# Heterochromatin composition, organization and MeCP2 post-translational modifications

Vom Fachbereich Biologie
der Technischen Universität Darmstadt
zur Erlangung des akademischen Grades
Doctor rerum naturalium (Dr. rer. nat.)
genehmigte

**Dissertation**

von

**Annika Schmidt**

aus Hilden

Darmstadt 2022

Erstgutachterin: Prof. Dr. M. Cristina Cardoso
Zweitgutachterin: Prof. Dr. Ulrike A. Nuber

Tag der Einreichung: 22.09.2022

Schmidt, Annika:
Heterochromatin composition, organization and MeCP2 post-translational modifications
Darmstadt, Technische Universität Darmstadt,
Jahr der Veröffentlichung der Dissertation auf TUprints: 2023
Tag der mündlichen Prüfung: 23.11.2022

# Contents

## Summary

In the cell nucleus, the DNA together with several other factors is packaged into chromatin which is organized into active and inactive compartments called euchromatin and heterochromatin. The maintenance of the silenced and compacted heterochromatin state is highly important for the genome stability and spatial organization of nuclear compartments and its misfunction can cause severe deficits in cellular processes.

Heterochromatin is comprised of facultative heterochromatin that becomes active in specific situations and constitutive heterochromatin that is mostly silenced and constantly compacted. Constitutive heterochromatin from (peri)centromeric regions of several chromosomes fuses during interphase in mouse cells making up heterochromatin clusters. As the mechanism of the formation of constitutive heterochromatin and the factors involved in the process are not fully understood, we investigated the composition of constitutive heterochromatin in mouse tissues. Therefore, we optimized a protocol for native and unbiased isolation of heterochromatin clusters from mouse tissues and combined it with a quantitative mass spectrometry approach. We identified and validated previously unknown proteins along with proteins known to be involved in constitutive heterochromatin cluster formation and function. We observed differing heterochromatin organization between tissues and identified heterochromatin proteins with distinct abundance between mouse brain and liver tissue. Especially MeCP2, ATRX, and histone H1 might play a role in the distinct heterochromatin organization between tissues, but also newly identified candidate heterochromatin proteins could be involved.

We found the methyl-CpG binding protein (MeCP2) enriched in the mouse brain. It binds to methylated cytosines on the DNA, recruits numerous chromatin-associated factors, and was shown to induce dose-dependent heterochromatin reorganization during terminal differentiation. Mutations in the *MECP2* gene were associated with the neurological disorder Rett syndrome and altered protein levels were reported to cause disease phenotypes. Thus, we quantified MeCP2 levels in mouse brain nuclei and a cellular system upon exogenous expression of *Mecp2*-GFP. We showed that the level of MeCP2 in mouse brains is similar to its level in low transfected cells in our cellular system. In addition, we estimated the MeCP2 protein concentrations in the cell nucleus and heterochromatin compartments. We established a system that enables us to adjust the MeCP2 level of *in cellulo* and *in vitro* experiments to the *in vivo* protein level, thus ensuring comparability of experimental results to the physiological situation.

Heterochromatin organization and its function are not only regulated by the proteins involved and their levels, but also by post-translational modifications of chromatin proteins. Hence, we isolated MeCP2 from the adult mouse brain and analyzed its modifications by mass spectrometry. We identified various post-translational modifications on MeCP2 and confirmed the influence of arginine methylation and to a lesser extent serine phosphorylation on MeCP2-mediated heterochromatin organization in a cellular system. Coexpression of arginine methyltransferases 1 and 6 lead to a decrease in heterochromatin clustering. Interestingly, we identified the Rett syndrome mutation R106 as a possible dimethylation site, implicating a role of MeCP2 post-translational modifications in the pathology of Rett syndrome.

In summary, our data contribute to a better understanding of the constitutive heterochromatin organization in cells and its role in disease development.

## Zusammenfassung

Im Zellkern ist die DNA zusammen mit verschiedenen anderen Faktoren zu Chromatin gepackt, das in aktive und inaktive Kompartimente unterteilt wird, die Euchromatin und Heterochromatin genannt werden. Die Aufrechterhaltung des stillgelegten und dicht gepackten Zustandes des Heterochromatins ist äußerst wichtig, da es für die Genomstabilität und die räumliche Organisation der Kernkompartimente von wesentlicher Bedeutung ist und eine Fehlfunktion zu schwerwiegenden Beeinträchtigungen zellulärer Prozesse führen kann.

Heterochromatin besteht aus fakultativem Heterochromatin, das in bestimmten Situationen aktiv wird, und konstitutivem Heterochromatin, das meist stillgelegt und dauerhaft verdichtet ist. Konstitutives Heterochromatin aus (peri)zentromerischen Regionen verschiedener Chromosomen lagert sich in der Interphase in Mauszellen zu Heterochromatinclustern zusammen. Da der Mechanismus der Bildung von konstitutivem Heterochromatin und die an diesem Prozess beteiligten Faktoren noch nicht vollständig geklärt sind, untersuchten wir die Zusammensetzung des konstitutiven Heterochromatins in Mausgeweben. Dazu optimierten wir ein Protokoll zur nativen und unvoreingenommenen Isolierung von Heterochromatinclustern aus Mausgeweben und kombinierten es mit einer quantitativen Massenspektrometriemethode. Wir identifizierten und validierten bisher unbekannte und bekannte an der Bildung und Funktion von Heterochromatinclustern beteiligte Proteine. Wir beobachteten eine variierende Organisation des Heterochromatins in verschiedenen Geweben und identifizierten Heterochromatinproteine mit unterschiedlicher Abundanz im Gehirn- und Lebergewebe der Maus. Insbesondere MeCP2, ATRX and Histon H1 könnten eine Rolle bei der gewebespezifisch unterschiedlichen Organisation des Heterochromatins spielen, aber auch neu identifizierte Kandidaten für Heterochromatinproteine sind möglicherweise beteiligt.

Wir identifizierten das Methyl-CpG-bindende Protein 2 (MeCP2) als angereichert im Gehirn der Maus. Es bindet an methylierte Cytosine auf der DNA, rekrutiert zahlreiche Chromatin-assoziierte Faktoren und es wurde gezeigt, dass es während der terminalen Differenzierung dosisabhängig Heterochromatin-Clusterbildung induziert. Mutationen im *MECP2*-Gen wurden mit der neurologischen Erkrankung Rett-Syndrom in Verbindung gebracht und es wurde beschrieben, dass auch Veränderungen der Proteinkonzentration Krankheitssymptome verursachen können. Daher haben wir die MeCP2-Konzentration in Zellkernen aus dem Mausgehirn sowie in einem zellulären System nach Expression von *Mecp2*-GFP quantifiziert. Wir konnten zeigen, dass die MeCP2-Konzentration im Mäusegehirn der Konzentration in niedrig transfizierten Zellen in unserem zellulären System ähnelt. Darüber hinaus haben wir die MeCP2-Proteinkonzentrationen im Zellkern und in Heterochromatin-Kompartimenten bestimmt. Wir haben ein System entwickelt, das es ermöglicht, das MeCP2-Level von *in cellulo* und *in vitro* Experimenten an die *in vivo* Proteinkonzentration anzupassen und so die Vergleichbarkeit der experimentellen Ergebnisse mit der physiologischen Situation sicherzustellen.

Die Organisation des Heterochromatins und seine Funktion werden nicht nur durch die beteiligten Proteine und ihre Konzentration, sondern auch durch posttranslationale Modifikationen von Chromatinproteinen reguliert. Daher isolierten wir MeCP2 aus dem Gehirn erwachsener Mäuse

und analysierten seine Modifikationen mittels Massenspektrometrie. Wir identifizierten verschiedene posttranslationale Modifikationen und bestätigten den Einfluss von Argininmethylierung und Serinphosphorylierung auf die MeCP2-vermittelte Heterochromatinorganisation in einem zellulären System. Die Koexpression der Arginin-Methyltransferasen 1 und 6 führte zu einer Abnahme der Heterochromatin-Clusterbildung. Interessanterweise identifizierten wir die Rett-Syndrom Mutation R106 als mögliche Dimethylierungsstelle, was auf eine Rolle der posttranslationalen Modifikationen von MeCP2 bei der Pathologie des Rett-Syndroms hindeutet.
Zusammenfassend lässt sich festhalten, dass unsere Daten zu einem besseren Verständnis der Organisation des konstitutiven Heterochromatins in Zellen und seiner Rolle bei der Entstehung von Krankheiten beitragen.

# Preface

This thesis addresses constitutive heterochromatin organization and composition as well as the role of the Methyl-CpG binding protein 2 (MeCP2) and its post-translational modifications in heterochromatin organization. The thesis is organized in three chapters. The first chapter deals with the proteomic composition of constitutive heterochromatin in mouse tissues, the second one focusses on the quantification of MeCP2 in the cell nucleus and heterochromatin compartments (parts thereof based upon Zhang & Romero, Schmidt *et al.* 2022 [1]) and the third chapter describes the role of MeCP2 post-translational modifications in MeCP2-mediated heterochromatin organization (based upon Schmidt *et al.*, 2022 [2]).
First, a general introduction gives an overview about heterochromatin organization and MeCP2 function (based upon Schmidt, Zhang *et al.* 2020 [3]). Then, each chapter comprises an individual introduction, material and methods, results and discussion as well as supporting material. At the end, all chapters are summarized in a general conclusion and outlook.

# 1 Introduction

## 1.1 Chromatin organization in the nucleus

To fit the 2 m long genomic DNA into the restricted space in the cell nucleus, the DNA is wrapped around a complex of histones and then further folded over several levels reaching its highest compaction in the metaphase chromosome. The hierarchical folding of the DNA and its organization in the 3D space involves several steps: First, the DNA double helix and histones build up the smallest unit of organization, the nucleosome. Then, several nucleosomes can form fibers and loop structures mediated by architectural chromatin proteins and further cluster to shape topologically associated domains. These can be organized into A/B chromatin compartments, which are roughly equivalent to active (A) and inactive (B) chromatin states. Finally, the DNA is compacted into individual chromosomes, that occupy distinct territories in the cell nucleus [4].

On the first level of DNA packaging, 146 base pairs of DNA sequence forming a double-stranded helix are wrapped around an octamer of core histones in about 1.7 turns, thus shaping a left-handed superhelix. The histone core octamer is assembled from a tetramer of two H3-H4 dimers which is flanked by two H2A-H2B dimers building up the nucleosome core particle [5–8]. To stabilize and further compact the DNA, the linker histone H1 binds at the entry-exit site of the DNA winded around the histone core [7], resulting in the formation of a nucleosome. This "Beads on a string" structure is the basis for further DNA folding into higher order structure and determines DNA accessibility for other factors [7, 9].

The expression of the canonical core histones is replication coupled to allow their direct incorporation into the nucleosomes of the newly synthesized DNA. The histones are synthesized during S-phase, transcription and translation are tightly regulated, and the genes are localized in intronless gene clusters containing several copies of the histone genes [10, 11]. From the structural point of view, canonical histones contain a common histone-fold domain consisting of three alpha helices separated by two loops [12] and an N-terminal domain [7]. Especially the N-terminal domain is extensively post-translationally modified and the modifications on the protruding histone tails were considered to build up the "histone code" modulating the regulation of DNA compaction and transcription [13, 14]. The post-translational modifications (PTMs) observed on histone tails were proposed to induce interaction affinities or create new binding sites for chromatin-associated proteins, to act in concert in the regulation of specific processes, and to be important to establish different chromatin states [13].

In addition to the cell-cycle dependent histones, several histone subtypes were described. They originate from genes outside of the clustered histone genes and can replace the canonical histones in specific situations independent of the cell cycle, thus contributing to the diversity of nucleosome function (reviewed in [11, 15]). As the histone subtypes differ from the canonical histones in sequence, they can change the nucleosome structure and stability and might recruit distinct interaction partners and chromatin complexes. For example, the histone variant H3.3 which differs from H3 by substitution of alanine to serine on position 31 in the N-terminal tail and 4 – 5 substitutions in the histone fold domain, was reported to decrease nucleosome stability [11, 16].

This effect was enhanced in nucleosomes carrying both H3.3 and H2A.Z, and they were reported to occur mainly on promotors and enhancers of transcriptionally active genes [16]. Moreover, histone subtypes can be enriched with specific post-translational modifications, thus creating distinct binding sites for interaction partners [11]. The H3.3 specific serine 31 carries a phosphorylation site [17], which was described to recruit or repel chromatin complexes [18] and to be involved in development and gene regulation [18–20]. Although H3.3 was found associated with gene activation and chromatin opening, other studies found it localizing at repetitive sequences and centromeres [21, 22]. Concluding, the types of histone variants and their combination within the nucleosome, their modifications, their localization in different cell types, and their interaction partners altogether contribute to nucleosome function and thus to chromatin organization.

Cytologically, chromatin can occur in two different types: The more open structured euchromatin and the highly compacted heterochromatin [23, 24]. Their differences in packaging density allow their visualization by microscopy upon DNA staining (see figure 1.1), where dense regions appear darker [23]. In mouse cells, (peri)centromeric heterochromatin regions from different chromosomes cluster during interphase forming densely stained regions, called chromocenters (see 1.2) ([25], reviewed in [26]). The open structure of euchromatin allows access of the replication machinery to the DNA more easily and comprises mainly actively transcribed genes. In contrast, heterochromatin occurs in regions of silenced genes and is characterized by high 5-methylcytosine levels and methyl-binding proteins [27]. Furthermore, both types carry different histone PTMs. In general, mainly acetylation and methylation marks are responsible for the

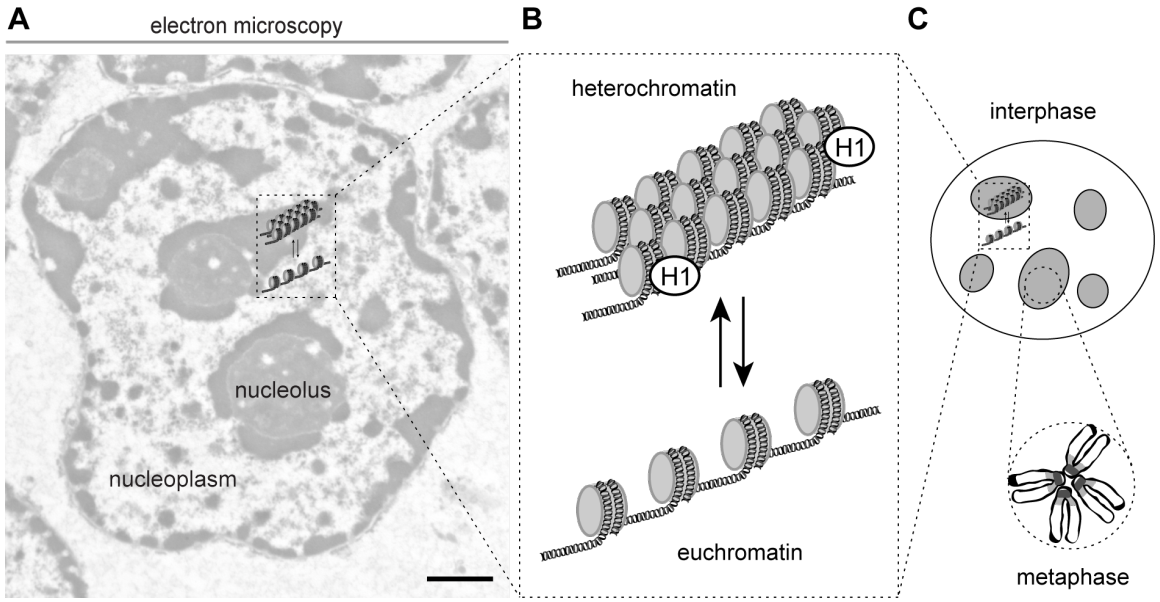



**Figure 1.1: Chromatin organization in the nucleus.** (A) Electron microscopy image of a mouse liver cell nucleus shows the electron dense heterochromatin in contrast to the more open state of euchromatin (taken from Jost *et al.*, 2012 [24]). Scale bar 0.5 µm. (B) Heterochromatin is densely packed and, together with many other factors, histone H1 is involved in its compaction. Euchromatin is the more open and transcriptionally active form of chromatin. (C) During interphase in mouse cells, the DNA dye DAPI dark stained heterochromatin regions of different chromosomes assemble to form heterochromatin clusters, also called chromocenters. This figure was adapted from Jost *et al.*, 2012 [24].

regulation of chromatin compaction. The acetylation of lysine neutralizes its positive charge, which might weaken histone DNA interactions resulting in a less compacted chromatin state [24]. Histone deacetylases are involved in transcriptional repression, deacetylate histones and often act as a part of corepressor complexes which mediate heterochromatin compaction [28–31]. Distinct methylation marks can be used to distinguish chromatin types, as heterochromatin is characterized by H3K9 methylation and euchromatin is commonly methylated on H3K4 [32–34]. Nevertheless, due to the many different PTMs occurring at the same time and all dynamically changeable, chromatin is mainly regulated by their interplay and their specific recruitment of protein complexes.

## 1.2 Heterochromatin

Heterochromatin, the condensed and transcriptionally inactive form of chromatin, can be further subdivided into facultative and constitutive heterochromatin [35]. Facultative heterochromatin can become active in specific situations as upon cell differentiation or during development [27, 36]. An example of its formation is the inactivated X chromosome in females [27, 36]. In contrast, constitutive heterochromatin is consistently silenced among cell types and localized at repetitive sequences occurring at pericentromeric regions and telomeres [27, 37]. Both heterochromatin types share the high DNA methylation levels and hypoacetylation of histones characteristic for heterochromatin but can be distinguished based on specific modifications on histone H3. While constitutive heterochromatin commonly carries H3K9 trimethylation sites recognized by heterochromatin protein 1 (HP1), facultative heterochromatin is typically trimethylated on H3K27, a binding site for polycomb proteins [24, 27, 36, 37].
In mouse cells, minor satellite sequences are located at the centromeric region and major satellite repeats at the pericentromeric region of metaphase chromosomes (figure 1.2) [38–40]. Centromeric and pericentromeric regions are AT-rich and evolutionary conserved, but the DNA tandem repeats vary in sequence and length between species [41]. Mouse minor satellite repeats consist of 120 bp long tandem repeats occupying 1-2% of the genome [39, 41], while the mouse major satellite repeats have a sequence of 234 bp which is constituted of four different 58-60 bp simple repeats and comprise 10% of the genome (figure 1.2) [41–43]. In addition to minor and major satellites, other satellite sequences with lower occurrence were identified and might localize at centromeric regions [41, 43–45].
As mentioned earlier, (peri)centromeric satellite DNA from different chromosomes clusters during interphase in mouse cells forming DNA dense regions called chromocenters [25, 26] (figure 1.2). It is still controversial whether the centromeric minor satellite sequences are part of the chromocenters or should be considered as associated sequences [26, 38, 45–49]. Chromocenter numbers were addressed in several studies revealing differences between cell types [50–53], changes during development [54] or upon cell differentiation [53, 55–57]. Interestingly, overexpression of a plasmid coding for the methyl-CpG binding protein 2 lead to decreased chromocenter numbers [55]. The most studied function of constitutive heterochromatin is the chromatin compaction and maintenance of the silent state of the genes and aberrant gene expression was linked to disease phenotypes [27, 58, 59]. Although the cytological properties and molecular architecture

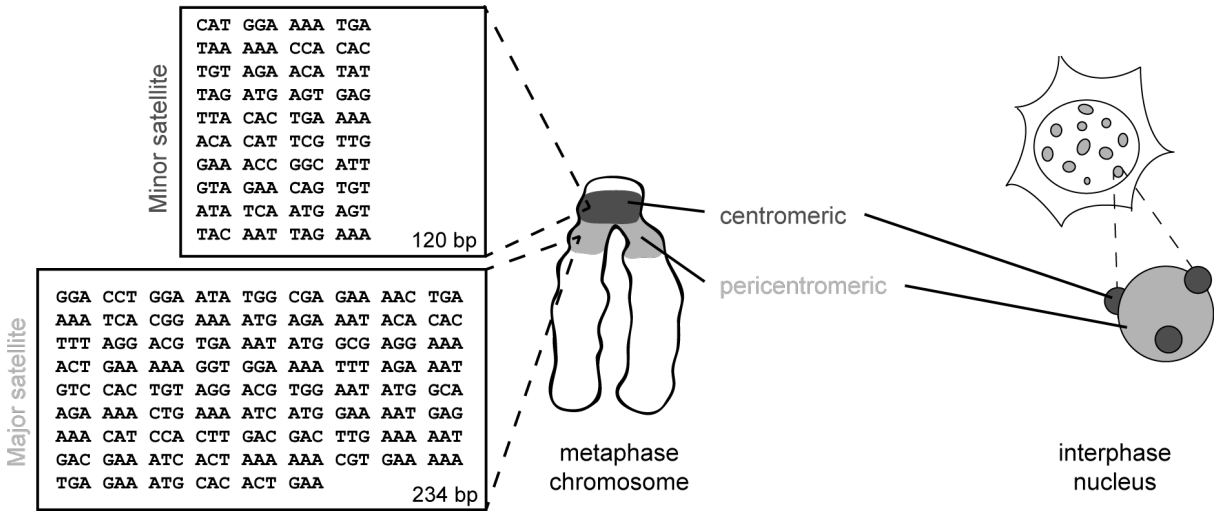



**Figure 1.2: Satellite repeats in mouse constitutive heterochromatin.** The 120 bp long minor satellite repeats are located at the centromeric region (depicted in dark gray), the 234 bp long major satellite repeats at the pericentromeric region (depicted in light gray) of the chromosomes in mice. During interphase, (peri)centromeric regions of different chromosomes cluster to form chromocenters. The pericentromeric regions build up the chromocenter core and the minor satellite regions are attached on its surface. It is still discussed whether the minor satellite repeats are part of the chromocenters.

of constitutive heterochromatin are studied extensively (reviewed in [26, 37]), its functions beyond transcriptional repression remain vague. Pericentromeric heterochromatin was described to ensure genome stability by preventing the activity of transposons and contributing to proper chromosome segregation [58, 60]. In addition, several studies highlight the involvement of chromocenters in chromatin organization during cell cycle [47], development [54] and differentiation [55–57] and disease [61–63]. In the cell nucleus, chromocenters tend to localize in the nuclear periphery or close to nucleoli [51, 53, 64, 65]. Interactions of heterochromatin regions with components of the nuclear lamina and nucleosomes were reported [66–68], suggesting a connection between heterochromatin positioning and gene expression regulation [69]. Furthermore, a role of heterochromatin in the nuclear spatial organization of compartments was proposed [26]. The rod photoreceptor cells of nocturnal animals show an inverted heterochromatin organization, characterized by localization of the heterochromatin in a big chromocenter in the nuclear center surrounded by euchromatin [70]. This inverted organization is established during postnatal development, caused by a loss of expression of lamin B receptor and lamin A/C genes [71] and thought to improve the transmission of light and thereby the vision of nocturnal animals [70]. Recently, Falk *et al.* analyzed these inverted nuclei in comparison to cells with conventional organization and found that attractions between heterochromatic regions are crucial for the establishment of compartmentalization [72]. The conventional nuclear organization additionally required interactions of heterochromatin with the lamina [72].

Although the DNA satellite repeat sequences were long considered nonfunctional, transcripts of repeats sequences were described in various contexts by now [37]. It was proposed that non-coding small RNAs are involved in proliferation during cell cycle [73], development and

differentiation [74–78], cellular senescence [79], cellular stress [80–84] and diseases like cancer [79, 85–87]. But, compared to the number of satellite repeats, the abundance of satellite RNAs is quite low, indicating either very little transcription rates or labile transcripts [37]. Thus, there are still many open questions concerning chromocenter function.

## 1.3 Heterochromatin isolation methods

On a very basic level, chromatin can be described as a complex of DNA wrapped around core histones, but in addition, there are various other proteins involved in building up a highly dynamic network of interactions. Thus, it is practically not possible to perfectly mimic the *in vivo* conditions of pH and ionic strength *in vitro*, as they show differences among individual chromatin components [88]. To elucidate the proteome of specific chromatin fractions, there are two basic concepts to retain as much as possible the *in vivo* interactions. On the one side, some methods aim to enrich the chromatin fraction in its native state and another option is using a crosslinking reagent to fix interactions before applying often harsh conditions for the enrichment. In addition, there are methods to isolate total chromatin (as well as euchromatin/ heterochromatin fractions) and methods to enrich for one specific locus. Table 1.1 shows an overview of chromatin enrichment methods previously applied to heterochromatin. Of note, many of the heterochromatin enrichment techniques listed are not applicable to tissues as they require live cells or genetic engineering of cells [89].

**Table 1.1:** Systems for chromatin isolation.

| Principle | Method | Biological system | Target | Crosslinking necessary | Applicable to tissues | References |
|---|---|---|---|---|---|---|
| Total chromatin isolation techniques | | | | | | |
| differential MNase digestion | Differential MNase | | euchromatin/ heterochromatin | no | no | [90, 91] |
| differential salt extraction | CHESS-DIA | | euchromatin/ heterochromatin | no | no | [92] |
| differential centrifugation in sucrose gradient | Gradient-Seq | | euchromatin/ heterochromatin | yes | yes | [93] |
| centrifugation with sucrose gradient | | mouse | constitutive heterochromatin | no | yes | [94, 95] |
| Locus-specific purification methods | | | | | | |
| DNA-based hybridisation | PICh, HyCCAPP | mouse, human, drosophila, yeast | telomeres, MaSat, MiSat, alpha satellites | yes | yes | [96–102] |
| immunopurification | QTIP | human | telomeres | yes | yes | [103, 104] |
| immunoprecipitation + TALE system | enChIP | mouse | telomeres | yes | no | [105] |

| | | | | | | |
|---|---|---|---|---|---|---|
| immunoprecipitation + CRISPR-dCas | CAPTURE, CLASP (*in vitro*) | human, drosophila | telomeres | yes | no | [106, 107] |
| proximity biotinylation | TRF1-BioID | human | telomeres | no | no | [108] |
| proximity biotinylation + CRISPR-dCas | CasID, C-BERST, CAPLOCUS | mouse, human | telomeres, MaSat, MiSat, alpha satellites | no | no | [109–111] |

*modified from Gauchier *et al.*, 2020 [88] and van Mierlo & Vermeulen, 2021 [89]. Abbreviations: CHESS-DIA - chromatin enriching salt separation coupled to data independent acquisition; PICh - Proteomics of isolated chromatin segments; HyCCAPP - hybridisation capture of chromatin-associated proteins for proteomics; QTIP - quantitative telomeric chromatin isolation protocol; enChIP - engineered DNA-binding molecule-mediated chromatin immunoprecipitation; CAPTURE - CRISPR affinity purification *in situ* of regulatory elements; CLASP - Cas9-locus-associated proteome; C-BERST - dCas9-APEX biotinylation at genomic elements by restricted spatial tagging.

In many studies, heterochromatin enrichment is followed by the analysis of the proteins by mass spectrometry which allows protein identification with high throughput and high sensitivity. Although mass spectrometry is a qualitative analysis method, the implementation of chemical labels, metabolic labeling systems, and external mass standards like AQUA peptides allow quantification of the results by comparison to a standard. In addition, label-free quantification methods were developed (reviewed in [112]).

## 1.4 Methyl-CpG binding protein 2

This section dealing with the Methyl-CpG binding protein 2 was modified from Schmidt & Zhang *et al.*, 2020 [3].

As a hallmark of heterochromatin, 5-methylcytosine marks block the access of several factors to the DNA and serve as a binding site for epigenetic readers. Methyl binding domain (MBD) proteins bind to 5-methylcytosine on the DNA via their MBD domains and recruit additional chromatin binding factors, thus contributing to transcriptional regulation [113, 114]. The methyl-CpG binding protein 2 (MeCP2) is the founding member of the methyl binding domain (MBD) protein family [115] and as well the most studied one.
The *MECP2* gene is highly conserved between species, located on the X chromosome, and mutations in the *MECP2* gene were linked to the neurological disorder Rett syndrome [116]. MeCP2 has two isoforms generated by alternative splicing that start either in exon 1 (MeCP2 e1) or exon 2 (MeCP2 e2) and, thus, result in differing N-terminal protein sequences [117, 118]. Both isoforms are considered functionally equivalent, although different expression patterns in developing and postnatal mouse brains were described [119, 120]. Nevertheless, MeCP2 e1 is the predominant isoform in brain [120] and was proposed to be more relevant in Rett syndrome [121–123]. As MeCP2 e2 was the first isoform identified and most studies are based on this isoform, we use the amino acid coordinates from MeCP2 e2 throughout this thesis.
Trypsin digestion of MeCP2 revealed at least six functionally different MeCP2 domains, namely N-terminal domain (NTD), MBD, intervening domain (ID), transcriptional repression domain

(TRD), C-terminal domain (CTD) alpha and beta [124] (see figure 1.4). Functionally characterized are the MBD, that is required for 5-methylcytosine binding, and the TRD, which was shown to interact with multiple proteins involved in transcriptional repression (see figure 1.3) [125–128]. Recently, the N-CoR/SMRT interacting domain (NID) was mapped additionally [129]. The MeCP2 domain structure is shown in figure 1.3.

### 1.4.1 MeCP2 DNA binding

Early studies on MeCP2 characterized it as a protein capable to bind to a single, symmetrically methylated CpG dinucleotide via the MBD domain spanning amino acids 89 – 162 and thereby overlapping approximately twelve base pairs of DNA [115, 130, 131]. Later studies indicated that the N-terminal domain (NTD) enhanced DNA binding affinity via the MBD [132], while the intervening domain (ID), TRD, and C-terminal domain (CTD) alpha showed methylation-independent DNA binding capabilities and CTD beta was proposed to bind to chromatin, but not to naked DNA [132, 133]. Furthermore, three AT-hook-like domains were identified within the ID, TRD, and CTD alpha domains (AT-hook 1, aa 184–195; AT-hook 2, aa 264–273; AT-hook 3, aa 341–364). The AT-hook motif is a short motif binding to the minor groove of AT-rich DNA via the core consensus amino acid sequence RGRP [134]. These methylation-independent DNA binding capabilities allow MeCP2 to bind to different sites on the DNA at the same time, thus, possibly contributing to genome-wide chromatin organization. Except for the MBD, MeCP2 was shown to be mostly an intrinsically disordered protein. Upon binding to DNA, though, increased secondary structure in the ID and TRD domains was observed [132]. The MBD is the only domain showing structurally conserved motifs, as it contains four beta-sheets and one alpha-helix building up a wedge shape with a beta-sheet face presenting positively charged amino acids for interaction with the DNA as determined by nuclear magnetic resonance analysis [135]. Accordingly, this domain showed only minor conformational changes as a result of DNA binding [132, 135]. The subsequent crystal structure of the MBD bound to the Bdnf gene promotor revealed that MBD mCpG interaction might involve five water molecules, leaving only three amino acids with direct contact with the DNA: D121, R111, and R133 [136]. In line with this study, these amino acids were found mutated in Rett syndrome patients and *MECP2* carrying these mutations showed significantly reduced MeCP2 DNA binding [62, 63, 136].
Dynamic structural analysis of MeCP2 using hydrogen/ deuterium exchange mass spectrometry (H/DX-MS), led to the proposal that the intrinsically disordered MeCP2 samples multiple conformational states, also during non-specific interaction with the DNA [137].
Using genome-wide chromatin immunoprecipitation-sequencing (ChIP-seq) analysis, MeCP2 was found to bind globally across the genome tracking mCpG density [138]. Furthermore, in purified nuclei from mouse brains, MeCP2 was shown to be expressed at near histone octamer levels [138]. These findings suggest that MeCP2 binds globally across the genome reducing transcriptional noise.
Nevertheless, MeCP2 was also described to bind to actively transcribed unmethylated DNA *in vivo* [139, 140] with only a minor portion of MeCP2-bound promoters being highly methylated [140]. A possible explanation would be that MeCP2 folds upon binding to DNA and scans

the DNA for suitable binding sites making use of its non-specific DNA binding sites [124, 132]. Thus, it would only bind non-specifically to active genes to scan the DNA for mCpG binding sites. Recently, MeCP2 was reported to bind not only mCpG but also mCpApC [141]. The patterns of mCpApC differ between neuronal cell types and may, thus, contribute to cell type-specific effects of MeCP2 [142, 143].

In addition to binding DNA and methylated cytosines, MeCP2 was proposed to bind to 5-hydroxymethylcytosine (5hmC) in mouse brain [144] and embryonic stem cells [145]. 5hmC is an oxidation product of 5mC and can be further oxidized to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) by TET (Ten-eleven-translocation) proteins, which might enable active DNA demethylation by different pathways (reviewed in [113]). In addition, 5hmC levels were reported to be differentially distributed between different tissues, about ten times lower than 5mC levels [146] and associated with actively expressed and developmentally regulated genes [147]. Nevertheless, these findings are highly debated, as the results are tissue and cell type dependent [144, 145], the recognition mechanism of 5hmC by MeCP2 is unclear and other studies hint at a binding affinity similar to binding unmethylated DNA [148–150].

A more indirect way for MeCP2 to repress transcription by DNA binding is the protection of MeCP2 bound 5mC against oxidation to 5hmC by TET enzymes by restricting their access to the methylated cytosine [151]. This was proposed to contribute to restricting transcriptional noise [138] and, in particular, repressing tandem repeat DNA expression [151] and L1 retrotransposition [152–154]. TET-mediated L1 activation was shown to be prevented by binding of MeCP2 to 5mC [154].

Summarizing, methylation-specific and unspecific MeCP2 DNA binding are both essential for its function in transcriptional repression and chromatin organization, and its multifunctional domain structure allows the protein to simultaneously bind to DNA and interact with other proteins.

### 1.4.2 MeCP2 protein-protein interactions

Interactions of MeCP2 with several proteins mediate and regulate its multiple functions in transcriptional regulation, chromatin organization and RNA splicing. An overview of interacting proteins, the interacting MeCP2 regions and the function of these interactions is presented in Figure 1.3 and Table 1.2.

One major mechanism by which MeCP2 represses transcription is by recruiting corepressor complexes to methylated DNA. One such complex contains mSin3A and histone deacetylases (HDACs), suggesting that transcriptional repression may in part rely on histone deacetylation [125, 126], e.g., by removing active chromatin marks. mSin3A was shown to be the direct MeCP2 binding partner, whereas HDAC showed a weaker binding affinity to MeCP2 and, thus, might bind via mSin3A [125]. Another corepressor complex reported to interact with MeCP2 is the N-CoR/SMRT interacting with a small region within the TRD domain, which was thus called NID. The data suggested that MeCP2 recruits N-CoR/SMRT to methylated DNA and that this MeCP2 bridge function is disturbed in Rett syndrome [129]. Interestingly, binding of Sin3A was not disrupted by NID mutations [129].

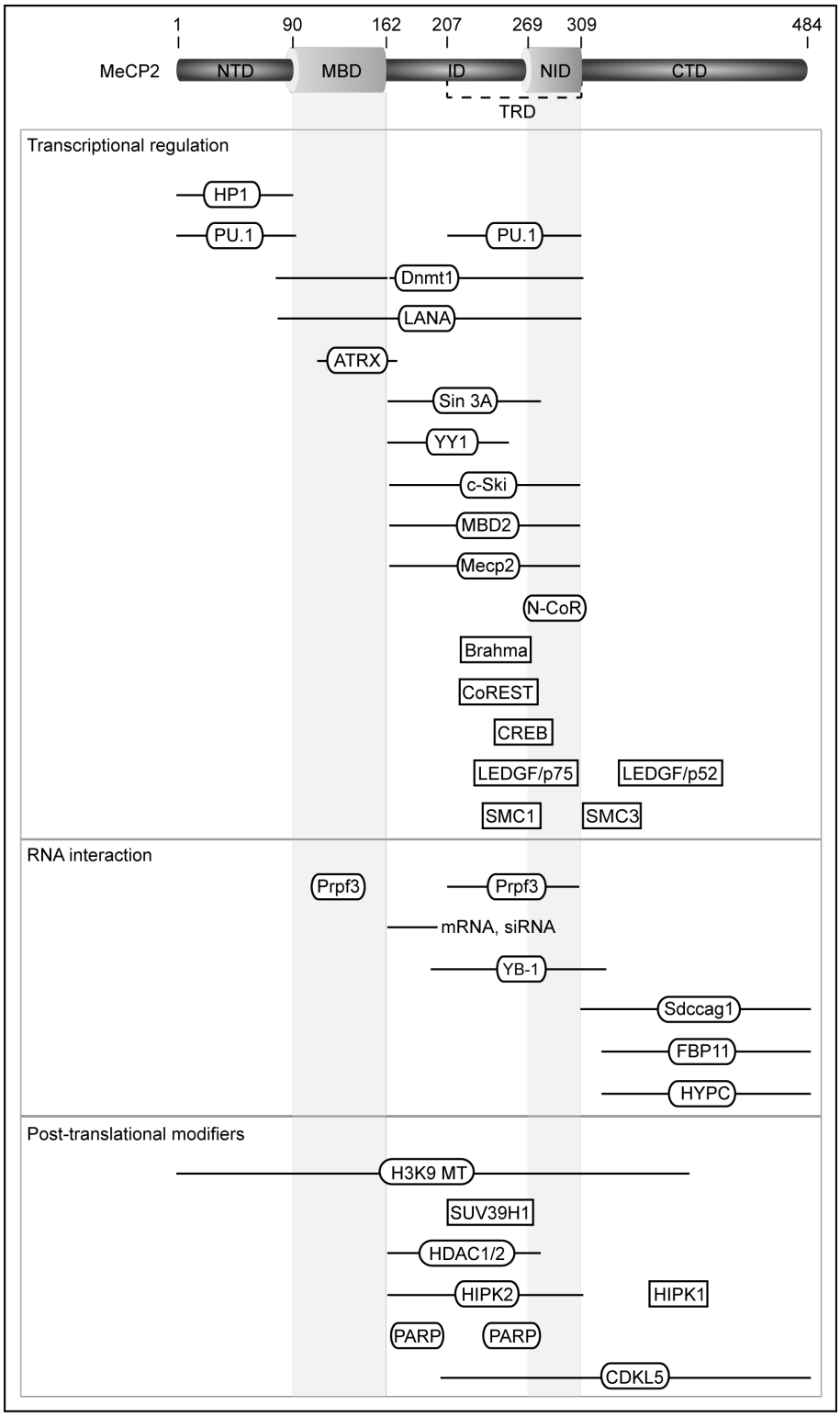



**Figure 1.3: Overview of MeCP2 interaction partners.** MeCP2 interaction partners, group by main function and ordered by where they interact within MeCP2, if known. References are given in Table 1.2. Rectangles indicate proteins with no mapped interaction region within MeCP2. NTD: N-terminal domain; MBD: methyl binding domain; ID: intervening domain; NID: N-CoR interacting domain; CTD:-C-terminal domain; TRD: transcriptional repression domain. Amino acid labeling according to mouse MeCP2 isoform e2. Protein domain structure generated using DOG 1.0 software [155]. This figure was taken from Schmidt & Zhang *et al.*, 2020 [3].

In addition to transcriptional repression, MeCP2 might also work as an activator, as it was found associated with the transcriptional activator CREB1 (cyclic AMP-responsive element-binding protein 1) at the promoter of an activated gene [139]. In gene expression analysis from mouse hypothalamus the gain of MeCP2 was shown to result in more transcriptional activation than repression, whereas MeCP2 loss lead to reverse effects [139]. These results are in line with a previous study, where only a minor portion of MeCP2 was found bound to methylated CpGs, but 63% of MeCP2 were bound to actively expressed promoters [140]. In other studies though, MeCP2 was found to track methylated CpGs genome wide [138], as described above.
As transcriptional activity is influenced by chromatin organization, these MeCP2 functions can hardly be separated. By interacting with the histone methyltransferase acting on histone H3 lysine 9, MeCP2 was reported to target histone methylation to methylated regions on the DNA [156]. As mentioned above, MeCP2 transcriptional repression involves recruitment of histone deacetylases and deacetylation of histones is likely followed by histone methylation [156], thus switching chromatin from an active to a repressive state. Histone methylation may result in recruitment of other proteins like heterochromatin protein 1 (HP1), thus reinforcing the repressed chromatin state [157, 158]. MeCP2 and HP1 were shown to interact [159] and both were reported to associate with SUV39H1 (suppressor of variegation 3-9 homolog 1) histone methyltransferase [128, 160], which methylates histone H3 lysine 9. In addition, MeCP2 might be involved in regulation of maintenance DNA methylation by DNMT1 (DNA methyltransferase 1), as the interaction of both proteins was also described [161]. DNMT1 interacts with HDAC1 and 2 [162, 163], and was shown to replace the mSin3A-HDAC complex upon MeCP2 binding [161].
Another mechanism by which MeCP2 modulates chromatin architecture could be oligomerization. In that regard, MeCP2 was shown to associate with itself and the methyl-CpG binding domain protein 2 (MBD2) [164]. Furthermore, MeCP2 associates with the chromatin remodeling protein ATRX (alpha-thalassemia/ mental retardation syndrome X-linked). Analysis of *Mecp2* null mouse brains showed delocalization of ATRX from heterochromatic foci, suggesting a MeCP2-dependent ATRX targeting to heterochromatic regions in mature neurons [165]. As the MeCP2 mediated ATRX targeting to heterochromatin took place only in mature neurons where MeCP2 is very abundant [165], this underscores the relevance of MeCP2 level for its function.
MeCP2 might also play a functional role in RNA splicing, as it binds to WW domains of the splicing factors FBP (formin-binding protein) 11 and HYPC (Huntington yeast partner C) via a proline-rich domain in the MeCP2 C-terminus [166, 167]. Genotype-phenotype studies on Rett syndrome frameshift mutations support the hypothesis that disruption of the proline-rich region in the MeCP2 C-terminus, thus abolishing its binding to FBP11 and HYPC, contributes to Rett phenotype [167]. In addition, association of MeCP2 with the Y box-binding protein 1 (YB-1), a conserved DNA and RNA binding protein [168], promotes exon inclusion in YB-1 responsive CD44-splicing reporter assays [168]. This leads to the proposal that misregulation of transcription as well as splicing might contribute to Rett syndrome [168].

**Table 1.2:** MeCP2 interaction partners and function upon interaction.

| | Interactor | MeCP2 function upon interaction | References |
|---|---|---|---|
| Transcriptional repression | HP.1 | repression, formation of subcellular silencing compartments | Agarwal *et al.*, 2007 [159] |
| | PU.1 | formation of repression complex, possibly recruitment of mSin3A-HDAC | Suzuki *et al.*, 2003 [169] |
| | Dnmt1 | association with MeCP2 contributes to maintenance methylation | Kimura & Shiota 2003 [161] |
| | LANA | MeCP2 directs LANA to chromocenters, might contribute to LANA-mediated repression | Matsumura *et al.*, 2010 [170], Krithivas *et al.*, 2002 [171] |
| | ATRX | targeting to heterochromatic regions in mature neurons, silencing of imprinted genes; possibly control of nucleosome positioning | Nan *et al.*, 2007 [165], Kernohan *et al.*, 2010 [172] |
| | Sin3A | transcriptional repression, corepression complex with HDAC and MeCP2 | Nan *et al.*, 1998 [125], Jones *et al.*, 1998 [126] |
| | YY1 | cooperation in repression | Forlani *et al.*, 2010 [173] |
| | c-Ski | transcriptional repression | Kokura *et al.*, 2001 [127] |
| | MBD2 | heterointeractions, might increase heterochromatin clustering | Becker *et al.*, 2013 [164] |
| | MeCP2 | homointeractions, might increase heterochromatin clustering | Becker *et al.*, 2013 [164] |
| | N-CoR | recruitment of N-CoR/SMRT to methylated DNA, bridge function of MeCP2 | Kokura *et al.*, 2001 [127], Lyst *et al.*, 2013 [129] |
| | Brahma | transcriptional repression | Harikrishnan *et al.*, 2005 [174] |
| | CoREST | transcriptional repression possibly involving REST, CoREST, MeCP2, SUV39H1 and HP1 | Lunyak *et al.*, 2002 [128] |
| | CREB | transcriptional activation | Chahrour *et al.*, 2008 [139] |
| | LEDGF/p75 | might differentially influence gene activation | Leoh *et al.*, 2012 [175] |
| | SMC1, SMC3 | interaction with MeCP2, ATRX, might promote repression by loop formation | Kernohan *et al.*, 2010 [172], Gonzales *et al.*, 2012 [176] |
| RNA interaction | Prpf3 | RNA binding, possibly involved in splicing | Long *et al.*, 2011 [177] |
| | mRNA, siRNA | not known | Jeffrey *et al.*, 2004 [178] |
| | YB-1 | RNA-dependent complex, regulation of splicing | Young *et al.*, 2005 [168] |
| | Sdccag1 | not known | Long *et al.*, 2011 [177] |
| | FBP11 | not known | Buschdorf & Stratling 2004 [167], Bedford *et al.*, 1997 [166] |
| | HYPC | not known | Buschdorf & Stratling 2004 [167] |
| post-translational modifications | H3K9 MT | targeting of histone methylation to methylated DNA | Fuks *et al.*, 2003 [156], Lunyak *et al.*, 2002 [128] |
| | SUV39H1 | association with MeCP2 might contribute to silencing by methylation of H3K9, creating HP1 binding sites | Lunyak *et al.*, 2002 [128] |
| | HDAC 1/2 | histone deacetylases form corepression complex with MeCP2 and Sin3A | Nan *et al.*, 1998 [125], Jones *et al.*, 1998 [126] |
| | HIPK2, HIPK1 | kinases might phosphorylate MeCP2 on S80 and S216 | Bracaglia *et al.*, 2009 [179], Lombardi *et al.*, 2017 [180] |
| | PARP | poly(ADP-ribosyl)ation reduces MeCP2 heterochromatin clustering ability | Becker *et al.*, 2016 [181] |
| | CDKL5 | association *in vitro*, phosphorylation of MeCP2 by CDKL5 unclear (opposing results in the two publications) | Mari *et al.*, 2005 [182], Lin *et al.*, 2005 [183] |

This table was taken from Schmidt & Zhang *et al.*, 2020 [3].

Although several MeCP2 interaction partners were identified so far, the whole network of protein-protein interactions, their interplay and the entire composition of MeCP2 silencing compartments require further investigation. Importantly, MeCP2 DNA binding and protein-protein interactions need to be studied in the context of post-translational modifications as these can abolish or enhance DNA and protein binding, thus, ultimately influencing chromatin organization.

### 1.4.3 MeCP2 post-translational modifications

Recently, several MeCP2 post-translational modifications (PTMs) were reported, mostly in large-scale proteomic studies focusing on mapping one specific PTM in the whole proteome. In Table 1.3, experimentally determined MeCP2 modifications are summarized, together with the species in which they were identified, the methods used for identification along with references. A more detailed list can be found on PhosphoSitePlus.org [184], including additional sites only available as curated datasets.

