Adam Hall, R. Mackie, F. Lutz, B. E. Gowen, M. Hendzel, P. T. Georgel and J. Ausio (2011). "MeCP2 binds to nucleosome free (linker DNA) regions and to H3K9/H3K27 methylated nucleosomes in the brain." Nucleic Acids Research.

Tian, F., A. Marini and R. Lipsky (2010). "NMDA receptor activation induces differential epigenetic modification of Bdnf promoters in hippocampal neurons." Amino Acids **38**(4): 1067-1074.

Ting, A. H., K. M. McGarvey and S. B. Baylin (2006). "The cancer epigenome - components and functional correlates." Genes & Development **20**(23): 3215-3231.

Tsumura, A., T. Hayakawa, Y. Kumaki, S.-i. Takebayashi, M. Sakaue, C. Matsuoka, K. Shimotohno, F. Ishikawa, E. Li, H. R. Ueda, J.-i. Nakayama and M. Okano (2006). "Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b." Genes to Cells **11**(7): 805-814.

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

Tweedie, S., J. Charlton, V. Clark and A. Bird (1997). "Methylation of genomes and genes at the invertebrate-vertebrate boundary." Molecular and Cellular Biology **17**(3): 1469-75.

Urdinguio, R., A. Fernandez, P. Lopez-Nieva, S. Rossi, D. Huertas, M. Kulis, C. Liu, C. Croce, G. Calin and M. Esteller (2010). "Disrupted microRNA expression caused by Mecp2 loss in a mouse model of Rett syndrome." Epigenetics **5**(7): 656-663.

Valinluck, V. and L. C. Sowers (2007). "Endogenous Cytosine Damage Products Alter the Site Selectivity of Human DNA Maintenance Methyltransferase DNMT1." Cancer Research **67**(3): 946-950.

Valinluck, V., H.-H. Tsai, D. K. Rogstad, A. Burdzy, A. Bird and L. C. Sowers (2004). "Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2)." Nucleic Acids Research **32**(14): 4100-4108.

Wakefield, R. I. D., B. O. Smith, X. Nan, A. Free, A. Soteriou, D. Uhrin, A. P. Bird and P. N. Barlow (1999). "The solution structure of the domain from MeCP2 that binds to methylated DNA." Journal of Molecular Biology **291**