**Table 1.3:** Summary of MeCP2 post-translational modifications.

| | Residue | Modification | Species | MS / other methods | References |
|---|---|---|---|---|---|
| NTD | K12 | ubi | human | x / - | Gonzales *et al.*, 2012 [176] |
| | S13 | phos | human, mouse | x / - | Gonzales *et al.*, 2012 [176], Humphrey *et al.*, 2013 [185], Shiromizu *et al.*, 2013 [186] |
| | S53 | phos | human | x / - | Shiromizu *et al.*, 2013 [186], Bian *et al.*, 2014 [187], Sharma *et al.*, 2014 [188] |
| | S68 | phos | mouse | x / - | Huttlin *et al.*, 2010 [189] |
| | S70 | phos | mouse, human | x / - | Huttlin *et al.*, 2010 [189], Mertins *et al.*, 2016 [190] |
| | S78 | phos | human, mouse, rat | x / - | Dephoure *et al.*, 2008 [191], Zanivan *et al.*, 2008 [192], Tweedie-Cullen *et al.*, 2009 [193] |
| | S80 | phos | human, mouse, rat | x / x | Zhou *et al.*, 2006 [194], Tao *et al.*, 2009 [195], Bracaglia *et al.*, 2009 [179] |
| | K82 | ubi | human | x / - | Gonzales *et al.*, 2012 [176] |
| | S86 | phos | mouse, human | x / x | Ebert *et al.*, 2013 [196], Mertins *et al.*, 2014 [197] |
| MBD | R115 | met | human | x / - | Geoghegan *et al.*, 2015 [198] |
| | S116 | phos | human | x / - | Dephoure *et al.*, 2008 [191], Kettenbach *et al.*, 2011 [199], Sharma *et al.*, 2014 [188] |
| | K119 | ubi, dimet | human | x / - | Gonzales *et al.*, 2012 [176], Jung *et al.*, 2008 [200] |
| | Y120 | phos | human, mouse | x / x | Dephoure *et al.*, 2008 [191], Bergo *et al.*, 2015 [201], D'Annessa *et al.*, 2018 [202] |
| | K130 | ubi | human | x / - | Wagner *et al.*, 2011 [203], Gonzales *et al.* 2012 [176] |
| | K135 | ubi | human | x / - | Gonzales *et al.*, 2012 [176] |
| | T148 | phos | mouse | x / - | Tao *et al.*, 2009 [195] |
| | S149 | phos | mouse, human | x / - | Tao *et al.*, 2009 [195], Olsen *et al.*, 2010 [204], Kettenbach *et al.*, 2011 [199] |
| | T160 | phos | mouse | x / - | Tweedie-Cullen *et al.*, 2009 [193] |
| | R162 | met | mouse, human | x / - | Guo *et al.*, 2014 [205], Larsen *et al.*, 2016 [206] |

| | | | | | |
|---|---|---|---|---|---|
| ID | 163 - 206 | PAR | human, mouse, rat | x / x | Jungmichel *et al.*, 2013 [207], Becker *et al.*, 2016 [181] |
| | S164 | phos | mouse | x / x | Tao *et al.*, 2009 [195], Tweedie-Cullen *et al.*, 2009 [193], Stefanelli *et al.*, 2016 [208] |
| | S166 | phos | mouse, human | x / - | Huttlin *et al.*, 2010 [189], Yi *et al.*, 2014 [209], Mertins *et al.*, 2014 [197] |
| | S178 | phos | human | x / - | Shiromizu *et al.*, 2013 [186] |
| | T184 | phos | human, mouse | x / - | Mertins *et al.*, 2014 [197] |
| | T203 | phos | human | x / - | Carrier *et al.*, 2016 [210] |
| | S204 | phos | human | x / - | Carrier *et al.*, 2016 [210] |
| | K210 | dimet | human | x / - | Jung *et al.*, 2008 [200] |
| | S216 | phos | human (mouse, rat) | x / x | Olsen *et al.*, 2010 [204], Kettenbach *et al.*, 2011 [199], Lombardi *et al.*, 2017 [180] |
| | K219 | acet | rat | x / - | Lundby *et al.*, 2012 [211] |
| | K223 | ubi | human | x / - | Akimov *et al.*, 2018 [212] |
| | K223 | SUMO | mouse | - / x | Cheng *et al.*, 2014 [213] |
| | T228*** | phos | human | x / - | Mertins *et al.*, 2014 [197] |
| | S229 | phos | human, rat (mouse) | x / x | Zhou *et al.*, 2006 [194], Chen *et al.*, 2009 [214], Gonzales *et al.*, 2012 [176] |
| | K233 | ubi | human | x / - | Gonzales *et al.*, 2012 [176] |
| | 244 - 275 | PAR | human, mouse, rat | x / x | Jungmichel *et al.*, 2013 [207], Becker *et al.*, 2016 [181] |
| | K249 | ubi | human | x / - | Gonzales *et al.*, 2012 [176] |
| | K256 | ubi | human | x / - | Gonzales *et al.*, 2012 [176] |
| | K267 | met | human | x / - | Wu *et al.*, 2015 [215] |
| NID | K271 | ubi | human | x / - | Gonzales *et al.*, 2012 [176] |
| | S274 | phos | mouse (human) | x / x | Tweedie-Cullen *et al.*, 2009 [193], Humphrey *et al.*, 2013 [185], Ebert *et al.*, 2013 [196] |
| | S292 | phos | mouse, rat | x / x | Humphrey *et al.*, 2013 [185], Liu *et al.*, 2015 [216] |
| | S295 | phos | mouse | x / - | Humphrey *et al.*, 2013 [185] |
| | K305 | ubi | human | x / - | Gonzales *et al.*, 2012 [176] |
| | K307 | ubi, acet | human | x / - | Gonzales *et al.*, 2012 [176] |
| | T308 | phos | mouse | - / x | Ebert *et al.*, 2013 [196] |
| CID | T311 | phos | mouse, human | x / - | Huttlin *et al.*, 2010 [189], Mertins *et al.*, 2014 [197], Parker *et al.*, 2015 [217] |
| | S313 | phos | human, mouse | x / - | Bian *et al.*, 2014 [187], Sharma *et al.*, 2014 [188], Parker *et al.*, 2015 [217] |
| | K321 | acet, ubi | human, mouse | x / - | Gonzales *et al.*, 2012 [176], Beli *et al.*, 2012 [218], Weinert *et al.*, 2013 [219] |
| | T327 | phos | human | x / - | Shiromizu *et al.*, 2013 [186] |
| | S341 | phos | mouse | x / - | Humphrey *et al.*, 2013 [185] |
| | K347 | met | human | x / x | Dhayalan *et al.*, 2011 [220], Wu *et al.*, 2015 [221] |
| | S357 | phos | human | x / - | Yang *et al.*, 2006 [222] |
| | S359 | phos | human | x / - | Yang *et al.*, 2006 [222], Bian *et al.*, 2014 [187] |
| | S360 | phos | human, mouse | x / - | Yang *et al.*, 2006 [222], Grimsrud *et al.*, 2012 [223], Humphrey *et al.*, 2013 [185] |
| | S393 | phos | human | x / - | Bian *et al.*, 2014 [187] |
| | S399 | phos | mouse, rat, human | x / - | Tao *et al.*, 2009 [195], Gonzales *et al.*, 2012 [176] |
| | S421 | phos | mouse, rat (human) | x / x | Zhou *et al.*, 2006 [194], Tao *et al.*, 2009 [195], Deng *et al.*, 2010 [224] |

| | | | | |
|---|---|---|---|---|
| S424 | phos | human, rat, mouse | x / x | Dephoure *et al.*, 2008 [191], Tao *et al.*, 2009 [195], Li *et al.*, 2011 [225] |
| T434 | gl | rat, mouse | x / - | Wang *et al.*, 2010 [226], Alfaro *et al.*, 2012 [227], Trinidad *et al.*, 2012 [228] |
| T441 | gl | mouse | x / - | Alfaro *et al.*, 2012 [227] |
| T443/ T444*** | gl | rat | x / - | Wang *et al.*, 2010 [226] |
| K447 | acet | human | x / - | Choudhary *et al.*, 2009 [229], Beli *et al.*, 2012 [218], Wu *et al.*, 2015 [215] |
| T477 | phos | human | x / - | Sharma *et al.*, 2014 [188] |
| S484 | phos | human, mouse | x / - | Kettenbach *et al.*, 2011 [199], Schweppe *et al.*, 2013 [230], Mertins *et al.*, 2014 [197] |

Modifications identified by mass spectrometry might have unclear localization. *modification numbering according to mouse MeCP2 isoform starting in exon 2 (mouse: 484 aa, human: 486 aa, rat: 492 aa) **references only exemplary (for more information see PhosphositePlus.org) ***residue numbering according to species mentioned as it differs from mouse. This table was taken from Schmidt & Zhang *et al.*, 2020 [3].

Most of the modifications were identified in large-scale studies and not further validated by any other assay. Furthermore, in most cases no additional information is available regarding their influence on MeCP2 function (e.g. [185–188, 198, 199, 203]). Many of these PTMs were mapped using a single cell line (e.g. [185, 188, 199]), and their existence *in vivo* has not been demonstrated. For these reasons, we will focus here on the more detailed studies providing validation and functional relevance of MeCP2 PTMs, in particular, within the context of chromatin.

The first phosphorylation (phos) site identified on MeCP2 was mapped to the CTD on serine 421. S421phos was found as an upshifted band on Western blot analysis upon membrane depolarization [194, 231, 232] and occurred exclusively in the brain, although MeCP2 was detected in many other tissues [233]. S421A/S424A double mutant mice showed better performance in hippocampal memory tests, enhanced long-term potentiation [224] and increased locomotor activity [195]. Analysis of MeCP2 S421A mice revealed an increased dendritic complexity, and defects in the response to novel experiences [234]. As global S421phos was observed upon membrane depolarization, this modification might not regulate the expression of specific genes, but rather be involved in modulating global response to membrane depolarization [234].

Together with S421phos, S80phos within the NTD is one of the most studied MeCP2 phosphorylation sites with functional characterization. In contrast to S421 phosphorylation, serine 80 was reported to be dephosphorylated upon membrane depolarization and S80A mutant mice show decreased locomotor activity [195]. The modification is highly enriched in the brain and ubiquitously distributed similarly to total MeCP2 [195]. S80A mutation decreased MeCP2 chromatin binding affinity, although the MeCP2 S80A protein levels and subcellular distribution did not differ relative to the wildtype MeCP2. Thus, it was suggested that the phosphorylation possibly fine-tunes chromatin association [195]. The homeodomain-interacting protein kinases 1 (HIPK1) and 2 (HIPK2) were proposed to be responsible for MeCP2 phosphorylation at serine 80 [179, 180].

Another MeCP2 phosphorylation site influencing chromatin binding affinity was identified on tyrosine 120 within the MBD domain of MeCP2. This tyrosine residue is substituted in a Rett syndrome patient by aspartic acid [235], which could mimic the phosphorylated state. *Mecp2*

Y120D mutation was found to cause a decrease in binding affinity of MeCP2 to heterochromatin [62]. This could be explained at the structural level by computational modeling indicating that MeCP2 Y120D drastically reduces MeCP2 affinity for DNA as compared to wildtype MeCP2 [202].
A conserved serine (S164) located at the beginning of ID just after the MBD, was shown to be abundantly phosphorylated in the brain in a developmentally regulated manner [208]. While the phospho-mimicking version S164D showed minor binding to chromatin in live-cell kinetic studies, the phospho-defective mutation S164A had the opposite effect [208]. These results could be explained by *in silico* modeling of the 3D structure of this phosphorylation site, revealing the addition of negative charge to the protein surface as a consequence of S164 phosphorylation, hence, decreasing DNA binding. Immunofluorescence analysis of wild type neurons versus *Mecp2* S164 mutants revealed that temporal regulation of S164 phosphorylation is required for proper nuclear size and neuronal dendritic branching [208].
In addition to phosphorylation, poly(ADP-ribosyl)ation (PAR) of MeCP2 at the ID and TRD domains was reported to occur *in vivo* in the mouse brain and to influence heterochromatin structure. The addition of this anionic modification within the two highly cationic MeCP2 protein domains responsible to bind DNA was proposed to lead to a general decrease in DNA binding affinity [181]. Concomitantly, poly(ADP-ribosyl)ation of MeCP2 was shown to reduce binding and clustering of pericentromeric heterochromatin in cell-based assays, suggesting a role of this PTM in MeCP2 chromatin architecture regulation [181].
Altogether, MeCP2 modifications have been shown to regulate its ability to bind and organize DNA/chromatin, as they change the molecular properties of the respective amino acids, which can be critical depending on the position of the residue within the MeCP2 domains. Yet, as mentioned above, most of the modifications identified in MeCP2 have not been functionally characterized and their role in Rett syndrome is unclear.

### 1.4.4 *MECP2* Rett mutations

Mutations in the *MECP2* gene were reported to cause Rett syndrome, a neurodevelopmental disorder affecting mainly girls and one of the most common causes of intellectual disability in females [116]. The disorder is characterized by normal development of the patients until the age of 7 to 18 months, when their development stagnates followed by a decline of higher brain functions and loss of acquired developmental skills [236, 237]. Common symptoms include communicative disfunction, social withdrawal, mental deficiency, loss of speech and purposeful hand movements, and breathing irregularities [237]. *MECP2* Rett mutations were identified in all protein domains (see RettBASE), but the severity of effects varies depending on the location (see [3]). While missense mutations are mostly found within the MBD domain, nonsense mutations frequently occur in the TRD (figure 1.4).
The most frequent *MECP2* missense mutation localized in the MBD is T158 which can be changed to T158M and T158A. On the level of protein function, it results in decreased protein stability and methyl-DNA binding ability [238, 239]. Also, R133C and R106W are common Rett syndrome mutations located in the MBD and were shown to reduce DNA binding [239, 240].

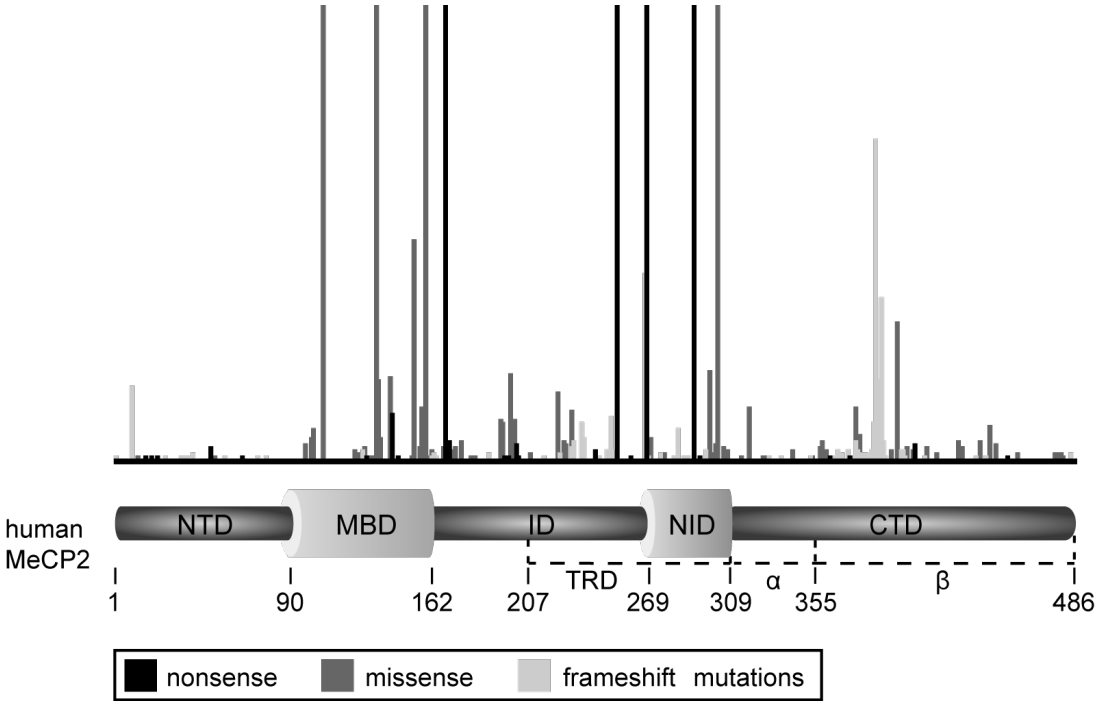



**Figure 1.4: *MECP2* mutation spectrum in Rett syndrome patients.** The frequency of missense mutations is shown in dark gray, of nonsense mutations in black and of frameshift mutations in light gray. The location of the mutations is indicated in a schematic representation of the domain structure of MeCP2. NTD: N-terminal domain, MBD: Methyl binding domain, ID: Intervening domain, NID: N-CoR interacting domain, TRD: Transcriptional repression domain, CTD: C-terminal domain.

While R133 is one of the three amino acids directly contacting the DNA, T158 and R106 are involved in the stabilization of MeCP2 DNA binding via the Asx-ST motif [136].

The frequent Rett mutations in ID and NID are nonsense mutations resulting in truncated protein versions. For instance, MeCP2 R168X (ID) and R270X (NID) both showed deficits in the compaction of nucleosomal arrays and reduced DNA binding [241–243], underscoring the importance of the ID and NID for methylation independent DNA binding and chromatin compaction. A more detailed summary of high frequency Rett syndrome related *MECP2* point mutations and their phenotypes are given in [3].

Moreover, differences in MeCP2 protein levels can cause disease phenotypes. Abnormal neurological phenotypes were reported for *Mecp2*-deficient mice [244, 245] but also for mice with mild *Mecp2* overexpression [246] or humans with *MECP2* duplications [247]. Of note, phenotypes caused by reduced MeCP2 levels could be rescued by expression of *Mecp2* [248–252].

### 1.4.5 MeCP2 in heterochromatin organization

MeCP2 is an epigenetic reader, accumulates at constitutive heterochromatic regions, and interacts with many corepressors as described above. In addition, MeCP2 might act as a histone H1 like chromatin linker. MeCP2 and H1 bind to the linker DNA entry-exit site of nucleosomes [7, 133] and MeCP2 binding to nucleosomes was described to produce a similar architectural motif as H1 on nucleosomes [253]. Furthermore, MeCP2 was able to replace histone H1 from chromatin

*in vitro* [253]. Both proteins showed similar mobility and competed for binding to nucleosomes *in vivo* [253]. A study of H1 levels in wild type and *Mecp2*-deficient neurons revealed that the H1 level increases in the absence of MeCP2, indicating that the proteins replace each other on the nucleosome binding site [138]. As MeCP2 in neurons was reported to be nearly as abundant as histone octamers, it was proposed that MeCP2 acts as a global repressor dampening transcriptional noise genome-wide rather than as a repressor of specific genes [138].
In addition, MeCP2 was shown to compact nucleosomal arrays *in vitro* [241]. This observation was independent of methyl DNA binding and mediated by ID, TRD, and CTD, as MeCP2 Rett mutants with truncations missing TRD and CTD could not compact and oligomerize nucleosomal arrays [132, 241, 242, 254]. This observation might at least in part explain the severe phenotypes of Rett syndrome patients caused by the nonsense mutations R168X and R270X.
Another important role of MeCP2 in heterochromatin organization was described by Brero *et al.*, who observed increasing MeCP2 levels upon myogenic differentiation in parallel with increased clustering of pericentromeric heterochromatin, visualized by lower heterochromatin cluster numbers [55]. Moreover, even ectopic expression of a construct coding for MeCP2-YFP could promote pericentromeric heterochromatin clustering in a dose-dependent manner in a cellular system [55]. While wild type embryonic stem cells showed increasing MeCP2 levels and heterochromatin clustering with ongoing differentiation, *Mecp2*-deficient embryonic stem cells were impaired in heterochromatin organization [57]. In addition, ectopic expression of MeCP2 fusion protein constructs with Rett mutations showed impaired heterochromatin clustering function and decreased heterochromatin accumulation [62]. Of note, retargeting of the MeCP2 Rett mutants with impaired chromatin binding ability to heterochromatic regions restored their clustering function [63].
Interestingly, MeCP2 was recently shown to undergo liquid-liquid phase separation in different conditions [1, 255, 256] and this mechanism was proposed to be involved in heterochromatin cluster formation [257–259].

# 2 Aims of the study

In the eukaryotic nucleus, inactive genes are packed into highly compacted heterochromatin. There are two types of heterochromatin, the facultative heterochromatin which becomes active in specific situations, and the constitutive heterochromatin which is constantly silenced. During the interphase in mouse cells, constitutive heterochromatin from (peri)centromeric regions of several chromosomes clusters to form chromocenters. The organization of these clusters can differ between cell types, during development or differentiation. Aberrant changes in heterochromatin structure and partial activation of silenced regions were associated with diseases. Thus, it is important to unravel the mechanisms and regulation of heterochromatin organization. Here we studied the proteomic composition of mouse chromocenters from brain and liver (figure 2.1A) to address the following questions:

Ia Which proteins are essential for constitutive heterochromatin formation and common among tissues?

Ib Which proteins shape the differing organization of constitutive heterochromatin in tissues?

Heterochromatin is characterized by high cytosine methylation levels and methyl-CpG binding proteins. The methyl-CpG binding protein 2 (MeCP2) is an epigenetic reader binding to 5-methylcytosine and is involved in transcriptional regulation and chromatin organization. It was found highly abundant in the brain and was shown to induce heterochromatin clustering in a dose-dependent manner. Mutations in the *MECP2* gene, either affecting its function or altering its level, were associated with the neurological disorder Rett syndrome. For these reasons, we aimed to analyze the MeCP2 function in chromatin organization in a level-dependent manner (figure 2.1B, C) addressing the following questions:

II What is the absolute concentration of MeCP2 in the nucleus and heterochromatin compartments in the mouse brain in comparison to cells?

III In addition to the MeCP2 level, how do post-translational modifications regulate MeCP2 function in chromatin organization?

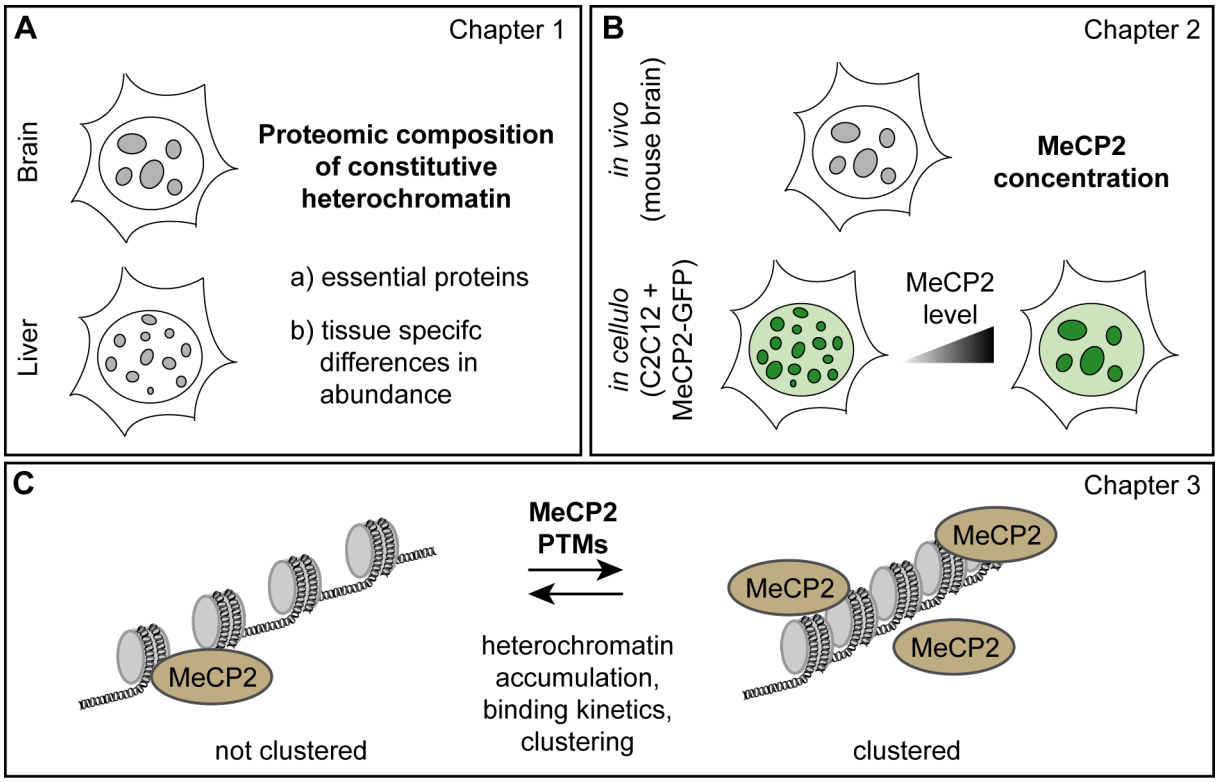



**Figure 2.1: Scheme illustrating the aims of the study.** (A) Analysis of the proteomic composition of constitutive heterochromatin in mouse tissues: Essential chromocenter proteins and tissue specific differences in abundance. (B) Quantification of MeCP2 in the cell nucleus and in heterochromatin compartments. (C) Identification of MeCP2 post-translational modifications and characterization of their functional consequences on heterochromatin organization.

# 3 The proteomic composition of constitutive heterochromatin in mouse tissues

## 3.1 Introduction

The DNA in the cell nucleus is not packed uniformly and chromatin can be broadly subdivided into two cytologically different types, the more open euchromatin and the highly compacted heterochromatin [23, 24]. Euchromatin contains the majority of actively expressed genes, whereas heterochromatin comprises mainly inactive genes and various repeat elements. The latter is further characterized by high cytosine methylation levels and methyl-CpG binding proteins, as well as specific histone modifications like H3K9 trimethylation and proteins recognizing these modifications [27]. Heterochromatin can be subdivided into facultative and constitutive heterochromatin [260]. Facultative heterochromatin can become transcriptionally active in different settings as upon cell differentiation or during development [27, 36]. Constitutive heterochromatin is considered more stable occurring at the same genetic regions in different cell types [37] and it forms mainly at repetitive sequences at (peri)centromeric regions and telomeres [27, 37]. In mouse cells, major satellite repeats are located at the pericentromeric, minor satellite sequences in the centromeric region of metaphase chromosomes [38–40]. In the interphase, the (peri)centromeric satellite DNA from different chromosomes clusters to form so-called chromocenters ([25], reviewed in [26]). Several studies highlight the involvement of chromocenters in chromatin organization [72] during cell cycle [47], development [54] and differentiation [55–57] and disease [61–63]. Although pericentromeric DNA satellite repeat sequences were first considered to be non-functional, several studies reported possible functions of transcription of constitutive heterochromatin (reviewed in [37, 84]). Its transcription was described in various contexts, including during cell cycle in proliferating cells [73], development and differentiation [74–78], cellular senescence [79], cellular stress [80–84] and diseases like cancer [79, 85–87]. Although pericentromeric regions are evolutionary conserved, the DNA tandem repeats showed variations in sequence and length between species (reviewed in [41]). The chromocenter number was differing between cell types [50–52], changing during development [54], upon cell differentiation [55–57] or upon overexpression of certain proteins [55].

To elucidate the heterochromatin organization in mouse tissues, we performed a microscopic analysis on mouse tissue sections and found differences between brain and liver tissue. We hypothesized that the distinct heterochromatin organization might originate from a differing quantitative proteomic composition. Thus, we analyzed the heterochromatin proteome in an unbiased manner and from tissues (*in vivo*) by adapting a protocol for chromocenter isolation based on a series of sucrose gradient centrifugations from Prusov & Zatsepina [94, 95]. The quantitative mass spectrometry analysis of pericentromeric heterochromatin fractions and whole nuclei resulted in the identification of common chromocenter proteins, but also quantitative differences in chromocenter protein abundance between the tissues. Validation of the protein hits by immunofluorescence staining and Western blots revealed differences in heterochromatin accumulation between the tissues.

## 3.2 Material & methods

### 3.2.1 Organ preparation

For nuclei and subsequent chromocenter isolation, 3-month-old C57BL/6 mice (Charles River Laboratories, Inc.) were sacrificed according to the animal care and use regulations (Government of Hessen, Germany) and the organs of interest were collected, washed with PBS, and frozen in liquid nitrogen.
The organs used for tissue sections and subsequent immunofluorescence staining were isolated as described above and fixed in 10% buffered formalin solution (#HT501128, Sigma-Aldrich) for 24 h. Tissues were sequentially dehydrated in 70% ethanol for 30 min, 70% ethanol for 45 min, 96% ethanol for 60 min and 96% ethanol for 45 min, twice in absolute ethanol for 45 min followed by xylol for 60 min and 30 min. The organs were embedded in paraffin (#CN49.2, Carl Roth), transferred to embedding cassettes, cooled down slowly, and sliced using a microtome into 6 µm slices.

### 3.2.2 Isolation of chromocenters from mouse organs

All steps for nuclei and chromocenter isolation were performed with solutions precooled on ice. For nuclei isolation the frozen mouse brains were crushed to powder and homogenized in 15 ml 0.25 M sucrose solution in buffer A (20 mM triethanolamine-HCl (pH 7.6), 30 mM KCl, 10 mM $MgCl_2$, 1 mM Dithiothreitol (DTT), 1 mM phenylmethylsulfonyl fluoride (PMSF) (#6367.1, Carl Roth)). After centrifugation for 10 min at 1000 xg, the supernatant was discarded and the pellet was resuspended in 2.5 M sucrose buffer (2.5 M sucrose in buffer A) to a final sucrose concentration of 2.1 M. The raw nuclei fraction was obtained by centrifugation for 30 min at 50000 xg using an SW28 rotor (Swinging-bucket rotor SW28, Beckman Coulter). The pellet was resuspended in 0.25 M sucrose buffer (0.25 M sucrose in buffer A) and centrifuged at 1000 xg.
The nuclei were counted and defined numbers ($7.1 \times 10^7$ nuclei/ml in 5 ml per tube) were used for chromocenter isolation modified from a protocol from Prusov and Zatsepina [94, 95]. First, the nuclei were resuspended in 20 ml buffer B (50 mM triethanolamine-HCl (pH 7.6), 5 mM $MgCl_2$, 0.2% Triton X-100), incubated for 5 min on ice and centrifuged at 1000 xg for 10 min. For washing, the pellet was resuspended in 20 ml buffer C (2 mM triethanolamine-HCl (pH 7.6), 0.5 mM $MgCl_2$) and centrifuged at 1000 xg for 10 min. The pellet was resuspended in 5 ml buffer D (2 mM triethanolamine-HCl (pH 7.6), 0.2 mM $MgCl_2$) to a concentration of $7.1 \times 10^7$ nuclei/ml and sonicated two times for 20 s at 20% power (250-450 Sonifier, BRANSON ultrasonic corporation). Each step was controlled on a light microscope. The sonicated fraction was diluted in buffer D to a volume of 10 ml, RNAseA was added to a final concentration of 1 mg/ml (#10109169001, Roche Life science products) and the samples were incubated rotating at 4 °C overnight. The suspension was layered on 5 ml of 0.5 M sucrose in buffer D and centrifuged at 400 xg for 10 min. Then, the supernatant was layered on 5 ml 1 M sucrose in buffer D and centrifuged at 2500 xg (without deceleration using the break function of the centrifuge) to sediment nucleoli. Then, the supernatant was transferred to a centrifugation tube for the SW28 rotor, layered on 5 ml 1 M sucrose in buffer D, and overlayed with 10 – 15 ml buffer D for centrifugation at 27000 xg for

25 min to obtain the crude chromocenter fraction. Subsequently, the chromocenter pellet was resuspended in 8 ml 0.5 M sucrose in buffer E (2 mM triethanolamine-HCl (pH 7.4), 0.05 mM $MgCl_2$) to be loaded on a sucrose gradient. For the gradient, 1.8 M sucrose (E3), 1.4 M sucrose (E2), and 1 M sucrose (E1) in buffer E were layered on top of each other in an SW28 centrifugation tube using 8 ml each, the sample was loaded on top and the tube was filled up with buffer E. Centrifugation at 32000 xg for 40 min without deceleration resulted in two chromocenter bands between the different sucrose concentrations (see scheme in figure 3.2). The chromocenter bands were extracted, diluted in buffer E in a new centrifugation tube, and centrifuged at 82000 xg for 25 min. The resulting chromocenter pellet was diluted in buffer E and stored at -80 °C.

### 3.2.3 Quantitative mass spectrometry

The sample preparation and mass spectrometry measurements described here were carried out by our cooperation partners Oliver Popp and Gunnar Dittmar (MDC Berlin), and the samples analyzed by mass spectrometry were prepared by Stephanie Meyer.
Briefly, the proteins from nuclei and chromocenter fractions were precipitated using methanol-chloroform precipitation (adapted from [261]). The pellet containing the proteins was dried and subsequently resuspended in denaturation buffer (6 M urea, 2 M thiourea, 10 mM HEPES-KOH pH 8) by sonication. 50 µg of each sample were reduced using 1 mM tris(2-carboxyethyl) phosphine (TCEP) and carbamidomethylated using 5.5 mM chloroacetamide. The digestion was performed with 0.5 µg sequencing grade endopeptidase Lys-C (Wako) for 3 h at room temperature, followed by dilution with four volumes of 50 mM ammonium-bicarbonate buffer and overnight incubation with 1 µg sequencing grade trypsin (#V5111, Promega Corporation). The reaction was stopped by adding trifluoracetic acid (TFA) to a final concentration of 1% and the peptides were purified on C18 stage tips. For dimethyl labeling 1/4 of each sample was pooled for medium-heavy labeling while 3/4 of each sample was individually labeled with a light labeling reagent (see figure 3.15). The peptides were dried using a vacuum concentrator, reconstituted in 100 mM triethylammonium bicarbonate (TEAB) and dimethyl-labeled in an automatic setup (see [262, 263]). Then, each light sample was pooled with 1/4 of the medium-heavy labeled master mix and desalted using C18 stage tips.
The samples were measured on a QExactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific) connected to a Proxeon nano-LC system (Thermo Fisher Scientific) using a high performance liquid chromatography tandem mass spectrometry (HPLC-MS/MS) method with data-dependent acquisition selecting the TOP 10 peaks for high-energy collisional dissociation (HCD) fragmentation. 5 µl sample were injected, loaded on a nano-LC column (0.074 mm x 250 mm, 3 µm Reprosil C18, Dr Maisch GmbH) and eluted using a 4 h gradient from 4% to 76% acetonitrile (solvent A: 5% acetonitrile, 0.1% formic acid; solvent B: 80% acetonitrile, 0.1% formic acid).

### 3.2.4 Mass spectrometry data analysis

The database search for protein identification described here was performed by our cooperation partner Jiaxuan Chen (Proteomics core facility, IMB Mainz) using MaxQuant software version 1.6.5.0 [264, 265]. The search was done using the Uniprot database [266] for *Mus musculus* (02/2019) and without matching between runs. Carbamidomethylation was set as a fixed modification, and methionine oxidation and lysine acetylation on the protein N-terminus were set as variable modifications. The multiplicity for quantification was set to two and a false discovery rate (FDR) of 0.01 was applied. The MaxQuant output protein groups table was filtered removing contaminants, false positive, and reverse hits.
From this protein groups table obtained from our cooperation partners, we considered only peptides with heavy/light ratio count of two or higher. The ratios were normalized to the most frequent value of the ratios according to Geiger *et al.*, 2011 [267]. Only the proteins reliably identified in all three replicates were selected for analysis and further filtered as described in the results. All calculations and plots were generated using the R software package [268].
The gene ontology analysis was performed using the Gene Ontology enrichment analysis and visualization tool (GOrilla, last database update in March 2021) [269]. The genes of interest were added as the unranked target list, the list of all identified genes was added as the background list and *Mus musculus* was selected as the organism of the input gene list. Gene ontology terms with a p-value lower than $5 \times 10^{-5}$ were considered and the terms were grouped according to the common gene ontology term within the diagram of the GOrilla output. Redundant gene ontology terms were removed manually.

### 3.2.5 (Immuno)fluorescence staining

Samples taken in each step of the nuclei and chromocenter isolation procedure were fixed in solution using 3.7% formaldehyde. Subsequently, a few drops of each sample were transferred to microscopy slides and dried at 80 °C. For demasking they were incubated for 3 – 5 min in 100 °C sodium citrate buffer (10 mM sodium citrate, pH 6), washed with PBS, stained with 1 µg/ml 4',6-diamidino-2-phenylindole (DAPI) for 10 min in the dark, washed and mounted with Mowiol 4-88 (#81381, Sigma-Aldrich; 4.3 M Mowiol 4-88 in 0.2 M Tris-HCl pH 8.5 with 30% glycerol) supplemented with 2.5% DABCO antifade (1,4-diazabicyclo[2.2.2]octan, #D27802, Sigma-Aldrich).
For immunofluorescence staining, the tissue slices were incubated at 60 °C for 2 h to melt the paraffin. Subsequently, they were incubated three times for 5 min in xylol for paraffin removal. Then, the tissue slices were rehydrated by sequential incubation for 5 min each in 96% ethanol, 90% ethanol, 80% ethanol, 70% ethanol, and three times in water. Antigen demasking was performed by treating the tissue slices with sodium citrate buffer at 100 °C or 30 min in an autoclave. The tissue slices were equilibrated for 15 min in PBS, permeabilized with 0.7% Triton X-100 in PBS two times for 15 min, and washed three times with PBS. The tissue slices were circled with a liquid-blocking pen and blocked with 4% BSA for 30 min. Incubation with primary antibodies (table 3.1) was performed overnight at 4 °C, followed by three times washing with PBST (0.1% Tween 20 in PBS) for 10 min and secondary antibody incubation for 1 h at room

temperature. The tissue slices were washed three times with PBST for 10 min, counterstained with 1 µg/ml DAPI, washed with PBS, and water, and mounted using Mowiol as described above.

### 3.2.6 Fluorescence microscopy & image analysis

All characteristics of the microscopy systems used including lasers/lamps, filters, objectives, and detection systems are listed in table 3.2.

#### 3.2.6.1 Microscopic analysis of chromocenter organization, subcellular localization, and heterochromatin accumulation

The confocal z-stack imaging of tissue slices was carried out on an UltraView VoX spinning disk microscopy system. Z-stacks of 0.3 µm were taken using a 60x objective.
For the analysis of heterochromatin organization and heterochromatin accumulation the nuclei and chromocenters were segmented based on the DAPI channel using Volocity software (Perkin Elmer). The detailed segmentation workflow is described in figure 3.16. The chromocenter number per nucleus and the individual chromocenter volumes were determined by the Volocity software. The heterochromatin accumulation was calculated as the ratio of the mean chromocenter intensity per nucleus versus the nucleoplasm intensity. The violin plots and the significance tests were created using the R software package [268]. The data visualization in a violin plot compared to a box plot is depicted in figure 3.17.
For the analysis of the subcellular localization of the proteins of interest, fluorescence intensities were measured and the images were processed using ImageJ software [270, 271].

#### 3.2.6.2 Microscopic examination of the chromocenter isolation procedure

Fluorescence images of nuclear fractions collected during the chromocenter isolation procedure were imaged on a Zeiss Axiovert 200 or a Zeiss Axioplan microscope with a 40x objective. The images were processed using ImageJ software [270, 271].

### 3.2.7 Cell culture

C2C12 mouse myoblast cells (see table 3.3) were grown in Dulbecco's modified Eagle Medium (DMEM) with high glucose (#D6429, Sigma-Aldrich) supplemented with 20% fetal bovine serum (#F7524, Sigma-Aldrich), 1x glutamine (#G7513, Sigma-Aldrich) and 1 µM gentamicin (#G1397, Sigma-Aldrich) at 37 °C and 5% $CO_2$ in a humidified incubator. Tests to check for potential mycoplasma contamination were performed regularly.

### 3.2.8 Western blot analysis

Mouse brain and liver nuclei as well as C2C12 mouse myoblast cells were lysed in lysis buffer (0.025 M Tris HCl (pH 8), 1 M NaCl, 0.05 M glucose, 0.01 M EDTA, 0.2% Tween 20, 0.2% Nonidet P40 Substitute (#74385, Sigma-Aldrich) supplemented with protease inhibitors (1 mM phenylmethylsulfonyl fluoride (PMSF) (#6367.1, Carl Roth), 10 µM E64 (#E3132, Sigma-Aldrich),

1 µM pepstatin A (#P5318, Sigma-Aldrich)) and mechanically disrupted. The protein concentration was determined by Pierce assay (PierceTM 660 nm Protein-Assay-Kit, #22662, Thermo Fisher Scientific) using a plate reader (Infinite M200 PRO, Tecan). Subsequently, all samples were diluted in Laemmli buffer (2% SDS, 50 mM Tris (pH 6.8), 10% glycerol, 0.01% bromophenol blue, 100 mM DTT) and incubated at 95 °C for 10 min. 30 µg of each sample were analyzed by SDS-PAGE and transferred to a nitrocellulose membrane using a semi-dry blotting system at 25 V for 35 - 90 min. The membranes were stained with Ponceau S solution (#P7170, Sigma-Aldrich) to check for successful transfer. Then, they were blocked with 5% low-fat milk in PBS for 30 min and incubated with primary antibodies (antibodies and dilutions are listed in table 3.1) in 3% low-fat milk in PBS overnight. The membranes were washed three times with 0.1% PBST (0.1% Tween 20 in PBS), incubated with secondary antibodies in 3% low-fat milk for 1 h, and washed again three times with PBST. The membranes were covered with ECL solution (Clarity Western ECL substrate, #1705061, Bio-Rad) and the chemiluminescence signal was detected using an Amersham AI600 imager (table 3.2).

## 3.3 Results

### 3.3.1 Pericentromeric heterochromatin characteristics differ between mouse brain and liver

First, we analyzed the heterochromatin organization in mouse brain and liver tissue. As described before, pericentromeric heterochromatin from several chromosomes forms heterochromatin clusters in the interphase in mouse cells, also called chromocenters [25]. Thus, we stained the DNA using the DNA dye DAPI on tissue slices from the mouse brain and liver. After confocal microscopy z-stack imaging, we segmented nuclei and chromocenters and compared the chromocenter numbers and volumes between tissues (figure 3.1).

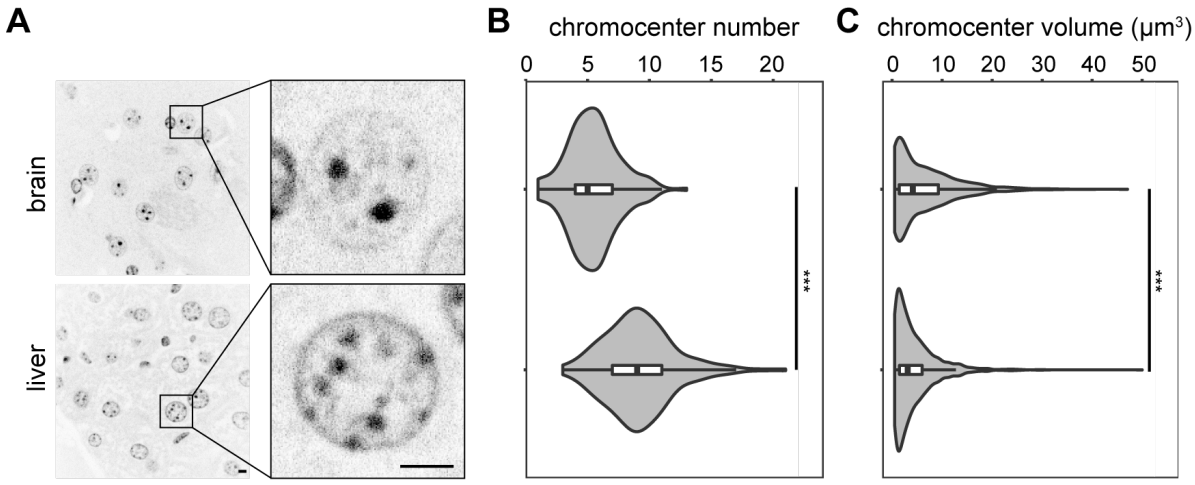



**Figure 3.1: Pericentromeric heterochromatin organization in mouse brain and liver tissue.** (A) Mouse brain and liver tissue stained with the DNA dye DAPI. The zoomed images on the right show an exemplary single nucleus for each tissue. Scale bars 5 µm. The violin plots depict the chromocenter number per nucleus (B) and the individual chromocenter volumes in $\mu m^3$ (C). The p-values were calculated using the Wilcoxon-Rank test. ***$p<0.001$. The statistics are summarized in table 3.4.

Mouse liver nuclei showed about 2-fold higher chromocenter numbers than mouse brain nuclei and respectively, smaller chromocenter volumes. The results indicate significant differences in heterochromatin organization between mouse brain and liver tissue and, thus, represent a good system to analyze tissue-specific differences in chromocenter composition.

### 3.3.2 Isolation of the heterochromatin fraction from mouse tissues and mass spectrometry analysis workflow

As the DNA staining of the mouse brain and liver tissue showed highly significant differences in heterochromatin organization, we hypothesized that these changes can be due to quantitative differences in the heterochromatin proteome. To analyze the pericentromeric heterochromatin fractions by mass spectrometry, we first needed to isolate these fractions from the tissues. Therefore, we adapted and modified a protocol from Prusov and Zatsepina [94, 95] to be used not only for the liver but also for brain tissue. Therefore, we added an RNaseA treatment and adjusted the number of nuclei used in the different steps. In comparison to previously used methods for

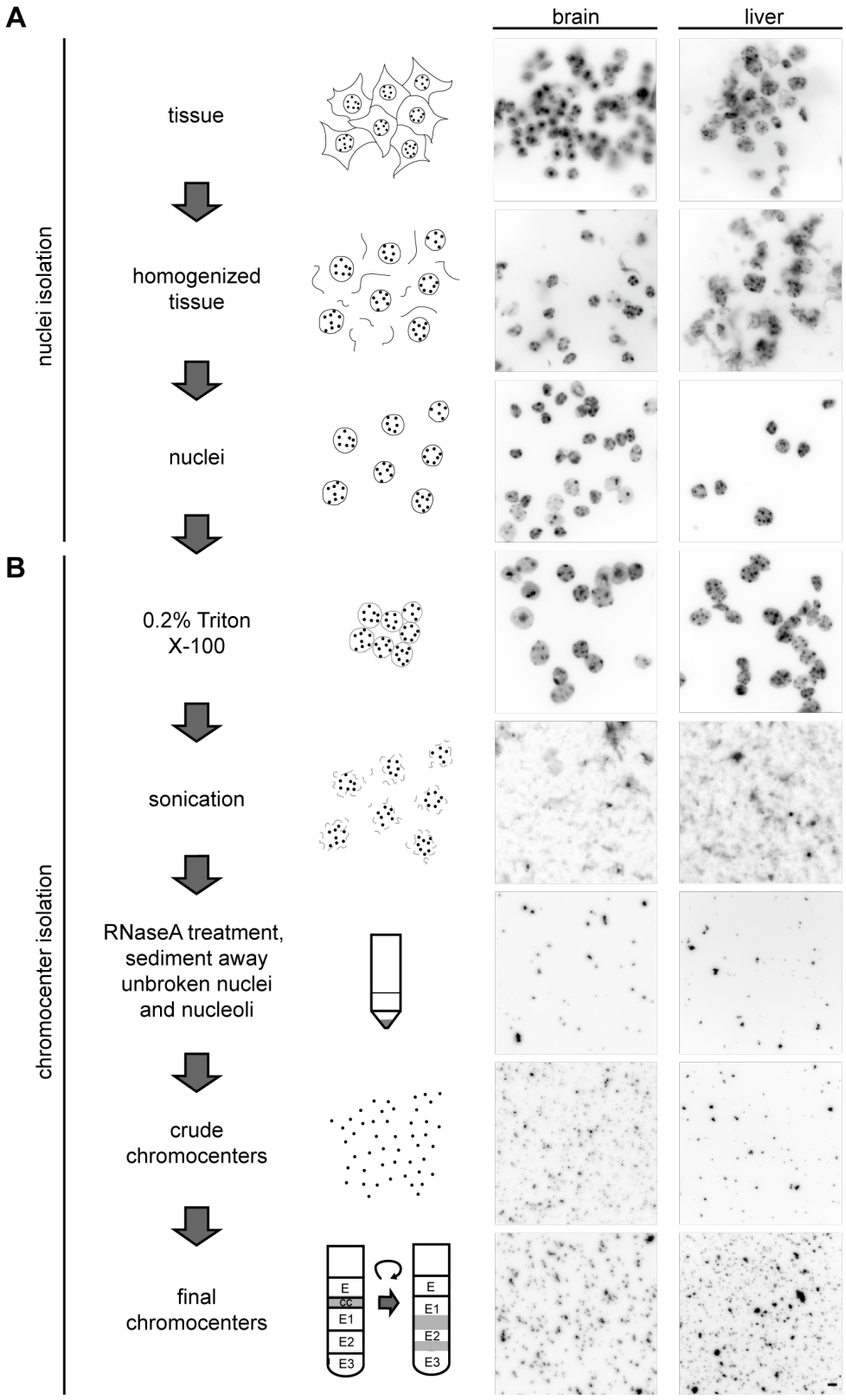



**Figure 3.2: Nuclei and chromocenter isolation from mouse brain and liver tissue.** Scheme of the workflow for nuclei (A) and chromocenter isolation (B) from mouse tissue with the corresponding images showing DNA (DAPI) staining of all steps for brain and liver. Scale bar 5 µm.

heterochromatin isolation (see table 1.1), this protocol is based on a series of sucrose gradient centrifugation steps and, thus, unbiased albeit at the cost of some possible contaminants from other subnuclear structures. The individual steps of the protocol and DNA-stained fluorescence microscopy images of all isolation steps are depicted in figure 3.2.
Briefly, nuclei isolation was performed by homogenizing the tissue and centrifugation in a high magnesium sucrose buffer. For chromocenter isolation, the nuclei were treated with Triton X-100 to permeabilize the nuclear membrane. Then, the magnesium concentration was reduced, keeping the chromocenters in a condensed state while other chromatin fractions dispersed. A sonication step and RNaseA treatment led to the disruption of other subnuclear structures. Unbroken nuclei and nucleoli were removed by centrifugation to obtain the crude chromocenter fraction. Finally, the chromocenters were loaded on a sucrose gradient for further purification. The nuclei and chromocenter fractions were processed for quantitative mass spectrometry by our cooperation partners as described in figure 3.15. After methanol chloroform precipitation, the proteins were denatured, reduced, alkylated, and digested with Lys-C and trypsin before they were subjected to dimethyl labeling and HPLC-MS/MS measurement. The following figures displaying mass spectrometry data are based on the nuclei and chromocenter isolation performed by Stephanie Meyer and subsequent mass spectrometry measurements performed by Oliver Popp and Gunnar Dittmar.
A detailed workflow of the data analysis including data filtering and selection of hits is described in figure 3.3. The data was filtered by removing contaminants, false positives, reverse hits, and proteins with low H/L ratio count (figure 3.3A). Then the H/L ratios were normalized to the most frequent value according to Geiger *et al.* [267] (figure 3.3B).
To identify proteins enriched in heterochromatin, we followed different selection strategies (figure 3.3C). Initially, we selected proteins that were identified reproducibly in all three biological replicates and enriched in chromocenters in comparison to the nuclei ($\log_2$(ratio chromocenter/ nuclei) $>0$). To obtain the quantitative differences between brain and liver chromocenters, the proteins enriched in chromocenters of both tissues were plotted against each other. In a subsequent step, we applied a validation-based cut-off which will be detailed below. Lastly, we were interested in the low abundant proteins identified in the chromocenters which were not reproducibly identified in the nuclei (see figure 3.3C). The results of the different hit selection steps will be described in detail below.

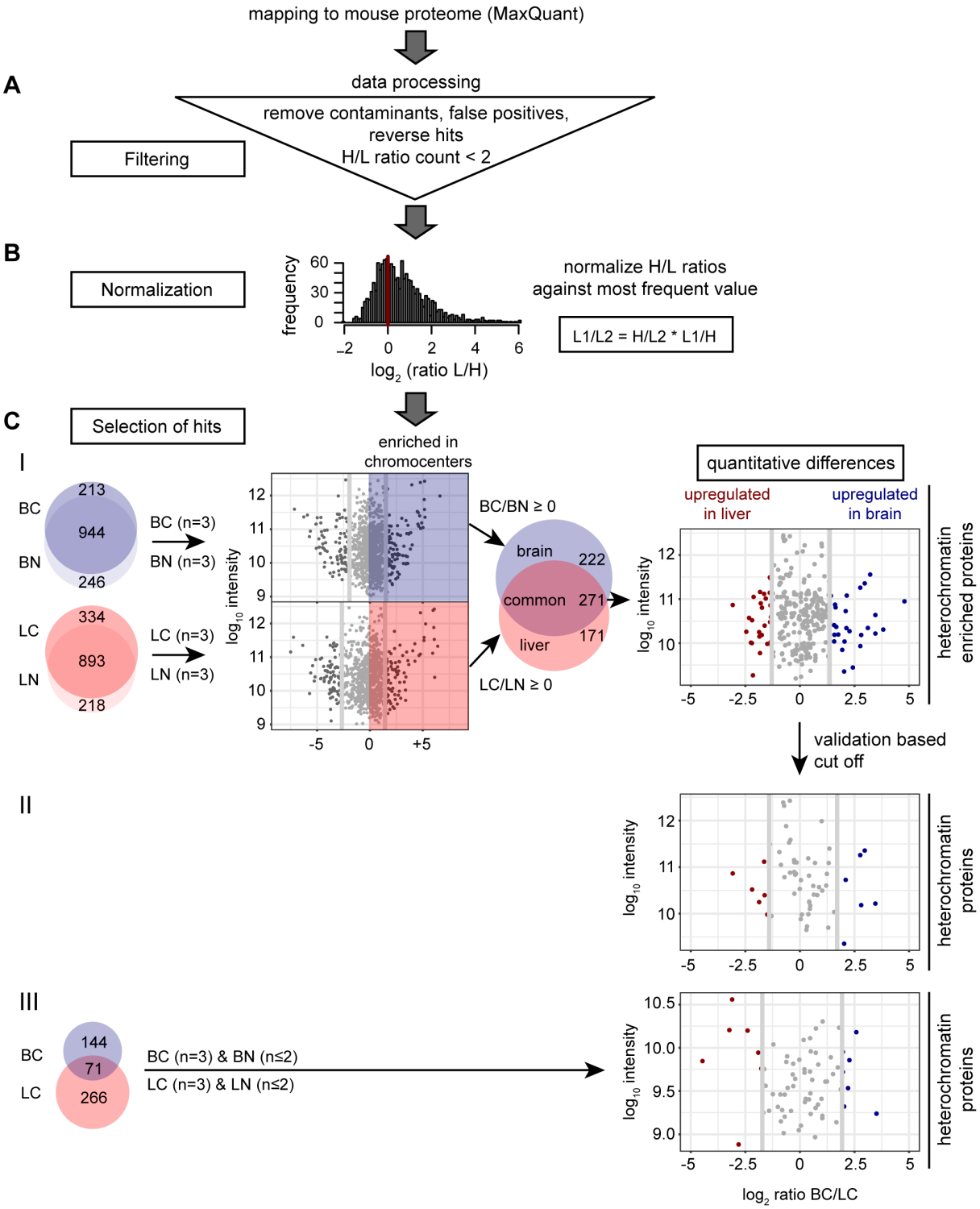



**Figure 3.3: Workflow for the data analysis of the quantitative mass spectrometry experiment.** The data was analyzed using MaxQuant software by mapping the identified peptides against the UniProt mouse proteome. (A) The data was filtered by removing contaminants, false positives, and reverse hits identified by MaxQuant. Proteins with a heavy/light (H/L) ratio count $\geq 2$ were considered for further analysis. (B) The H/L ratios were normalized against the most frequent value and the different fractions were compared to each other by dividing one heavy/light ratio by the other (H/L1 / H/L2 = H/L1 x L2/H = L2/L1). (C) The proteins identified in all three biological replicates (n = 3) were considered for analysis and several filtering steps were applied. (I) The proteins identified in chromocenters

**Figure 3.3 (*previous page*):** were plotted against the nucleus proteins for the brain and liver, then the common proteins enriched in chromocenters were compared between the tissues. (II) A validation-based cut-off was applied to the proteins from (I). (III) Proteins identified in three replicates in the chromocenters ($n = 3$) and two or less ($n \leq 2$) replicates in nuclei were plotted for brain versus liver. BC: brain chromocenter, BN: brain nuclei, LC: liver chromocenter, LN: liver nuclei.

### 3.3.3 Proteomic analysis of heterochromatin enriched proteins

First, we plotted the chromocenter proteins versus the nucleus proteins identified reproducibly in three biological replicates for each tissue to determine those proteins enriched in the chromocenters (figure 3.3C).

To identify quantitative differences of proteins enriched in heterochromatin between brain and liver, we plotted the $\log_2$ ratios of brain chromocenter proteins over liver chromocenter proteins against the $\log_{10}$ intensities (figure 3.4). The upper and lower 10% of the proteins were considered as up- or downregulated in brain or liver chromocenters (figure 3.4 and 3.18).

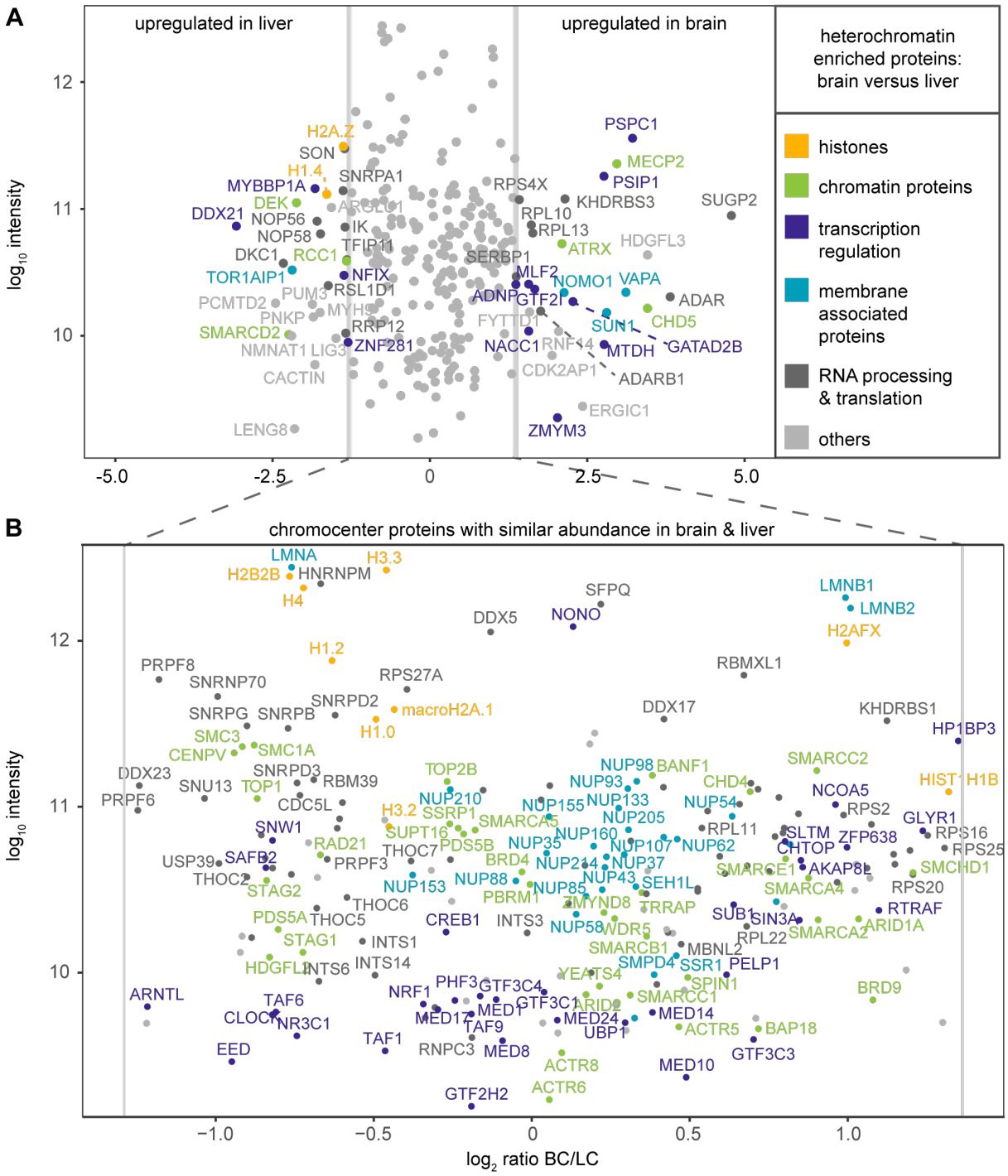



**Figure 3.4: Heterochromatin enriched proteins in mouse tissues identified by quantitative mass spectrometry.** All proteins identified in three biological replicates and enriched in the chromocenters compared to the nuclei (see table 3.5) were used as input (figure 3.3 C.I). (A) The dot plot shows the $\log_{10}$ intensity plotted against the $\log_2$ ratio of brain chromocenter (BC) versus liver chromocenter (LC) proteins. Proteins with equal abundance are depicted in gray, and proteins enriched in the brain or liver (upper or lower 10%) are labeled and color-coded. The color code indicates the protein function manually assigned based on the UniProt webpage functional information. Histones are labeled in orange, chromatin proteins in green, proteins involved in transcriptional regulation in purple, membrane associated proteins in cyan, proteins involved in RNA processing or RNA-binding proteins in dark gray, and proteins not fitting into the categories in light gray. (B) Rescaled version of the dotplot shown in (A) focusing on the heterochromatin enriched proteins with similar abundance between brain and liver. Color coding is described in (A).

The proteins enriched in heterochromatin and upregulated in the brain comprised the chromatin proteins MeCP2, ATRX, and CHD5, as well as several transcriptional regulators. The histones H2A.Z and H1.4 were enriched in the liver. In addition, the chromatin proteins DEK, RCC1, and SMARCD2, and ZNF281 which plays a role in transcriptional regulation were upregulated in the liver. Proteins involved in RNA processing and translation or those which did not match any of the functional categories were found in both tissues.
The proteins enriched in chromocenters with equal abundance in the brain and liver comprised histones, chromatin proteins, and proteins involved in the regulation of transcription, but also various membrane and RNA associated ones. Of note, many components of the nuclear pore complex were identified. In addition, RNA associated proteins with functions in RNA processing, spliceosome and ribosome formation were found in the fraction of conserved chromocenter proteins.
Next, we performed a gene ontology analysis of all the proteins that we found enriched in chromocenters using the GOrilla tool [269]. The results for biological process, molecular function, and cellular component are shown in figure 3.5 and resemble those of the manually categorized proteins from figure 3.4.
The GO analysis of the biological processes revealed that many genes on our input list were associated with metabolic processes, including the regulation of gene expression and DNA-templated transcription. The biological processes with the highest enrichment values were nucleosome and nucleus organization, chromosome segregation, but also nuclear pore organization, protein-DNA complex disassembly, and nuclear export. For the category of molecular function, the GO terms chromatin and nucleosome binding as well as structural molecule activity, and constituent of the ribosome and of the nuclear pore showed the highest enrichments. Cellular components with the highest enrichment values were the cohesin complex, THO complex, nucleosome, nuclear pore, and ribosomal subunit.
Summarizing the results of the GO analysis, we found many proteins associated with typical GO terms for heterochromatin proteins like nucleosome organization, gene expression, and nucleic acid binding. In addition, many proteins associated with nuclear pores, the ribosome, and RNA were identified, indicating that we copurified several membrane associated and nucleolar proteins.

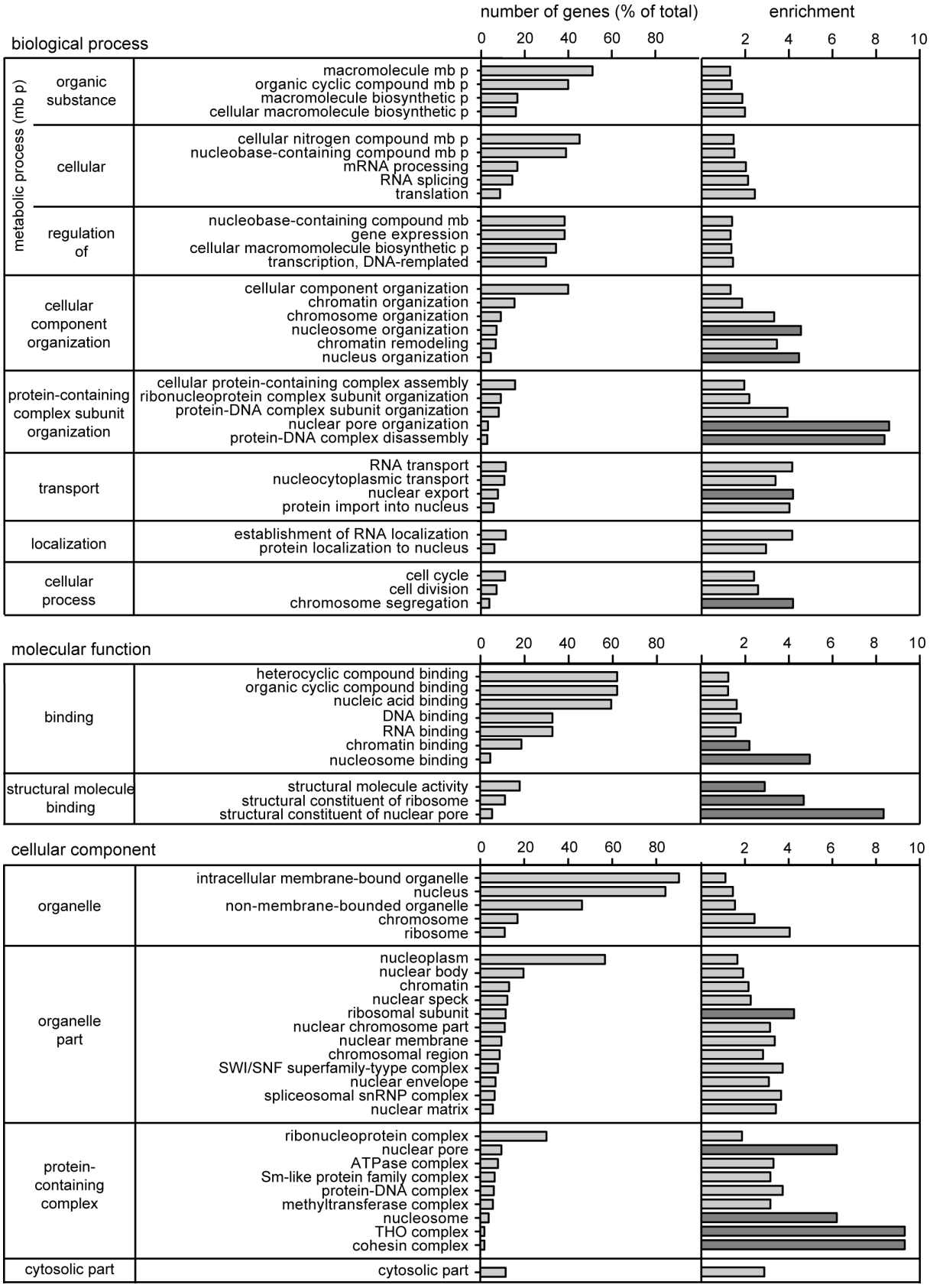



**Figure 3.5: Gene ontology (GO) analysis of the identified heterochromatin enriched proteins.** All proteins identified in three biological replicates and enriched in the chromocenters compared to the nuclei were used as input (figure 3.3C.I). The protein list was subjected to the GOrilla tool [269] for gene

**Figure 3.5 (*previous page*):** ontology analysis in the categories biological process, molecular function, and cellular component. The proteins enriched in heterochromatin were added as the target list, all identified proteins (without cut-off or filtering) as the background list. GO terms with a p-value lower than $5 \times 10^{-5}$ were considered. The GO terms were grouped according to a common gene ontology term within the diagram of the GOrilla output and redundant terms were removed manually. Plotted is the number of genes in the GO term (b) as % of total genes in the target list (n) and the enrichment calculated by the GOrilla tool as (b/n)/(B/N) with b: number of genes in the target list associated with specific GO term; n: total number of genes in the target list; B: number of genes in the background list associated with specific GO term; N: total number of genes in the background list. The five proteins with the highest enrichment in each category are highlighted in dark gray.

### 3.3.4 Validation of heterochromatin enriched protein hits

The quantitative mass spectrometry analysis of heterochromatin fractions revealed several proteins associated with the nuclear membrane, especially nuclear pore proteins, but also many ribosomal proteins and those involved in RNA processing. Thus, we validated the mass spectrometry results by immunofluorescence staining on tissue slices to distinguish between heterochromatin proteins and potential contaminants. We made use of specific antibodies directed against selected protein hits representing all functional categories defined in previous sections and tested for colocalization of the chromocenters stained with DAPI and the antibody signal. In parallel, we performed Western blot analysis to confirm the specificity of the antibodies and the presence of the proteins in the nuclei isolated from the tissues. The validations of proteins found upregulated in the brain or liver are shown in figure 3.6. The results of the protein hits with similar abundance in the brain and liver with localization at the chromocenters are shown in figure 3.7, whereas the ones not localizing at chromocenters are shown in figure 3.8 and 3.9.
Among the proteins upregulated in the brain, we validated the chromatin proteins ATRX and MeCP2 as well as GATAD2B which is involved in transcriptional regulation. While ATRX and MeCP2 colocalized with the DAPI signal at the chromocenters, GATAD2B showed nuclear localization in the brain, but chromocenter localization in liver tissue (figure 3.6A). The Western blots showed much stronger signals for brain nuclei than for liver nuclei, thus confirming the quantitative differences determined by mass spectrometry (figure 3.6B). The antibody testing on the Western blots was performed using C2C12 cell lysate in addition to the mouse brain and liver nuclei, as the lysate contains the cytoplasmic proteins in addition to the nuclear ones.
From the proteins upregulated in the liver, we selected histone H1, SMARCD2, and ZNF281 for validation. Histone H1 localized at the chromocenters. The chromatin protein SMARCD2 did not show any signal on brain tissue and gave a diffuse nucleus staining on the liver tissue. The zinc finger protein ZNF281 localized in the nucleus in the brain and liver (figure 3.6A). The Western blot of histone H1 showed a larger smear for liver nuclei, possibly caused by post-translational modifications of histone H1 occurring specifically in the liver. The blots for SMARCD2 and ZNF281 showed weak signals for the liver and hardly any signal for the brain (figure 3.6B), mimicking the quantitative differences observed in the immunofluorescence stainings and the quantitative mass spectrometry analysis.

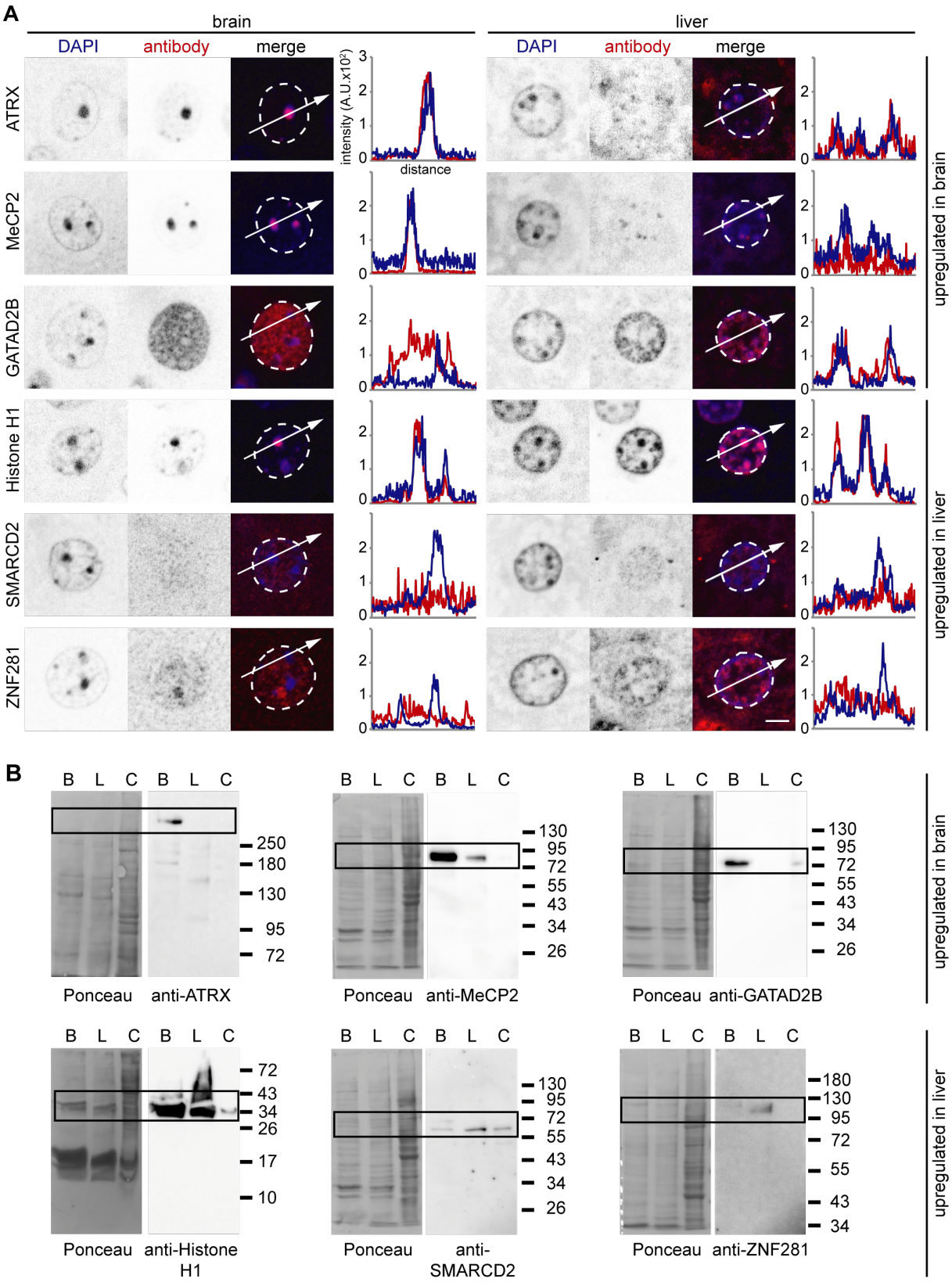



**Figure 3.6: Validation of candidate heterochromatin enriched proteins with different abundance in brain and liver tissue.** (A) Immunofluorescence staining on mouse brain and liver tissue slices using antibodies specific to the identified proteins. The nuclear outlines are marked in white on the

**Figure 3.6 (*previous page*):** merged channel image. Line plots of the fluorescence intensity in arbitrary units (A.U.) plotted against the distance depict the colocalization of the antibody staining (red) with the DNA counterstain (DAPI, blue). Scale bar 5 µm. (B) Western blots of brain nuclei lysate (B), liver nuclei lysate (L), and whole mouse myoblast cell lysate (C) using the same antibodies as in A. The protein transfer is shown by Ponceau S staining on the left, and the antibody signals with chemiluminescence detection are shown on the right. The molecular weight marker indicates the protein weight in kDa and the black boxes mark the bands of interest. Overview images are shown in figure 3.19.

To validate the protein hits with similar abundance in chromocenters between tissues, we selected proteins of all the functional groups defined previously. The histones H4, H3.3, macroH2A.1, and H2AX as well as the topoisomerase I and II immunofluorescence stainings colocalized with the DAPI staining at the chromocenters in the brain and liver tissue (figure 3.7A). While the Western blot signal for histone H4 was weak for the nuclei, the Western blots tested for histone H3.3, macroH2A.1, histone H2AX and topoisomerase I and II showed intense bands for the nuclei isolated from tissues (figure 3.7B).
SMARCB1 and SMARCA2 are components of the mating-type switching (SWI)/sucrose non-fermenting (SNF) ATP-dependent chromatin remodeling complex. This complex was reported to mediate epigenetic regulation, as it is actively involved in chromatin remodeling, and roles in transcriptional regulation were described (reviewed in [272]). Interestingly, we identified several SMARC and ARID1 and 2 proteins, all components of these complexes [272], indicating their association with chromocenters. SIN3A was shown to associate with known heterochromatin proteins [125, 126], but the association might not be stable [273]. Nevertheless, only Smarca2 showed a weak localization at the chromocenters in liver tissue, while SMARCA2 in the brain, SMARCB1, and SIN3A did not localize at chromocenters (figure 3.8A).
As we found many nuclear pore proteins in our analysis, we stained for NUP153 in addition to Lamin B and Lamin A/C as membrane associated proteins. All of them are localized at the nuclear membrane and partly to the nucleoplasm. For the group of RNA associated proteins, we analyzed the spliceosome component SNRNP70 and the splicing factor MBNL2. SNRNP70 localized in the nucleus and MBNL2 in the nucleus and in brain tissue additionally in the cytoplasm (figure 3.8B, figure 3.21). Both proteins show accumulations outside of the chromocenters, possibly at nuclear speckles.
All Western blots of the proteins with similar abundance in the brain and liver that are not localizing at chromocenters showed clear signals for the tissue nuclei. Anti-NUP153 was an exception, as it was not suited for the use on Western blots according to the manufacturer's information (figure 3.9).

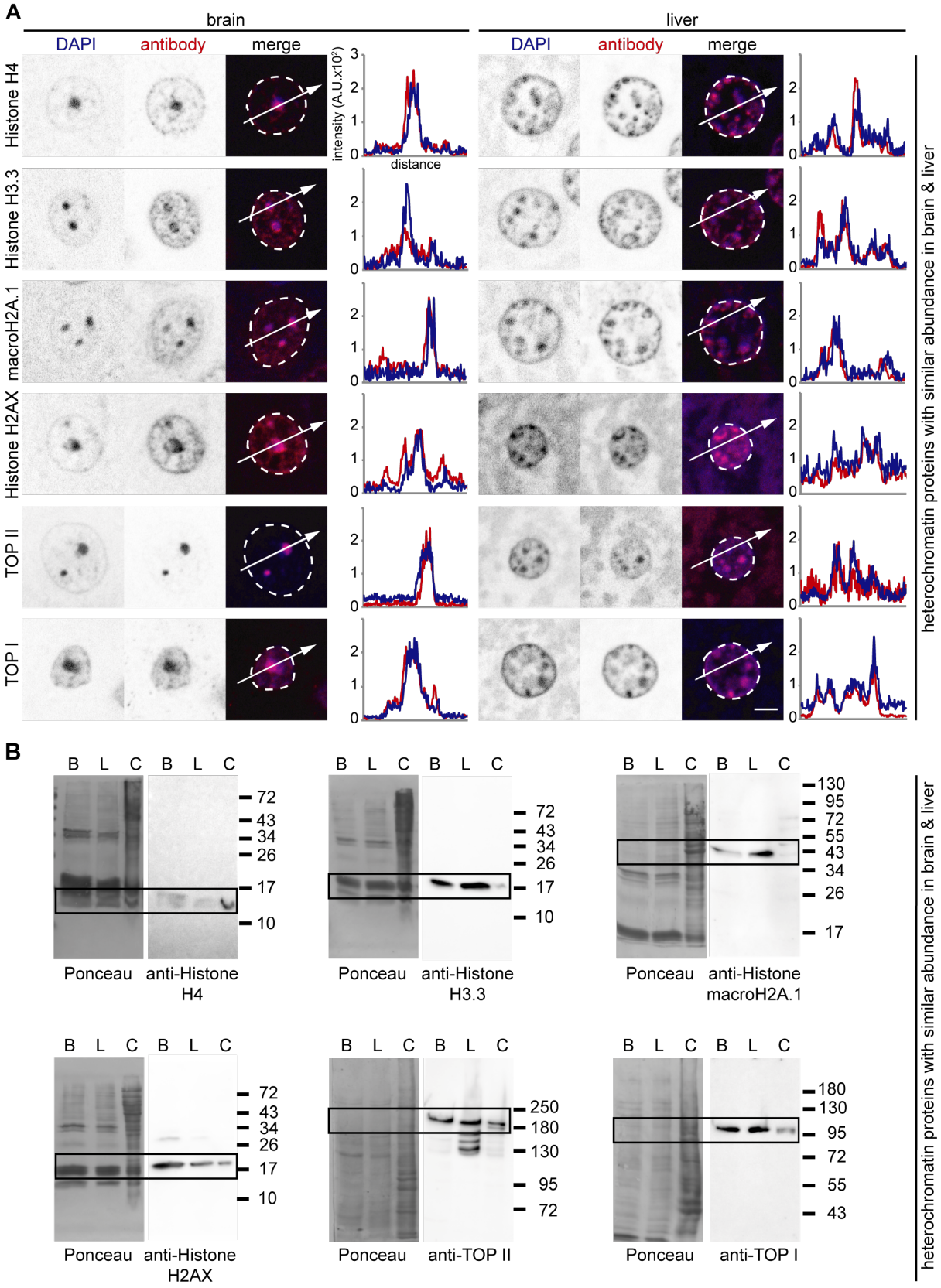



**Figure 3.7: Validation of heterochromatin enriched proteins with similar abundance in brain and liver tissue showing chromocenter localization.** (A) Immunofluorescence staining on mouse brain and liver tissue slices using antibodies specific to the identified proteins. The nuclear outlines are

**Figure 3.7 (*previous page*):** shown in white on the merged channel image. Line plots of the fluorescence intensity in arbitrary units (A.U.) plotted against the distance depict the colocalization of the antibody staining (red) with the DNA counterstain (DAPI, blue). Scale bar 5 µm. (B) Western blots of brain nuclei lysate (B), liver nuclei lysate (L), and whole mouse myoblast cell lysate (C) using the same antibodies as in A. The protein transfer is shown by Ponceau S staining on the left, and the antibody signals with chemiluminescence detection are shown on the right. The molecular weight marker indicates the protein weight in kDa and the black boxes mark the bands of interest. Overview images are shown in figure 3.20. The TOPII staining and the histone H2AX staining were performed by Stephanie Meyer, and the TOPI staining and imaging were performed by Katalina Gagova.

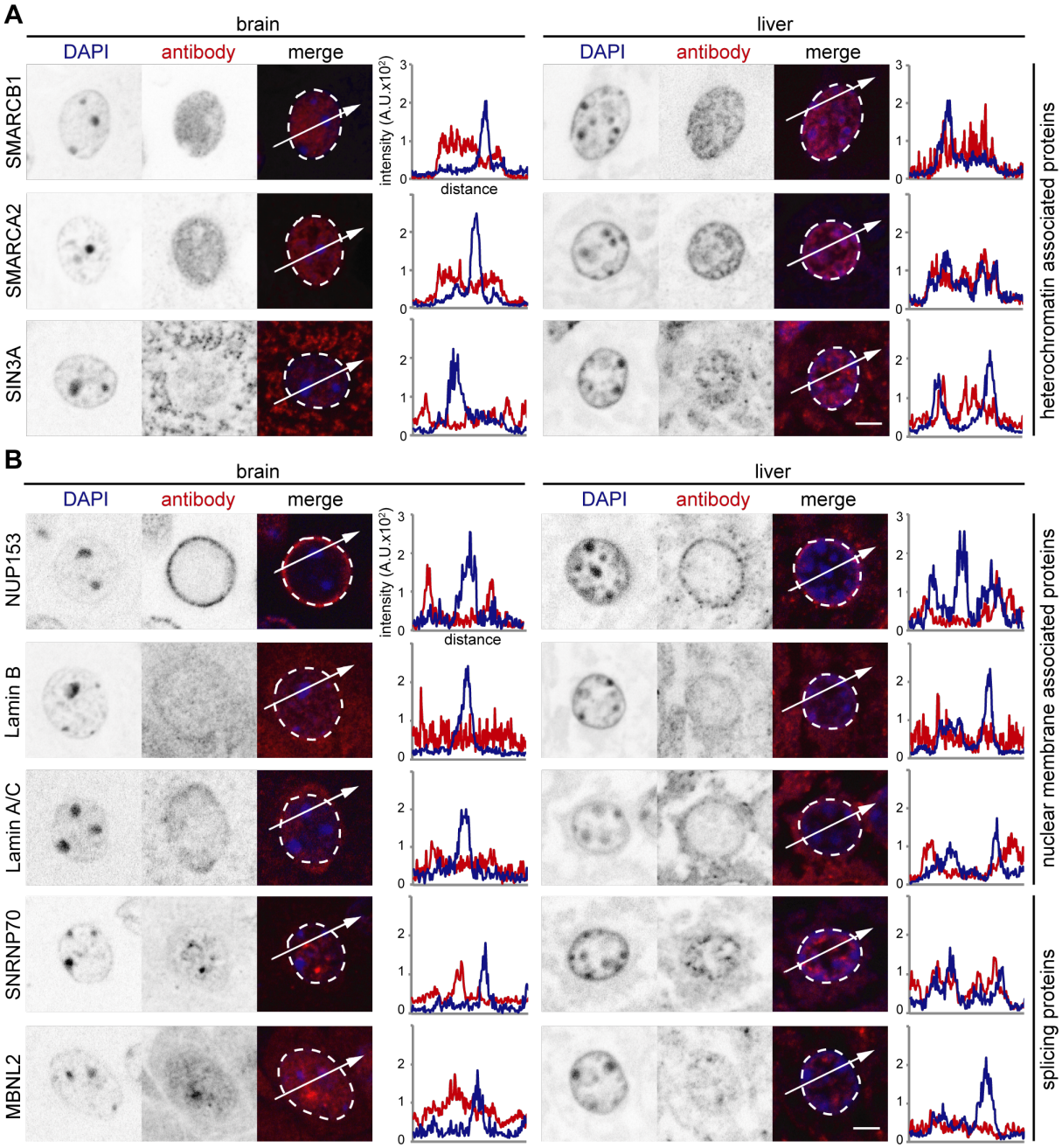



**Figure 3.8: Validation of candidate heterochromatin interacting proteins with similar abundance in brain and liver tissue not localizing at the chromocenters.** (A) Immunofluorescence staining on mouse brain and liver tissue slices using specific antibodies to the identified proteins. The nuclear outlines are marked in white on the merged channel image. Line plots of the fluorescence intensity in arbitrary units (A.U.) plotted against the distance depict the colocalization of the antibody staining (red) with the DNA counterstain (DAPI, blue). Scale bar 5 µm. Overview images are shown in figure 3.21.

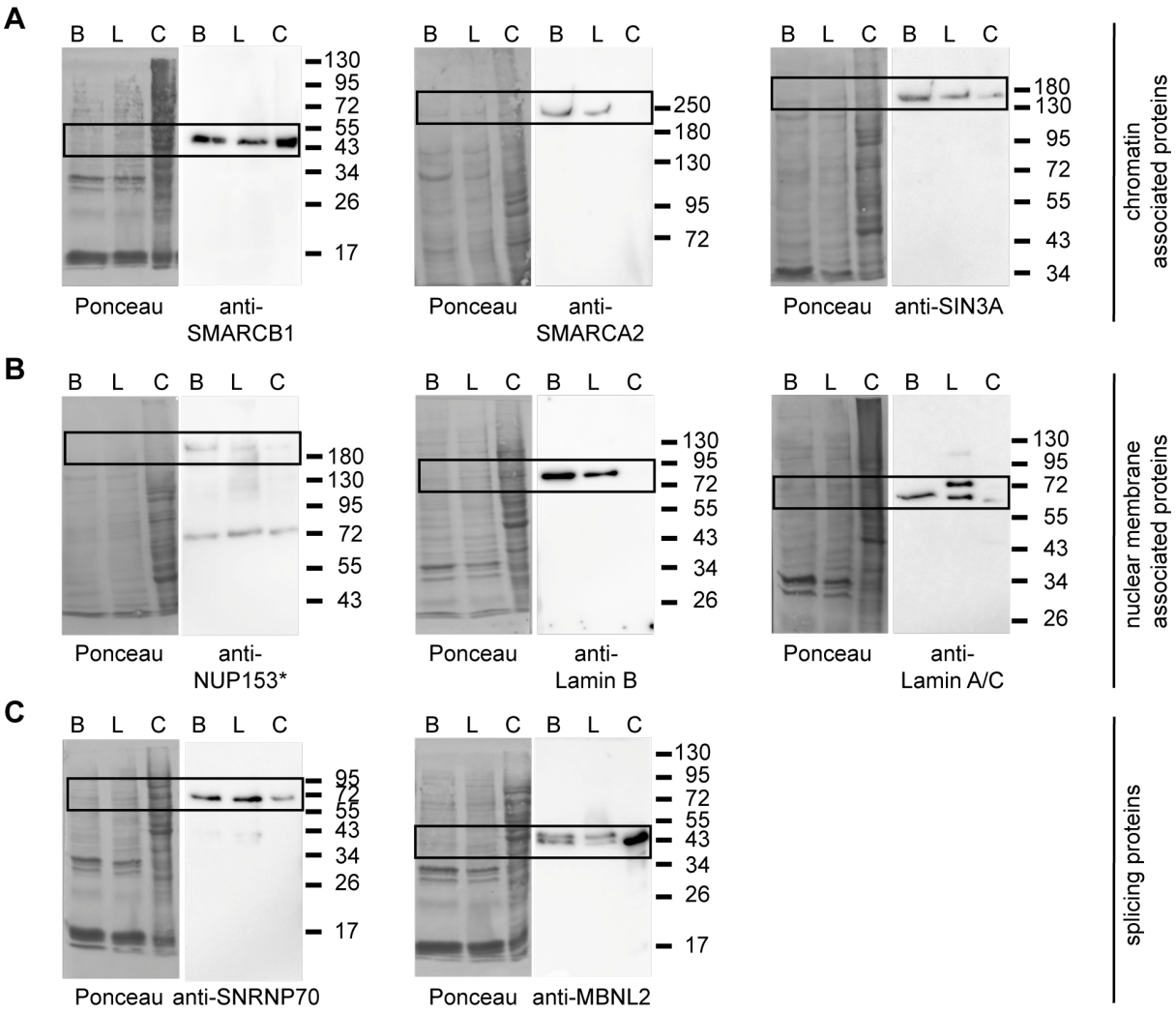



**Figure 3.9: Validation of candidate heterochromatin interacting proteins with similar abundance in brain and liver by Western blot analysis.** Shown are Western blots of chromatin associated proteins (A), nuclear membrane associated proteins (B) and splicing proteins (C). Western blots of brain nuclei lysate (B), liver nuclei lysate (L), and whole mouse myoblast cell lysate (C) using the same antibodies as in figure 3.8. The protein transfer is shown by Ponceau S staining on the left, and the antibody signals with chemiluminescence detection are shown on the right. The molecular weight marker indicates the protein weight in kDa and the black boxes mark the bands of interest. *The antibody for NUP153 is not suitable for Western blot detection according to the manufacturer's instructions.

Next, we measured the heterochromatin accumulation of the validated proteins which localized at the chromocenters. Therefore, we segmented nuclei and chromocenters (see figure 3.16) and calculated the accumulation as ratio of the average chromocenter intensity per nucleus and the nucleus intensity (figure 3.10).

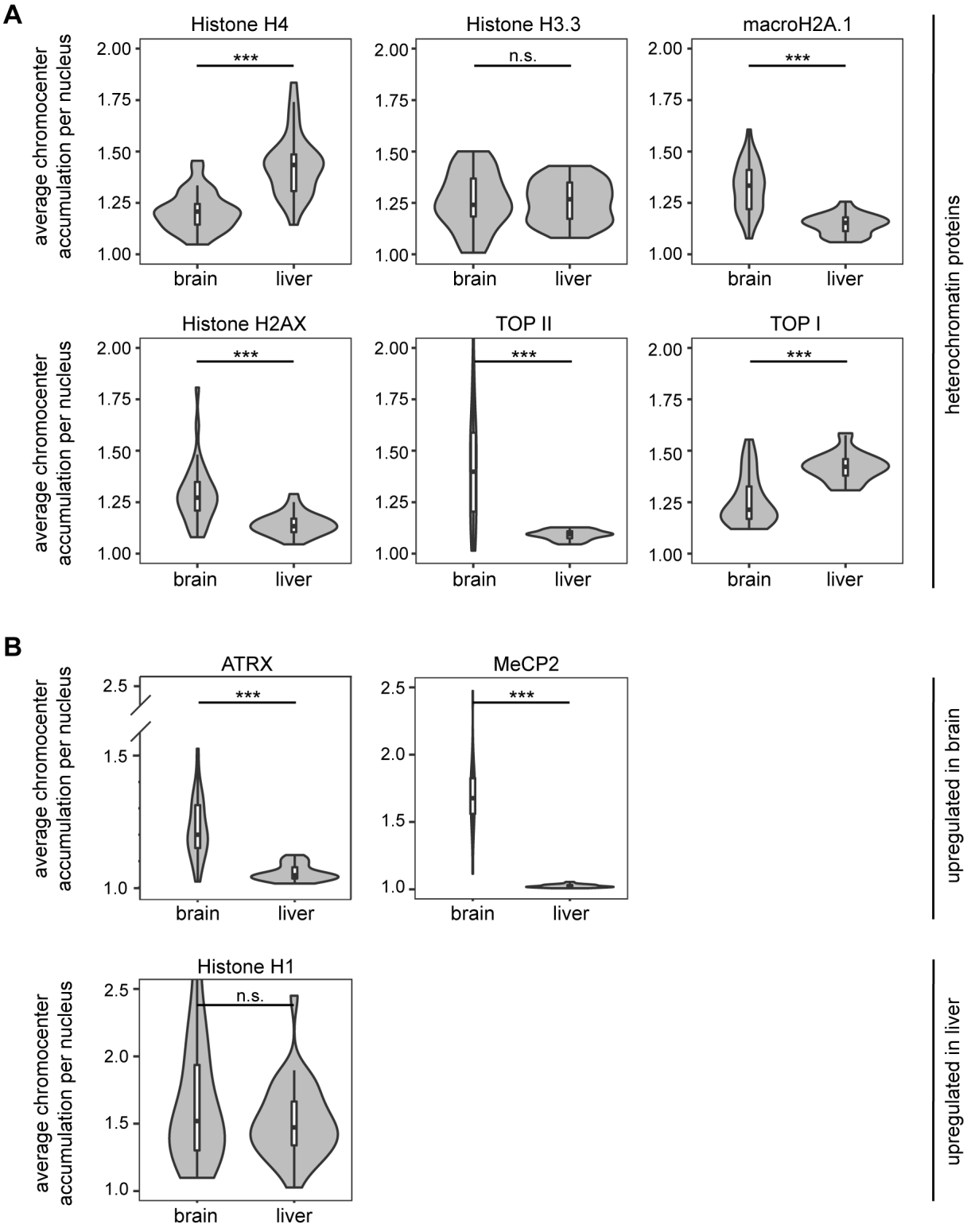



**Figure 3.10: Chromocenter accumulation analysis of heterochromatin proteins identified in brain and liver.** Violin plots depict the average chromocenter accumulation per nucleus calculated as the ratio of chromocenter intensity versus nucleoplasm intensity after nucleus and chromocenter segmentation using Volocity software (see figure 3.16). The chromocenter accumulation was determined for all validated protein hits localizing at the chromocenters, the heterochromatin proteins with equal abundance in the brain and liver (A), and the proteins upregulated in the brain or liver (B). The p-values were calculated using the Wilcoxon-Rank test. ***$p<0.001$, ** $p<0.05$, n.s.: not significant. The statistics are summarized in table 3.4.

The heterochromatin accumulation analysis showed similar values for histone H3.3 in both tissues but revealed quantitative differences in the accumulation in the brain and liver for the histones H4, macroH2A.1, and H2AX and the topoisomerases I and II (figure 3.10A). Histone H4 and topoisomerase I accumulation at chromocenters was higher in the liver, whereas macroH2A.1, histone H2AX, and topoisomerase II accumulation was higher in the brain. Of note, the ratios obtained from the quantitative mass spectrometry were not the same for all proteins within the group of similar abundance in both tissues (figure 3.4). Topoisomerase I and histone H4 showed a tendency for a higher abundance in the liver, whereas the histone variant H2AX showed a tendency for a higher abundance in the brain, all in line with the results from the accumulation analysis. The histone variant macroH2A.1 showed a higher chromocenter accumulation in the brain, while the mass spectrometry results indicated a similar protein amount in brain and liver tissue. Thus, further analysis is necessary to evaluate the abundance of macroH2A.1 in mouse tissues. In addition, the chromocenter accumulation analysis showed higher values for topoisomerase II in the brain compared to the liver, although the mass spectrometry analysis did not show this tendency. In the case of topoisomerase II, the results of both methods are not directly comparable, as the mass spectrometry analysis specifically identified the isoform topoisomerase II beta whereas the antibody used for detection via immunological methods is directed against both topoisomerase IIa and IIb (see table 3.1). Thus, higher chromocenter accumulation measured for topoisomerase II might be based on signals for both isoforms, topoisomerase IIa and IIb. Of note, we measured the accumulation of the proteins by comparing chromocenter intensity versus nucleoplasm intensity so that proteins with higher abundance in the nucleoplasm show lower values. Furthermore, we cannot rule out that we lose proteins during the chromocenter isolation procedure so the results from mass spectrometry analysis and immunofluorescence staining do not show the same protein amounts. In addition, our analysis approach does not account for cell-type specific differences within the tissues, as we used whole tissues for the experiments.
The validated proteins MeCP2 and ATRX upregulated in the isolated chromocenters from brain tissue showed also a higher chromocenter accumulation in brain tissue examined by immunofluorescence staining (figure 3.10B). Histone H1, whose isoform H1.4 was upregulated in liver tissue, revealed similar chromocenter accumulation between the tissues in immunofluorescence staining. As other isoforms of histone H1 were identified with similar abundance in the brain and liver by quantitative mass spectrometry, the measured accumulation represents the average accumulation of all isoforms present in the chromocenters and detected by the histone H1 specific antibody. Isoform-specific antibodies would be necessary to quantify the heterochromatin accumulation of individual isoforms.
We conclude that histones, the topoisomerases I and II, MeCP2, and ATRX localized at the chromocenters, whereas other tested proteins localized in the nucleus. The membrane-associated proteins and splicing proteins did not specifically localize at chromocenters in mouse brain and liver tissue. They might interact with the chromocenters, especially as many chromocenters are localized close to the nuclear membrane and nucleoli, but they are not heterochromatin-specific proteins.

### 3.3.5 Proteomic analysis of heterochromatin proteins

Our approach for chromocenter isolation is advantageous in comparison to other enrichment procedures, as it does not introduce any bias towards a specific protein. Nevertheless, it does not allow for stringent washing steps to remove loosely bound proteins. We identified many known heterochromatin proteins in the isolated chromocenter fractions, but the analysis also revealed proteins of other functional groups like membrane-associated proteins and proteins involved in RNA processing and translation. The immunofluorescence stainings confirmed that these proteins did not show subcellular localization at the chromocenters. Thus, they represent potential contaminants copurified with the chromocenters or associated with them due to the proximity in the cell nucleus. For this reason, we decided to apply a validation-based cut-off to the results to filter out possible contaminants and yield a more robust selection of heterochromatin proteins. Therefore, we filtered the list of chromocenter enriched proteins from figure 3.4 keeping all validated proteins with chromocenter localization and removing all proteins with a lower $\log_2$ ratio of brain chromocenters versus liver chromocenters. The plot based on the filtered protein list is shown in figure 3.11. In total, we identified 271 proteins enriched in the chromocenters of the brain and liver (see figure 3.3C). Applying the validation-based cut-off resulted in a list of 58 proteins enriched in the chromocenters that represent candidates for new constitutive heterochromatin proteins (figure 3.11).
The color coding based on the functional annotation of the identified proteins revealed several histones, many chromatin proteins as well as some proteins involved in transcriptional regulation. Although some membrane associated and RNA associated proteins were still present, the big group of nucleoporins and ribosomal and splicing proteins were excluded by the validation-based filtering. The filtering resulted in explicit changes in the GO analysis output (figure 3.12). Within the category of biological process, chromatin organization was the GO term with the highest number of associated genes from our target list, and nucleosome positioning, chromatin assembly or disassembly, and negative regulation of DNA recombination were the GO terms with the highest enrichment. Most of the proteins were associated with nucleic acid binding and DNA binding. The highest enrichment for molecular function was observed for the GO terms nucleosomal DNA binding, nucleosome binding, and chromatin DNA binding. All proteins present in the filtered list belonged to the intracellular organelle part, most of them localizing in the nucleus, non-membrane bound organelles, and chromosomal parts. The top five enriched GO terms for cellular components were DNA packaging complex, nucleosome, protein-DNA complex, nuclear chromosome, and heterochromatin.
Summarizing, all GO terms resemble protein functions related to chromatin packaging in nucleosomes and heterochromatin fractions. We conclude that our unbiased approach for heterochromatin isolation in combination with a rigorous validation-based cut-off represents a system for reliable identification of heterochromatin proteins.

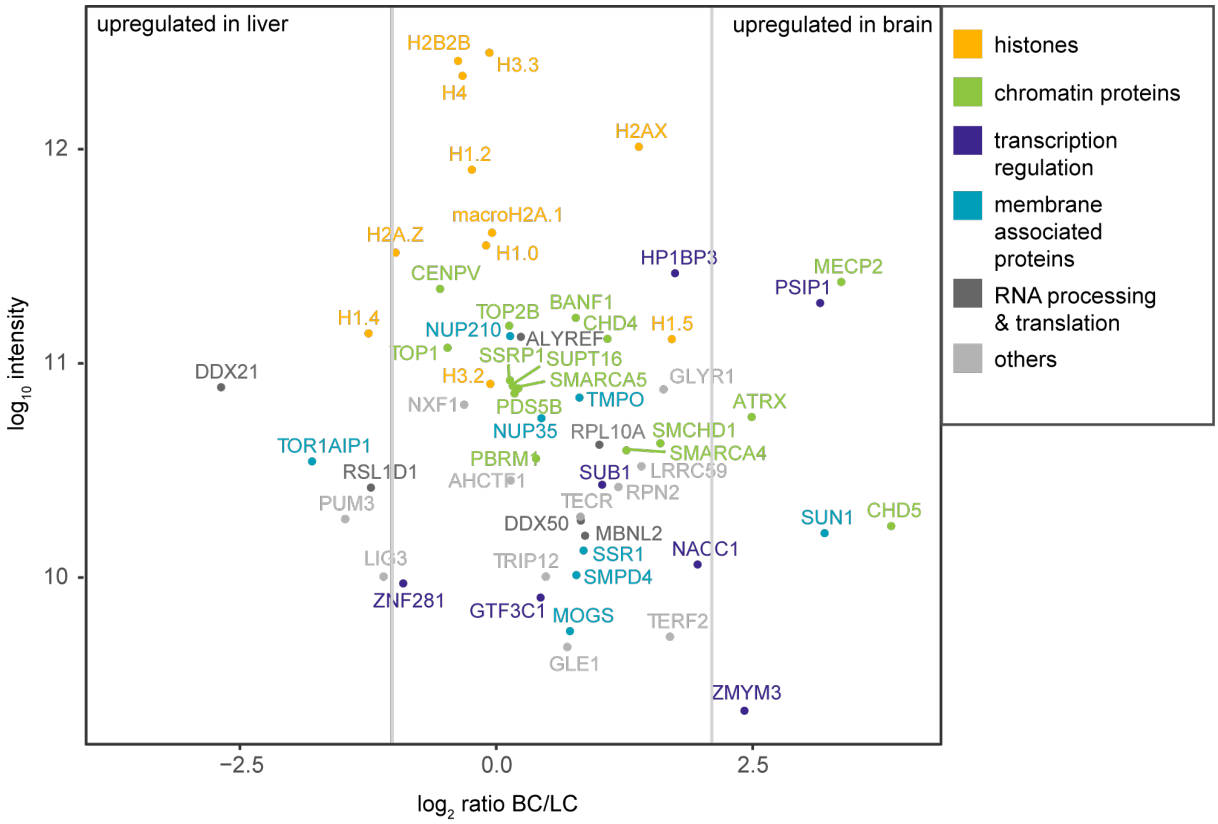



**Figure 3.11: Heterochromatin proteins identified by mass spectrometry after the validation-based cut-off.** This plot contains the proteins that localized at the chromocenters in the validation stainings and all other proteins with similar or higher chromocenter enrichment. The dot plot shows the $\log_{10}$ intensity plotted against the $\log_2$ ratio of brain chromocenter (BC) versus liver chromocenter (LC) proteins. Vertical lines indicate the three categories: proteins enriched in the liver, proteins with similar abundance in the brain and liver chromocenters, and proteins enriched in the brain. The color code indicates the protein function manually assigned based on the UniProt webpage functional information. Histones are labeled in orange, chromatin proteins in green, proteins involved in transcriptional regulation in purple, membrane associated proteins in cyan, proteins involved in RNA processing or RNA-binding proteins in dark gray, and proteins not fitting into the categories in light gray.

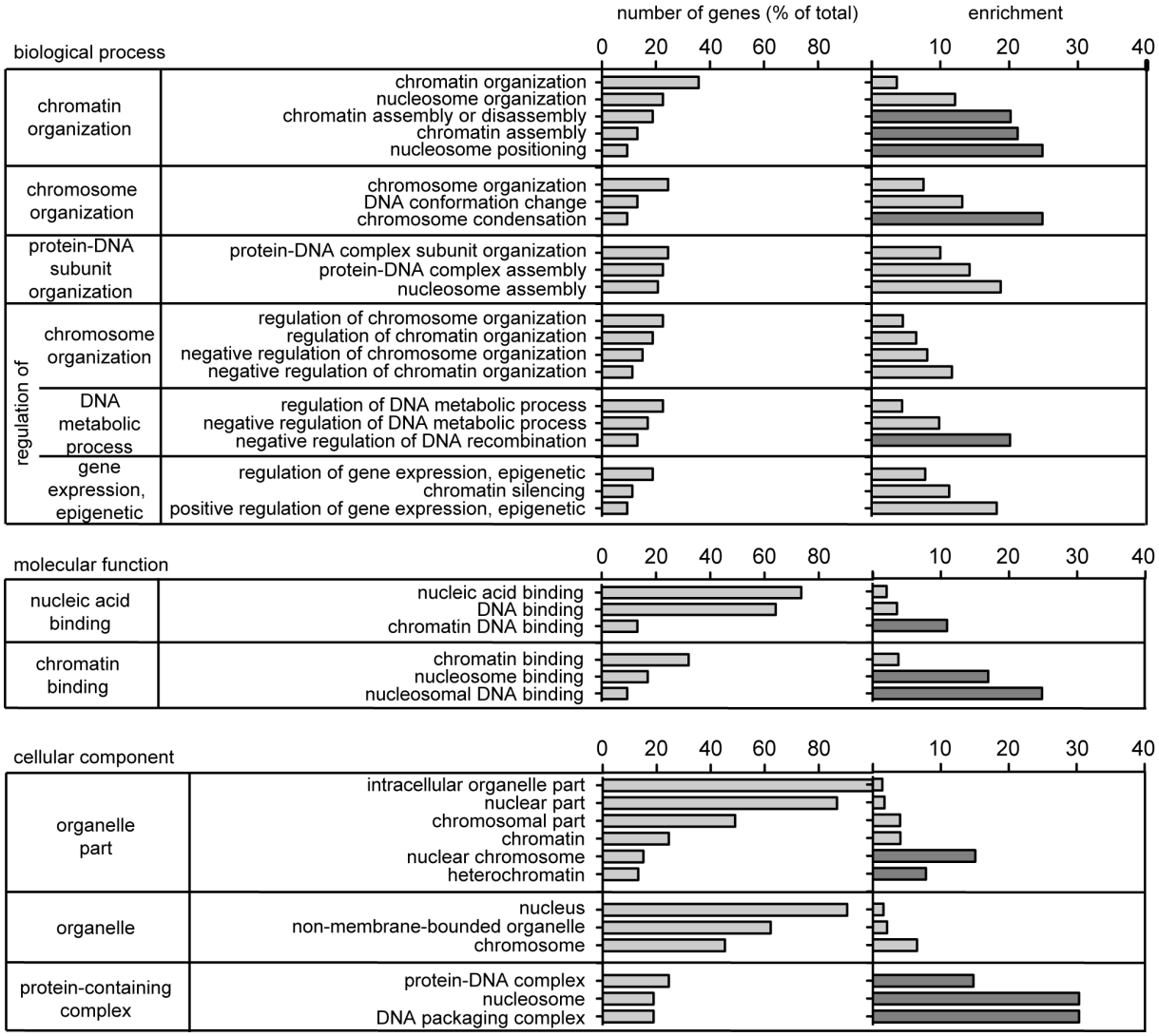



**Figure 3.12: Gene ontology (GO) analysis of heterochromatin proteins selected after validation-based cut-off.** The protein list was subjected to GOrilla tool [269] for gene ontology analysis in the categories biological process, molecular function, and cellular component. The proteins enriched in heterochromatin were added as the target list and all identified proteins (without cut-off or filtering) as the background list. GO terms with a p-value lower than $5 \times 10^{-5}$ were considered. The GO terms were grouped according to the highest common gene ontology term within the diagram of the GOrilla output and redundant terms were removed manually. Plotted is the number of genes in the GO term (b) as % of total genes in the target list (n) and the enrichment calculated by the GOrilla tool as (b/n)/(B/N) with b: number of genes in the target list associated with specific GO term; n: total number of genes in the target list; B: number of genes in the background list associated with specific GO term; N: total number of genes in the background list. The five proteins (biological function, cellular component) or three proteins (molecular function) with the highest enrichment in the categories are highlighted in dark gray.

### 3.3.6 Proteomic analysis of low abundant heterochromatin proteins

Initially, we decided to filter out all protein hits that were not reproducibly identified in all three biological replicates of the nuclei and chromocenter fractions. Of note, we might have lost low abundance heterochromatin proteins with this approach, as they could have been enriched in the chromocenter fractions due to the isolation procedure but might not be reproducibly identified in the less enriched nuclei fraction. For this reason, we decided to plot also the reproducibly identified (in the three biological replicates) heterochromatin proteins, which were though not identified in all three replicates of the nuclei (figure 3.13). Within the group of low abundant heterochromatin proteins, only a few histones and chromatin proteins were identified. Instead, many proteins involved in transcriptional regulation and many proteins with very specific functions not fitting into the applied functional categories were identified. RNA processing and translation-related proteins were present with low abundance.

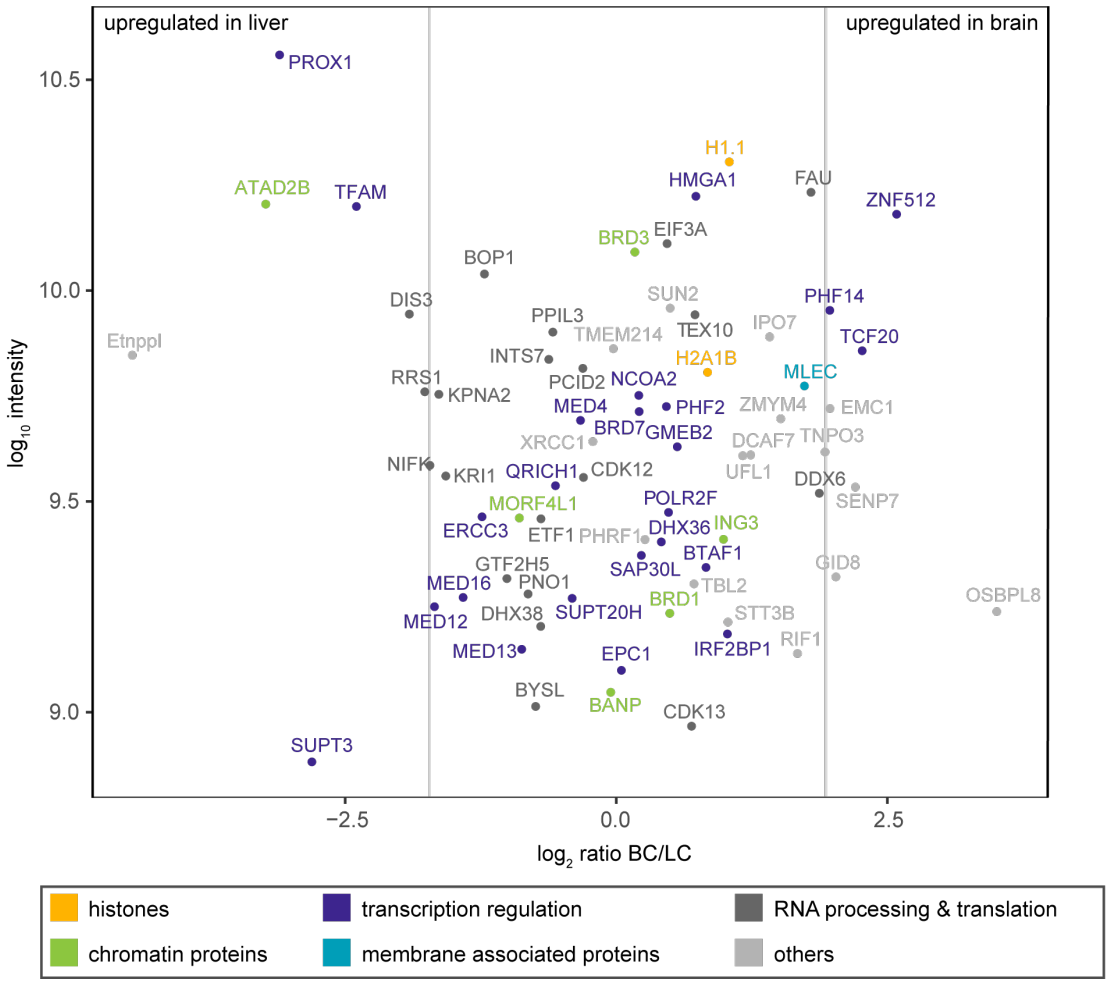



**Figure 3.13: Heterochromatin proteins exclusively identified in the chromocenters by quantitative mass spectrometry.** This plot contains the proteins identified in the chromocenter fraction of all three replicates, but not reproducibly identified in the nucleus fractions (in two or fewer replicates). The dot plot shows the $\log_{10}$ intensity plotted against the $\log_2$ ratio of brain chromocenter (BC) versus liver chromocenter (LC) proteins. Vertical lines indicate the three categories: proteins enriched in the liver, proteins with similar abundance in the brain and liver chromocenters, and proteins enriched in the brain. The color code indicates the protein function manually assigned based on the UniProt webpage functional information. Histones are labeled in orange, chromatin proteins in green, proteins involved in transcriptional regulation in purple, membrane associated proteins in cyan, proteins involved in RNA processing or RNA-binding proteins in dark gray, and proteins not fitting into the categories in light gray.

## 3.4 Discussion

Constitutive heterochromatin is involved in the maintenance of genome stability [58, 60] and is required for the nuclear spatial organization in compartments [72]. Disfunction of heterochromatin was associated with cancer progression [58]. Recently, constitutive heterochromatin was reported to constitute membraneless liquid-like compartments concentrating factors within and excluding others, thus regulating nuclear DNA metabolism [257–259, 274]. In mouse cells, constitutive heterochromatin can be found at pericentromeric regions, which cluster during interphase forming chromocenters [25, 26]. The organization of heterochromatin in chromocenters differs between cell types [50–53] and changes upon differentiation [55–57]. Thus, we wanted to address whether the observed changes in chromocenter organization are related to the proteomic composition. We aimed to identify conserved chromocenter proteins, on the one hand, and specific differences in protein abundance between cells with differing heterochromatin organization, on the other hand.

Mouse brain tissue showed higher heterochromatin clustering represented by lower chromocenter numbers and larger volumes in comparison to liver tissue upon staining of tissue sections using the DNA dye DAPI (figure 3.1). Thus, the question was raised whether there are differences in the proteomic composition of chromocenters from these tissues.

To isolate the chromocenter proteins, we adapted a protocol for the isolation of chromocenters from mouse liver published by Prusov and Zatsepina [94, 95] and adjusted it to be used with liver and brain tissue. In contrast to other heterochromatin enrichment methods relying on immunoprecipitation of a specific heterochromatin protein or a specific gene locus (see 1.1), our approach is not biased towards a specific previously identified heterochromatin component. In addition, our protocol does not rely on protein crosslinking prior to the isolation procedure, thus yielding native heterochromatin. Of note, the heterochromatin protein 1 (HP1), which was used for heterochromatin enrichment in previous studies (e.g., [275]), was not found reproducibly enriched in chromocenters in our study. It was partially enriched in nuclei, partially in chromocenters, and thus filtered out during the analysis procedure. In line with these results, HP1 isoforms were reported to show different localization dependent on the species, cell type, cell cycle, and the differentiation status [159, 276–280]. In mouse cells, HP1$\alpha$ and $\beta$ colocalized with heterochromatin, whereas HP1$\gamma$ was mainly distributed in the cell nucleus and predominantly associated with euchromatin [276, 278, 280]. Nevertheless, different HP1 isoforms were found associated with active genes and involved in transcriptional activation and development [281], indicating that the protein acts as an adaptor with diverging functions depending on the context. In addition, HP1$\alpha$ and $\beta$ were not found in several differentiated cells, e.g. terminally differentiated nucleated erythrocytes of non-mammalian vertebrates, on a subset of neurons, and in liver parenchyma [278, 279]. Thus, as we analyzed the whole cell population of brain and liver tissue, the reduced accumulation of HP1 at chromocenters might be explained by cell type and differentiation-specific differences in protein distribution.

For the mass spectrometry analysis, we used a dimethyl labeling approach which allowed quantitative comparison of different samples. To base our analysis only on the most reliable protein

hits, we considered only proteins reproducibly identified in all three biological replicates and enriched in chromocenters in comparison to the full nucleus proteome for further analysis. Thus, we cannot rule out that we lost heterochromatin interacting proteins which were not exclusively localized at the chromocenters, bound dynamically to heterochromatin, or were separated from the chromocenters during the isolation procedure. The proteomic analysis led to the identification of several histones, chromatin proteins, and proteins involved in transcriptional regulation. Furthermore, membrane associated proteins, proteins involved in ribosome or spliceosome formation, and proteins not fitting into any of the categories were identified. Chromocenters were described to localize close to nucleoli and the nuclear membrane [51, 64, 65], indicating heterochromatin association with these regions and explaining the copurification. In addition, nucleoporins and lamins were described to interact with heterochromatin and proposed to be involved in heterochromatin organization [71, 282–285]. Previous proteomic studies of heterochromatin yielded similar proteins, including histones, proteins involved in transcription and RNA processing, chromatin remodelers, RNA binding proteins, and nuclear pore proteins [90, 275, 286]. Furthermore, gene ontology analysis revealed proteins with the GO classifications like nucleic acid binding, chromosome, chromatin, DNA packaging, and chromatin assembly [90].

To validate the mass spectrometry analysis results, we determined the subcellular localization of representative protein hits by immunofluorescence stainings on tissue sections. Based on the protein hits with confirmed chromocenter accumulation *in situ*, we applied a validation-based cut-off to the proteins identified by mass spectrometry to obtain a list of heterochromatin-specific proteins. The subsequent GO analysis confirmed successful data filtering, as terms like chromatin assembly and nucleosome positioning were highly enriched, while terms associated with nucleoporins and nucleoli were absent.

Using this method, we identified chromocenter proteins conserved between species and tissue-specific differences in chromocenter protein abundance. We identified a list of 58 proteins that might be involved in chromocenter-specific functions and additional 213 potential heterochromatin interacting proteins. A graphic summary of the subcellular localization of the validated hits is shown in figure 3.14.

Besides the core histones, we identified the topoisomerases I (TOP I) and IIb (TOP IIb) as well as the histone variants macroH2A.1, H2AX, and H3.3 as being upregulated in the chromocenters. DNA topoisomerases are essential enzymes catalyzing the transient cleavage and subsequent resealing of DNA to induce or remove coils of the DNA helix either by inducing a single-strand break (TOP I) or a double-strand break (TOP II) [287]. On the one hand, topoisomerases are involved in opening chromatin making it accessible for processes like transcription [288, 289], on the other hand, they were described to be required for gene silencing of heterochromatic regions and chromosome condensation [290–293]. Further studies are necessary to elucidate the role of the topoisomerases in pericentromeric heterochromatin in mammals, as most studies were performed in Drosophila and Arabidopsis and the functions might be species, cell type, and differentiation-specific.

Comparing the structure of histone variants to the canonical histone H2A, macroH2A contains an additional large macro domain at its C-terminus and H2A.X contains a C-terminal Ser-Gln-Glu-

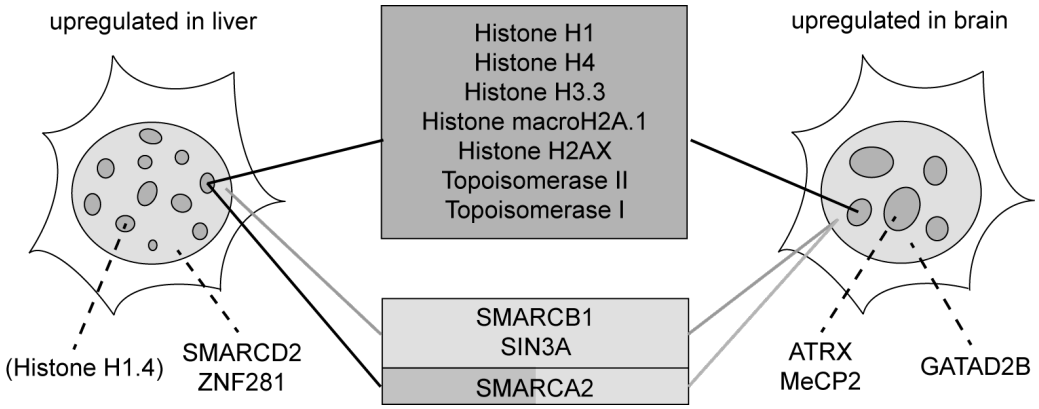



**Figure 3.14: Schematic summary of the heterochromatin proteins validated.** The cell schemes mimic the differing heterochromatin organization in mouse brain and liver tissue. From the proteins upregulated in the liver, Histone H1.4 localized at the chromocenters and SMARCD2, and ZNF281 localized in the nucleus. From the proteins upregulated in the brain, ATRX and MeCP2 localized are at the chromocenters and GATAD2B is localized in the nucleus. For the proteins with similar abundance in both tissues, all histones, histone variants, and the topoisomerases I and II localized at the chromocenters, SMARCB1, SIN3A in the nuclei, and SMARCA2 localized at the chromocenters or in the nucleus depending on the tissue.

Tyr motif critical for its function [11]. MacroH2A.1 is considered to be involved in gene silencing, and higher-order chromatin compaction and associated with heterochromatin repeats [11, 294–296], in line with our results. The histone variant H2AX plays an important role in DNA damage response as its phosphorylation on serine 139, resulting in $\gamma$H2AX, occurs upon DNA damage [297], recruits DNA repair proteins and is commonly used as a marker for DNA double-strand breaks [298]. Although H2AX was described to distribute randomly on chromatin [297], another study found H2AX overrepresented at heterochromatin due to a higher frequency of H2AX in heterochromatin than in euchromatin [299]. In addition, H2AX might be involved in maintaining genome stability [300] and in chromatin remodeling and inactivation of sex chromosomes in male mice [301]. Thus, we propose that H2AX accumulates at chromocenters in mouse brain and liver tissue, possibly playing a role in chromatin organization. Histone H3.3 differs from the canonical H3 proteins by a few amino acid substitutions, one localized in the N-terminal tail, the others in the histone-fold domain [11]. The variant replaces the canonical H3 at active genes and promotors but was also found at repetitive heterochromatin regions such as telomeres and pericentromeric heterochromatin [21]. The death domain-associated protein (DAXX) was shown to specifically interact with H3.3 via the unique motif of the histone variant. In addition, it was reported to build a complex with ATRX which is involved in H3.3 deposition at pericentromeric heterochromatin [21, 302–304]. Due to the many factors involved in H3.3 deposition, its function in biological processes is still vague [305].

Among the proteins enriched in chromocenters and displaying a tissue-specific difference in abundance, we identified ATRX as enriched in the brain. ATRX was reported to bind to tandem repeat sequences and to be involved in chromatin organization and gene silencing [306–311]. Interestingly, coimmunoprecipitation experiments indicated an interaction of ATRX with the epigenetic

reader MeCP2 [165, 172], which localized at heterochromatin and was enriched in brain tissue in our experiments. MeCP2 binds to methylcytosine on the DNA and modulates transcriptional regulation [126, 131, 139, 312] and chromatin organization (reviewed in [3]). MeCP2 levels differ between cell types [233, 313] and increase upon differentiation [55, 233, 313, 314] which results in higher heterochromatin clustering [55, 57, 159]. Interestingly, ATRX and MeCP2 were proposed to regulate the expression of certain genes by chromatin structure organization [315] and ATRX might contribute to MeCP2-mediated heterochromatin organization during neural differentiation [316].

In mouse liver, we identified histone H1.4 as upregulated and enriched in chromocenters. Histone H1.4 is one of five main type histones H1 including H1.1-H1.5, which are expressed in a replication-dependent manner in somatic cells during S-phase [15]. H1.2-5 are present in most somatic tissues [317, 318] while H1.1 expression was described in several mouse tissues, including liver and brain [317, 319, 320]. The deletion of either H1.2, H1.3, or H1.4 suggested that the different variants can compensate for the lack of each other [321], but the differing distribution of the variants points toward variant-specific functions [15, 322]. Sequential inactivation of all three variants resulted in a lower H1 to nucleosome ratio, reduced DNA packaging, and changes in gene expression [323, 324]. Interestingly, it was proposed that histone H1 and MeCP2 compete for binding to a common chromatin binding site, as MeCP2 accelerated H1 exchange and both showed similar nucleosome binding motifs [253]. This idea was supported by reports describing the amount of H1 and MeCP2 in the nucleus of different cells. While the amount of H1 in neuronal nuclei was 0.45 molecules per nucleosome, glia cells had one H1 molecule per nucleosome [325]. On the contrary, neuronal cells showed very high MeCP2 levels with 0.5 MeCP2 molecules per nucleosome in comparison to low MeCP2 amounts in glia cells (ca. 0.06 molecules/ nucleosome) [138]. In addition, Skene *et al.* found a 2-fold increase in histone H1 levels in MeCP2 deficient neuronal nuclei, suggesting that MeCP2 replaces histone H1 in neurons [138]. Thus, the interplay of histone H1 and MeCP2 might contribute to the differing chromocenter organization observed between the brain and liver in our study, especially as both proteins were found to be involved in heterochromatin organization.

In addition to the validated chromocenter proteins, we identified many other candidate chromocenter proteins. This list of protein candidates possibly involved in chromocenter formation and function can serve as a basis for further studies aiming to identify novel heterochromatin components. The generation of fluorescently tagged fusion proteins could complement immunofluorescence stainings for *in situ* analysis of protein localization, as antibodies are not available for all proteins of interest. Moreover, the expression of plasmids coding for fusion proteins in a cellular system would allow a protein-level dependent analysis of chromocenter clustering, as described before for MeCP2 [55].

In this study, we showed that an unbiased heterochromatin isolation procedure with subsequent quantitative mass spectrometry analysis in combination with antibody-based validation experiments can be used to elucidate the mouse heterochromatin proteome. We identified 58 candidate chromocenter proteins that might be involved in chromocenter formation and function. Furthermore, we showed that the heterochromatin organization differs between mouse brain and liver

tissue, and we identified candidate proteins that might mediate the differing organization. We propose that MeCP2, ATRX, and histone H1 play important roles in the distinct pericentromeric heterochromatin organization in these tissues.

## 3.5 Supplementary material

### Supplementary tables

**Table 3.1:** Primary and secondary antibody characteristics.

| Reactivity | Host | Dilution | Application | Catalog / clone | Company / reference |
|---|---|---|---|---|---|
| Anti-ATRX (D-5) | mouse | 1:50 | IF / WB | sc-55584 | Santa Cruz Biotechnology |
| Anti-Gatad2b | rabbit | 1:200 | IF / WB | AB-2641884 | invitrogen |
| Anti-Histone H1 (EPR6536) | rabbit | 1:200 / 1:1000 | IF / WB | ab134914 | abcam |
| Anti-Histone macro H2A.1 | rabbit | 1:200 / 1:500 | IF / WB | 07-219 | upstate |
| Anti-Histone H2AX | rabbit | 1:200 | IF | A300-083A | Bethyl |
| Anti-Histone H2AX | rabbit | 1:1000 | WB | ab20669 | abcam |
| Anti-Histone H3.3 (EPR17899) | rabbit | 1:1000 | IF / WB | ab176840 | abcam |
| Anti-Histone H4 | rabbit | 1:200 | IF / WB | ab7311 | abcam |
| Anti-Lamin A/C (XB 10) | mouse | Undiluted TCSN | IF / WB | - | gift from Brian Burke |
| Anti-Lamin B (X223) | mouse | Undiluted TCSN | IF / WB | 65147C | Progen |
| Anti-Mbnl2 (3B4) | mouse | 1:50 | IF / WB | sc-136167 | Santa Cruz Biotechnology |
| Anti-MeCP2 (4H7) | rat | undiluted TCSN | IF / WB | - | Jost *et al.*, 2011 [326] |
| Anti-mSin3A (K20) | rabbit | 1:50 | IF / WB | sc-994 | Santa Cruz Biotechnology |
| Anti-Nup153 (QE5) | mouse | 1:100 | IF / WB | ab24700 | abcam |
| Anti-SMARCA2 / BRM | rabbit | 1:100 / 1:1000 | IF / WB | ab15597 | abcam |
| Anti-SMARCB1 / BAF47 (D8M1) | rabbit | 1:250 / 1:1000 | IF / WB | 91735S | Cell Signaling Technology |
| Anti-Smarcd2 (EPR20860-251) | rabbit | 1:100 / 1:1000 | IF / WB | ab220164 | abcam |
| Anti-Topoisomerase I (EPR5375) | rabbit | 1:100 | IF / WB | ab109374 | abcam |
| Anti-Topoisomerase II alpha + beta (TOP2B) (EPR5377) | rabbit | 1:100 / 1:1000 | IF / WB | ab109524 | abcam |
| Anti-U1 snRNP70 (C-3) | mouse | 1:50 / 1:100 | IF / WB | sc-390899 | Santa Cruz Biotechnology |
| Anti-Znf281 (D-8) | mouse | 1:50 / 1:100 | IF / WB | sc-166933 | Santa Cruz Biotechnology |
| Anti-mouse IgG Cy3 | donkey | 1:500 | IF | 715-166-151 | Jackson ImmunoResearch Laboratories, inc. |
| anti-mouse IgG Cy5 | donkey | 1:250 | IF | 715-175-150 | Jackson ImmunoResearch Laboratories, inc. |
| Anti-mouse IgG HRP | sheep | 1:5000 | WB | NA 931 | Amersham Pharmacia Biotech |

| | | | | | |
|---|---|---|---|---|---|
| anti-rabbit IgG Cy3 | donkey | 1:500 | IF | 711-165-152 | Jackson ImmunoResearch Laboratories, inc. |
| anti-rabbit IgG Cy5 | donkey | 1:400 | IF | 711-175-152 | Jackson ImmunoResearch Laboratories, inc. |
| Anti-rabbit IgG HRP | goat | 1:10000 | WB | A0545 | Sigma-Aldrich, Inc. |
| Anti-rat IgG Cy5 | donkey | 1:250 | IF | 712-175-153 | Jackson ImmunoResearch Laboratories, inc. |

IF: Immunofluorescence; WB: Western blot; HRP: horseradish peroxidase; TCSN: Tissue culture supernatant.

**Table 3.2:** Imaging and flow cytometer characteristics.

| System / Company | Lasers / Lamps | Filters (ex. & em. (nm)) | Objectives / lenses | Detection system | Application |
|---|---|---|---|---|---|
| Widefield Axiovert 200 / Zeiss | HBO100 mercury lamp | DAPI (300-400 & 410-510) | 40x Plan-Neofluar NA 1.4 Oil Ph3 | 12-bit AxioCam mRM | Fluorescence imaging chromocenter isolation steps |
| Widefield Axioplan / Zeiss | HBO100 mercury lamp | DAPI (450-490 & 515-565) | 40x Plan-Neofluar NA 1.3 Oil Ph3 | 12-bit AxioCam mRM | Fluorescence imaging chromocenter isolation steps |
| UltraView VoX spinning disk on an inverted Nikon Ti-E microscope / Perkin Elmer | Solid state diode lasers (405 nm, 561 nm, 640 nm) | 405/568/640 405: 415–475 561: 580–650 640: 664–754 | 60x Plan-Apochromat NA 1.45 Oil | cooled 14-bit Hamamatsu® C9100-50 EMCCD | Confocal z-stack imaging of tissue slices |
| Amersham AI600 imager / GE Healthcare | Chemiluminescence | - | - | 16-bit Peltier cooled Fujifilm Super CCD | Western blot imaging |

ex.: excitation; em.: emission.

**Table 3.3:** Cell line characteristics.

| Name | Species | Type | Genotype | References |
|---|---|---|---|---|
| C2C12 | *Mus musculus* | myoblast | wild type | Yaffe & Saxel, 1977 [327] |

**Table 3.4:** Plot statistics.

| Figure | Sample | n | Median | Mean | StDev | 95% CI | p-value |
|---|---|---|---|---|---|---|---|
| 3.1B | brain | 223 | 5 | 5.46 | 2.06 | 0.01 | <2.2e-16 |
| | liver | 224 | 9 | 9.16 | 2.86 | 0.01 | |
| 3.1C | brain | 223 | 4.12 | 6.30 | 6.21 | 0.01 | 1.01e-09 |
| | liver | 224 | 3.1 | 4.48 | 4.31 | 0.01 | |
| 3.10A | Histone H4 brain | 33 | 1.21 | 1.21 | 0.10 | - | 8.706e-10 |
| | Histone H4 liver | 32 | 1.43 | 1.43 | 0.15 | - | |
| | Histone H3.3 brain | 31 | 1.24 | 1.26 | 0.14 | - | 0.8102 |
| | Histone H3.3 liver | 33 | 1.27 | 1.26 | 1.10 | - | |
| | macro H2A.1 brain | 35 | 1.33 | 1.32 | 0.12 | - | 8.099e-11 |
| | macro H2A.1 liver | 32 | 1.15 | 1.15 | 0.05 | - | |
| | Histone H2AX brain | 34 | 1.27 | 1.29 | 0.15 | - | 1.809e-07 |
| | Histone H2AX liver | 31 | 1.13 | 1.14 | 0.05 | - | |
| | TopII brain | 35 | 1.40 | 1.42 | 0.28 | - | 2.346e-07 |
| | TopII liver | 32 | 1.09 | 1.09 | 0.02 | - | |
| | TopI brain | 30 | 1.21 | 1.26 | 0.12 | - | 3.962e-07 |
| | TopI liver | 31 | 1.42 | 1.42 | 0.07 | - | |
| 3.10B | ATRX brain | 30 | 1.20 | 1,22 | 0.11 | - | 1.106e-11 |
| | ATRX liver | 32 | 1.05 | 1.06 | 0.03 | - | |
| | MeCP2 brain | 30 | 1.68 | 1.69 | 0.26 | - | <2.2e-16 |
| | MeCP2 liver | 32 | 1.02 | 1.02 | 0.01 | - | |
| | Histone H1 brain | 30 | 1.52 | 1.66 | 0.47 | 0.01 | 0.3934 |
| | Histone H1 liver | 31 | 1.47 | 1.52 | 0.28 | - | |

n: number of cells; StDev: standard deviation; CI: confidence interval; p-value: comparison brain to liver.

**Table 3.5:** Proteins enriched in heterochromatin identified by mass spectrometry.

| protein name | Protein ID | $\log_{10}$ intensity | $\log_2$ ratio BC/LC | protein function | enrichment category |
|---|---|---|---|---|---|
| ACTA2 | P62737 | 10.57 | 1.07 | others | common |
| ACTR5 | Q80US4 | 9.67 | 0.47 | chromatin | common |
| ACTR6 | A0A0R4J009 | 9.23 | 0.06 | chromatin | common |
| ACTR8 | Q8R2S9 | 9.52 | 0.09 | chromatin | common |
| AHCTF1 | Q8CJF7 | 10.43 | -0.25 | others | common |
| AKAP8L | Q9R0L7 | 10.64 | 0.86 | transcription regulation | common |
| ALYREF | O08583 | 11.10 | -0.15 | RNA processing & translation | common |
| ARID1A | A2BH40 | 10.32 | 1.03 | chromatin | common |
| ARID2 | E9Q7E2 | 9.87 | 0.17 | chromatin | common |
| ARNTL | Q9WTL8 | 9.79 | -1.22 | transcription regulation | common |
| BANF1 | O54962 | 11.19 | 0.38 | chromatin | common |
| BAP18 | Q9DCT6 | 9.66 | 0.72 | chromatin | common |
| BRD4 | Q9ESU6 | 10.61 | -0.03 | chromatin | common |
| BRD9 | A0A0R4J175 | 9.84 | 1.08 | chromatin | common |
| CCDC47 | Q9D024 | 10.02 | 1.18 | others | common |
| CD2BP2 | Q9CWK3 | 9.93 | 0.40 | RNA processing & translation | common |
| CDC5L | Q6A068 | 11.07 | -0.73 | transcription regulation | common |
| CENPV | Q9CXS4 | 11.32 | -0.94 | chromatin | common |

| protein name | Protein ID | $\log_{10}$ intensity | $\log_2$ ratio BC/LC | protein function | enrichment category |
|---|---|---|---|---|---|
| CHD4 | Q6PDQ2 | 11.09 | 0.69 | chromatin | common |
| CHTOP | Q9CY57 | 10.68 | 0.85 | transcription regulation | common |
| CLOCK | O08785 | 9.75 | -0.82 | transcription regulation | common |
| CMAS | Q99KK2 | 10.92 | -0.47 | others | common |
| CPSF4 | E0CXT7 | 9.79 | -0.31 | RNA processing & translation | common |
| CREB1 | Q01147 | 10.24 | -0.27 | transcription regulation | common |
| CSTF3 | Q99LI7 | 10.68 | -0.26 | RNA processing & translation | common |
| DDOST | O54734 | 10.43 | 0.77 | membrane | common |
| DDX17 | Q501J6 | 11.53 | 0.42 | RNA processing & translation | common |
| DDX23 | D3Z0M9 | 11.13 | -1.24 | RNA processing & translation | common |
| DDX5 | Q61656 | 12.05 | -0.13 | RNA processing & translation | common |
| DDX50 | Q99MJ9 | 10.24 | 0.43 | RNA processing & translation | common |
| DNAJA1 | P63037 | 10.57 | -0.68 | others | common |
| DNAJA2 | Q9QYJ0 | 10.61 | 0.37 | others | common |
| DNAJB4 | Q9D832 | 9.69 | -1.22 | others | common |
| EED | Q921E6 | 9.46 | -0.95 | transcription regulation | common |
| FARSA | Q8C0C7 | 9.81 | 0.27 | RNA processing & translation | common |
| GLE1 | Q8R322 | 9.65 | 0.30 | others | common |
| GLYR1 | Q922P9 | 10.85 | 1.24 | transcription regulation | common |
| GNL3L | Q6PGG6 | 9.89 | 0.48 | others | common |
| GTF2H2 | Q9JIB4 | 9.19 | -0.19 | transcription regulation | common |
| GTF3C1 | Q8K284 | 9.88 | 0.04 | transcription regulation | common |
| GTF3C3 | Q3TMP1 | 9.60 | 0.70 | transcription regulation | common |
| GTF3C4 | Q8BMQ2 | 9.86 | -0.16 | transcription regulation | common |
| H1F0 | P10922 | 11.53 | -0.49 | histones | common |
| H2AFX | P27661 | 11.99 | 1.00 | histones | common |
| H2AFY | Q9QZQ8 | 11.59 | -0.43 | histones | common |
| H3F3C | P02301 | 12.43 | -0.46 | histones | common |
| HDGFL2 | Q3UMU9 | 10.09 | -0.83 | chromatin | common |
| HIST1H1B | P43276 | 11.09 | 1.32 | histones | common |
| HIST1H1C | P15864 | 11.88 | -0.63 | histones | common |
| HIST2H2BB | Q64525 | 12.39 | -0.77 | histones | common |
| HIST2H3C2 | P84228 | 10.88 | -0.45 | histones | common |
| HIST4H4 | P62806 | 12.32 | -0.72 | histones | common |
| HP1BP3 | Q3TEA8 | 11.40 | 1.35 | chromatin | common |
| HSD17B12 | O70503 | 9.64 | 0.08 | others | common |
| INTS1 | Q6P4S8 | 10.19 | -0.54 | transcription regulation | common |
| INTS10 | Q8K2A7 | 9.73 | -0.34 | RNA processing & translation | common |
| INTS14 | Q8R3P6 | 9.98 | -0.50 | RNA processing & translation | common |
| INTS3 | Q7TPD0 | 10.24 | -0.01 | transcription regulation | common |
| INTS6 | Q6PCM2 | 9.95 | -0.67 | RNA processing & translation | common |
| ISY1 | Q69ZQ2 | 10.21 | -0.89 | RNA processing & translation | common |
| KHDRBS1 | Q60749 | 11.52 | 1.12 | RNA processing & translation | common |

| protein name | Protein ID | $\log_{10}$ intensity | $\log_2$ ratio BC/LC | protein function | enrichment category |
|---|---|---|---|---|---|
| LAS1L | A2BE28 | 10.00 | 0.19 | RNA processing & translation | common |
| LMNA | P48678 | 12.44 | -0.76 | membrane associated | common |
| LMNB1 | P14733 | 12.26 | 0.99 | membrane associated | common |
| LMNB2 | P21619 | 12.20 | 1.01 | membrane associated | common |
| LRRC59 | Q922Q8 | 10.50 | 1.02 | others | common |
| MBNL2 | Q8C181-4 | 10.17 | 0.47 | RNA processing & translation | common |
| MCM3AP | Q9WUU9 | 9.95 | -0.14 | others | common |
| MED1 | Q925J9 | 9.84 | -0.11 | transcription regulation | common |
| MED10 | Q9CXU0 | 9.37 | 0.49 | transcription regulation | common |
| MED14 | A2ABV5 | 9.76 | 0.38 | transcription regulation | common |
| MED17 | Q8VCD5 | 9.78 | -0.30 | transcription regulation | common |
| MED24 | Q99K74 | 9.71 | 0.08 | transcription regulation | common |
| MED8 | Q9D7W5 | 9.59 | -0.09 | transcription regulation | common |
| MOGS | Q80UM7 | 9.73 | 0.33 | membrane associated | common |
| MYL12B | Q3THE2 | 9.70 | 0.03 | others | common |
| MYL6 | Q60605 | 10.24 | 0.45 | others | common |
| NCOA5 | Q91W39 | 11.01 | 0.96 | transcription regulation | common |
| NONO | Q99K48 | 12.09 | 0.13 | transcription regulation | common |
| NOSIP | Q9D6T0 | 10.22 | -0.92 | others | common |
| NR3C1 | P06537 | 9.62 | -0.74 | transcription regulation | common |
| NRF1 | Q3UXF4 | 9.81 | -0.34 | transcription regulation | common |
| NSUN2 | Q1HFZ0 | 9.83 | -0.21 | RNA processing & translation | common |
| NUP107 | Q8BH74 | 10.71 | 0.29 | membrane associated | common |
| NUP133 | Q8R0G9 | 10.99 | 0.28 | membrane associated | common |
| NUP153 | E9Q3G8 | 10.59 | -0.38 | membrane associated | common |
| NUP155 | Q99P88 | 10.94 | 0.05 | membrane associated | common |
| NUP160 | Q9Z0W3 | 10.76 | 0.20 | membrane associated | common |
| NUP205 | B9EJ54 | 10.86 | 0.31 | membrane associated | common |
| NUP210 | Q9QY81 | 11.10 | -0.26 | membrane associated | common |
| NUP214 | Q80U93 | 10.70 | 0.24 | membrane associated | common |
| NUP35 | Q8R4R6 | 10.72 | 0.05 | membrane associated | common |
| NUP37 | Q9CWU9 | 10.63 | 0.23 | membrane associated | common |
| NUP43 | P59235 | 10.50 | 0.22 | membrane associated | common |
| NUP54 | Q8BTS4 | 10.94 | 0.63 | membrane associated | common |
| NUP58 | Q8R332 | 10.35 | 0.14 | membrane associated | common |
| NUP62 | Q63850 | 10.80 | 0.46 | membrane associated | common |
| NUP85 | Q8R480 | 10.46 | 0.17 | membrane associated | common |
| NUP88 | Q8CEC0 | 10.55 | -0.05 | membrane associated | common |
| NUP93 | Q8BJ71 | 11.11 | 0.30 | membrane associated | common |
| NUP98 | Q6PFD9 | 11.15 | 0.33 | membrane associated | common |
| NXF1 | Q99JX7 | 10.78 | -0.71 | others | common |
| PBRM1 | Q8BSQ9 | 10.53 | -0.01 | chromatin | common |
| PDS5A | Q6A026 | 10.26 | -0.80 | chromatin | common |

| protein name | Protein ID | $\log_{10}$ intensity | $\log_2$ ratio BC/LC | protein function | enrichment category |
|---|---|---|---|---|---|
| PDS5B | Q4VA53 | 10.84 | -0.22 | chromatin | common |
| PELP1 | Q9DBD5 | 9.99 | 0.62 | transcription regulation | common |
| PHF3 | B2RQG2 | 9.83 | -0.24 | transcription regulation | common |
| PPIL2 | Q9D787 | 10.12 | -0.92 | others | common |
| PRPF3 | Q922U1 | 10.68 | -0.65 | RNA processing & translation | common |
| PRPF6 | Q91YR7 | 10.98 | -1.25 | RNA processing & translation | common |
| PRPF8 | Q99PV0 | 11.77 | -1.18 | RNA processing & translation | common |
| RAD21 | Q61550 | 10.71 | -0.67 | chromatin | common |
| RAE1 | Q8C570 | 10.79 | 0.31 | RNA processing & translation | common |
| RANBP2 | Q9ERU9 | 11.38 | 0.18 | others | common |
| RBM15 | Q0VBL3 | 10.59 | -0.76 | RNA processing & translation | common |
| RBM39 | Q8VH51-2 | 11.16 | -0.69 | RNA processing & translation | common |
| RBMXL1 | Q91VM5 | 11.79 | 0.67 | RNA processing & translation | common |
| RNPC3 | Q3UZ01 | 9.61 | -0.19 | RNA processing & translation | common |
| RPL10A | A0A3B2WBL1 | 10.60 | 0.61 | RNA processing & translation | common |
| RPL11 | Q9CXW4 | 10.87 | 0.54 | RNA processing & translation | common |
| RPL12 | P35979 | 10.65 | 0.17 | RNA processing & translation | common |
| RPL21 | O09167 | 10.74 | 1.20 | RNA processing & translation | common |
| RPL22 | P67984 | 10.28 | 0.68 | RNA processing & translation | common |
| RPL23 | P62830 | 10.97 | 0.56 | RNA processing & translation | common |
| RPL23A | P62751 | 10.96 | 0.84 | RNA processing & translation | common |
| RPL27 | P61358 | 10.82 | 0.77 | RNA processing & translation | common |
| RPL30 | P62889 | 10.51 | 0.53 | RNA processing & translation | common |
| RPL31 | P62900 | 10.84 | 0.80 | RNA processing & translation | common |
| RPL35A | O55142 | 10.65 | 1.15 | RNA processing & translation | common |
| RPL5 | P47962 | 11.06 | 0.78 | RNA processing & translation | common |
| RPL9 | P51410 | 10.70 | 0.60 | RNA processing & translation | common |
| RPN1 | Q91YQ5 | 10.65 | 1.03 | others | common |
| RPN2 | Q9DBG6 | 10.40 | 0.80 | others | common |
| RPS11 | P62281 | 10.87 | 0.80 | RNA processing & translation | common |
| RPS13 | P62301 | 10.63 | 1.06 | RNA processing & translation | common |
| RPS14 | P62264 | 10.64 | 0.68 | RNA processing & translation | common |
| RPS15 | P62843 | 10.61 | 0.75 | RNA processing & translation | common |
| RPS15A | P62245 | 10.49 | 0.53 | RNA processing & translation | common |
| RPS16 | P14131 | 10.83 | 1.25 | RNA processing & translation | common |
| RPS19 | Q9CZX8 | 10.89 | 1.07 | RNA processing & translation | common |
| RPS2 | P25444 | 10.95 | 0.99 | RNA processing & translation | common |
| RPS20 | P60867 | 10.59 | 1.21 | RNA processing & translation | common |
| RPS23 | P62267 | 10.54 | 0.97 | RNA processing & translation | common |
| RPS25 | P62852 | 10.75 | 1.31 | RNA processing & translation | common |
| RPS27A | P62983 | 11.71 | -0.39 | RNA processing & translation | common |
| RPS3 | P62908 | 11.11 | 0.72 | RNA processing & translation | common |
| RPS3A | P97351 | 11.14 | 0.69 | RNA processing & translation | common |

| protein name | Protein ID | $\log_{10}$ intensity | $\log_2$ ratio BC/LC | protein function | enrichment category |
|---|---|---|---|---|---|
| RPS5 | Q91V55 | 10.48 | 0.36 | RNA processing & translation | common |
| RPS6 | P62754 | 10.71 | 1.15 | RNA processing & translation | common |
| RPS7 | P62082 | 11.02 | 0.66 | RNA processing & translation | common |
| RTRAF | Q9CQE8 | 10.38 | 1.10 | transcription regulation | common |
| RUVBL1 | P60122 | 11.04 | 0.03 | transcription regulation | common |
| RUVBL2 | Q9WTM5 | 11.12 | 0.42 | transcription regulation | common |
| SAFB2 | Q80YR5 | 10.63 | -0.84 | transcription regulation | common |
| SART1 | Q9Z315 | 10.87 | -0.62 | RNA processing & translation | common |
| SEH1L | Q8R2U0 | 10.52 | 0.33 | membrane associated | common |
| SENP3 | Q9EP97 | 9.72 | 0.71 | others | common |
| SF3B5 | Q923D4 | 10.69 | -0.84 | RNA processing & translation | common |
| SFPQ | Q8VIJ6 | 12.22 | 0.22 | RNA processing & translation | common |
| SIN3A | Q60520 | 10.32 | 0.85 | transcription regulation | common |
| SLTM | Q8CH25-2 | 10.79 | 0.80 | transcription regulation | common |
| SMARCA2 | Q6DIC0 | 10.32 | 0.91 | chromatin | common |
| SMARCA4 | Q3TKT4 | 10.57 | 0.88 | chromatin | common |
| SMARCA5 | Q91ZW3 | 10.86 | -0.18 | chromatin | common |
| SMARCB1 | Q9Z0H3 | 10.22 | 0.36 | chromatin | common |
| SMARCC1 | P97496 | 9.86 | 0.31 | chromatin | common |
| SMARCC2 | Q3UID0 | 11.22 | 0.90 | chromatin | common |
| SMARCE1 | O54941 | 10.69 | 0.80 | chromatin | common |
| SMC1A | Q9CU62 | 11.37 | -0.88 | chromatin | common |
| SMC3 | Q9CW03 | 11.36 | -0.92 | chromatin | common |
| SMCHD1 | Q6P5D8 | 10.60 | 1.21 | chromatin | common |
| SMPD4 | Q6ZPR5 | 9.99 | 0.39 | membrane associated | common |
| SNRNP70 | Q62376 | 11.66 | -0.99 | RNA processing & translation | common |
| SNRPB | P27048 | 11.47 | -0.77 | RNA processing & translation | common |
| SNRPB2 | Q9CQI7 | 10.83 | -0.86 | RNA processing & translation | common |
| SNRPD1 | P62315 | 11.13 | 0.06 | RNA processing & translation | common |
| SNRPD2 | P62317 | 11.55 | -0.62 | RNA processing & translation | common |
| SNRPD3 | P62320 | 11.14 | -0.74 | RNA processing & translation | common |
| SNRPF | P62307 | 10.93 | -0.61 | RNA processing & translation | common |
| SNRPG | P62309 | 11.49 | -0.90 | RNA processing & translation | common |
| SNU13 | Q9D0T1 | 11.05 | -1.04 | RNA processing & translation | common |
| SNW1 | A0A0B4J1E2 | 10.80 | -0.82 | transcription regulation | common |
| SPIN1 | Q61142 | 9.97 | 0.49 | chromatin | common |
| SSB | P32067 | 11.44 | 0.20 | others | common |
| SSR1 | A0A286YCT4 | 10.10 | 0.46 | membrane associated | common |
| SSRP1 | Q08943 | 10.90 | -0.26 | chromatin | common |
| STAG1 | Q9D3E6 | 10.12 | -0.72 | chromatin | common |
| STAG2 | O35638 | 10.56 | -0.84 | chromatin | common |
| SUB1 | P11031 | 10.41 | 0.64 | transcription regulation | common |
| SUPT16 | G3X956 | 10.87 | -0.23 | chromatin | common |

| protein name | Protein ID | $\log_{10}$ intensity | $\log_2$ ratio BC/LC | protein function | enrichment category |
|---|---|---|---|---|---|
| TAF1 | Q80UV9 | 9.53 | -0.46 | transcription regulation | common |
| TAF6 | Q62311 | 9.76 | -0.81 | transcription regulation | common |
| TAF9 | Q8VI33 | 9.75 | -0.19 | transcription regulation | common |
| TCERG1 | Q8CGF7 | 11.02 | -0.60 | transcription regulation | common |
| TECR | Q9CY27 | 10.26 | 0.43 | others | common |
| TERF2 | O35144 | 9.70 | 1.30 | others | common |
| THOC1 | Q8R3N6 | 10.63 | -0.82 | RNA processing & translation | common |
| THOC2 | B1AZI6 | 10.58 | -0.90 | RNA processing & translation | common |
| THOC5 | A0A0R4J0J6 | 10.39 | -0.68 | RNA processing & translation | common |
| THOC6 | Q5U4D9 | 10.45 | -0.58 | RNA processing & translation | common |
| THOC7 | Q7TMY4 | 10.67 | -0.38 | RNA processing & translation | common |
| TMPO | Q61033 | 10.77 | 0.82 | membrane associated | common |
| TOP1 | Q04750 | 11.05 | -0.87 | chromatin | common |
| TOP2B | Q64511 | 11.15 | -0.27 | chromatin | common |
| TRIP12 | G5E870 | 9.98 | 0.09 | others | common |
| TRRAP | A0A1D5RLL4 | 10.48 | 0.35 | chromatin | common |
| UBP1 | Q811S7 | 9.70 | 0.30 | transcription regulation | common |
| USP39 | Q3TIX9 | 10.66 | -0.99 | RNA processing & translation | common |
| WDR18 | Q4VBE8 | 10.38 | 0.35 | others | common |
| WDR5 | P61965 | 10.33 | 0.26 | chromatin | common |
| WRNIP1 | Q91XU0 | 10.62 | -0.32 | others | common |
| YEATS4 | Q9CR11 | 9.92 | 0.21 | chromatin | common |
| YTHDC1 | E9Q5K9 | 10.42 | 0.12 | RNA processing & translation | common |
| ZFP638 | A0A0N4SV80 | 10.76 | 1.00 | transcription regulation | common |
| ZMYND8 | A2A483 | 10.36 | 0.23 | chromatin | common |
| ADAR | Q99MU3 | 10.31 | 3.82 | RNA processing & translation | brain |
| ADARB1 | Q91ZS8-2 | 10.19 | 1.76 | RNA processing & translation | brain |
| ADNP | Q9Z103 | 10.40 | 1.37 | transcription regulation | brain |
| ATRX | Q61687 | 10.73 | 2.10 | chromatin | brain |
| CDK2AP1 | O35207 | 9.84 | 1.94 | others | brain |
| CHD5 | E9PYL1 | 10.22 | 3.46 | chromatin | brain |
| ERGIC1 | Q9DC16 | 9.44 | 2.43 | others | brain |
| FYTTD1 | Q91Z49 | 10.19 | 1.58 | others | brain |
| GATAD2B | Q8VHR5 | 10.27 | 2.27 | transcription regulation | brain |
| GTF2I | Q9ESZ8 | 10.37 | 1.67 | transcription regulation | brain |
| HDGFL3 | Q9JMG7 | 10.64 | 3.46 | others | brain |
| KHDRBS3 | Q9R226 | 11.08 | 2.15 | RNA processing & translation | brain |
| MECP2 | Q9Z2D6-2 | 11.36 | 2.97 | chromatin | brain |
| MLF2 | Q99KX1 | 10.41 | 1.57 | transcription regulation | brain |
| MTDH | Q80WJ7 | 9.93 | 2.77 | transcription regulation | brain |
| NACC1 | Q7TSZ8 | 10.04 | 1.57 | transcription regulation | brain |
| NOMO1 | Q6GQT9 | 10.34 | 2.13 | membrane associated | brain |
| PSIP1 | Q99JF8 | 11.26 | 2.77 | transcription regulation | brain |

| protein name | Protein ID | $\log_{10}$ intensity | $\log_2$ ratio BC/LC | protein function | enrichment category |
|---|---|---|---|---|---|
| PSPC1 | Q8R326 | 11.56 | 3.22 | transcription regulation | brain |
| RNF14 | Q9JI90 | 10.04 | 2.05 | others | brain |
| RPL10 | Q6ZWV3 | 10.87 | 1.61 | RNA processing & translation | brain |
| RPL13 | P47963 | 10.81 | 1.63 | translation regulation | brain |
| RPS4X | P62702 | 11.07 | 1.42 | RNA processing & translation | brain |
| SERBP1 | Q9CY58-2 | 10.47 | 1.37 | RNA processing & translation | brain |
| SUGP2 | Q8CH09 | 10.95 | 4.79 | RNA processing & translation | brain |
| SUN1 | Q9D666-3 | 10.18 | 2.81 | membrane associated | brain |
| VAPA | Q9WV55 | 10.34 | 3.11 | membrane associated | brain |
| ZMYM3 | Q9JLM4 | 9.35 | 2.03 | transcription regulation | brain |
| ARGLU1 | Q3UL36 | 11.01 | -1.57 | others | liver |
| CACTIN | Q9CS00 | 9.77 | -1.83 | others | liver |
| DDX21 | Q9JIK5 | 10.86 | -3.08 | transcription regulation | liver |
| DEK | Q7TNV0 | 11.05 | -2.12 | chromatin | liver |
| DKC1 | Q9ESX5 | 10.57 | -2.33 | RNA processing & translation | liver |
| H1.4 | P43274 | 11.12 | -1.64 | histones | liver |
| H2A.Z | P0C0S6 | 11.49 | -1.38 | histones | liver |
| IK | Q9Z1M8 | 10.86 | -1.35 | RNA processing & translation | liver |
| LENG8 | D3YWS8 | 9.27 | -2.15 | others | liver |
| LIG3 | P97386 | 9.98 | -1.49 | others | liver |
| MYBBP1A | Q7TPV4 | 11.16 | -1.83 | transcription regulation | liver |
| MYH9 | Q8VDD5 | 10.18 | -1.75 | others | liver |
| NFIX | E9PUH7 | 10.48 | -1.37 | transcription repression | liver |
| NMNAT1 | Q9EPA7 | 10.00 | -2.20 | others | liver |
| NOP56 | Q9D6Z1 | 10.90 | -1.79 | RNA processing & translation | liver |
| NOP58 | Q6DFW4 | 10.80 | -1.74 | RNA processing & translation | liver |
| PCMTD2 | B0R0C7 | 10.26 | -2.45 | others | liver |
| PNKP | G5E8N7 | 10.15 | -1.85 | others | liver |
| PUM3 | Q8BKS9 | 10.25 | -1.87 | others | liver |
| RCC1 | Q8VE37 | 10.59 | -1.32 | chromatin | liver |
| RRP12 | Q6P5B0 | 10.02 | -1.34 | RNA processing & translation | liver |
| RSL1D1 | Q8BVY0 | 10.40 | -1.62 | RNA processing & translation | liver |
| SMARCD2 | Q99JR8 | 10.01 | -2.24 | chromatin | liver |
| SNRPA1 | P57784 | 11.14 | -1.38 | RNA processing & translation | liver |
| SON | H9KV00 | 11.47 | -1.36 | RNA processing & translation | liver |
| TFIP11 | Q9ERA6 | 10.60 | -1.32 | RNA processing & translation | liver |
| TOR1AIP1 | Q921T2 | 10.52 | -2.19 | membrane associated | liver |
| ZNF281 | Q99LI5 | 9.95 | -1.30 | transcription regulation | liver |

Filtered protein list obtained from the MaxQuant search against the UniProt database for *Mus musculus*. BC: brain chromocenter; LC: Liver chromocenter. Protein function and enrichment category according to the classification in figure 3.4.

## Supplementary figures

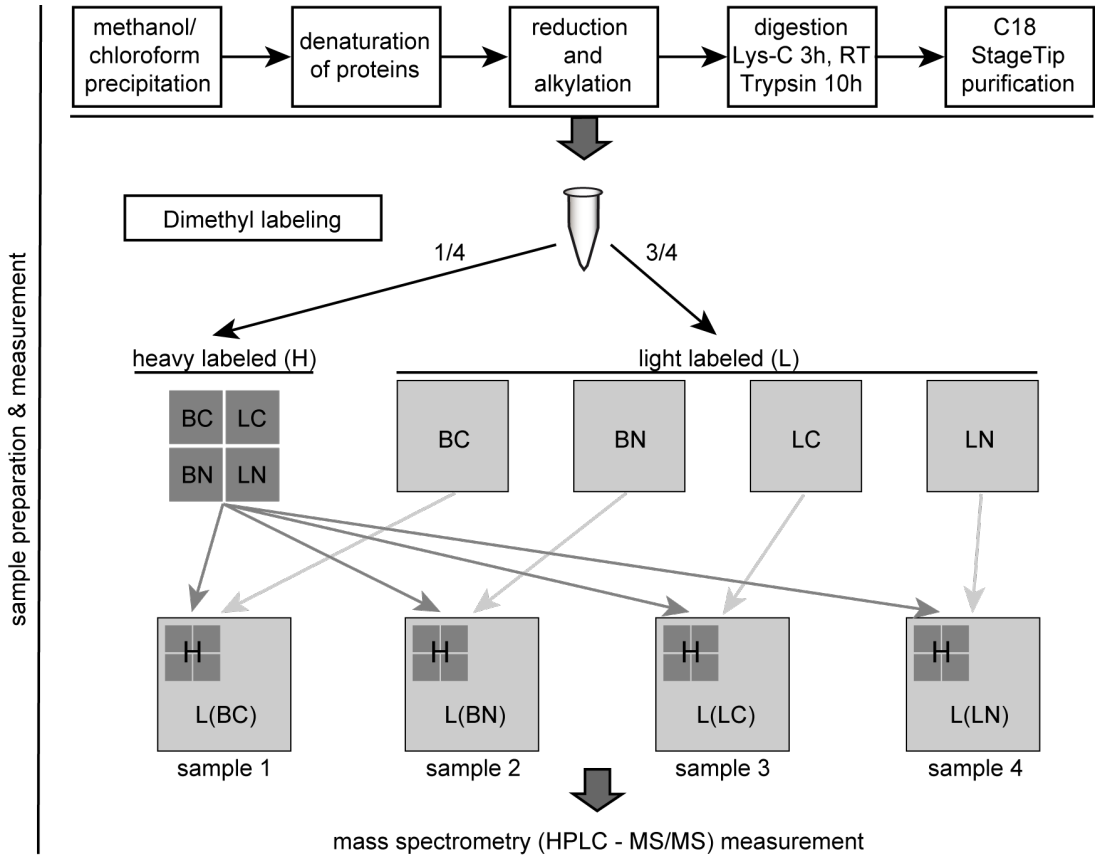



**Figure 3.15: Workflow of the sample preparation and dimethyl labeling for quantitative mass spectrometry analysis.** All steps described here were performed by our cooperation partners at MDC Berlin (Oliver Popp, Gunnar Dittmar). (A) Sample preparation for mass spectrometry comprising protein precipitation, denaturation, reduction & alkylation, digestion with Lys-C and trypsin, and C18 StageTip purification. (B) Scheme of the dimethyl labeling procedure. One-fourth of each sample was pooled into a master mix and the proteins were chemically labeled with a heavy label. The remaining proteins of the individual samples were separately labeled with a light label. One-fourth of the master mix was spiked into each of the light-labeled samples as a reference. BC: brain chromocenters, BN: brain nuclei, LC: liver chromocenters, LN: liver nuclei. Then, all samples were subjected to HPLC-MS/MS measurement.

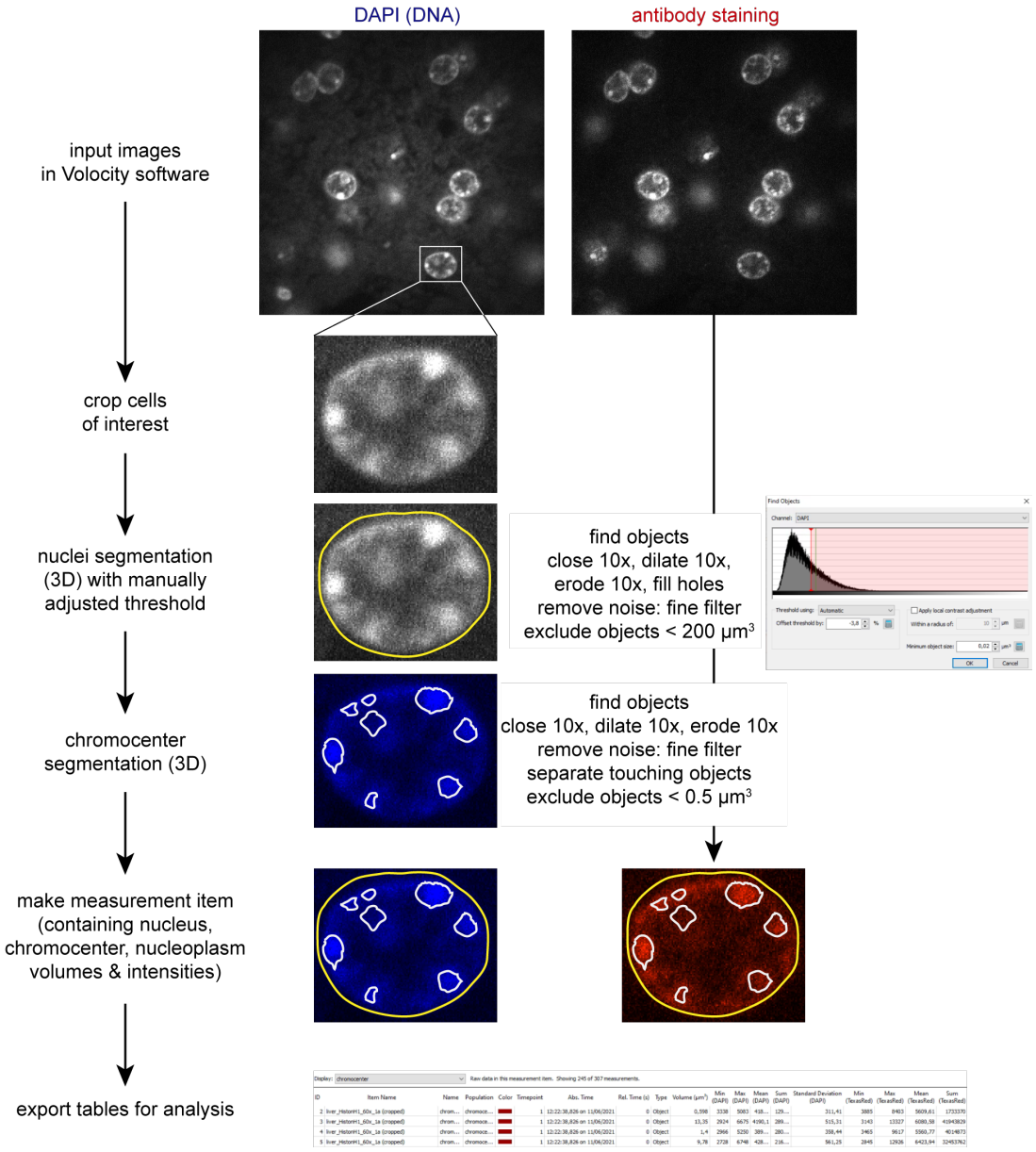



**Figure 3.16: Image analysis workflow for nuclei and chromocenter segmentation on confocal z-stack images.** The confocal z-stacks were obtained on an Ultra-Vox spinning disk system automatically saving the images in the Volocity software (Perkin Elmer). Based on the DAPI channel, individual cells were cropped, nuclei and chromocenters were segmented with the parameters described and a measurement item was created for each cell. The results were exported as tables for further analysis using the R software package.

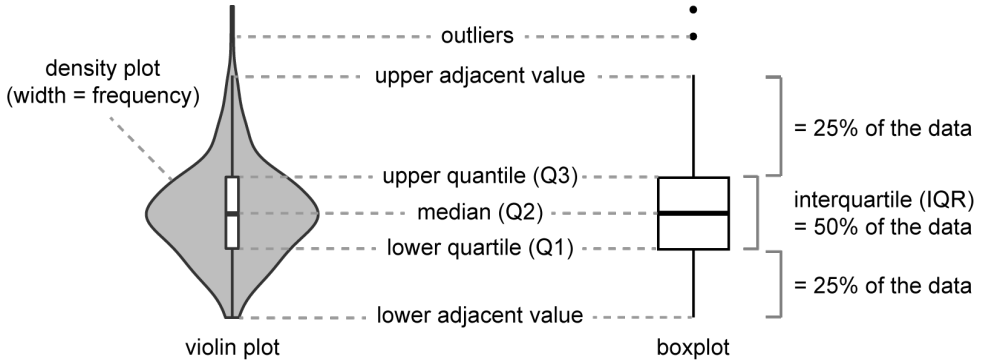



**Figure 3.17: Data visualization in a violin plot.** A violin plot is a combination of a density plot and a box plot. The density plot depicts the frequency of the observations and the boxplot shows the median, upper and lower quantile, upper and lower adjacent value, and outliers. The adjacent values represent the minimum and maximum values excluding outliers.

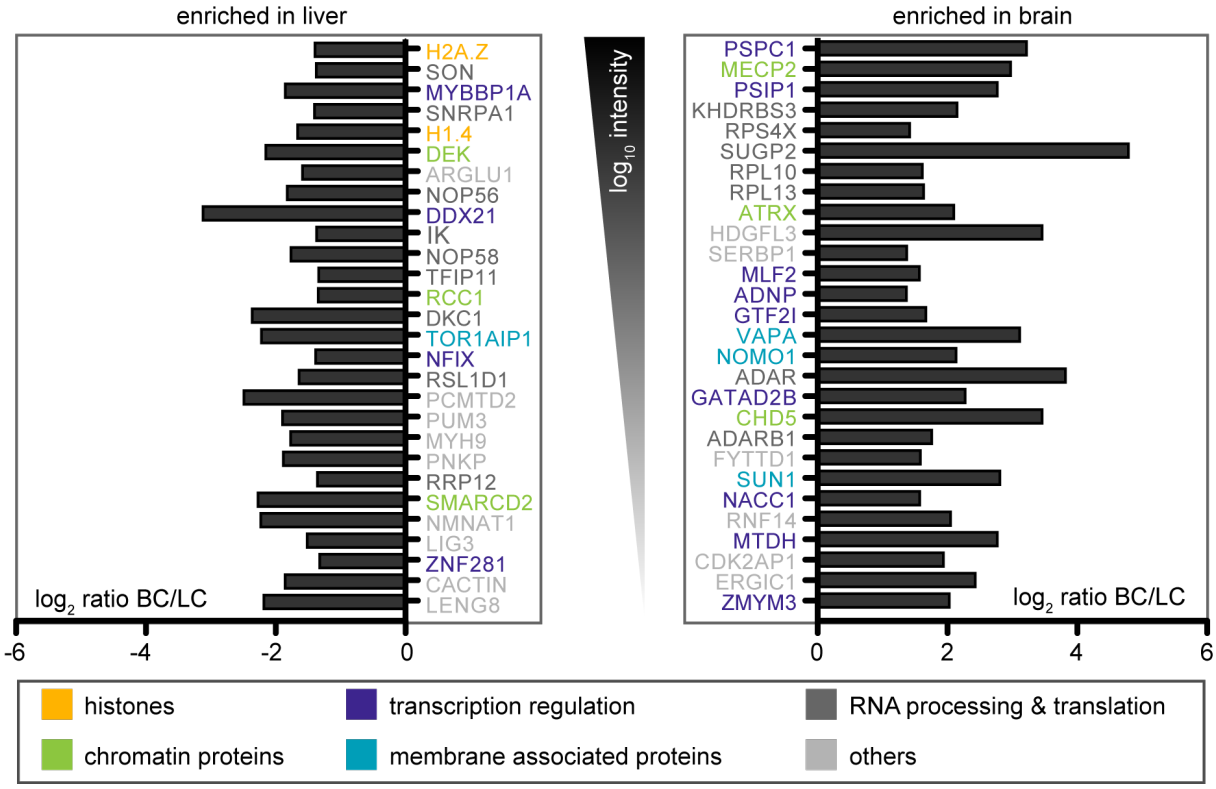



**Figure 3.18: Heterochromatin interacting proteins upregulated in brain or liver identified by quantitative mass spectrometry analysis without cut-off.** (see figure 3.4) Bar diagram showing the $\log_2$ ratio of brain chromocenter (BC) versus liver chromocenter (LC) proteins ordered according to descending $\log_{10}$ intensity. The color code indicates the protein function manually assigned based on the UniProt webpage functional information. Histones are labeled in orange, chromatin proteins in green, proteins involved in transcriptional regulation in purple, membrane associated proteins in cyan, proteins involved in RNA processing or RNA-binding proteins in dark gray, and proteins not fitting into the categories in light gray.

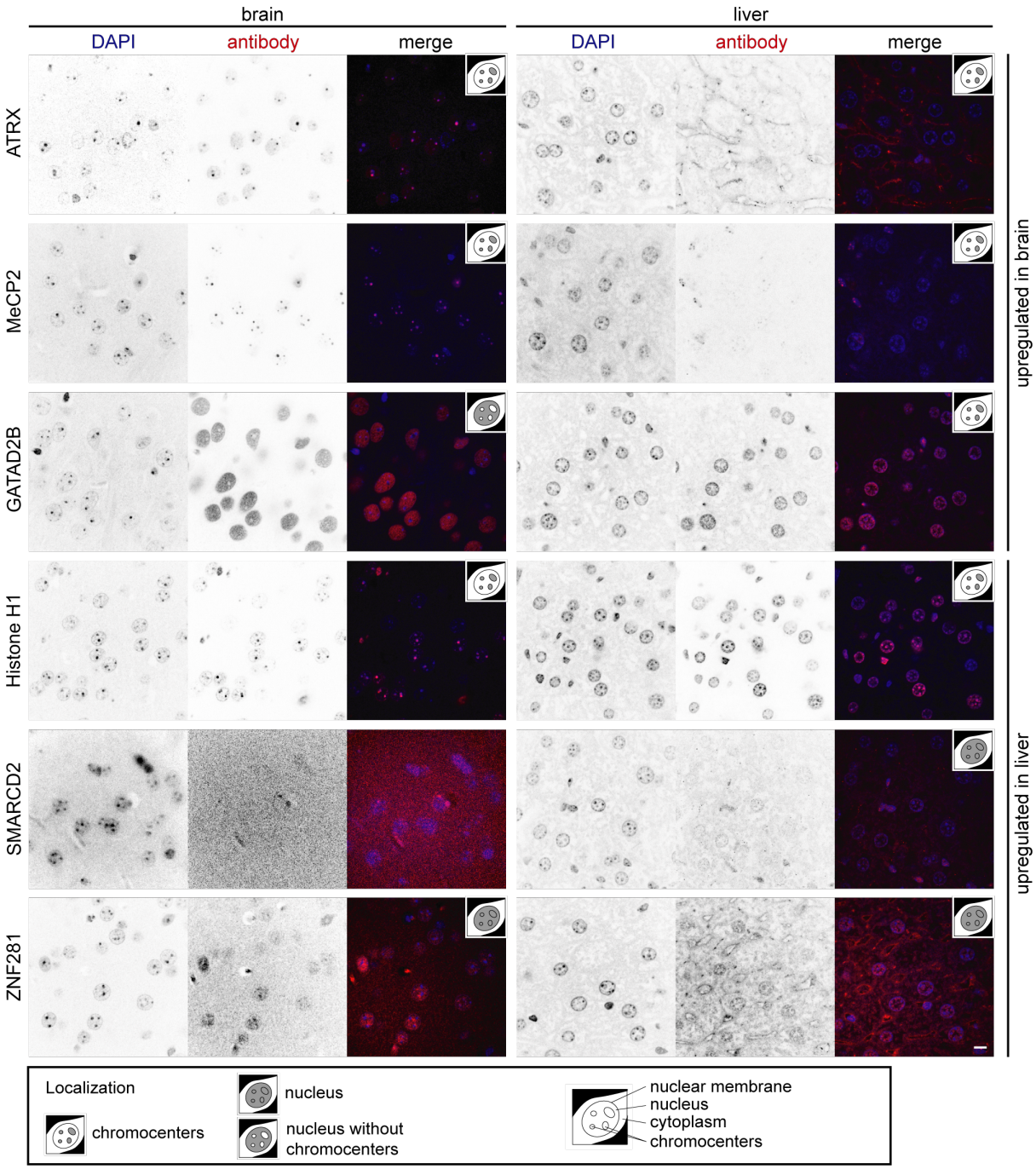



**Figure 3.19: Overview images of the validation stainings of heterochromatin enriched proteins with different abundance in brain and liver tissue.** The montages show the subcellular localization of the proteins in mouse brain and liver tissue with the DAPI DNA staining in blue, and the antibody staining in red. The schemes represent the subcellular localization described in the legend. Scale bar 5 µm.

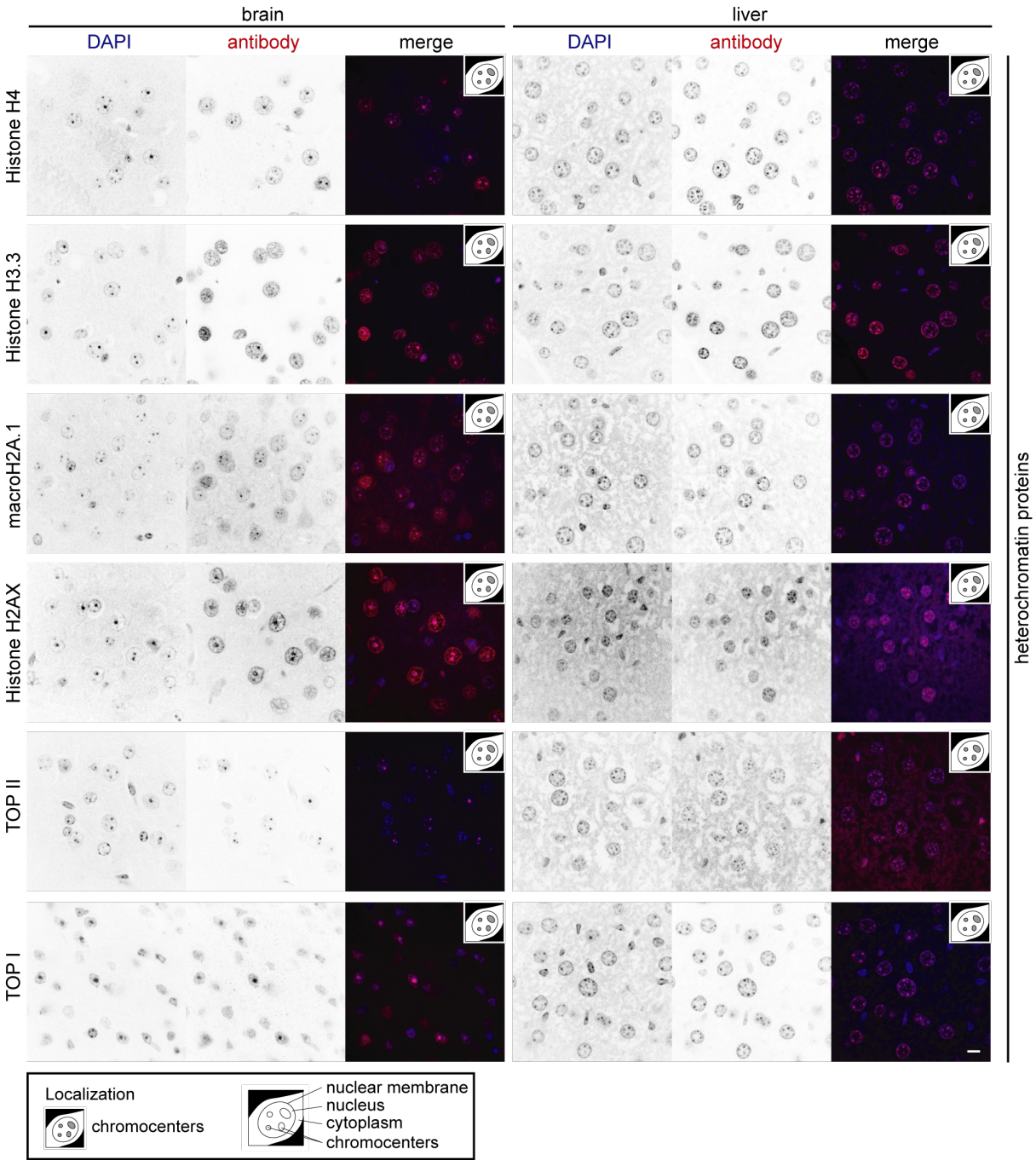



**Figure 3.20: Overview images of the validation stainings of heterochromatin enriched proteins with similar abundance in brain and liver tissue showing chromocenter localization.** The montages show the subcellular localization of the proteins in mouse brain and liver tissue with the DAPI DNA staining in blue and the antibody staining in red. The schemes represent the subcellular localization described in the legend. Scale bar 5 µm.

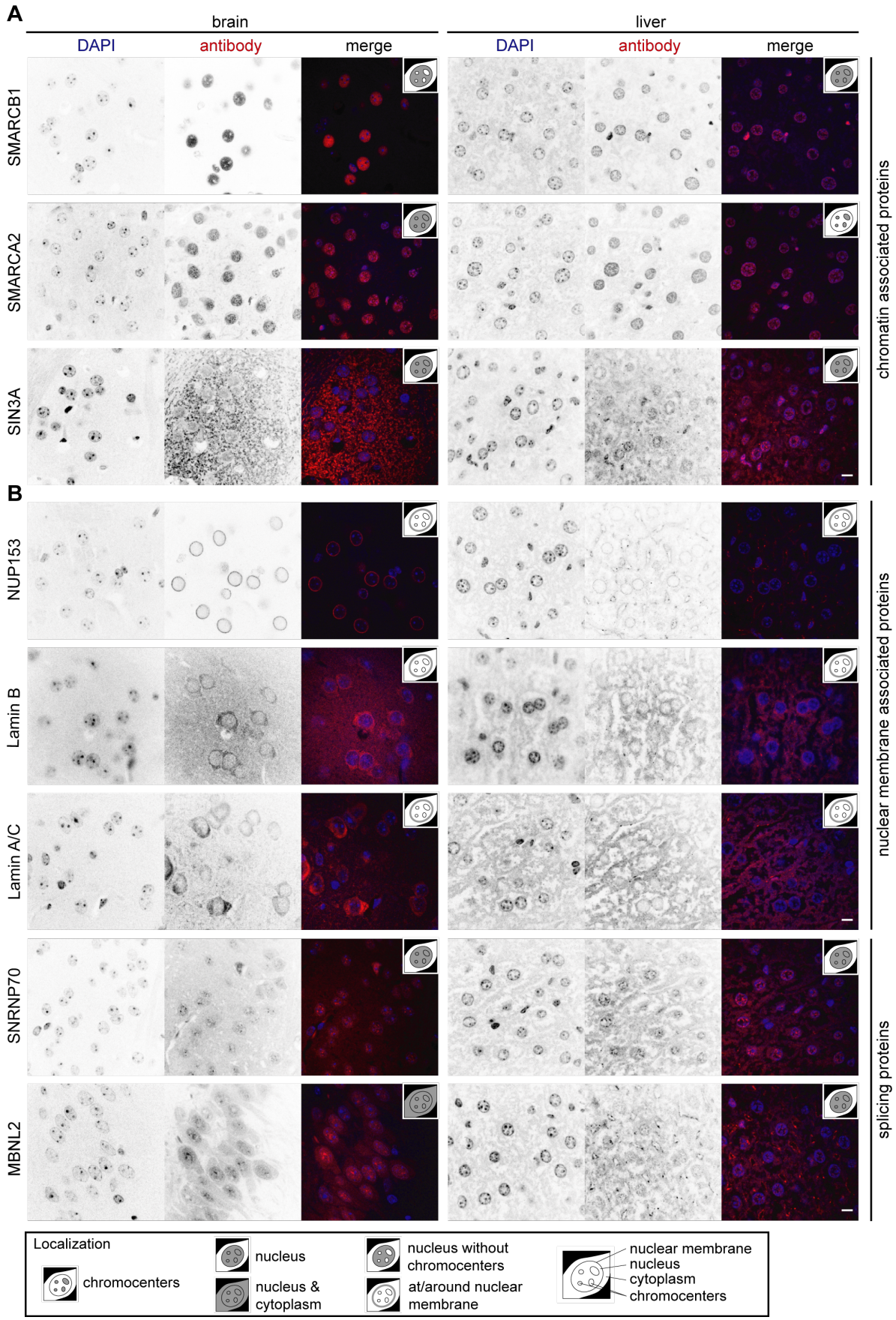



**Figure 3.21: Overview images of the validation stainings of candidate heterochromatin**

**Figure 3.21 (*previous page*): interacting proteins with similar abundance in brain and liver tissue not localizing at the chromocenters.** The montages show the subcellular localization of the proteins in mouse brain and liver tissue with the DAPI DNA staining in blue, and the antibody staining in red. The schemes represent the subcellular localization described in the legend. Scale bar 5 µm.

# 4 Quantification of MeCP2 in the cell nucleus and heterochromatin compartments

The validation and calibration of a cellular system mimicking MeCP2 *in vivo* levels described in this chapter is modified from Zhang & Romero, Schmidt *et al.*, 2022 [1].

## 4.1 Introduction

Heterochromatin, the more silent and transcriptionally inactive form of chromatin, is characterized by high DNA compaction and cytosine methylation levels. Methylated cytosines are recognized by a family of methyl-CpG binding proteins such as the Methyl-CpG binding protein 2 (MeCP2). MeCP2 was shown to bind to a single symmetrically methylated CpG pair [115, 130] and to be involved in transcriptional regulation [125, 126, 131, 139] and chromatin organization (reviewed in [3]).

Ectopic expression of MeCP2 was reported to induce dose-dependent heterochromatin clustering [55] and MeCP2 protein levels were shown to increase during differentiation [55, 233, 313, 314]. Immunofluorescence stainings revealed tissue- and cell type-specific differences in MeCP2 protein levels [233, 313], but no correlation between protein and RNA levels was observed [313]. MeCP2 was found to be present in high levels in most neurons, but absent from microglial cells [233, 313]. The absolute quantification of MeCP2 in neuronal and glial cell nuclei separated by fluorescence-activated cell sorting (FACS) revealed that MeCP2 molecules in neurons are nearly as abundant as nucleosomes [138]. Thus, the authors proposed that MeCP2 might rather act as a global chromatin regulator dampening transcription in a genome-wide manner than acting as a regulator of specific genes [138].

Mutations on the *MECP2* gene located on the X chromosome were identified as the main cause of Rett syndrome, a human neurological disorder [116]. In addition, abnormal levels of MeCP2 were linked to Rett syndrome-like phenotypes. In fact, abnormal neurological phenotypes were reported for *Mecp2* deficient mice [244, 245], mice with a 50% *Mecp2* deficiency [328], in mice with mild overexpression of *Mecp2* [246] and for duplication of *MECP2* in male humans [247]. Interestingly, phenotypes originated from reduced MeCP2 levels could be rescued by targeted expression of *MECP2* [248–252].

As already small changes in MeCP2 levels can cause neurological symptoms in mice and humans, it is important to carefully adjust the MeCP2 level in experiments performed *in vitro* or *in cellulo*. Only the tight control of all experimental conditions allows a comparison to the physiological situation and thus, meaningful results. For this reason, we aimed to develop a system to compare protein levels between different experimental setups. We developed a method to quantify MeCP2 in mouse brain tissue (*in vivo*) and in (transfected) mouse myoblast cells (*in cellulo*). Then, we utilized our method to compare the levels between the two systems and calculated the protein concentration in the nucleus and its subcompartment heterochromatin to be applied in *in vitro* experiments. This method can be extrapolated and applied to any other system combinations.

## 4.2 Material & methods

### 4.2.1 Protein purification from insect cells

The protein production and purification from SF9 cells (Invitrogen) was performed as described before [164]. In brief, Sf9 cells (table 4.2) were cultivated in EX-CELL 420 Serum-Free Medium (#24420G, Sigma-Aldrich) supplemented with 10% fetal calf serum under shaking at 27 °C. For the recombinant baculovirus production, the SF9 cells were transfected using polyethylenimine (#23966, Polysciences) following the manufacturer's instructions. Baculovirus (P1 stock) in the medium was collected by centrifugation and the supernatant containing the virus was used to infect new SF9 cells to get higher titers of baculovirus (P2). This step was repeated once (P3) before new SF9 cells were infected with the virus (P3) and incubated at 27 °C for 4 days for protein production. Cell pellets containing MeCP2-GFP were collected by centrifugation at 1200 rpm for 5 min and stored at -80 °C. The cells were thawed on ice, resuspended in cold lysis buffer (0.025 M Tris HCl (pH 8), 1 M NaCl, 0.05 M glucose, 0.01 M EDTA, 0.2% Tween 20, 0.2% NP-40 substitutive and protease inhibitors (1 mM phenylmethylsulfonyl fluoride (PMSF) (#6367.1, Carl Roth), 10 µm E64 (#E3132, Sigma-Aldrich), 1 µM pepstatin A (#P5318, Sigma-Aldrich)) and sheared by syringe treatment. The lysate was cleared by centrifugation and loaded onto GFP-binding protein coupled to beads [329]. After binding, the beads were washed with lysis buffer and the proteins were eluted using 4 M $MgCl_2$. The buffer was exchanged to PBS and the proteins were stored frozen until use.

### 4.2.2 Nuclei isolation from mouse brain tissue

Three-month-old C57BL/6 mice (Charles River Laboratories, Inc.) were sacrificed and dissected and the organs were collected, washed with PBS, and frozen in liquid nitrogen. For nuclei isolation the frozen mouse brains were crushed to powder and homogenized in 0.25 M sucrose solution (20 mM triethanolamine-HCl (pH 7.6), 30 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 1 mM PMSF). After centrifugation for 10 min at 1000 xg, the supernatant was discarded and the pellet resuspended in sucrose buffer to a final sucrose concentration of 2.1 M. The raw nuclei fraction was obtained by centrifugation for 30 min at 50000 xg. The pellet was resuspended in 0.25 M sucrose solution and centrifuged at 1000 xg. Nuclei were resuspended in PBS and counted using a microscope. Subsequently, the nuclei were lysed by resuspension in Laemmli buffer (2% SDS, 50 mM Tris (pH 6.8), 10% glycerol, 0.01% bromophenol blue, 100 mM DTT) and incubated at 95 °C for 10 min.

### 4.2.3 Mammalian cell culture and transfection

All cell lines were free of mycoplasma contamination and are listed in table 4.2. C2C12 mouse myoblast cells were cultured in Dulbecco's modified Eagle Medium (DMEM) high glucose (#D6429, Sigma-Aldrich) supplemented with 20% fetal bovine serum (FCS) (#F7524, Sigma-Aldrich), 1x L-glutamine (#G7513, Sigma-Aldrich), and 1 µM gentamicin (#G1397, Sigma-Aldrich). The transfection was performed using Neon Transfection System (Thermo Fisher Scientific) accord-

ing to the manufacturer's instructions. Cells were harvested for Western blot / FACS sorting or fixed for immunofluorescence staining 20 h after transfection if not stated differently.

### 4.2.4 Flow cytometry

C2C12 myoblasts transiently transfected with pMeCP2G (pc1121) were harvested 20 h after transfection, resuspended in PBS, and separated according to their transfection level by fluorescence-activated cell sorting (FACS) on the S3 Cell Sorter (Bio-Rad Laboratories) with a 488 nm laser and a $525 \pm 30$ nm emission filter. Cells were plotted for $\log_{10}$ GFP sum intensity and divided into 40 bins. The first 11 bins were defined as negative cells (background autofluorescence) by comparison to the untransfected C2C12 control cells. Cells in bins 13 to 21 were defined as low expressing, cells in bins 24 to 32 as high expressing (plots with the gates for negative, low, and high transfected cells are shown in figure 4.1). Low and high MeCP2 expressing cells were sorted and collected as pellets for Western blot.

### 4.2.5 Western blot

Cell pellets collected from FACS (for low and high *Mecp2* expressing cells) or directly from culture (for untransfected cells) were lysed in lysis buffer (see 4.2.1), mechanically disrupted, and boiled at 95 °C in Laemmli buffer (see 4.2.2). Samples were run on 8% SDS-PAGE and transferred to a nitrocellulose membrane. After blocking with 5% low-fat milk in PBS the membranes were incubated with anti-MeCP2 rat monoclonal antibodies 4H7 and 4G10 cell culture supernatants from hybridoma cell lines overnight followed by anti-rat IgG Cy3 secondary antibody diluted 1:1000 for 1 h (Table 4.3). Fluorescent signals were detected using an Amersham Imager (Table 4.4). The bands on the Western blots were quantified using the software ImageJ [270] (https://imagej.nih.gov/ij/) by selecting single lanes, plotting intensities along the lanes, selecting the peaks corresponding to the bands of interest and measuring the bands' relative intensity (area under the peak minus background signal).

### 4.2.6 Immunofluorescence staining on cells

For immunostaining, cells were grown on gelatin-coated coverslips and fixed with ice-cold methanol for 6 min. After washing, cells were permeabilized with 0.5% Triton X-100, washed with PBST (0.01% Tween 20), blocked with 0.1% fish skin gelatin for 20 min and incubated with primary antibody anti-MeCP2 rabbit polyclonal (Table 4.3) 1:250 for 2 h. After washing with PBST (0.1% Tween 20), cells were incubated 1 h with secondary antibody anti-rabbit IgG Cy5 (Table 4.3) 1:400, followed by another washing step with PBST (0.1% Tween 20). Samples were counterstained with 1 µg/ml DAPI and mounted in Mowiol.

### 4.2.7 Immunofluorescence staining on tissues

Three-month-old C57BL/6 mice (Charles River Laboratories, Inc.) were sacrificed and dissected and the organs were collected. After washing with PBS, tissues were fixed in 10% buffered

formalin solution (#HT501128, Sigma-Aldrich) for 24 h. Afterwards, they were sequentially dehydrated in 70% ethanol for 30 min, 70% ethanol for 45 min, 96% ethanol for 60 min, and 96% ethanol for 45 min, twice in absolute ethanol for 45 min followed by xylol for 60 min and 30 min. Organs were embedded in paraffin, transferred to embedding cassettes, cooled down, and sliced on a microtome into 6 µm slices.
For paraffin removal, the slides were incubated 2 h at 60 °C and subsequently incubated three times 5 min in xylol. The tissue slices were rehydrated by sequential incubation for 5 min each in 96% ethanol, 90% ethanol, 80% ethanol, 70% ethanol, and three times in water. For antigen unmasking the slides were treated with sodium citrate buffer (10 mM sodium citrate, pH 6) at 100 °C or 30 min in an autoclave. The samples were equilibrated for 15 min in PBS, permeabilized with 0.7% Triton X-100 in PBS two times for 15 min and washed three times with PBS. The tissue slices were surrounded with a liquid-blocking pen and blocked with 4% BSA for 30 min. Incubation with primary antibodies (Table 4.3) was performed overnight at 4 °C, followed by three times washing with PBST (0.1% Tween 20 in PBS) for 10 min and secondary antibody incubation for 1 h at room temperature. Samples were washed three times with PBST for 10 min, counterstained with 1 µg/ml DAPI, washed with PBS, and water, and mounted using Mowiol.

### 4.2.8 Imaging and image analysis

The immunofluorescence stainings of fixed C2C12 cells were imaged using an Axiovert 200 microscope (Table 4.4) with 63x Plan-Apochromat 1.4 NA oil immersion objective. Segmentation of nuclei, heterochromatin, and intensity measurements were performed in the software ImageJ 2.0 [270] (https://imagej.nih.gov/ij/). Heterochromatin segmentation was based on the normalization of pixel intensities to a local maximum intensity. First, individual pixel intensities were calculated in squares of 30 x 30 pixels, being the maximum of these 900 pixels the local maximum. To avoid "dark squares" becoming thresholded, only the pixels with intensities higher than 8/42 of the cell maximum (higher local maximum on the cell) were considered for further steps. Thresholding was applied by giving a value of 1 (white) to pixels with intensities higher than 21/42 (for heterochromatin compartments) or 37/42 (for heterochromatin core) of the local maximum intensity. These thresholded images were used to generate individual and total heterochromatin (core) ROIs, that were subsequently subtracted to generate the nucleoplasm ROI for each nucleus (script published in [1]). To categorize the cells equivalently to the cells used in Western blot, we measured the total intensity of the nucleus and applied an equivalent binning system for the FACS sorting.
Additional imaging of the same samples was performed on a confocal microscope Leica TCS SPE-II equipped with a 63x/1.30 ACS APO Oil CS objective (Table 4.4) as z-stacks with 0.2 µm interval. For this subset of images, 3D segmentation of nuclei and heterochromatin compartments was performed using the Volocity software. Briefly, nuclei segmentation was done based on the DAPI channel by finding objects, dilating three times, filling holes, eroding three times, and choosing objects by size.
The stained tissue slices for quantification of MeCP2 *in vivo* were imaged on an Ultra-View VoX spinning disk on an inverted Nikon Ti-E microscope in 3 µm z-stacks. The nuclear volumes were

obtained by 3D segmentation with manual thresholding using the Volocity software. Briefly, nuclei were segmented based on the DAPI channel including the following steps: find objects, close 10x, dilate 10x, erode 10x, fill holes, remove noise with a fine filter and exclude objects $< 200\,\mu m^3$. To obtain the ratios of nuclear subcompartments to the nucleus intensities, the mean intensity values for MeCP2 were measured in the nucleus and three manually selected ROIs in the heterochromatin and the nucleoplasm on 2D confocal images.

### 4.2.9 Quantification of MeCP2 concentrations

To calculate the MeCP2 nuclear concentration, we used the number of molecules obtained from Western blot and the volumes of the nuclei obtained from the confocal images (see figure 4.3A). These concentration values together with the ratios of mean intensity in the compartments in comparison to the whole nucleus were used to calculate the local MeCP2 concentrations in the nuclear subcompartments (see figure 4.3B). Although the intensity ratios for the mouse myoblast cells obtained from compartment volumes and compartment areas were comparable, the information used here derives from the areas as the sample number was larger. For the quantification in mouse brain tissue, the ratios obtained by manual ROI selection on confocal images were used.

## 4.3 Results

As abnormal MeCP2 levels were reported to cause neuronal disfunction and Rett syndrome-like phenotypes [244–247, 328], we aimed to develop a procedure to calibrate the MeCP2 level in individual cells and subcellular compartments. Applying the same MeCP2 protein amount in different experimental systems would allow a comparison of *in cellulo* and *in vitro* results to physiological conditions *in vivo*.

For the quantification of MeCP2 in a cellular system, mouse myoblast cells were transfected with a plasmid coding for MeCP2-GFP and sorted according to their nuclear GFP intensity 20 h after transfection. The cells were separated into three levels: negative (untransfected), low and high MeCP2 (figure 4.1A, D). As myoblast cells have very low endogenous MeCP2 protein levels, they are a suitable system to manipulate the levels of MeCP2 and analyze the effects [55]. For absolute MeCP2 quantification, the sorted cells were collected for Western blot. The concentration of MeCP2-GFP purified from insect cells was determined in comparison to a BSA standard calibration series on an SDS-PAGE stained with Coomassie (figure 4.1C). Subsequently, the MeCP2-GFP standard was used to quantify the protein level of the sorted cells by Western blot using MeCP2-specific antibodies and fluorescence-coupled secondary antibodies (figure 4.1D). The number of MeCP2 molecules per nucleus in low-level expressing cells was five times higher than in untransfected cells, while the number of MeCP2 molecules per nucleus in high-expressing cells was ten times higher than in low-expressing cells.

Next, we wanted to compare the absolute MeCP2 levels obtained from FACS sorting and Western blot quantification to MeCP2 in cells observed by microscopy (figure 4.1D). Therefore, mouse myoblast cells transfected with MeCP2-GFP were fixed 20 h after transfection, stained against MeCP2, and counterstained with the DNA dye DAPI. The cells were imaged by widefield fluorescence microscopy and nuclei, heterochromatin fractions as well as its core regions were segmented based on the DAPI channel. The same binning system as for the FACS sorting was applied using the full histogram of nuclear GFP intensity of the cells (figure 4.1D). The MeCP2 mean heterochromatin intensity per area correlated with the classification bins (figure 4.1B). In addition, the average heterochromatin cluster area per nucleus increased with higher MeCP2 levels (figure 4.1B), in line with previous studies [55].

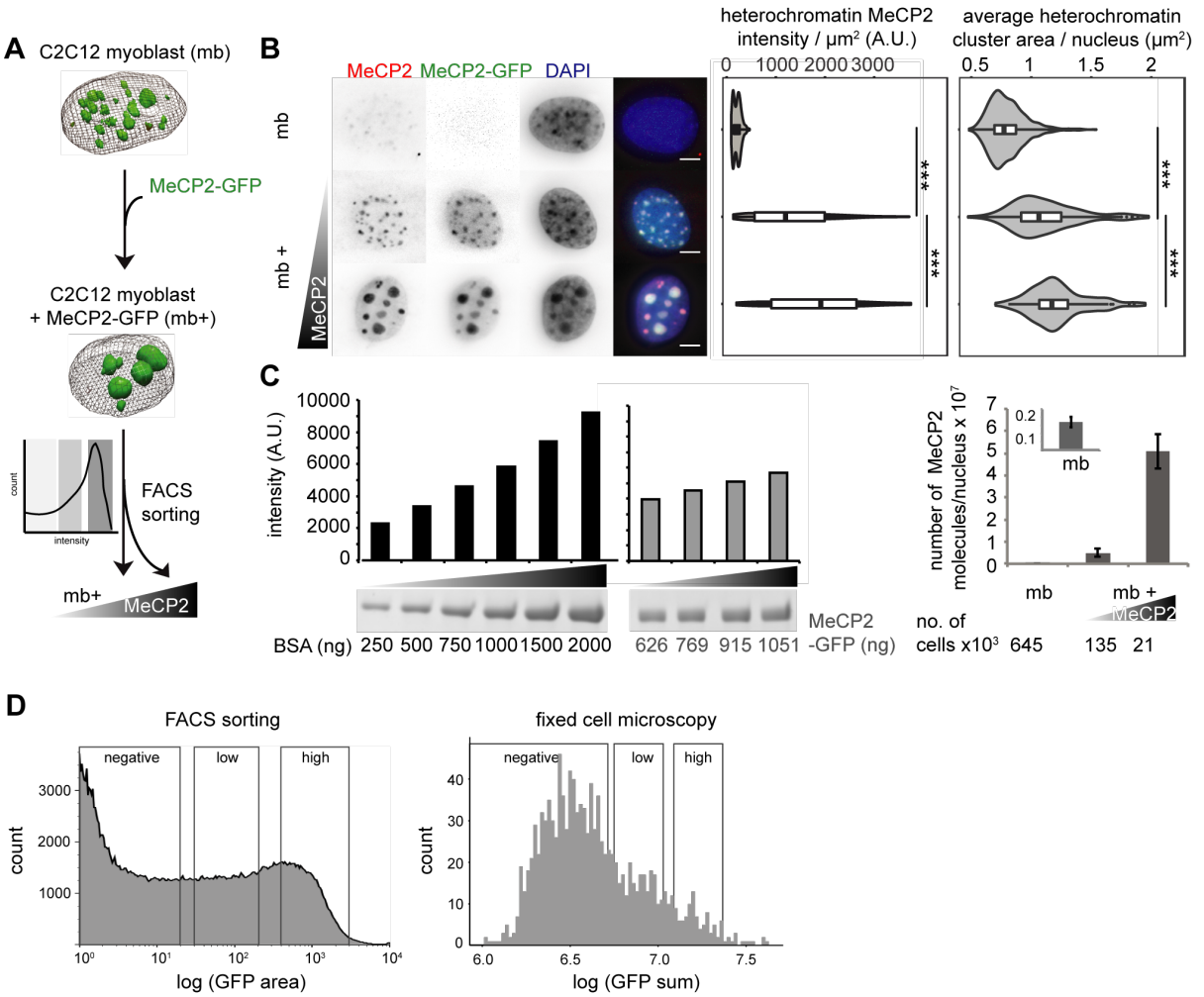



**Figure 4.1: Validation and calibration of a cellular system mimicking *in vivo* MeCP2 physiological behavior.** (A) Scheme of the experiment: C2C12 myoblast (mb-) were transfected with a plasmid encoding for MeCP2-GFP. After 20 h, transfected cells (mb+) were sorted into two categories, low and high expressing, according to the GFP intensity using fluorescence-activated cell sorting (FACS). (B) Immunofluorescence staining showing MeCP2 levels in mouse myoblasts before and after transfection of MeCP2-GFP. Cells were counterstained with DAPI. Scale bars 5 µm. Boxplots show the MeCP2 heterochromatin mean intensity and the mean heterochromatin cluster area of untransfected, low, and high MeCP2 expressing myoblasts of three independent replicates (***$p < 0.001$, Wilcoxon test). Statistics can be found in Table 4.5. (C) Quantification of total MeCP2 in mouse myoblasts. The concentration of the MeCP2-GFP standard was determined by SDS-PAGE and Coomassie staining in comparison to a BSA standard series. The MeCP2 standard was used to quantify the number of MeCP2 proteins in untransfected, low and high expressing FACS sorted mouse myoblasts by Western blot against MeCP2 for three independent replicates (average values ± standard deviation). Full gels and blots are shown in figure 4.4. (D) Instrument-independent calibration of the quantification of MeCP2. To allow comparison of intensities between instruments, we classified the cells using a population curve after transfection plotting counts versus the log of the GFP intensity, considering three windows: negative, defined from the measurements obtained from untransfected cells, and then dividing the positively transfected cells in 29 bins, in which 2-10 correspond to low MeCP2 levels and 13-21 to high MeCP2 levels. The figure was modified from Zhang & Romero, Schmidt *et al.*, 2022 [1].

The quantification of MeCP2 in mouse brain tissue was performed in a similar way as the quantification in mouse myoblast cells. First, nuclei were isolated from mouse brain tissue and lysed in Laemmli buffer. The MeCP2-GFP standard row and three biological replicates of mouse brain nuclei lysates were analyzed by SDS-PAGE and transferred to a membrane. MeCP2 protein levels were detected using MeCP2 specific antibodies in combination with fluorescently tagged secondary antibodies (figure 4.2B). The number of MeCP2 molecules per cell was within the same range of low-level MeCP2 expressing myoblast cells (figure 4.2C). The MeCP2 staining on mouse brain tissue revealed intense MeCP2 staining at the heterochromatin clusters and showed a tendency for less and bigger heterochromatin clusters compared to low expressing myoblast cells (figure 4.2A). The absolute quantification on Western blot was performed using full brain protein lysates, thus considering all brain regions and cell types, whereas the staining shows only one exemplary cell.

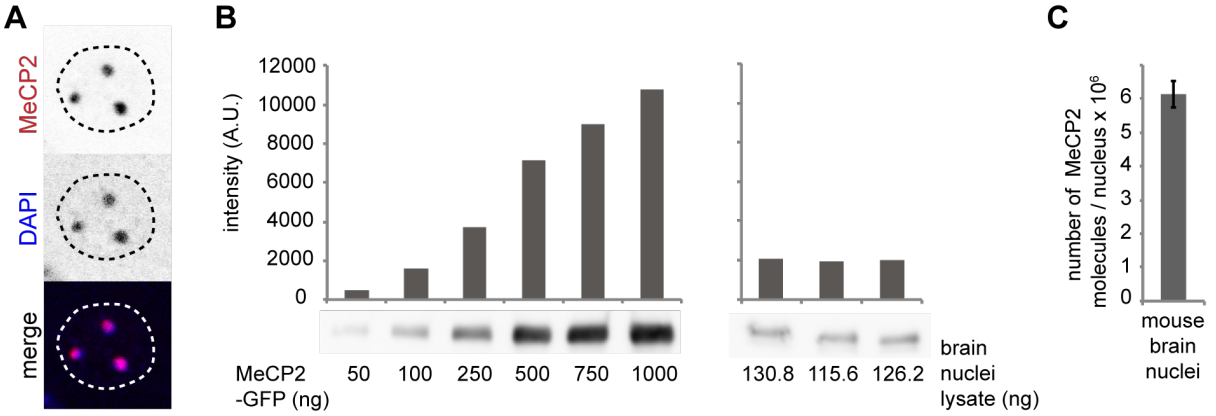
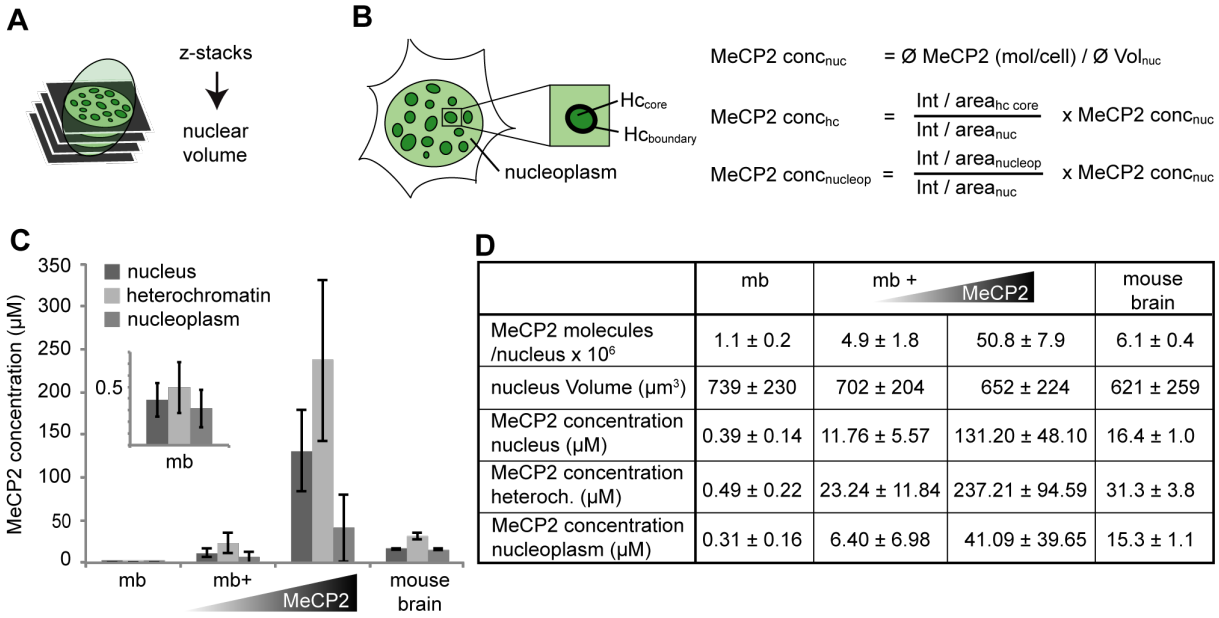



**Figure 4.2: Quantification of MeCP2 levels in mouse brain (*in vivo*).** (A) Immunofluorescence staining of MeCP2 on mouse brain tissue, the DNA was counterstained with DAPI. The dotted line indicates the nuclear outline. (B) Quantification of total MeCP2 in mouse brain. The concentration of the MeCP2-GFP standard was determined by SDS-PAGE and Coomassie staining in comparison to a BSA standard series before (see figure 4.2C). The plots show the relative intensities of the Western blot bands of the MeCP2-GFP standard row and three biological replicates of mouse brain nuclei used for quantification. The protein amount in ng is indicated under the Western blot images. Full Western blots are shown in figure 4.5. (C) The number of MeCP2 molecules per nucleus in mouse brain tissue was plotted as an average of the three biological replicates with standard deviation.

To compare the MeCP2 amount determined for cells and brain tissue to levels used in *in vitro* experiments, it was necessary to calculate MeCP2 concentrations. Therefore z-stack images of fixed cells were taken, and the nuclear volumes were calculated based on nucleus segmentation on the DAPI channel (figure 4.3A). The nucleus MeCP2 concentration was calculated from the number of molecules and the nuclear volume. We obtained a concentration of 0.39 µM in untransfected myoblast cells, 11.76 µM in low and 131.20 µM in high-expressing cells, and 16.4 µM in mouse brain nuclei. Thus, the MeCP2 concentration in mouse brain nuclei was only slightly higher than in low-expressing myoblast cells. The protein concentrations in the nuclear subcompartments heterochromatin and nucleoplasm were calculated from the nucleus MeCP2 concentration and the mean intensity ratios of subcompartments per nucleus (figure 4.3B-D). As MeCP2 ac-

cumulates at heterochromatin, its concentration at heterochromatin clusters was higher than in the nucleus. While untransfected myoblasts showed a heterochromatin MeCP2 concentration of 0.49 µM, low expressing cells and mouse brain nuclei showed a higher concentration of 23.24 µM and 31.3 µM, whereas high expressing myoblasts had a concentration of 237.21 µM.

*In vivo*, the MeCP2 concentration in heterochromatin was two times higher than the nucleus concentration using the imaging conditions described. Thus, it should be considered for every experiment whether the nucleus or the heterochromatin protein concentration mimics the physiological conditions. Concluding, we showed that our method can be used to ensure similar protein level conditions across different experimental systems, thus making them comparable between each other and to the physiological conditions.




| | mb | mb + MeCP2 | | mouse brain |
|---|---|---|---|---|
| MeCP2 molecules /nucleus x $10^6$ | 1.1 ± 0.2 | 4.9 ± 1.8 | 50.8 ± 7.9 | 6.1 ± 0.4 |
| nucleus Volume (µm³) | 739 ± 230 | 702 ± 204 | 652 ± 224 | 621 ± 259 |
| MeCP2 concentration nucleus (µM) | 0.39 ± 0.14 | 11.76 ± 5.57 | 131.20 ± 48.10 | 16.4 ± 1.0 |
| MeCP2 concentration heteroch. (µM) | 0.49 ± 0.22 | 23.24 ± 11.84 | 237.21 ± 94.59 | 31.3 ± 3.8 |
| MeCP2 concentration nucleoplasm (µM) | 0.31 ± 0.16 | 6.40 ± 6.98 | 41.09 ± 39.65 | 15.3 ± 1.1 |

**Figure 4.3: Absolute MeCP2 quantification *in cellulo* (mouse myoblasts) and *in vivo* (mouse brain tissue).** (A) Scheme of the calculation of the nuclear volume based on thresholded z-stacks. (B) Scheme of the cell segmentation and the calculations of the MeCP2 concentration in the nucleus and its subcompartments heterochromatin and nucleoplasm. Nucleus concentration ($conc_nuc$) was obtained from the total protein amount per nucleus obtained from Western blot analysis (Figure 4.1C and 4.2BC) divided by the nucleus volume. Subcompartment concentrations ($conc_{hc}$ or $conc_{nucleop}$) were calculated from the intensity ratios between the correspondent subcompartments (heterochromatin core region for $conc_{hc}$, nucleoplasm for $conc_{nucleop}$) versus nucleus multiplied by the nucleus concentration. (C) Average MeCP2 concentrations in the nucleus and its subcompartments heterochromatin and nucleoplasm in untransfected, low, high expressing mouse myoblasts and in mouse brain (average ± standard deviation). (D) Table showing the results of total MeCP2 quantification (as average ± standard deviation) in untransfected, low, high expressing myoblasts and mouse brain tissue. The calculated values of the number of molecules, nucleus volume (µm³), and MeCP2 concentration (µM) for each condition are listed. The figure was modified from Zhang & Romero, Schmidt *et al.*, 2022 [1].

## 4.4 Discussion

Abnormal MeCP2 levels, evoked by overexpression as well as loss of expression, were reported to cause neuronal disfunction and Rett syndrome-like phenotypes [244–247, 328]. For this reason, it is important to carefully adjust the MeCP2 concentrations used for *in cellulo* and *in vitro* experiments to a range comparable to physiological protein levels. Thus, we developed a procedure to calibrate the MeCP2 level in single cells and subcellular compartments. The method was applied to different experimental platforms to compare the MeCP2 levels across various systems.

As immunofluorescence staining and the absolute quantification of MeCP2 in mouse myoblast cells confirmed their low MeCP2 levels, they were used as the cellular model system to manipulate MeCP2 levels by ectopic expression of a plasmid coding for MeCP2-GFP. Expression of *Mecp2-GFP* in mouse myoblast cells at low levels yielded a MeCP2 level comparable to the protein amount determined for mouse brain nuclei and can, thus, be used to mimic *in vivo* physiological conditions. The numbers of MeCP2 molecules per nucleus obtained from our experiments are in line with the levels reported in previous studies for transfected mouse myoblast cells [181] and mouse brain nuclei [138]. The high level expressing myoblast cells in our study showed about eight times higher MeCP2 levels than mouse brain nuclei. Skene *et al.* performed fluorescence-activated cell sorting (FACS) of mouse brain nuclei based on NeuN staining. The number of MeCP2 molecules per nucleus in unsorted brain nuclei ($6 \times 10^6$) was 2.6-fold lower than in FACS-sorted neurons ($15.5 \times 10^6$) and 2.4-fold higher than in FACS-sorted glia cells [138]. As the MeCP2 amount was reported to be cell type, tissue type and differentiation level specific [55, 233, 313, 314], high *Mecp2* expressing myoblasts might still be in the physiological range of the MeCP2 amount in highly differentiated cells in specific brain regions.

In addition to the absolute protein quantification *in vivo* and *in cellulo* by Western blot, we applied the cell sorting gates to microscopy images. The cells showed a level-dependent increase of the MeCP2 signal in heterochromatin and dose-dependent heterochromatin clustering, as reported before [55]. Thus, these results confirmed the compatibility of our FACS-sorting-based protein quantification method with microscopic data. We used the same binning system to classify MeCP2 levels in living cells [1].

In comparison to previous studies [138, 181], we did not only determine the number of MeCP2 molecules per nucleus, but also the protein concentrations in the nucleus and its subcompartments heterochromatin and nucleoplasm. Especially the MeCP2 concentration in heterochromatin is of importance for *in vitro* experiments, as MeCP2 highly accumulates at heterochromatin in a level-dependent manner and induces dose-dependent heterochromatin clustering [55].

Recently, it was proposed that the mechanism of heterochromatin clustering might be based on liquid-liquid phase separation events [257–259]. Of note, also MeCP2 was reported to undergo phase separation, initially shown only in the presence of DNA, nucleosomal arrays, or crowding agents [255, 256]. The systematic analysis of the phase separation properties of MeCP2 revealed that it is dependent on many parameters including protein concentration, salt concentration, crowding agents, and (methyl)DNA [1]. We could show that MeCP2 was able to form liquid-like

droplets in physiological protein concentrations either with low salt or in presence of crowders [1], underscoring the importance of protein level calibration for *in vitro* experiments.
Concluding, our calibration system can serve as a basis for further MeCP2-related studies carried out under physiological conditions either mimicking MeCP2 concentration in the full nucleus or the heterochromatin. In addition, it enables the comparison of different experimental setups including Western blot results, fixed and live cell microscopy as well as *in vitro* systems.

## 4.5 Supplementary material

### Supplementary tables

**Table 4.1:** Plasmid characteristics.

| Name | pc number* | Fluorescent protein | Gene species | Promotor | References |
|---|---|---|---|---|---|
| pFB-MeCP2G | pc1571 | GFP | *Rattus norvegicus* | polyhedrin | Jost *et al.*, 2011 [326]; Becker *et al.*, 2013 [164] |
| pMeCP2G | pc1121 | EGFP | *Rattus norvegicus* | CMV | Brero *et al.*, 2005 [55] |

*pc number: Plasmid collection number.

**Table 4.2:** Eukaryotic cell line characteristics.

| Name | Species | Type | Genotype | References |
|---|---|---|---|---|
| SF9 | *Spodoptera frugiperda* | ovarian tissue | wild type | Vaughn *et al.*, 1977 [330] |
| C2C12 | *Mus musculus* | myoblast | wild type | Yaffe & Saxel., 1977 [327] |

**Table 4.3:** Primary and secondary antibody characteristics.

| Reactivity | Host | Dilution | Application | Catalog / clone | Company / reference |
|---|---|---|---|---|---|
| Anti-MeCP2 | rabbit | 1:250 | IF | - | Jost *et al.*, 2011 [326] |
| Anti-MeCP2 | rat | Undiluted TCSN | WB, IF | Clones 4H7 & 4G10 | Jost *et al.*, 2011 [326] |
| Anti-rabbit IgG Cy5 | donkey | 1:400 | IF | #711-175-152 | Jackson ImmunoResearch |
| Anti-rat IgG Cy5 | donkey | 1:250 | IF | #712-175-153 | Jackson ImmunoResearch |
| Anti-rat IgG Cy3 | donkey | 1:1000 | WB | #712-165-153 | Jackson ImmunoResearch |

WB: Western blot; IF: Immunofluorescence; TCSN: Tissue culture supernatant.

**Table 4.4:** Imaging and flow cytometer characteristics.

| System / Company | Lasers / Lamps | Filters (ex. & em. (nm)) | Objectives / lenses | Detection system | Application |
|---|---|---|---|---|---|
| S3e Cell Sorter / Bio Rad Laboratories, Herkules, CA, USA | 488 nm | GFP (em.: 525/30) | - | FSC with PMT; SSC with PMT; Fluorescence detectors with PMTs | Quantification of MeCP2 concentration in cells |

| | | | | | |
|---|---|---|---|---|---|
| Widefield microscope Axiovert 200 / Zeiss, Germany | HBO100 mercury lamp | DAPI (300-400 & 410-510); GFP (473-491 & 506-534); Cy5 (590-650 & 663-738) | oil immersion 63X Plan-Apochromat (1.4 NA) | 12-bit AxioCam mRM | Quantification heterochromatin intensities and areas |
| Confocal microscope Leica TCS SPE-II / Wetzlar, Germany | Multicolor solid-state Laser module RYBV 405 nm / 25 mW 488 nm / 10 mW 635 nm / 18 mW | em.: DAPI: 410-600; GFP: 493-540; Cy5: 640-783 | oil immersion 63X ACS APO CS (1.3 NA) | Leica SP-Detector adjustable in the range of 430 - 750 nm | Quantification heterochromatin intensities, volume and areas |
| Ultra-View VoX spinning disk microscope/ PerkinElmer, UK | Solid state diode lasers (405 nm, 561 nm, 640 nm) | 405/568/640** 405: 415–475 561: 580–650 640: 664–754 | oil immersion 60x PlanApochromat (NA 1.45) | cooled 14-bit Hamamatsu® C9100-50 EMCCD | Confocal z-stack imaging of tissue slices |
| Amersham AI600 imager/ GE Healthcare, Chicago, IL, USA | White light (trans) Chemiluminescence, fluorescence | Cy2: 525BP20, Cy3: 605BP40, Cy5: 705BP40 | - | 16-bit Peltier cooled Fujifilm Super CCD | Western blot, SDS-PAGE imaging |

*ex.: Excitation; em.: Emission.; FCS: Forward Scatter; PMT: Photomultiplier tube(s); SSC: Side scatter; **dichroic specification.

**Table 4.5:** Plot statistics (main figures).

| Figure | Sample | $n_a$ | $n_b$ | Median | Mean | StDev | 95% CI | p-value |
|---|---|---|---|---|---|---|---|---|
| 4.1B.1 | Myoblasts | 3 | 786 | 216.56 | 213.51 AU | 75.49 | 5.29 | - |
| | Myoblasts + MeCP2-GFP (low) | 3 | 210 | 1197.81 | 1305.00 AU | 856.84 | 116.56 | <2.2e-16 |
| | Myoblasts + MeCP2-GFP (high) | 3 | 91 | 1915.16 | 1871.52 AU | 1024.32 | 213.32 | 5.94e-06 |
| 4.1B.2 | Myoblasts | 3 | 786 | 0.77 | 0.80 $\mu m^2$ | 0.15 | 0.01 | - |
| | Myoblasts + MeCP2-GFP (low) | 3 | 210 | 1.06 | 1.11 $\mu m^2$ | 0.28 | 0.04 | <2.2e-16 |
| | Myoblasts + MeCP2-GFP (high) | 3 | 91 | 1.17 | 1.22 $\mu m^2$ | 0.24 | 0.05 | <2.2e-16 |

$n_a$: number of replicates; $n_b$: number of cells; StDev: standard deviation; CI: confidence interval; p-value: in comparison to wild type MeCP2.

## Supplementary figures

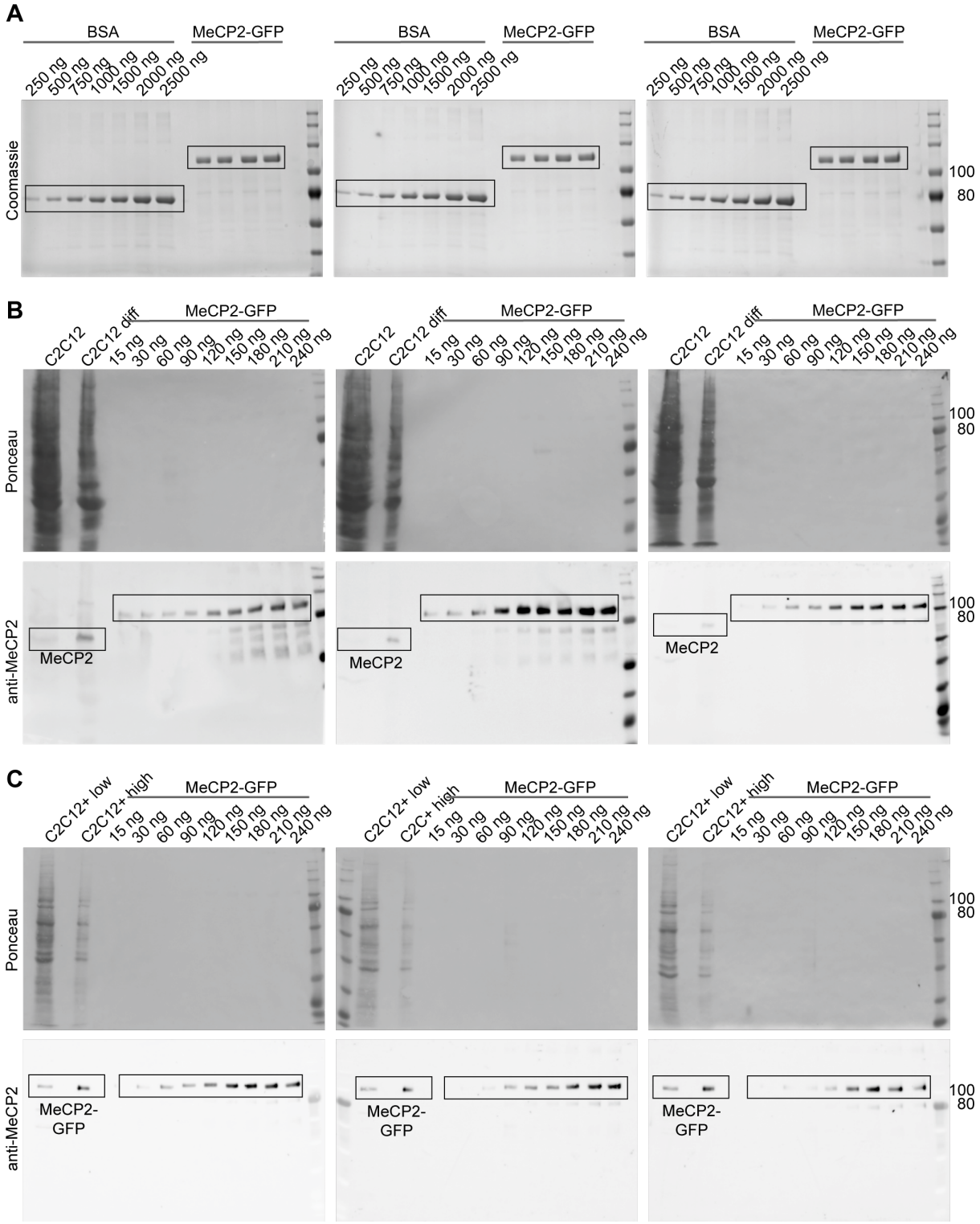



**Figure 4.4: SDS-PAGE and Western blot for MeCP2 quantification in cells.** (A) Full Coomassie-stained gels were used for quantification of the MeCP2 standard against the BSA standard curve (three replicates). Rectangles highlight the bands used for quantification. The 80 kDa and 100 kDa

**Figure 4.4 (*previous page*):** bands of the molecular weight marker are labelled. (B) Full membranes were stained for total protein with Ponceau S stain and incubated with MeCP2-specific antibodies for MeCP2 quantification in myoblasts (differentiated myoblasts (C2C12 diff) as control, three replicates). Rectangles highlight the bands used for quantification. The 80 kDa and 100 kDa bands of the molecular weight marker are labelled. C2C12 + low/high: C2C12 myoblast cells with low/high levels of MeCP2-GFP. (C) Full membranes were stained for total protein with Ponceau S stain and incubated with MeCP2-specific antibodies for MeCP2 quantification in transfected and FACS-sorted MeCP2-GFP positive myoblasts (three replicates). Rectangles highlight the bands used for quantification. The 80 kDa and 100 kDa bands of the molecular weight marker are labelled. The figure was taken from Zhang & Romero, Schmidt *et al.*, 2022 [1].

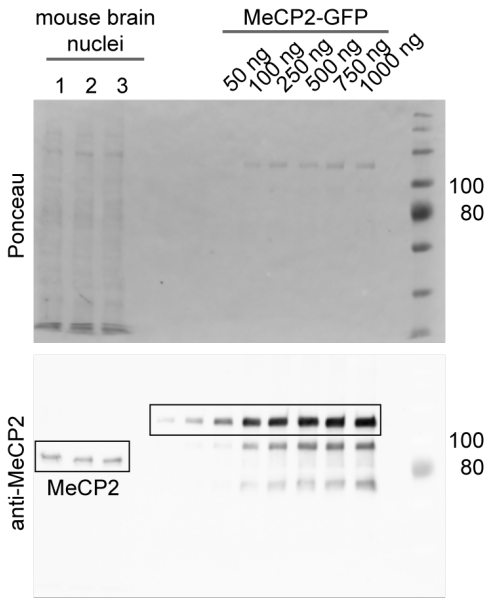



**Figure 4.5: Western blot for MeCP2 quantification in mouse brain tissue.** Full membrane stained for total protein with Ponceau S stain and incubated with MeCP2 specific antibodies for MeCP2 quantification in mouse brain nuclei (three biological replicates). Rectangles highlight the bands used for quantification. The 80 kDa and 100 kDa bands of the molecular weight marker are labelled.

# 5 MeCP2 heterochromatin organization is modulated by arginine methylation and serine phosphorylation

This chapter is modified from Schmidt *et al.*, 2022 [2].

## 5.1 Introduction

The methyl-CpG-binding protein 2 (MeCP2) is the founding member of the methyl-CpG binding domain (MBD) protein family and specifically binds to methylated CpGs via its MBD. As DNA methylation is mainly found on the less transcriptionally active heterochromatin, MeCP2 is prominently localized *in vivo* at pericentromeric chromatin regions, which contain highly methylated major satellite DNA repeats [115]. In addition to the MBD, MeCP2 contains a transcriptional repression domain (TRD) [312], the interdomain region (ID), and, more recently, the N-CoR/SMRT interacting domain (NID) has also been mapped [129]. MeCP2 binds to multiple interaction partners via these regions (reviewed in [3]). Several of the interacting partners are components of transcriptional repression complexes, for example Sin3A, HDAC and N-CoR [125–129]. MeCP2 might also be involved in transcriptional activation as it associates with CREB1 [139]. Aside from the MBD, which shows structurally conserved motifs, MeCP2 was reported to be an intrinsically disordered protein [124].

In mouse cells, pericentromeric heterochromatin from different chromosomes forms densely packed chromatin clusters in interphase called chromocenters ([25], see review [24]). Increased MeCP2 levels, either occurring during cell differentiation or upon exogenous expression of fusion protein constructs, cause large-scale reorganization of heterochromatin, which can be visualized as fusion events of heterochromatin clusters in mouse cells [55, 57, 159]. As constitutive heterochromatin has been shown to organize chromosomes within the cell nucleus [72], its reorganization has a potential impact on the general chromosome distribution. Recently, we and others proposed that heterochromatin cluster fusion events might be mediated by liquid-liquid phase separation [257, 258], as MeCP2 was shown to undergo phase separation under physiological conditions [1, 255, 256]. MeCP2 shows characteristic properties of phase separating proteins including intrinsically disordered regions and multivalency, and it was reported to interact with itself and several other interaction partners via regions outside of the MBD [164].

Mutations in the *MECP2* gene were linked to Rett syndrome, a human neurological disorder affecting mainly females, that is associated with intellectual disability among other symptoms [116]. *MECP2* mutations in males can lead to a wide spectrum of phenotypes ranging from mild intellectual impairment to severe neonatal encephalopathy and premature death [331]. Missense mutations in the MBD domain of *MECP2* affect heterochromatin accumulation due to reduced DNA binding ability, but also heterochromatin clustering [62]. The clustering function of some mutations could be rescued by retargeting MeCP2 to heterochromatin [63].

Importantly, MeCP2 is post-translationally modified and although many modifications have been identified, only a few were validated and functionally characterized (reviewed in [3, 332]). The phosphorylation of serine 421 in the C-terminal domain of MeCP2 was identified upon neuronal activity and stress exclusively in the brain, indicating a specific function under this

condition [194, 195]. Serine 80 phosphorylation in the N-terminal domain of MeCP2 was found in mouse and rat brains [195]. Serine to alanine mutated knock-in mice of both modification sites were reported to display opposing phenotypes, as S421A mice show increased, whereas S80A mice show decreased locomotor activity. In line with these results, membrane depolarization in cortical neurons results in dephosphorylation of serine 80 and phosphorylation of serine 421. Interestingly, the S80A mutation results in a decrease of MeCP2 chromatin binding affinity to *Pomc* and *Gtl2* promoters evaluated by ChIP-qPCR but did not lead to significant changes in gene transcription [195]. Besides, MeCP2 was found to be poly(ADP-ribosyl)ated in mouse brain tissue at ID and TRD, and this led to decreased DNA binding and heterochromatin clustering [181].

In this study, we aimed to identify post-translational modifications of MeCP2 from mouse brain (*in vivo*) and determine whether these modifications are involved in MeCP2 chromatin binding and clustering. 23% of the MeCP2 protein is composed of positively charged amino acids and we found only a few modified arginines compared to many modified lysines. In addition, we identified several phosphorylated serine and threonine residues, including the previously reported S80 and S421. We show that arginine methylation and to a much lesser extent also serine phosphorylation affect heterochromatin accumulation and binding kinetics and MeCP2 heterochromatin clustering function. In addition, coexpression of MeCP2 variants and the protein arginine methyltransferase 6 (PRMT6) reveals differences in heterochromatin clustering.

## 5.2 Material & methods

### 5.2.1 Nuclei isolation from mouse brains

3-month-old C57BL/6 mice (Charles River Laboratories, Inc.) were sacrificed according to the animal care and use regulations (Government of Hessen, Germany), and the organs were collected from the sacrificed animals, washed with PBS, and frozen in liquid nitrogen. For nuclei isolation the frozen mouse brains were crushed to powder and homogenized in 0.25 M sucrose solution (20 mM triethanolamine-HCl (pH 7.6), 30 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 1 mM PMSF). After centrifugation for 10 min at 1000 xg, the supernatant was discarded and the pellet resuspended in sucrose buffer to a final sucrose concentration of 2.1 M. The raw nuclei fraction was obtained by centrifugation for 30 min at 50000 xg. The pellet was resuspended in 0.25 M sucrose solution and centrifuged at 1000 xg. During the procedure, samples were taken after resuspension of the tissue, after homogenization and after nuclei isolation, fixed with 3.7% formaldehyde in solution for 15 min, dropped on slides, dried, and counterstained with 4',6-diamidino-2-phenylindole (DAPI) for microscopic examination of the individual steps.

### 5.2.2 Protein enrichment

For the MeCP2 enrichment from mouse brain tissue, we made use of its natural hepta-histidine tag for protein pull-down with Ni-IDA beads (His60 Ni Superflow resin, Clontech Laboratories, Inc.). First, $10^7$ mouse brain nuclei in PBS were pelleted by centrifugation at 1000 xg for 10 min. The nuclei were resuspended in buffer B (0.2% Triton X-100, 50 mM triethanolamine-HCl (pH 7.6), 5 mM $MgCl_2$), incubated on ice for 10 min and centrifuged at 1000 xg for 10 min. The supernatant was discarded and the pellet was washed three times by resuspension in 100 µl buffer C (2 mM triethanolamine-HCl (pH 7.6), 0.5 mM $MgCl_2$) and centrifugation for 10 min at 1000 xg. The pellet was resuspended in 500 µl 1 M NaCl equilibration buffer (50 mM sodium phosphate, 20 mM imidazole, pH 7.4), followed by sonication 3 x 20 s (250-450 Sonicator, BRANSON ultrasonic corporation) with microscopic control after each step. Subsequently, the lysate was diluted using 500 µl equilibration buffer without NaCl and added to the Nickel-Iminodiacetic acid (Ni-IDA) beads for incubation overnight at 4 °C with rotation. The beads were washed with 300 mM NaCl equilibration buffer, then with wash buffer (50 mM sodium phosphate, 300 mM NaCl, 40 mM imidazole, pH 7.4). The Ni-IDA beads were then resuspended in Laemmli buffer (2% SDS, 50 mM Tris (pH 6.8), 10% glycerol, 0.01% bromophenol blue, 100 mM DTT), incubated at 95 °C for 10 min and separated using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

### 5.2.3 Mass spectrometry

The samples to be analyzed by mass spectrometry were analyzed by SDS-PAGE and the gel was stained with Coomassie staining solution (5% aluminium sulfate-(14)-(18)-hydrate, 10% ethanol p.a., 0.02% CBB-G250 (Coomassie brilliant blue), 2% orthophosphoric acid; [333]) overnight. The in-gel tryptic digestion was performed as described before [334]. Briefly, the gel was destained

using Coomassie destaining solution (10% ethanol p.a., 2% orthophosphoric acid, LC-MS grade) two times for 10 min, equilibrated in water (MS grade), the bands of interest were excised, cut to small cubes and dried using a vacuum concentrator. For destaining the gel pieces were covered with destaining solution (40 mM ammonium bicarbonate, 50% acetonitrile, LC-MS grade), incubated at 37 °C for 30 min with shaking, and the solution was removed. Destaining was repeated at least two times and the gel pieces were dried using a vacuum concentrator. For trypsin digestion the gel pieces were covered with 12.5 ng/µl trypsin (sequencing grade modified trypsin, #V5111, Promega Corporation) in 40 mM ammonium bicarbonate and incubated at 37 °C with shaking overnight. The peptides were eluted by adding elution solution (50% acetonitrile, 0.5% trifluoroacetic acid, LC-MS grade), incubation for 20 min in an ultrasonic bath, transfer of the peptide solution to a new tube, and drying using a vacuum concentrator. Samples were resuspended in 20 µl buffer (0.1% formic acid in 2% acetonitrile, LC-MS grade), incubated in an ultrasonic bath for 5 min, and transferred to HPLC vials. Subsequent drying of the samples in a vacuum concentrator allowed storage at room temperature in the dark until the measurement.

The HPLC-MS/MS measurement was performed with the setup described before [334]. Briefly, an UPLC HSS T3 column and an UPLC Symmetry C18 trapping column for LC were used in combination with the nanoACQUITY gradient UPLC pump system (Waters) coupled to an LTQ Orbitrap Elite mass spectrometer (Thermo Fisher Scientific). The LTQ Orbitrap Elite was operated in a data-dependent mode using Xcalibur software either in collision-induced dissociation (CID) TOP20 or in TOP10 with high-energy collisional dissociation (HCD) and CID fragmentation for every precursor ion. For elution of the peptides a linear gradient from 5 – 30% for 60 min (CID TOP20) or 150 min (TOP10 HCD, CID) of buffer B (0.1 formic acid in acetonitrile, UPLC/MS grade) was applied, followed by a step gradient from 30 – 85% acetonitrile for 5 min at a flow rate of 400 nl/min.

Data analysis was performed using Proteome discoverer 1.3 (Thermo Fisher Scientific) with SEQUEST [335] and MaxQuant (version 2.0.3.0) with Andromeda [265] algorithms searching against the complete UniProt database (UniProt Consortium, 2021) for *Mus musculus*. A maximum of two missed tryptic cleavages was accepted and methionine oxidation, N-terminal acetylation, N-terminal pyroglutamate, lysine acetylation, lysine ubiquitination, lysine and arginine mono-methylation or di-methylation and serine/ threonine/ tyrosine phosphorylation were set as variable modifications. To identify all methylation and dimethylation sites, the search was repeated including either only lysine/ arginine methylation or dimethylation. The MaxQuant search was run with default parameters having matching between runs enabled.

### 5.2.4 Plasmids

All plasmids used in this study are listed in table 5.1. The plasmids coding for MeCP2-GFP and MeCP2-GFP mutants were kindly provided by Jana Frei. The PRMT plasmids were kindly provided by Chris Aßmann and Uta-Maria Bauer.

### 5.2.5 Cell culture and transfection

C2C12 mouse myoblast cells (female) and MTF mouse tail fibroblast (male) MeCP2 -/y cells were grown in Dulbecco's modified Eagle Medium (DMEM) with high glucose (#D6429, Sigma-Aldrich) supplemented with 20% (C2C12) or 10% (MTF -/y) fetal bovine serum (#F7524, Sigma-Aldrich), 1x glutamine (#G7513, Sigma-Aldrich) and 1 µM gentamicin (#G1397, Sigma-Aldrich) at 37 °C and 5% $CO_2$ in a humidified incubator. Mycoplasma tests were performed regularly and all cell lines are listed in table 5.2. C2C12 cells were tested for the ability to differentiate to myotubes, and MTF -/y cells were proven to be MeCP2 negative by immunofluorescence staining.

Transient transfections were performed using the Neon transfection System (Thermo Fisher Scientific) according to the manufacturer's instructions. For cotransfections of MeCP2 mutants and PRMTs, a plasmid amount ratio of 1:5 (2 µg, 10 µg) was used. After transfection, cells were seeded on gelatin-coated coverslips and grown at 37 °C and 5% $CO_2$ in a humidified incubator. 7 h after transfection, cells were washed with PBS, and transcription was induced by adding medium supplemented with 2 µM tetracycline. 20 h after induction, cells were washed with PBS and fixed either with ice-cold methanol for 6 min (MTF –/y) or with 3.7% formaldehyde (C2C12) for 15 min.

### 5.2.6 Western blot analysis

Mouse brain nuclei were lysed in Laemmli buffer (see above), mechanically disrupted, and incubated for 5 min at 95 °C. The samples were analyzed by SDS-PAGE and transferred to a nitrocellulose membrane using a semi-dry blotting system at 25 V for 35 min. For detection of post-translational modifications, membranes were blocked with 5% BSA in TBST (0.1% Tween 20 in TBS) for 1 h and incubated with anti-PTM antibodies diluted in TBST (antibodies and dilutions are listed in table 5.3) at 4 °C with shaking overnight. Membranes were washed three times with TBS, blocked for 30 min with 5% low-fat milk in PBS, and incubated overnight with anti-MeCP2 monoclonal rat antibody mix (4H7, 4G10, 4E1 undiluted, [326]). After three washing steps with 0.1% PBST (0.1% Tween 20 in PBS), membranes were incubated with Cy3-conjugated anti-rat IgG secondary antibodies diluted 1:1000 in 3% milk in PBS for 1 h, washed three times with PBST and the fluorescent signal for MeCP2 was detected using Amersham AI600 imager (see table 5.4). Subsequently, membranes were incubated with HRP-coupled secondary antibodies (either rabbit or mouse IgG) diluted in 3% milk for 1 h, washed three times with PBST, stained with ECL solution (Clarity Western ECL substrate, #1705061, Bio-Rad) and the chemiluminescence signal for the PTMs detected using an Amersham AI600 imager.

All other Western blots were performed similarly to MeCP2 detection. Briefly, membranes were blocked for 30 min with 5% low-fat milk in PBS, incubated with primary antibody diluted in 5% low-fat milk overnight at 4 °C shaking, washed three times with PBST, incubated 1 h at room temperature with secondary antibody diluted in 3% low-fat milk, washed three times with PBST and signals were detected using an Amersham imager.

### 5.2.7 Immunofluorescence staining

After fixation either with ice-cold methanol for 6 min (MTF −/y) or with 3.7% formaldehyde (C2C12) for 15 min and washing with PBS, cells were permeabilized with 0.7% Triton X-100 in PBS for 20 min and washed three times with 0.01% PBST. Cells were either directly stained with DAPI or blocked with 2% BSA in PBST for 20 min and incubated with primary antibodies diluted in 2% BSA in PBST for 2 h (primary and secondary antibodies with their respective dilutions are listed in table 5.3). After washing three times with 0.1% PBST, secondary antibodies in 2% BSA were applied for 1 h in the dark, followed by three times washing with 0.1% PBST, 12 min DAPI staining in the dark, washing with PBST and water and mounting in Mowiol 4-88 (#81381, Sigma-Aldrich; 4.3 M Mowiol 4-88 in 0.2 M Tris-HCl pH 8.5 with 30% glycerol) supplemented with 2.5% DABCO antifade (1,4-diazabicyclo[2.2.2]octan, #D27802, Sigma-Aldrich).

### 5.2.8 Fluorescence microscopy & image analysis

All characteristics of the microscopy systems used including lasers/lamps, filters, objectives, detection, and incubation systems are listed in table 5.4.

#### 5.2.8.1 Microscopic analysis of subcellular localization

Fluorescence and DIC images of transfected C2C12 cells were taken using a Nikon Eclipse TiE2 system with a 40x air Plan Apo λ DIC objective. Fluorescence images of transfected MTF -/y cells were acquired using a confocal microscope Leica TCS SPEII and intensities were measured using ImageJ [270, 271].

#### 5.2.8.2 Microscopic analysis of heterochromatin accumulation

Fluorescence images of transfected C2C12 cells for calculation of heterochromatin accumulation were acquired on a Zeiss Axiovert 200 microscope. Image segmentation was performed using an ImageJ macro described previously [1]. First, the cell nuclei were segmented semi-manually based on the DAPI intensity and a difference of gaussian blur filter was applied. For heterochromatin segmentation individual pixel intensities were calculated, local maxima were determined in squares of 30 x 30 pixels and pixel intensities were binned into 42 bins with the local maximum defining the intensity of the highest bin per square. Heterochromatin masks were obtained by thresholding the pixel intensities based on their respective bins, taking all pixels with bins $\geq$21 for total and bins $\geq$37 for core heterochromatin cluster regions. The nucleoplasm area was calculated by subtracting the total heterochromatin cluster areas from the nucleus area. The heterochromatin accumulation of MeCP2 mutants for each individual heterochromatin cluster was calculated by dividing the mean intensity of the heterochromatin cluster core by the mean intensity of the nucleoplasm (modified from [1], see figure 5.9). To analyze the protein level-dependency of accumulation differences, the cells were divided into low and high protein levels based on their GFP intensity as described before [1]. Briefly, the $\log_{10}$ GFP sum intensity of the cells was plotted and divided into 40 bins. In comparison to the GFP intensity of untransfected

cells, bins 1 to 11 were defined as negative, while cells in bins 13 to 21 were considered as low and cells in bins 24 to 32 as high expressing.

#### 5.2.8.3 Microscopic analysis of heterochromatin clustering

High-content screening microscopy of transfected C2C12 cells was performed using a Perkin Elmer Operetta imaging system and analyzed using the supplier's software Harmony (Version 3.5.1, Perkin Elmer Life Sciences, UK). Briefly, cell nuclei were segmented based on the DAPI channel image considering nuclei with a size of 110 to 250 $\mu m^2$ and roundness coefficient $>0.8$ that are not touching the edges of the image. Heterochromatin segmentation was also performed using the DAPI channel image, identifying high-intensity spots within the nuclei. Spots with a spot-to-region intensity ratio of at least 0.35, an area of at least $5\,px^2$, and a relative spot intensity of 0.0253 were considered. The nucleus and heterochromatin masks generated based on the DAPI channel image were used to segment the images of the other channels and the intensity and morphology properties of nuclei and heterochromatin clusters were measured (see figure 5.10). MeCP2 intensity bins based on GFP intensity were defined as described for the Axiovert images. Not all images obtained were considered for analysis, as the cell numbers per replicate and condition were different caused by differing transfection efficiencies. Thus, either the number of images (three biological replicates) was reduced to achieve comparable cell numbers, or the number of cells (two biological replicates) was adjusted to achieve exactly the same number of cells per condition and replicate.
For the cell cycle analysis, the frequencies of the DAPI sum intensities per nucleus were plotted as histograms. The DAPI intensities of the different samples were normalized to the corresponding G1 peak as described before [336]. The intervals for the G1, S and G2/ M phase were set manually based on the histogram and the percentages of the cells within each interval were plotted as bar diagrams.

#### 5.2.8.4 Fluorescence recovery after photobleaching analysis of heterochromatin binding kinetics

Live cell imaging for fluorescence recovery after photobleaching (FRAP) experiments was carried out on a confocal microscope Leica SP5 II with an HCX PL APO 63x oil lambda blue objective equipped with an ACU live cell chamber at 37 °C, 5% $CO_2$ and 60% humidity. MTF -/y cells were transfected with MeCP2 mutation constructs, seeded on gelatin-coated glass bottom p35 plates followed by induction with 2 µM tetracycline 7 h later and DNA staining with 100 nM SiR-DNA (SiR-Hoechst) (#SC007, Spirochrome) in presence of 5 µM verapamil (included in #SC007). 16 h later heterochromatin clusters were bleached with a 488 nm argon laser at 100% intensity for 2 s, and confocal images were taken with a frame size of 256 x 256 with 200 Hz and a pinhole of 1 AU in time intervals of 1.5 s. For analysis, the image series were registration corrected using ImageJ plugin StackReg (correction based on GFP channel) or HyperStackReg (correction based on DNA staining). The mean fluorescence intensities of the (pre- and post-bleach) bleached and unbleached region were background subtracted for each time point. For single normalization, intensities were normalized to the mean of the prebleach intensities. Fluorescence recovery

curves were fitted in ImageJ and t-half values and mobile fractions were obtained from the fitted curves. At least ten cells were analyzed for each construct and the means of the fitted curves, t-half values, and mobile fractions were plotted.

## 5.3 Results

### 5.3.1 MeCP2 is post-translationally modified in mouse brain

To identify post-translational modifications (PTMs) of MeCP2, we first isolated nuclei from mouse brain tissue using a sucrose buffer in combination with centrifugation. Then, we enriched MeCP2 from mouse brain nuclei using its natural hepta-histidine sequence localized in its C-terminal domain (see figure 5.1A, B) for protein pull down with Ni-IDA beads (figure 5.11). Subsequently, the enriched proteins were separated by SDS-PAGE, in-gel digested using trypsin, and subjected to mass spectrometry analysis using a TOP10 shotgun method with a combination of HCD and CID fragmentation for improved sequence coverage and modification identification. The mass spectrometry analysis of MeCP2 from the mouse brain yielded a sequence coverage of 60.1% and a series of PTMs, including lysine methylation and acetylation, arginine methylation, and phosphorylation on serines and threonines. Figure 5.1B depicts the MeCP2 protein sequence with the modifications identified.

The peptides on which the modifications were identified and the number of identifications obtained from automated data analysis using either Proteome Discoverer or MaxQuant software are listed in table 5.6. We decided to functionally characterize especially arginine methylation sites, as 7% of the MeCP2 amino acids are arginines, but we found only 11.4% of the arginines modified. In addition, a large number of *MECP2* mutations identified in Rett syndrome patients affect arginine residues. The arginine methylation sites identified on MeCP2-E2 isoform (starting in exon 2) namely R91, R106, R162, R167 were selected for further analysis and were validated by manual inspection of the spectra (see figures 5.12, 5.13, 5.14). The arginines R162 and R167 are located on the same peptide and were both identified as monomethylated and R162 also as dimethylated in the automated analysis. It was not possible to unambiguously determine the localization of the methylation sites on this peptide from the spectra (see figures 5.13, 5.14).

As the MeCP2 sequence comprises 13.2% of lysines and 7% of arginines, it is likely that some regions could not be covered in our measurements due to the generation of very short peptides by trypsin which cuts after lysine and arginine. In addition, there is a long sequence without any lysines and arginines in the C-terminus of MeCP2, containing the hepta-histidine sequence and a proline-rich region (see figure 5.1B) that was, thus, not accessible. Of note, it was reported that post-translational modifications like methylation lower the efficiency of trypsin-mediated cleavage. Therefore, we tried to increase sequence coverage using other enzymes for digestion but were unable to cover the missing regions (data not shown).

The results were validated by Western blot detection of mono and dimethyl arginine, serine / threonine / tyrosine phosphorylation as well as serine 80 and serine 421 phosphorylation on mouse brain nuclei extracts (figure 5.1C, 5.15). The same membranes were incubated with anti-MeCP2 antibodies to validate that the PTM signal was specific to MeCP2. We could show that MeCP2 is monomethylated as well as symmetrically and asymmetrically dimethylated on arginines. In addition, it is phosphorylated on serines, but not on tyrosines. The published MeCP2 phosphorylation sites S80 and S421 [194, 195] were also detected on MeCP2 isolated from mouse brains (figure 5.1C). Figure 5.1D shows a scheme of the arginine methylation

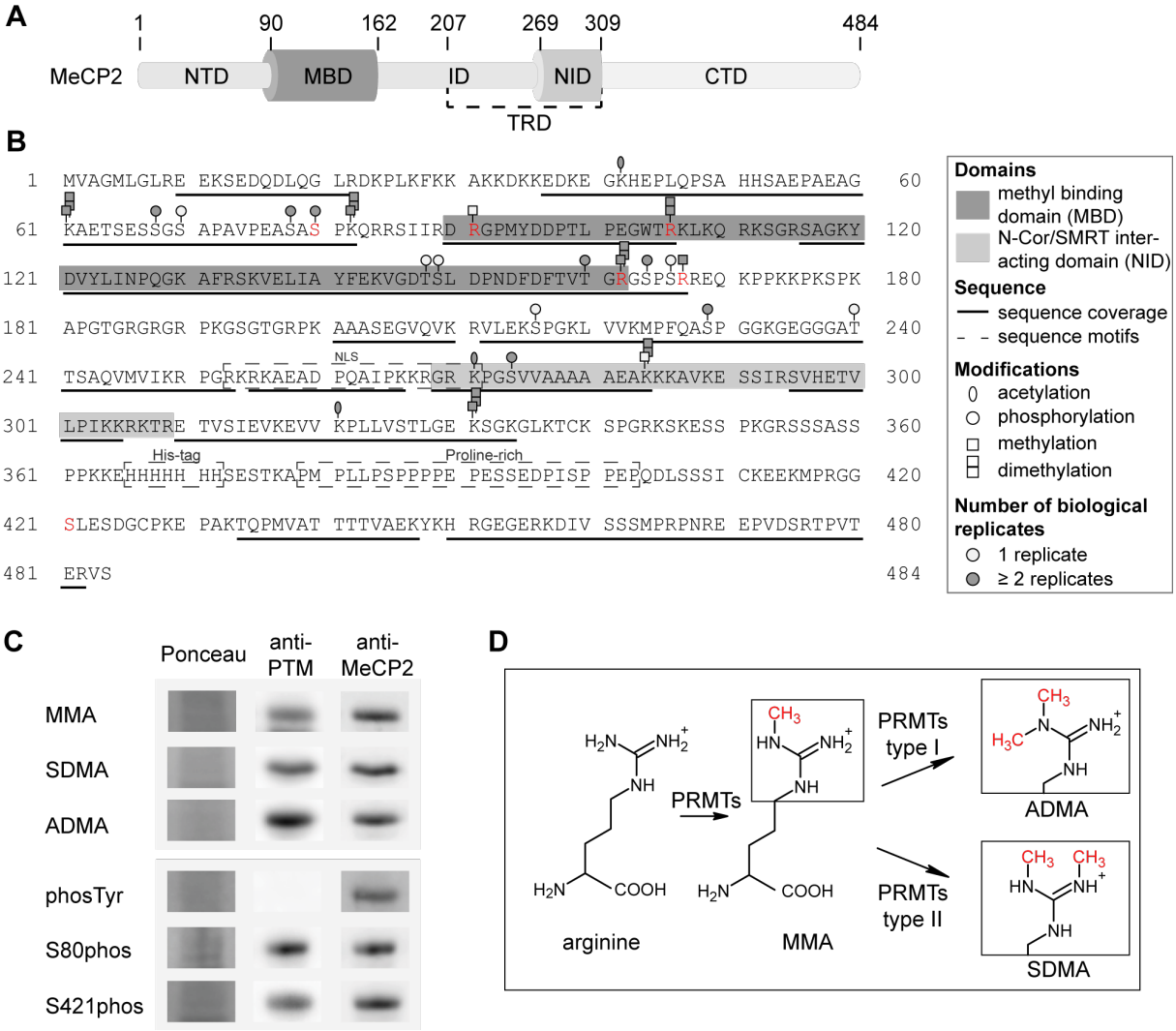



**Figure 5.1: Post-translational modifications of MeCP2 from mouse brain identified by mass spectrometry analysis.** (A) MeCP2 domain structure comprising N-terminal domain (NTD), methyl-CpG binding domain (MBD), intervening domain (ID), N-CoR/SMRT interacting domain (NID), transcriptional repression domain (TRD), and C-terminal domain (CTD). (B) MeCP2 protein sequence with modifications identified by mass spectrometry. MBD and NID, sequence coverage, sequence motifs, PTMs (acetylation, phosphorylation, methylation, dimethylation), and the number of biological replicates from independent experiments are marked as indicated. Modifications selected for further validation are marked in red. The software-based annotation of spectra and modification sites shown was obtained from Proteome Discoverer and MaxQuant. The arginine methylation sites selected for further validation were additionally manually inspected (see figures 5.12, 5.13, 5.14). The location of modifications identified on the same peptide might be uncertain. (C) Western blot analysis of mouse brain nuclei extracts tested for monomethyl arginine (MMA), symmetric dimethyl arginine (SDMA), asymmetric dimethyl arginine (ADMA), tyrosine phosphorylation (phosTyr), S80 and S421 phosphorylation (phos) and reprobed with MeCP2 specific antibodies. The Ponceau S stain visualizes the total proteins on the membrane and the full membranes of the Western blots are shown in figure 5.15. (D) Scheme of the arginine methylation reaction: protein arginine methyltransferases (PRMTs) can catalyze the monomethylation of arginine and subsequently the dimethylation, which can occur either on the same nitrogen (ADMA, catalyzed by type I PRMTs) or on the unmodified nitrogen (SDMA, catalyzed by type II PRMTs). The figure was modified from Schmidt *et al.*, 2022 [2].

reaction catalyzed by protein arginine methyltransferases (PRMTs). First, arginine residues can be monomethylated by PRMT enzymes of class I and II, subsequent asymmetric dimethylation is catalyzed by PRMT enzymes type I, and symmetric dimethylation by PRMT enzymes of type II. Of note, the arginine keeps its charge in the methylated state but shows a different charge distribution due to the bulky methylation groups.

### 5.3.2 Arginine methylation and serine phosphorylation site mutations do not influence MeCP2 subcellular localization with exception of MeCP2 R106 mutations

To investigate the functional consequences of MeCP2 PTMs, we generated recombinant MeCP2 proteins tagged with GFP and altered at PTM sites. While most published work utilizes the amino acid substitutes aspartate (D) to mimic phosphorylation or alanine (A) to prevent phosphorylation, there are no commonly used substitutions for methylated arginines. Thus, we decided to mutate the arginines identified to be methylated to lysine (K) to retain the positive charge, to glutamine (Q) to sterically mimic a methylated arginine, and to leucine (L) to obtain methyl groups similar to a methylated arginine (figure 5.16C, D). In addition, arginine substitutions to glutamine and leucine were identified in Rett syndrome patients. Of note, none of these mutations is an ideal mimic for methylated or unmethylated arginine. In methylated arginine, the positive charge of the arginine is kept with the addition of the sterically hindering additional methyl group(s). These amino acid substitutions can, thus, only partially mimic these changes, either the positive charge (K), the methyl group (L), or a polar and sterically larger side chain (Q). MeCP2 arginine R106 was detected as methylated in our proteomic screen (figure 5.1B) and is found mutated to tryptophan (R106W) or glutamine (R106Q), with very low frequency also to glycine (R106G) and leucine (R106L) in Rett syndrome patients (online RettBASE, [337]). We, therefore, generated recombinant MeCP2 with the R106 mutated to produce lysine, glutamine, leucine, tryptophan, and glycine, thus including the previously explained substitutions for arginine methylated sites (K, Q, L) and all reported R106 Rett syndrome mutations (W, Q, G, L).

The MeCP2-GFP plasmids point mutated for modified sites (figure 5.2A) were transfected into male mouse tail fibroblasts MTF-/y (figure 5.16A, B), which are MeCP2 null cells, and C2C12 female myoblast cells (figure 5.2B-D), which have a very low to undetectable level of MeCP2 [1] and, thus, can be used as a functional MeCP2 null system. In the following, the constructs are abbreviated as 3K (triple R91K R162K R167K), 3Q (triple R91Q R162Q R167Q), and 3L (triple R91L R162L R167L).

First, we analyzed the subcellular localization of the altered MeCP2 proteins and compared them with the wild type MeCP2 protein. Like wild type MeCP2, the MeCP2 proteins with 3K, 3Q, 3L substitutions (figure 5.2B, figure 5.16A) as well as all the phosphorylation site altered proteins (figure 5.2C, figure 5.16B) were enriched at heterochromatin, visualized as dense DAPI stained DNA regions in the images. All R106 altered MeCP2 proteins, as reported earlier for MeCP2 R106W and R106Q [240, 338], tend to lose their heterochromatin enrichment and mislocalize to the negatively charged RNA-enriched nucleolar compartment (figure 5.2D). The nucleolar compartment is visualized in the DIC images where it appears as prominent large structures

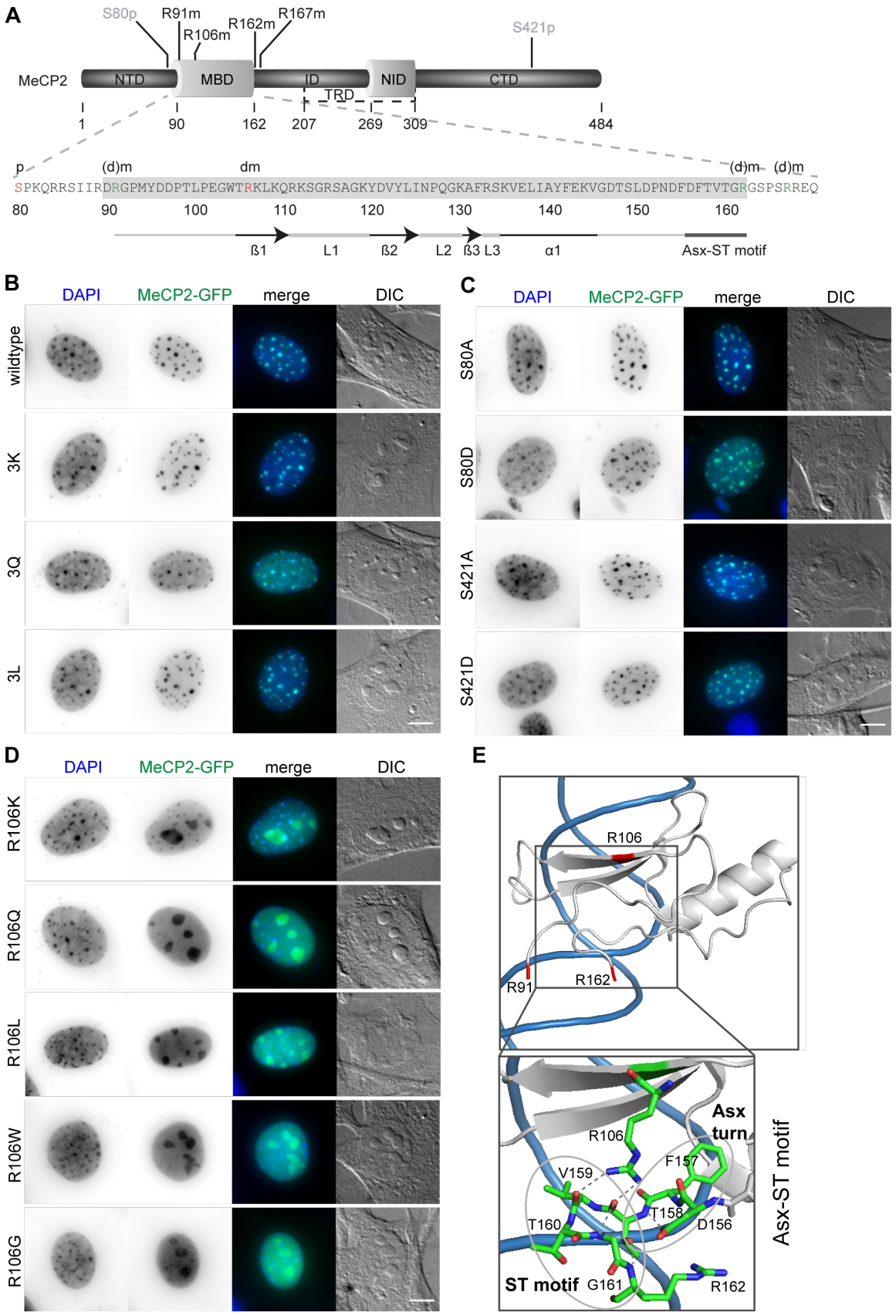



**Figure 5.2: Subcellular localization of MeCP2-GFP mutant constructs transfected in C2C12 mouse myoblast cells.** (A) MeCP2 domain structure and MBD amino acid sequence with the modifi-

**Figure 5.2 (*previous page*):** cation sites selected for functional validation are indicated in red (S80phos, R106dimet) or green (R91, R162, R167, referred to as 3x). (B, C, D) MeCP2 wild type and 3K, 3Q, 3L mutations (B) as well as S80 and S421 mutations (C) colocalized with DAPI dense heterochromatic regions, whereas R106 mutations lose their DAPI colocalization and mislocalized to the nucleoli (D). Scale bar 5 µm. (E) X-ray structure of the MeCP2 methyl-binding domain (MBD, shown in gray) in complex with methylated DNA (shown in blue) with the methylated arginine sites R91, R106, R162 highlighted in red (structure information from [136]; PDB accession code 3C2I). The enlarged image shows the Asx-ST motif stabilizing the MBD binding to methylated DNA. The Asx turn composed of D156, F157, and T158 is stabilized by a hydrogen bond between the carboxylate side chain of D156 and T158 main chain nitrogen. The ST motif of the amino acids 158 to 161 comprises two hydrogen bonds, one between the side chain hydroxyl group of T158 and the main chain nitrogens of G161 and R162, the second one between the main chain carbonyl group of T158 and G161 main chain nitrogen [136]. The structural data was generated and color-coded using PyMOL software. The figure was taken from Schmidt *et al.*, 2022 [2].

within the cell nucleus. Interestingly, it was described that peptides with a high occurrence of arginines tend to localize at the negatively charged nucleoli [339]. While MeCP2 R106K still shows some heterochromatin localization, the other R106 mutant proteins localized nearly exclusively within the nucleoli and showed higher intensities in the nucleoplasm compared to the wild type (figure 5.2D). These results might be explained by the role of R106 in MBD binding to the DNA. It was reported that MeCP2 binding to methyl-CpG on the DNA is mediated by direct contact of the three amino acids D121, R111, and R133 and might involve five water molecules [136]. Arginine 106 stabilizes the Asx-ST motif, a motif stabilizing MeCP2 DNA interaction ([136], see figure 5.2E). Interestingly, also the frequent Rett syndrome missense mutation T158M is localized in this motif. Both missense mutations occur very frequently and reduce DNA binding, emphasizing the importance of this motif for proper methyl-CpG binding and MeCP2 function.

### 5.3.3 MeCP2 arginine methylation and serine phosphorylation site mutants accumulate differently in heterochromatin

To quantitatively analyze the heterochromatin accumulation of the MeCP2 mutant constructs, we performed a cellular DNA/chromatin binding assay. C2C12 cells were transfected with the mutant constructs, fixed, and counterstained with the DNA dye DAPI. After imaging the cells, the nuclei were segmented semi-manually and heterochromatin compartments were segmented using a self-developed macro in ImageJ/Fiji ([1], figure 5.9). By calculating the ratio of the mean GFP intensity in the heterochromatin to the mean GFP intensity in the nucleoplasm, we obtained heterochromatin accumulation values (figure 5.3A). As MeCP2 levels might have an influence on its degree of heterochromatin accumulation, the cells were classified into low and high MeCP2 levels according to their mean nuclear fluorescence intensity as described before [1]. The quantitative analysis of heterochromatin accumulation revealed a slightly higher accumulation of MeCP2 3K than MeCP2 wild type. MeCP2 3Q showed a lower accumulation than wild type MeCP2 and MeCP2 3L an even lower accumulation compared to the MeCP2 3Q (figure 5.3B). The heterochromatin accumulation values of the phosphorylation site mutants were all very similar, with only the phospho mimic mutants S80D showing a slightly lower accumulation

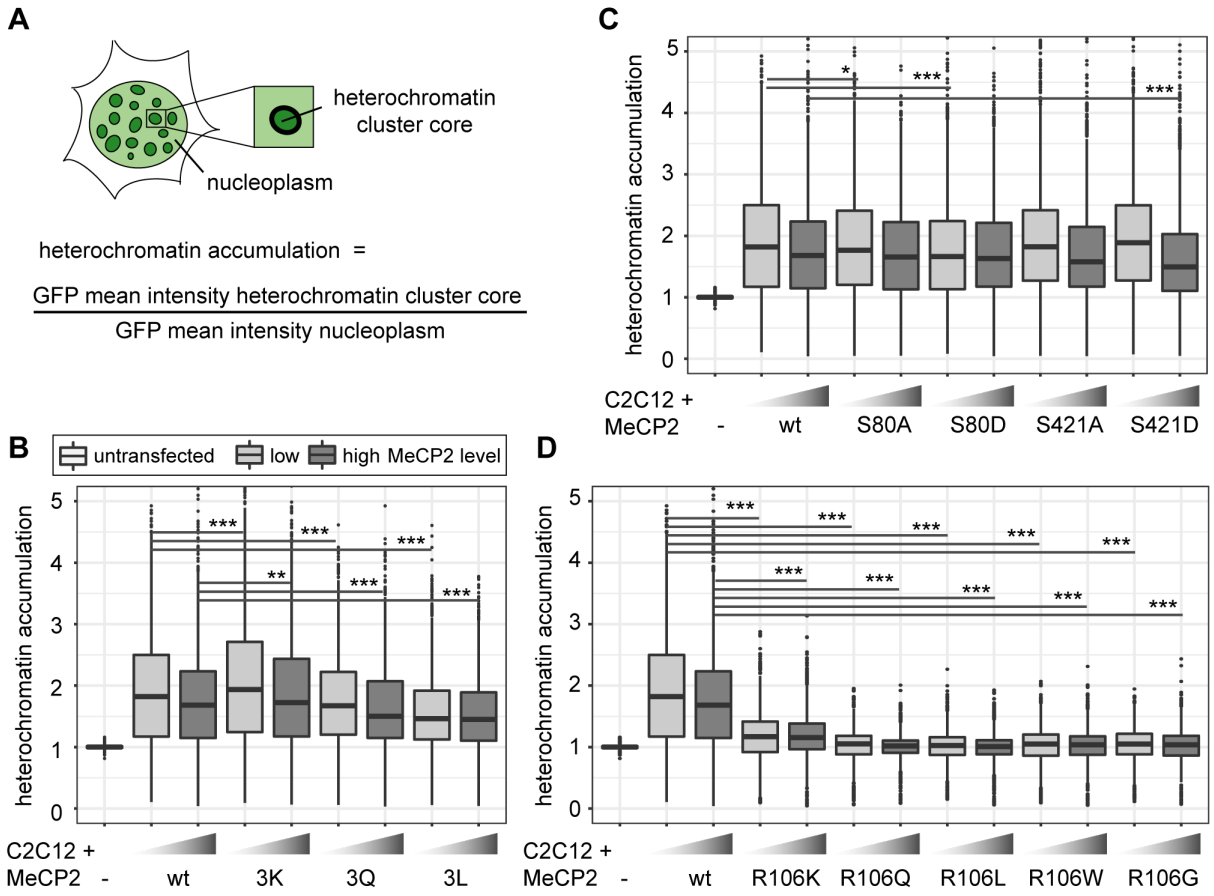



**Figure 5.3: Comparative analysis of heterochromatin accumulation of MeCP2 mutant constructs transfected in C2C12 mouse myoblast cells.** (A) Heterochromatin accumulation was calculated as the ratio of GFP mean intensity at heterochromatin versus nucleoplasm. The boxplots depict the heterochromatin accumulation of mutant MeCP2 including R91, R162, R167 (B), S80 and S421 (C), and R106 (D) constructs for low and high MeCP2 levels in the cells. Three biological replicates, statistical significance calculated using Wilcoxon-Rank test. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.001$. P-values and n-values are summarized in table 5.5. The figure was taken from Schmidt *et al.*, 2022 [2].

at low levels and S421D a lower accumulation at high levels (figure 5.3C). In line with the subcellular localization, the heterochromatin accumulation was drastically reduced in all R106 mutants (figure 5.3D). Only R106K still showed some heterochromatin accumulation with a ratio clearly above one. Overall, all constructs show a lower accumulation in case of high protein levels, which might hint at a saturation effect of MeCP2 binding to chromatin at high protein levels.

### 5.3.4 MeCP2 arginine methylation and serine phosphorylation site mutations affect its heterochromatin binding kinetics

Next, we wanted to know whether the MeCP2 modification site mutants show differences in heterochromatin binding kinetics. Therefore, MTF-/y cells were transfected with the different constructs, and heterochromatin compartments were photobleached using a focused laser microbeam on a confocal microscope. The fluorescence recovery was measured by taking images

before and every 1.5 s after photobleaching (figure 5.4A). Curve fitting of the intensity values over time allowed for calculation of fluorescence recovery half times and mobile fractions (figure 5.4, figure 5.17).

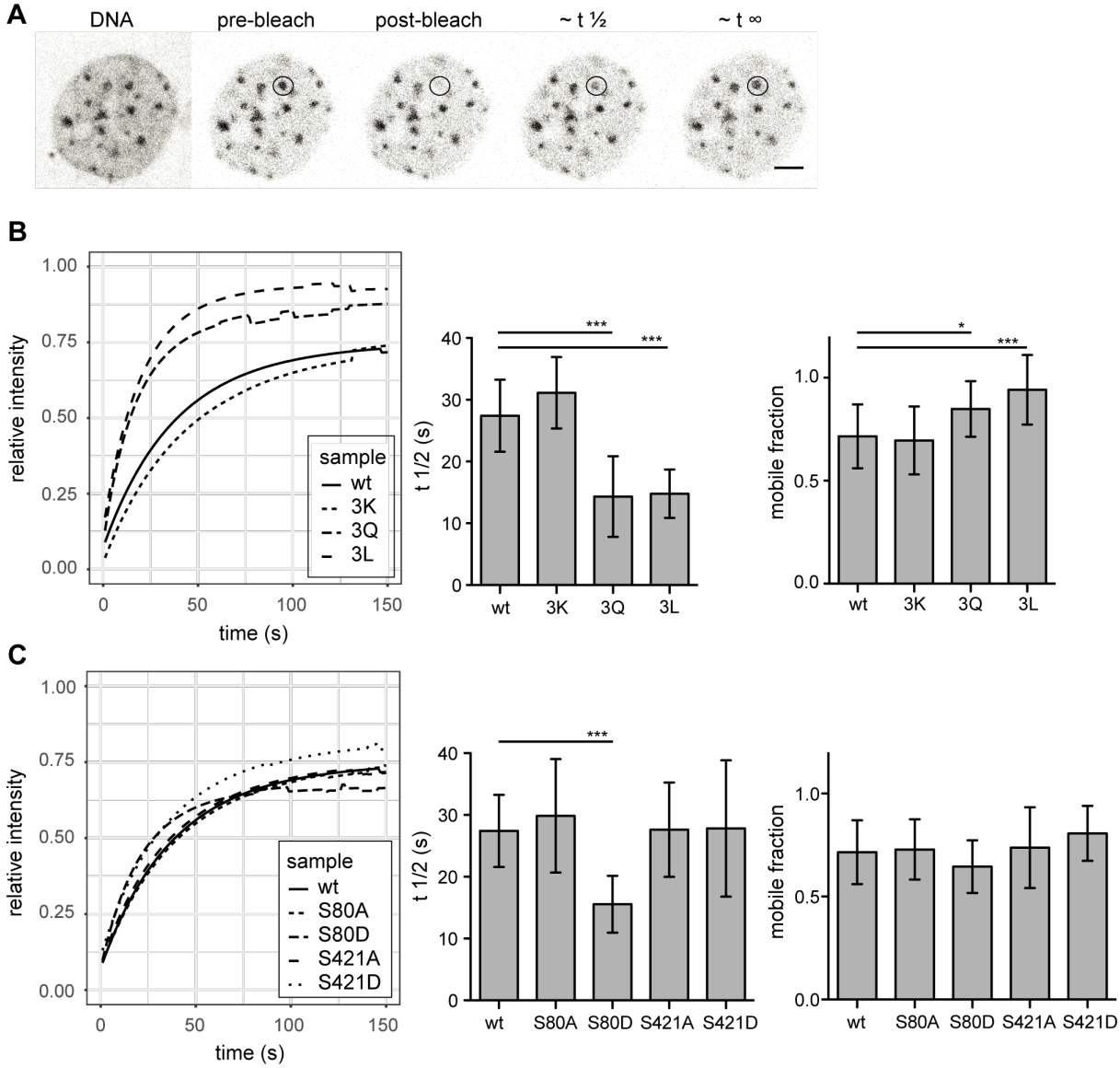



**Figure 5.4: Analysis of the dynamics of MeCP2 mutants in cells.** (A) Exemplary MTF *Mecp2* -/y cell transfected with wild type MeCP2 expression construct stained with the cell-permeable DNA dye SIR Hoechst and GFP signal is shown pre-bleaching, post-bleaching, and during fluorescence recovery. Scale bar 5 µm. (B) Fitted time curves for fluorescence recovery after photobleaching and bar diagrams showing the recovery half times (t 1/2) and mobile fractions for MeCP2 wild type (wt) and 3K, 3Q, 3L mutations including R91, R162, R167 (B) as well as for S80 and S421 mutations (C). P-values calculated by Wilcoxon-Rank test. ** $p < 0.05$, *** $p < 0.001$. P-values and n-values are summarized in table 5.5, single recovery curves with standard deviation are plotted in figure 5.17. The figure was modified from Schmidt *et al.*, 2022 [2].

The comparison of the recovery half times of wild type MeCP2 with those of the triple mutants revealed that the MeCP2 3Q and 3L mutants recover much faster than the wild type (figure

5.4B). The MeCP2 3K mutant, retaining the positive charge, showed similar kinetics as wild type MeCP2, emphasizing the importance of the positive charge for chromatin binding. These results are in line with the heterochromatin accumulation results, which depicted a slightly higher accumulation for 3K, but a lower one for 3Q and 3L constructs compared to the wild type (figure 5.3B). For the phosphorylation mutants only S80D showed faster recovery kinetics compared to wild type MeCP2, but no significant changes in the mobile fraction (figure 5.4C). This result also agrees with the heterochromatin accumulation data (figure 5.3C). As the R106 mutants were shown to hardly localize or accumulate at heterochromatin, their recovery times were too fast to be measured under similar conditions as wild type MeCP2 and were, thus, not analyzed.

### 5.3.5 Arginine methylation and serine phosphorylation site mutations of MeCP2 influence its heterochromatin clustering

To investigate the influence of MeCP2 modifications on chromatin organization, we performed a cellular chromatin clustering assay. As reported before, many small heterochromatin clusters tend to fuse to build fewer bigger clusters with increasing MeCP2 protein levels and cellular differentiation [55]. Thus, we aimed to analyze the cellular heterochromatin clustering in two ways, by observing the heterochromatin cluster number and the corresponding area (figure 5.5A), which should develop in an inverse manner. The transfected C2C12 cells were imaged on a high-content screening microscope and nuclei and heterochromatin clusters were segmented (figure 5.10). Depending on the GFP mean nuclear intensity, cells were classified into low and high MeCP2 levels and heterochromatin cluster numbers and areas were plotted (figure 5.5).
In comparison to wild type MeCP2, MeCP2 3K showed a higher heterochromatin clustering function represented by lower cluster numbers and larger cluster areas. MeCP2 3L showed a tendency to reduce heterochromatin clustering, while the results with MeCP2 3Q were not significantly different (figure 5.5B). These findings correlate well with those obtained for heterochromatin accumulation and heterochromatin binding kinetics. 3K showed higher heterochromatin accumulation, while 3Q and 3L showed lower accumulation and faster heterochromatin binding kinetics (figures 5.3B, 5.4B).
Regarding the phosphorylation site mutants, S421D showed the most striking clustering function difference to wild type MeCP2 represented by lower cluster numbers and larger cluster areas (figure 5.5C). In the other functional assays, though, MeCP2 S421D showed lower accumulation than wild type at high protein levels but did not show any significant differences in the binding kinetics (figures 5.3C, 5.4C). Hence, it is unclear how this amino acid substitution affects heterochromatin binding in relation to heterochromatin organization.
For MeCP2 arginine 106, all mutant constructs tested were associated with higher heterochromatin cluster numbers and smaller cluster areas than wild type MeCP2 (figure 5.5D), possibly because of their lack of heterochromatin accumulation.

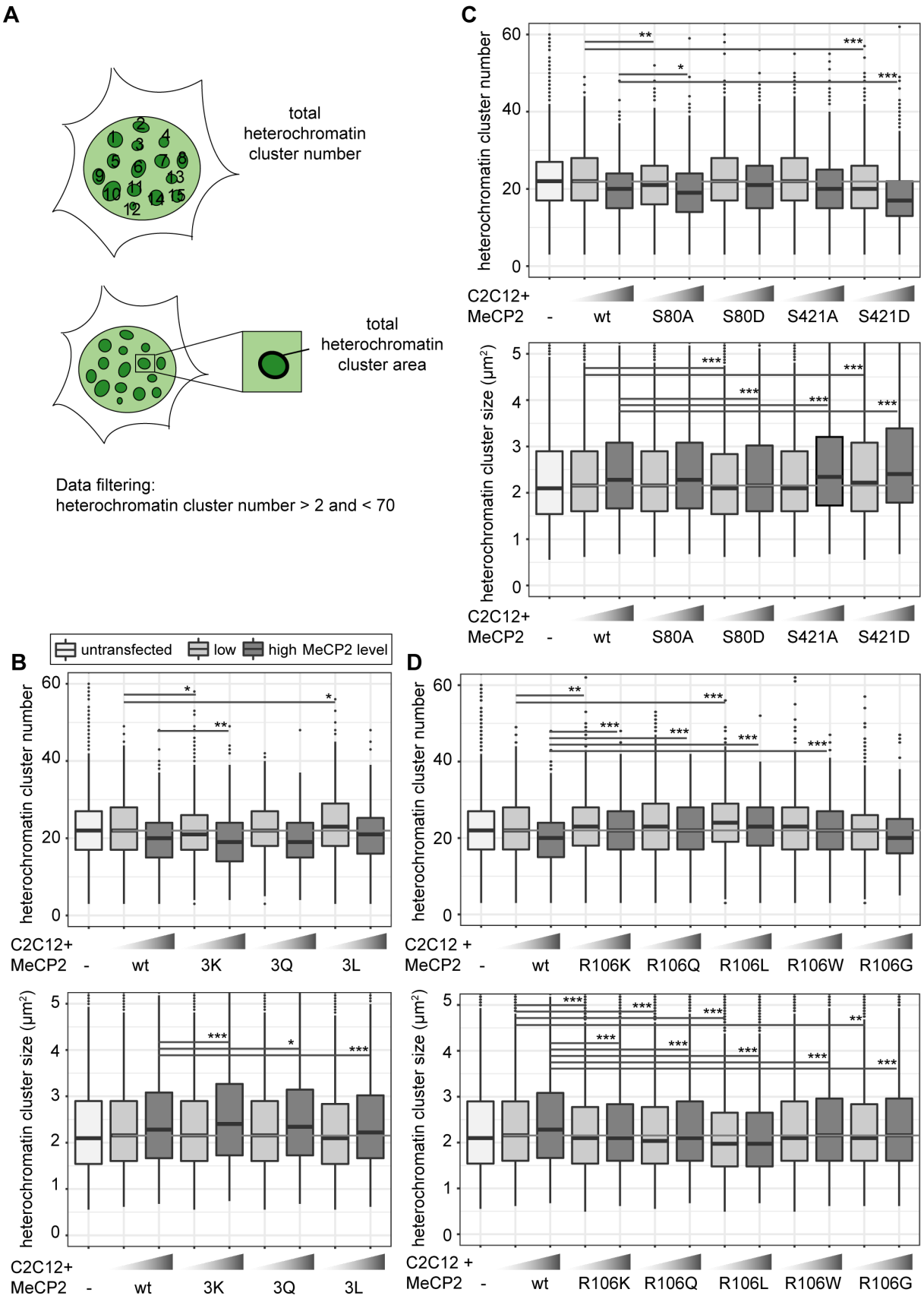



**Figure 5.5: Heterochromatin clustering and cluster size of cells transfected with MeCP2 mutant constructs.** Shown are the heterochromatin cluster numbers and areas (scheme in A) obtained

**Figure 5.5 (*previous page*):** for the arginine mutants including R91, R162, R167 (B), arginine R106 (D), and S80, S421 phosphorylation (C) mutant constructs in C2C12 cells from high-content screening microscopy after segmentation of nuclei and heterochromatin clusters. Cells were divided into low and high MeCP2 levels based on their nuclear GFP signal. Three biological replicates, statistical significance calculated using Wilcoxon-Rank test. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.001$, n.s. not significant. P-values and n-values are summarized in table 5.5. The figure was taken from Schmidt *et al.*, 2022 [2].

### 5.3.6 Protein arginine methyltransferases affect MeCP2 induced heterochromatin remodeling

As we observed changes in MeCP2 heterochromatin accumulation, clustering, and binding kinetics for the constructs mutated for arginine methylation sites, we tested whether these changes are due to the mutations inserted or a consequence of arginine methylation. Therefore, we performed coexpression experiments of protein arginine methyltransferases (*PRMTs*) with *Mecp2* to test whether the PRMTs affect the MeCP2 heterochromatin clustering function. We confirmed transfected cell viability by cell cycle profiling (figure 5.18). We made use of recombinant PRMTs with a Myc-tag that could be used for detection. The PRMTs tested comprised three enzymes that catalyze mono- and asymmetric dimethylation on arginines namely PRMT1, 4, 6, as well as PRMT5 which catalyzes mono- and symmetric arginine dimethylation. We chose PRMT1, as it is the most common arginine methyltransferase responsible for about 85% of all arginine methylations [340]. In addition, PRMT1 and PRMT6 preferentially methylate arginines in glycine and arginine-rich (GAR) motifs [340, 341] and two of the sites identified, R91 and R162, are localized adjacent to lysines. PRMT4 and PRMT1 can cooperate in gene regulation [342] but cannot substitute each other in all contexts [343]. PRMT5 is the predominant type II PRMT catalyzing symmetric arginine methylation and was associated with transcriptional repression [344]. It was reported before that the subcellular localization of the PRMT enzymes is highly dependent on the cell type and the target proteins [343]. Thus, to test for their subcellular localization, C2C12 cells were transfected with the PRMTs, fixed, and stained using an antibody against the Myc tag. Of the four PRMTs tested, PRMT1 localized in the nucleus and to a lesser extent in the cytoplasm, while PRMT6 localized exclusively in the nucleus (figure 5.6A, B), which is in line with previous studies [343, 345]. PRMT4 and PRMT5 were localized in the cytoplasm and, thus, not considered in further experiments (figure 5.19A, B). None of the PRMTs had an influence on MeCP2 localization (figure 5.6A, B, figure 5.19A, B).

First, we tested whether the PRMTs alone have an influence on heterochromatin clustering in C2C12 cells by plotting heterochromatin cluster numbers and areas (figure 5.6C, D). The values for heterochromatin cluster numbers and areas were obtained from high-content microscopy data after segmentation of nuclei and heterochromatin clusters (figure 5.10) and binning of the cells into low and high protein levels based on their nucleus mean fluorescence intensities. The presence of PRMT1 and PRMT6 resulted in higher heterochromatin cluster numbers and smaller heterochromatin cluster areas compared to untransfected cells, meaning they counteract the clustering of heterochromatin compartments. This effect was observed independent of the

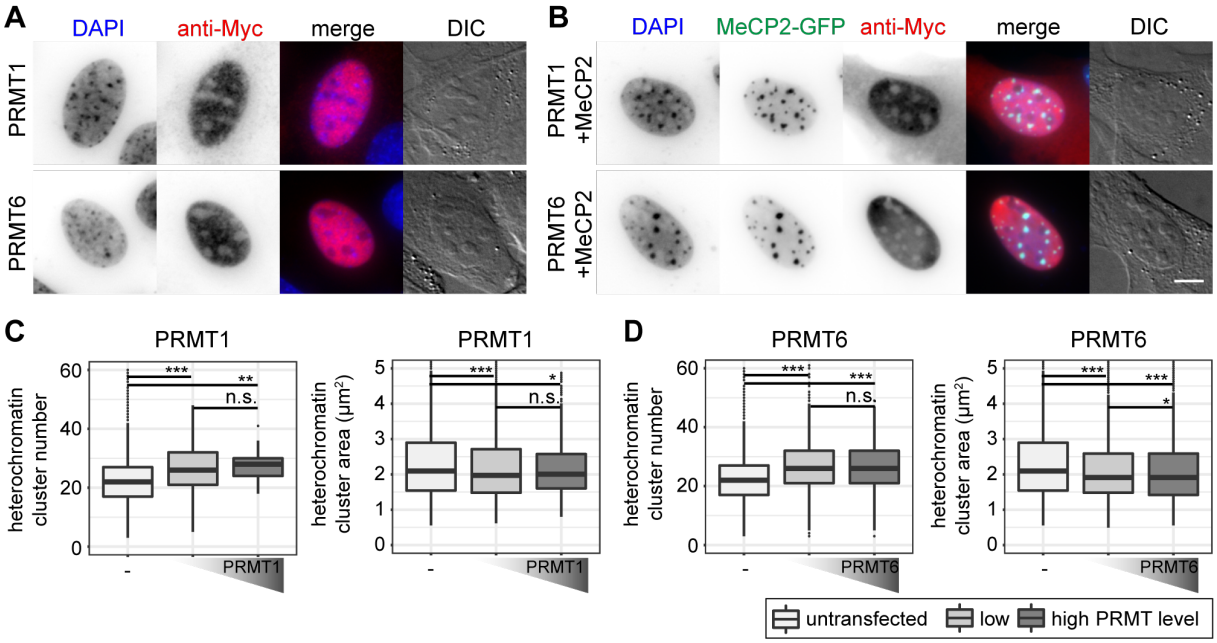



**Figure 5.6: Subcellular localization of PRMT1 and 6 constructs transfected in C2C12 mouse myoblast cells in the absence (A) and presence (B) of MeCP2.** Scale bar 5 µm. Boxplots depict the heterochromatin clustering in C2C12 cells without and with low and high levels of nuclear PRMT1 (C) and 6 (D) represented by the number and size of the heterochromatin clusters obtained from high-content screening microscopy. Two biological replicates, statistical significance calculated using Wilcoxon-Rank test. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.001$, n.s. not significant. P-values and n-values are summarized in tables 5.5, 5.7. The figure was taken from Schmidt *et al.*, 2022 [2].

PRMT level, as there was no difference in heterochromatin cluster numbers and areas between low and high PRMT levels (figure 5.6C, D)

Next, C2C12 cells were cotransfected with *PRMT1* and *PRMT6* together with *Mecp2* wild type or mutant constructs, cells were fixed and stained for the Myc tag and subsequently imaged on a high-content microscope (figure 5.10). The heterochromatin cluster numbers (figure 5.7) and areas (figure 5.20) were plotted as heatmaps for each PRMT in combination with the MeCP2 triple mutants.

Introduction of PRMT1 and PRMT6 together with MeCP2 wild type into cells resulted in a significantly decreased MeCP2 heterochromatin clustering shown by higher heterochromatin cluster numbers and smaller cluster areas (figure 5.7, figure 5.20). While MeCP2 at high levels could still cluster heterochromatin in the presence of PRMT1, there was nearly no clustering in the presence of PRMT6.

The cotransfection experiments of MeCP2 and PRMTs revealed differences in heterochromatin clustering between MeCP2 modification site mutants and MeCP2 wild type. Comparing the triple mutations to MeCP2 wild type in presence of PRMT1, MeCP2 3Q showed a similar heterochromatin cluster number distribution as the wild type, while the heatmaps of 3K and 3L differed (figure 5.7A). MeCP2 3K and 3L in presence of PRMT1 showed lower heterochromatin numbers in high levels compared to wild type, but the changes observed were not statistically

significant. The presence of PRMT6 blocked the ability of MeCP2 wild type to induce heterochromatin clustering, and the same was observed for MeCP2 3Q (figure 5.7B, figure 5.20B). MeCP2 3K, though, showed lower heterochromatin cluster numbers than wild type MeCP2 when cotransfected with PRMT6 and, thus, shows higher heterochromatin clustering. Thus, MeCP2 3K was able to reverse the negative effect of PRMT6 on heterochromatin clustering and the R to K substitution prevents its methylation by PRMTs. MeCP2 3L at high levels increased heterochromatin clustering (shown by lower heterochromatin cluster numbers and larger areas) in presence of PRMT6, but its clustering function was still impaired compared to its expression without PRMT6 (figure 5.7B, figure 5.20B). Thus, we conclude that the presence of PRMTs influences the heterochromatin clustering function of MeCP2 and its triple mutants. The differences in clustering are specifically pronounced comparing the positively charged lysine mutation with the uncharged but still polar glutamine and the non-polar leucine, emphasizing the importance of the positive charge for the MeCP2 heterochromatin clustering function.
The cotransfection experiments of MeCP2 R106 mutants and PRMT1 and PRMT6 revealed that the mutation of R106 decreases the MeCP2 heterochromatin clustering function showing higher cluster numbers and smaller areas (figure 5.7C, D, figure 5.20C, D). The R106 mutants could not counteract the reduced clustering function in the presence of PRMT1 and there was no clear difference in heterochromatin clustering between R106K, Q, and L mutants. In the presence of PRMT6 MeCP2 R106Q showed no significant heterochromatin clustering, while the presence of R106K and R106L increased heterochromatin clustering when cotransfected with PRMT6. Thus, the influence of PRMT6 on the R106 mutant heterochromatin clustering shows similar tendencies as for the 3x mutants but is less pronounced.

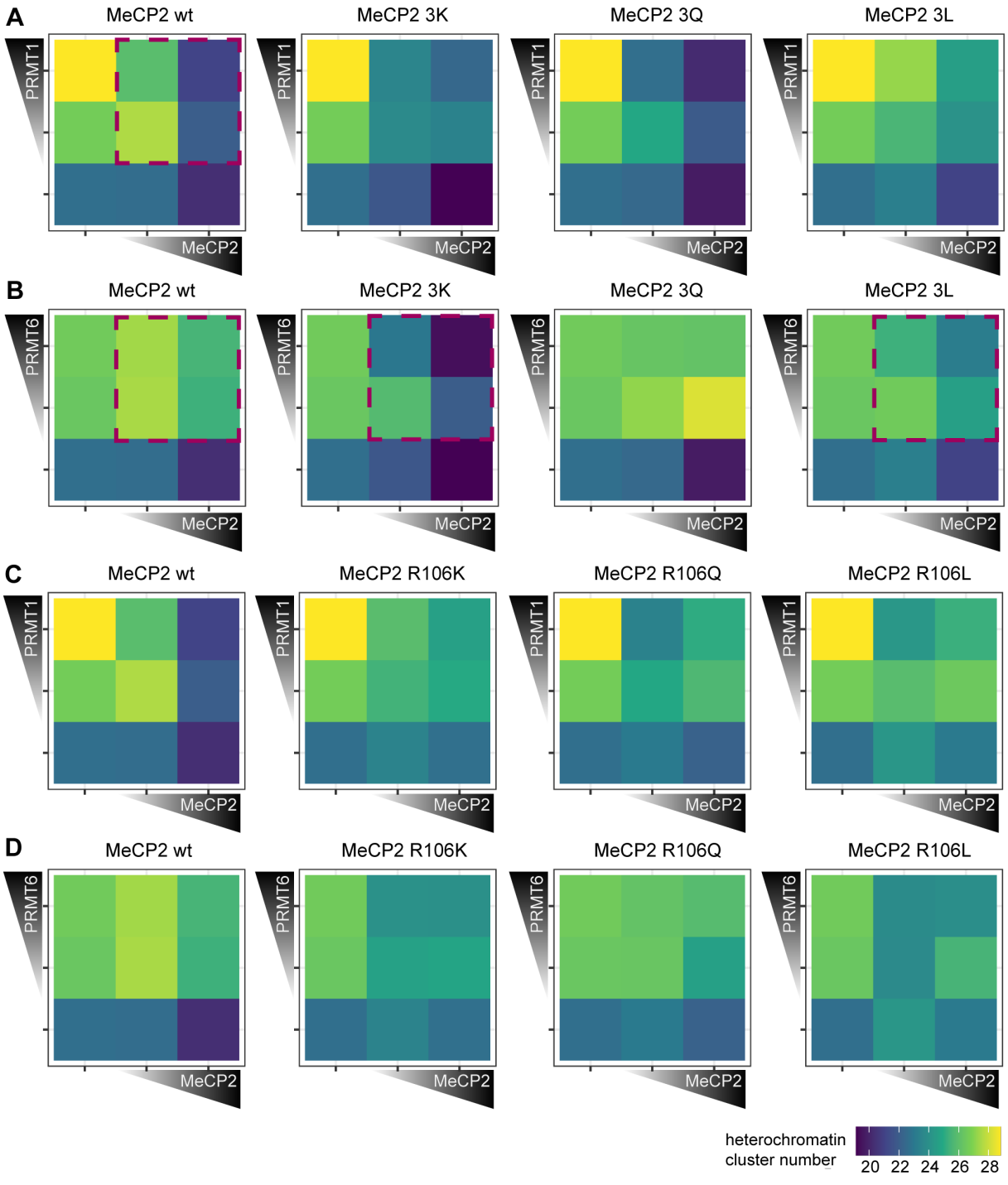



**Figure 5.7: Heterochromatin clustering of cells transfected with MeCP2 mutant constructs in the presence of PRMT 1 and 6.** Heatmaps show the heterochromatin cluster numbers obtained by high-content screening microscopy of C2C12 cells cotransfected with MeCP2 wild type or triple mutants and PRMT1 (A), wild type or triple mutants and PRMT6 (B), wild type or R106 mutants and PRMT1 (C) and wild type or R106 mutants and PRMT6 (D). Cells were binned for low and high fluorescence intensity according to their MeCP2 and PRMT levels. Heterochromatin cluster numbers are shown as means of at least 26 cells from at least two biological replicates. The p-values representing the statistical significance calculated using the Wilcoxon-Rank test are listed in table 5.5. Red boxes mark the most important observations. The figure was taken from Schmidt *et al.*, 2022 [2].

## 5.4 Discussion

MeCP2 is post-translationally modified and some of these modifications might influence its transcriptional regulation and protein-protein interactions [332], as well as its chromatin clustering abilities [181]. Although many modifications have been reported, only a few of them were functionally characterized or identified *in vivo*. In this study, we show that MeCP2 isolated from adult mouse brains is post-translationally phosphorylated on serines and threonines, methylated and acetylated on lysines, and methylated on arginines. Although MeCP2 is rich in arginines (comprising 7.2% of the amino acids), we identified only a few of them as modified by methylation. One possible explanation is the removal of the modification during the experimental procedure. Although arginine methylation is considered a rather stable and permanent modification, recent reports argue for the existence of arginine demethylation enzymes [346, 347]. On the one hand, studies involving drug treatments revealed rapid changes in arginine methylation [348–350]. On the other hand, candidate proteins catalyzing active arginine demethylation are discussed, among them, several lysine demethylases [347, 351, 352]. Thus, although we might not have identified all possible arginine methylation sites, we conclude that arginine methylation on MeCP2 occurs mostly on a few specific sites and should be tightly regulated.

As MeCP2 serine 80 and 421 phosphorylation sites are well validated but not functionally characterized in the context of heterochromatin organization, we also analyzed their heterochromatin accumulation, clustering properties, and binding kinetics (see figure 5.8). The phosphomimicking mutant S80D showed faster heterochromatin binding kinetics than wild type MeCP2, indicating reduced heterochromatin binding. By contrast, Tao *et al.* observed a decrease in MeCP2 chromatin binding affinity of the phospho-null MeCP2 S80A mutant to *Pomc* and *Gtl2* promoters by ChIP-qPCR [195]. As with this method, the authors measured MeCP2 binding to selected genomic regions and not the overall MeCP2 heterochromatin binding kinetics, we used a different method to elucidate overall (hetero)chromatin association. Thus, our results contribute to understanding the function of MeCP2 S80 phosphorylation in global heterochromatin binding. MeCP2 S80 plays a role in heterochromatin association, but not in its clustering. In contrast, the serine 421 phosphorylation mimicking mutant S421D showed increased heterochromatin clustering (with lower heterochromatin cluster numbers and larger cluster areas) compared to wild type MeCP2. As S421 phosphorylation was found exclusively in the brain upon neuronal stimulation [194, 195], we propose that clustering of heterochromatin compartments plays a role in this process.

To observe the consequences of *PRMT1* and *PRMT6*, we introduced plasmids coding for PRMT1 and 6 alone as well as together with MeCP2 in mouse myoblast cells. Solely the expression of *PRMT1* and *PRMT6* reduced the overall heterochromatin cluster size concomitantly increasing their numbers. PRMT1 mostly acts as a coactivator of transcription [340, 353], whereas PRMT6 mostly acts as transcriptional repressor [354–357], but it was also reported to contribute to gene activation [358]. Interestingly, cells from cancer patients showed higher *PRMT1* and *PRMT6* expression than cells from healthy tissue, which seemed to be beneficial for tumor growth [359]. Furthermore, *PRMT6* upregulation was found to correlate with DNA hypomethylation in mESCs

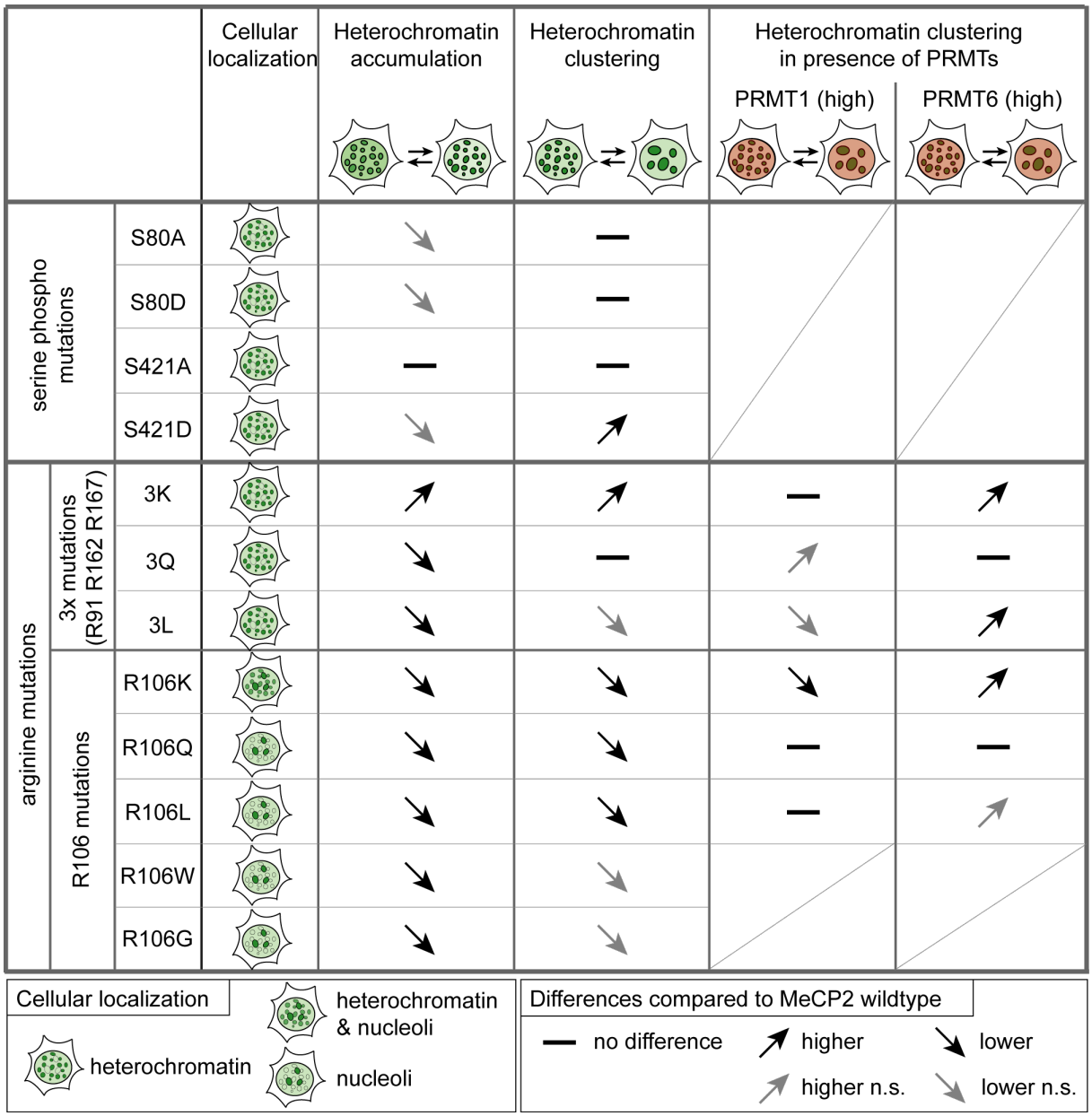



**Figure 5.8: Summary of the functional characterization of MeCP2-induced heterochromatin organization.** Functional differences of the MeCP2 PTM mutants in comparison to the wild type protein are represented as arrows showing an increase (pointing up) and decrease (pointing down), if not statistically significant (n.s.) in gray, and no difference is marked by a line. Triple mutations stand for R91/R162/R167 mutations. Heterochromatin accumulation and clustering are MeCP2 dose-dependent. For the heterochromatin clustering in presence of PRMT1 and 6, only high PRMT levels were considered. The figure was taken from Schmidt *et al.*, 2022 [2].

and MCF7 cells [360], providing a possible explanation for the reduced heterochromatin clustering we observed upon *PRMT6* overexpression. Coexpression of *Mecp2* wild type and *PRMT1* and *PRMT6* lead to significantly decreased MeCP2-mediated heterochromatin clustering (indicated by higher cluster numbers and smaller areas), suggesting that a high degree of arginine methylation in the cells drastically impairs MeCP2 clustering function. This effect was more

pronounced for PRMT6 than for PRMT1, although PRMT1 is considered responsible for the majority of cellular arginine methylation [361]. Thus, PRMT6 might have a higher affinity for MeCP2 than PRMT1. Moreover, *PRMT1* and *PRMT6* expression was reported to depend on the MeCP2 level in a neuroblastoma cell line, thus suggesting a positive gene regulatory interaction between MeCP2 and these two genes [362].

The overall changes in heterochromatin clustering upon *PRMT* overexpression could be explained either by overall higher arginine methylation levels or by increased arginine methylation levels of MeCP2. To distinguish between these possibilities we made use of the MeCP2 arginine point mutations, which are not modifiable by the PRMTs. For this purpose, we generated MeCP2 constructs mutated for the modification sites identified (R91, R162, R167) substituting the arginine with lysine (3K), glutamine (3Q), or leucine (3L). Although none of these substitutions can truly mimic the different arginine (modification) states, lysine retains the positive charge as (methylated) arginine, the polar glutamine should sterically mimic a methylated arginine and the unpolar leucine should mimic the methyl groups of methylated arginine (figure 5.16C, D). Furthermore, arginine substitutions by glutamine and leucine were also found in Rett syndrome patients. Arginine substitution with the positively charged lysine increased heterochromatin accumulation and clustering in comparison to wild type MeCP2. In contrast, substitutions with glutamine and leucine reduced heterochromatin accumulation and lead to faster heterochromatin binding kinetics effectively reducing the t 1/2 to half of the one obtained with wild type MeCP2. These results indicate the importance of the positive charge on R91, R162, and R167 for MeCP2 heterochromatin accumulation and clustering but especially for heterochromatin binding kinetics. Of note, MeCP2 3Q seems to be the best MeCP2 mimic for heterochromatin clustering emphasizing that its polarity and steric properties are more similar to those of wild type MeCP2. The absence of charge and polarity clearly impacts all functional properties of MeCP2 tested here as seen with the MeCP2 3L mutant. Thus, we hypothesize that MeCP2 is methylated in the brain at any given time at least at one of the arginine methylation sites identified and that this modification partially changes the positive charge distribution. Thus, methylated arginines might show similar properties as polar amino acids. In fact, it was described that arginine methylation alters the charge distribution to more diffuse (especially in the case of dimethylation) but still positive electrostatic properties [363, 364]. In addition, methylation changes arginine shape and reduces the number of possible hydrogen bonds [346, 365]. Cotransfection experiments of MeCP2 triple mutants and various PRMTs revealed differences in heterochromatin clustering between the mutants and the wild type protein, further strengthening the evidence for arginine methylation on one or more of the identified modification sites. In the presence of PRMT6, MeCP2 3K and 3L showed significantly higher heterochromatin clustering function represented by lower heterochromatin cluster numbers and larger areas than wild type MeCP2. MeCP2 3Q hardly induced any clustering, similar to MeCP2 wild type. MeCP2 3K showed higher clustering abilities (lower heterochromatin numbers and larger areas) than wild type MeCP2 and was able to reverse the effect of PRMT6 alone, which induced reduced clustering of heterochromatin. This effect might not be a direct result of the positive charge on heterochromatin clustering, but rather an indirect one as MeCP2 3K cannot be methylated by PRMTs on the mutated arginines. In comparison

to the wild type protein, the MeCP2 3L mutant showed higher clustering when PRMT6 was introduced, although it showed decreased clustering without PRMT6. These results suggest a very high arginine methylation level of wild type MeCP2 in the coexpression experiment, which decreases its clustering ability to such an extent that even the MeCP2 3L mutant clusters more than the methylated wild type protein. From the increased heterochromatin clustering functions of the 3K and 3L mutant, which cannot be methylated on the substituted arginines, we conclude that MeCP2 gets methylated by PRMT6 on these sites. Of note, arginine methylation catalyzed by the PRMTs 1 and 6 often takes place on arginines flanked by one or more glycines in so-called glycine and arginine-rich (GAR) motifs [364] and MeCP2 R91 and R162 are localized adjacent to glycines. Although our results emphasize that MeCP2 gets methylated on arginines and in consequence shows reduced heterochromatin clustering abilities, we cannot exclude that consequences of high arginine methylation levels also indirectly impact MeCP2 heterochromatin clustering. Examples could be the modification of MeCP2 interacting proteins or other proteins involved in heterochromatin clustering, e.g. histones. Interestingly, MBD2, another member of the methyl-CpG binding protein family, was shown to undergo arginine methylation, which resulted in reduced DNA binding and reduced functionality in transcriptional repression [366].

MeCP2 R106 was identified as dimethylated in our mass spectrometry analysis and is commonly mutated in Rett syndrome patients to tryptophan (W) and glutamine (Q), in very few cases also to glycine (G) and leucine (L). MeCP2 R106W is a frequent Rett syndrome mutation causing severe phenotypes [367] and the less common R106Q mutation was described to cause "classic" Rett syndrome [368–370], but there is insufficient clinical information reported for individuals with R106Q, R106G and R106L for a comparison of phenotypes (see RettBASE,[337]). R106W was reported to abolish DNA binding, while R106Q reduced it [240, 338]. Accordingly, our experiments showed that all R106 mutants mainly lost heterochromatin accumulation and mislocalized to the negatively charged nucleoli due to the high amount of positively charged amino acids in MeCP2 [339]. The reduced DNA binding and accumulation can be explained by the location of R106 close to the Asx-ST motif, which stabilizes MeCP2 DNA binding ([136], figure 5.3E). H/DX experiments revealed similar dynamic protein behavior of MeCP2 R106W and wild type protein [137] and circular dichroism spectra of R106W/Q showed no major changes in secondary structure compared to wild type MeCP2 [240, 338]. Instead, molecular modeling of the MeCP2 R106W/Q structures pointed towards local changes of hydrogen bonds and salt bridges [338, 371, 372], which might cause changes in DNA binding as R106 is part of a ß-strand in the MBD structure, stabilizes the Asx-ST motif and is buried and not exposed to the surrounding [371, 372]. Thus, MeCP2 R106W/Q mutations have not been found to induce changes in overall MeCP2 structure but rather result in smaller local changes in amino acid interactions. The heterochromatin clustering was highly impaired in the MeCP2 R106 mutants as well. The reason might be the lack of binding to pericentromeric heterochromatin regions as the clustering abilities of some Rett mutants could be rescued by repositioning of the proteins to the heterochromatin regions [63]. In addition, we recently showed that heterochromatin clustering *in vivo* can be modeled by *in vitro* phase separation [1]. The minimal basis for MeCP2 liquid-liquid phase separation was electrostatic self-interaction, but also DNA promoted *de novo* phase separation of MeCP2

in physiological salt conditions [1]. Thus, DNA binding, as well as oligomerization via its ID and TRD domain [164] is involved in MeCP2 heterochromatin clustering. From the cotransfection experiments of PRMTs and MeCP2 R106 mutants, it could be hypothesized that R106 is more likely to be methylated by PRMT6 than by PRMT1, as PRMT6 presence affected the clustering by MeCP2 R106 mutants. Altogether, our results demonstrate that post-translational modifications of MeCP2, in particular arginine methylation and to a lesser extent serine phosphorylation, play an essential role in modulating MeCP2 function in heterochromatin organization.

## 5.5 Supplementary material

### Supplementary tables

**Table 5.1:** Plasmid characteristics.

| Name | pc number* | Fluorescent protein | (Gene) species | Promoter | References |
|---|---|---|---|---|---|
| pmMeCP2G wt | 4701 | EGFP | *Mus musculus* synthetic | CAG TRE3G | Schmidt *et al.*, 2022 [2] |
| pmMeCP2G R91K R162K R167K (3K) | 4702 | EGFP | *Mus musculus* synthetic | CAG TRE3G | Schmidt *et al.*, 2022 [2] |
| pmMeCP2G R91Q R162Q R167Q (3Q) | 4703 | EGFP | *Mus musculus* synthetic | CAG TRE3G | Schmidt *et al.*, 2022 [2] |
| pmMeCP2G R91L R162L R167L (3L) | 4704 | EGFP | *Mus musculus* synthetic | CAG TRE3G | Schmidt *et al.*, 2022 [2] |
| pmMeCP2G S80A | 4705 | EGFP | *Mus musculus* synthetic | CAG TRE3G | Schmidt *et al.*, 2022 [2] |
| pmMeCP2G S80D | 4706 | EGFP | *Mus musculus* synthetic | CAG TRE3G | Schmidt *et al.*, 2022 [2] |
| pmMeCP2G S421A | 4707 | EGFP | *Mus musculus* synthetic | CAG TRE3G | Schmidt *et al.*, 2022 [2] |
| pmMeCP2G S421D | 4708 | EGFP | *Mus musculus* synthetic | CAG TRE3G | Schmidt *et al.*, 2022 [2] |
| pmMeCP2G R106K | 4718 | EGFP | *Mus musculus* synthetic | CAG TRE3G | Schmidt *et al.*, 2022 [2] |
| pmMeCP2G R106Q | 4719 | EGFP | *Mus musculus* synthetic | CAG TRE3G | Schmidt *et al.*, 2022 [2] |
| pmMeCP2G R106L | 4724 | EGFP | *Mus musculus* synthetic | CAG TRE3G | Schmidt *et al.*, 2022 [2] |
| pmMeCP2G R106W | 4721 | EGFP | *Mus musculus* synthetic | CAG TRE3G | Schmidt *et al.*, 2022 [2] |
| pmMeCP2G R106G | 4722 | EGFP | *Mus musculus* synthetic | CAG TRE3G | Schmidt *et al.*, 2022 [2] |
| phPRMT1-pcDNA3.1** | 4795 | - | *Homo sapiens* | CMV | Schmidt *et al.*, 2022 [2] |
| pmPRMT4-pcDNA3.1** | 4796 | - | *Mus musculus* | CMV | Schmidt *et al.*, 2022 [2] |
| phPRMT5-pcDNA3.1** | 4797 | - | *Homo sapiens* | CMV | Schmidt *et al.*, 2022 [2] |
| phPRMT6-pcDNA3.1*** | 4798 | - | *Homo sapiens* | CMV | Stein *et al.*, 2012 [354] |

*pc number: plasmid collection number. **plasmid naming according to Cardoso lab plasmid collection. ***originally published as pcDNA3.1 hPRMT6.

**Table 5.2:** Cell line characteristics.

| Name | Species | Type | Genotype | References |
|---|---|---|---|---|
| C2C12 | *Mus musculus* | myoblast | wild type | Yaffe & Saxel, 1977 [327] |
| MTF -/y | *Mus musculus* | tail fibroblast | Mecp2 -/y | Guy 2001 [244] |

**Table 5.3:** Primary and secondary antibody characteristics.

| Reactivity | Host | Dilution | Application | Catalog / clone | Company / reference |
|---|---|---|---|---|---|
| Anti-MeCP2 | rabbit | 1:500 | WB | - | Jost *et al.*, 2011 [326] |
| Anti-MeCP2 | rat | Undiluted TCSN | IF / WB | 4H7, 4G10, 4E1 | Jost *et al.*, 2011 [326] |
| Anti-mono methyl arginine (MMA) | rabbit | 1:1000 | WB | #8015 | Cell Signaling Technology |
| Anti-symmetric dimethyl arginine (SDMA) | rabbit | 1:1000 | WB | #13222 | Cell Signaling Technology |
| Anti-asymmetric dimethyl arginine (ADMA) | rabbit | 1:1000 | WB | #13522 | Cell Signaling Technology |
| Anti-phospho MECP2 Ser80 | rabbit | 1:1000 | WB | #P21953 | Molecular Probes, Inc. |
| Anti-phospho MECP2 Ser421 | rabbit | 1:1000 | WB | #PA5-35396 | Thermo Fisher Scientific |
| Anti-phospho tyrosine | mouse | 1:500 | WB | #3630 | Clontech Laboratories, Inc. |
| Anti-Myc tag | mouse | 1:1000 | IF | #ab32 / 9E10 | abcam |
| Anti-rat IgG Cy3 | donkey | 1:300 / 1:1000 | IF / WB | #712-165-153 | Jackson ImmunoResearch Laboratories, inc. |
| Anti-mouse IgG Cy5 | donkey | 1:250 | IF | #715-175-150 | Jackson ImmunoResearch Laboratories, inc. |
| Anti-rabbit IgG HRP | goat | 1:10000 | WB | #A0545 | Sigma-Aldrich, Inc. |
| Anti-mouse IgG HRP | sheep | 1:5000 | WB | #NA 931 | Amersham Pharmacia Biotech |

IF: Immunofluorescence; WB: Western blot; HRP: horseradish peroxidase; TCSN: Tissue culture supernatant.

**Table 5.4:** Imaging system characteristics.

| Microscope/ Company | Lasers/ lamps | Filters (ex. & em. (nm)) | Objectives/ lenses | Detection system | Incubation system | Application |
|---|---|---|---|---|---|---|
| Nikon CREST Eclipse TiE2, Nikon, Tokyo, Japan | SPECTRA X LED 470/24 nm (196 mW), 550/15 nm (260 mW), 640/30 nm (231 mW) | em.: Quad-bandpass (432/25 nm; 515/25 nm; 595/25 nm; 730/70 nm) | 40x air Plan Apo λ DIC (0.95 NA, 250 µm WD) | Nikon Qi2 751600 16.25 MPx | - | Fluorescence imaging + DIC |
| Confocal microscope Leica TCS SPEII, Wetzlar, Germany | Multicolor solid-state laser module RYBV 405 nm / 25 mW 488 nm / 10 mW | em.: DAPI: ex. 360/40, em. LP 425 GFP: ex. 470/40, em.: LP515 | oil immersion 63X ACS APO CS (1.3 NA) | Leica SP-Detector adjustable in the range of 430 - 750 nm | - | Fluorescence imaging of MTF -/y cells |
| Widefield Axiovert 200 /Zeiss, Germany | HBO100 mercury lamp | DAPI (300-400 & 410-510); GFP (473-491 & 506-534); | oil immersion 63x Plan-Apochromat (1.4 NA) | 12-bit AxioCam mRM | - | Fluorescence imaging for calculation of hete-rochromatin accumula-tion |
| Operetta high content screening microscopy/ PerkinElmer Life Sciences, UK | Xenon fiber-optic light source, 300 W, 360 – 640 nm continuous spectrum | DAPI: 360-400 & 410–480 GFP: 460-490 & 500–550 Cy5: 620-640 & 650-760 | 40x air (0.95 NA) long WD | 14-bit Jenoptik CMOS | - | high content screening microscopy |
| Confocal microscope Leica SP5 II, Wetzlar, Germany | 488 nm Argon ion laser, 633 nm HeNe gas laser 20 mW | AOBS beam splitter | HCX PL APO 63x / 1.4-0.6 oil lambda blue | 2 HyD Hybrid Detectors | ACU live cell chamber (Olym-pus), inverse DMI 600 stand | FRAP |
| Amersham AI600 imager/GE Healthcare, Chicago, IL, USA | White light (trans) Chemi-luminescence, fluorescence | Cy2: 525BP20, Cy3/ EtBr: 605BP40, Cy5: 705BP40 | - | 16-bit Peltier cooled Fujifilm Super CCD | - | Western blot, SDS-PAGE imaging |

ex.: excitation; em.: emission; WD: working distance; LP: long pass; BP: band pass.

**Table 5.5:** Plot statistics (main figures).

| Figure | Sample | n | Median | Mean | StDev | 95% CI | p-value |
|---|---|---|---|---|---|---|---|
| 5.3B | C2C12 untransfected | 200 | 1.00 | 1.00 | 0.03 | - | - |
| | MeCP2 wt (low) | 78 | 1.82 | 1.48 | 0.79 | 0.01 | - |
| | MeCP2 wt (high) | 100 | 1.68 | 1.46 | 0.70 | - | - |
| | MeCP2 3K (low) | 94 | 1.94 | 1.61 | 1.03 | 0.01 | 4.17e-05 |
| | MeCP2 3K (high) | 81 | 1.72 | 1.66 | 0.76 | 0.01 | 1.06e-03 |
| | MeCP2 3Q (low) | 92 | 1.67 | 1.43 | 0.65 | - | 7.55e-07 |
| | MeCP2 3Q (high) | 75 | 1.50 | 1.37 | 0.61 | - | 5.35e-06 |
| | MeCP2 3L (low) | 112 | 1.46 | 1.22 | 0.53 | - | $<$2.2e-16 |
| | MeCP2 3L (high) | 97 | 1.45 | 1.22 | 0.46 | - | $<$2.2e-16 |
| 5.3C | MeCP2 S80A (low) | 91 | 1.76 | 1.47 | 0.84 | 0.01 | 0.04144 |
| | MeCP2 S80A (high) | 64 | 1.66 | 1.45 | 0.73 | 0.01 | 0.4674 |
| | MeCP2 S80D (low) | 82 | 1.66 | 1.38 | 0.76 | 0.01 | 3.66e-07 |
| | MeCP2 S80D (high) | 69 | 1.63 | 1.36 | 0.72 | 0.01 | 0.487 |
| | MeCP2 S421A (low) | 132 | 1.82 | 1.51 | 0.77 | - | 0.9957 |
| | MeCP2 S421A (high) | 88 | 1.58 | 1.41 | 0.71 | - | 0.06161 |
| | MeCP2 S421D (low) | 112 | 1.89 | 1.37 | 0.71 | - | 0.1606 |
| | MeCP2 S421D (high) | 73 | 1.50 | 1.35 | 0.71 | 0.01 | 2.37e-06 |
| 5.3D | MeCP2 R106K (low) | 49 | 1.17 | 0.87 | 0.26 | - | $<$2.2e-16 |
| | MeCP2 R106K (high) | 61 | 1.15 | 1.00 | 0.29 | - | $<$2.2e-16 |
| | MeCP2 R106Q (low) | 64 | 1.05 | 0.80 | 0.16 | - | $<$2.2e-16 |
| | MeCP2 R106Q (high) | 55 | 1.02 | 0.82 | 0.12 | - | $<$2.2e-16 |
| | MeCP2 R106L (low) | 63 | 1.02 | 0.80 | 0.18 | - | $<$2.2e-16 |
| | MeCP2 R106L (high) | 74 | 1.01 | 0.79 | 0.15 | - | $<$2.2e-16 |
| | MeCP2 R106W (low) | 86 | 1.05 | 0.80 | 0.19 | - | $<$2.2e-16 |
| | MeCP2 R106W (high) | 55 | 1.04 | 0.83 | 0.18 | - | $<$2.2e-16 |
| | MeCP2 R106G (low) | 53 | 1.05 | 0.86 | 0.19 | - | $<$2.2e-16 |
| | MeCP2 R106G (high) | 54 | 1.04 | 0.83 | 0.19 | - | $<$2.2e-16 |
| 5.4B T 1/2 | MeCP2 wt | 17 | 26.97 | 27.40 | 5.83 | 0.90 | - |
| | MeCP2 3K | 11 | 31.52 | 31.13 | 5.79 | 0.11 | 0.1003 |
| | MeCP2 3Q | 16 | 14.28 | 14.33 | 6.50 | 0.10 | 3.56e-06 |
| | MeCP2 3L | 17 | 15.19 | 14.80 | 3.92 | 0.06 | 5.742e-08 |
| 5.4B mobile fraction | MeCP2 wt | 17 | 0.69 | 0.72 | 0.16 | - | - |
| | MeCP2 3K | 11 | 0.69 | 0.70 | 0.17 | - | 0.8691 |
| | MeCP2 3Q | 16 | 0.81 | 0.85 | 0.13 | - | 0.01422 |
| | MeCP2 3L | 17 | 0.91 | 0.94 | 0.17 | - | 0.001197 |
| 5.4C T 1/2 | MeCP2 S80A | 10 | 27.26 | 29.86 | 9.17 | 0.19 | 0.7486 |
| | MeCP2 S80D | 16 | 16.36 | 15.56 | 4.59 | 0.07 | 2.383e-07 |
| | MeCP2 S421A | 13 | 28.19 | 27.61 | 7.63 | 0.14 | 0.8047 |
| | MeCP2 S421D | 12 | 23.83 | 27.82 | 11.04 | 0.20 | 0.7438 |
| 5.4C mobile fraction | MeCP2 S80A | 10 | 0.69 | 0.73 | 0.15 | - | 0.615 |
| | MeCP2 S80D | 16 | 0.625 | 0.65 | 0.13 | - | 0.1708 |
| | MeCP2 S421A | 13 | 0.72 | 0.74 | 0.20 | - | 0.7534 |
| | MeCP2 S421D | 12 | 0.785 | 0.81 | 0.13 | - | 0.07998 |
| 5.5B number | C2C12 untransfected | 6787 | 22 | 22.68 | 8.12 | 0.01 | - |
| | MeCP2 wt (low) | 951 | 22 | 22.59 | 7.83 | 0.02 | - |
| | MeCP2 wt (high) | 704 | 20 | 20.32 | 7.21 | 0.02 | - |
| | MeCP2 3K (low) | 851 | 21 | 21.74 | 7.89 | 0.02 | 0.01342 |
| | MeCP2 3K (high) | 549 | 19 | 19.12 | 7.24 | 0.02 | 0.002595 |
| | MeCP2 3Q (low) | 493 | 22 | 22.44 | 6.96 | 0.02 | 0.9712 |
| | MeCP2 3Q (high) | 310 | 19 | 19.72 | 6.42 | 0.02 | 0.2815 |
| | MeCP2 3L (low) | 983 | 23 | 23.31 | 8.09 | 0.02 | 0.02703 |

| Figure | Sample | n | Median | Mean | StDev | 95% CI | p-value |
|---|---|---|---|---|---|---|---|
| | MeCP2 3L (high) | 604 | 21 | 21.03 | 7.30 | 0.02 | 0.05115 |
| 5.5B area | C2C12 untransfected | 6787 | 2.10 | 2.37 | 1.16 | - | - |
| | MeCP2 wt (low) | 951 | 2.16 | 2.39 | 1.19 | - | - |
| | MeCP2 wt (high) | 704 | 2.28 | 2.57 | 1.26 | - | - |
| | MeCP2 3K (low) | 851 | 2.16 | 2.41 | 1.21 | - | 0.1795 |
| | MeCP2 3K (high) | 549 | 2.40 | 2.69 | 1.35 | - | 1.33e-12 |
| | MeCP2 3Q (low) | 493 | 2.16 | 2.40 | 1.14 | - | 0.2373 |
| | MeCP2 3Q (high) | 310 | 2.34 | 2.59 | 1.22 | - | 0.01024 |
| | MeCP2 3L (low) | 983 | 2.10 | 2.35 | 1.11 | - | 0.2251 |
| | MeCP2 3L (high) | 604 | 2.22 | 2.51 | 1.21 | - | 0.004491 |
| 5.5C number | MeCP2 S80A (low) | 1271 | 21 | 21.51 | 7.72 | 0.01 | 0002977 |
| | MeCP2 S80A (high) | 793 | 19 | 19.52 | 7.64 | 0.02 | 0.02728 |
| | MeCP2 S80D (low) | 1414 | 22 | 22.71 | 8.26 | 0.01 | 0.6859 |
| | MeCP2 S80D (high) | 886 | 21 | 20.98 | 8.14 | 0.02 | 0.1388 |
| | MeCP2 S421A (low) | 2403 | 22 | 22.59 | 8.06 | 0.01 | 0.7858 |
| | MeCP2 S421A (high) | 1440 | 20 | 20.23 | 7.53 | 0.01 | 0.5655 |
| | MeCP2 S421D (low) | 1105 | 20 | 20.83 | 8.12 | 0.02 | 1.22e-07 |
| | MeCP2 S421D (high) | 840 | 17 | 18.31 | 7.54 | 0.02 | 3.86e-10 |
| 5.5C area | MeCP2 S80A (low) | 1271 | 2.16 | 2.38 | 1.15 | - | 0.815 |
| | MeCP2 S80A (high) | 793 | 2.28 | 2.56 | 1.27 | - | 0.4688 |
| | MeCP2 S80D (low) | 1414 | 2.10 | 2.33 | 1.13 | - | 2.02e-10 |
| | MeCP2 S80D (high) | 886 | 2.16 | 2.49 | 1.31 | - | 7.54e-14 |
| | MeCP2 S421A (low) | 2403 | 2.10 | 2.40 | 1.16 | - | 0.5981 |
| | MeCP2 S421A (high) | 1440 | 2.34 | 2.63 | 1.30 | - | 4.46e-06 |
| | MeCP2 S421D (low) | 1105 | 2.22 | 2.53 | 1.29 | - | $<$2.2e-16 |
| | MeCP2 S421D (high) | 840 | 2.40 | 2.78 | 1.49 | - | $<$2.2e-16 |
| 5.5D number | MeCP2 R106K (low) | 1368 | 23 | 23.46 | 7.62 | 0.01 | 0.004146 |
| | MeCP2 R106K (high) | 567 | 22 | 22.71 | 7.20 | 0.02 | 1.41e-09 |
| | MeCP2 R106Q (low) | 1434 | 23 | 23.09 | 8.75 | 0.01 | 0.2508 |
| | MeCP2 R106Q (high) | 375 | 22 | 22.15 | 7.80 | 0.03 | 2.59e-05 |
| | MeCP2 R106L (low) | 1050 | 24 | 24.24 | 8.26 | 0.02 | 1.72e-06 |
| | MeCP2 R106L (high) | 167 | 23 | 23.10 | 8.00 | 0.04 | 5.29e-06 |
| | MeCP2 R106W (low) | 1450 | 23 | 22.90 | 8.46 | 0.01 | 0.3898 |
| | MeCP2 R106W (high) | 738 | 22 | 21.83 | 7.75 | 0.02 | 4.07e-05 |
| | MeCP2 R106G (low) | 1106 | 22 | 22.04 | 7.66 | 0.01 | 0.08368 |
| | MeCP2 R106G (high) | 334 | 20 | 20.57 | 6.88 | 0.02 | 0.398 |
| 5.5D area | MeCP2 R106K (low) | 1368 | 2.10 | 2.32 | 1.09 | - | 2.54e-10 |
| | MeCP2 R106K (high) | 567 | 2.10 | 2.37 | 1.11 | - | $<$2.2e-16 |
| | MeCP2 R106Q (low) | 1434 | 2.03 | 2.32 | 1.14 | - | 4.29e-16 |
| | MeCP2 R106Q (high) | 375 | 2.10 | 2.37 | 1.15 | - | $<$2.2e-16 |
| | MeCP2 R106L (low) | 1050 | 1.97 | 2.21 | 1.02 | - | $<$2.2e-16 |
| | MeCP2 R106L (high) | 167 | 1.97 | 2.22 | 1.08 | - | $<$2.2e-16 |
| | MeCP2 R106W (low) | 1450 | 2.10 | 2.41 | 1.22 | - | 0.4714 |
| | MeCP2 R106W (high) | 738 | 2.16 | 2.42 | 1.21 | - | $<$2.2e-16 |
| | MeCP2 R106G (low) | 1106 | 2.10 | 2.36 | 1.13 | - | 0.005653 |
| | MeCP2 R106G (high) | 334 | 2.16 | 2.44 | 1.19 | - | 2.98e-12 |
| 5.6C number | PRMT1 (low) | 126 | 26 | 26.71 | 8.91 | 0.05 | 1.09e-07 |
| | PRMT1 (high) | 21 | 28 | 27.62 | 5.71 | 0.08 | 0.001045 |
| 5.6C area | PRMT1 (low) | 126 | 1.97 | 2.24 | 1.11 | - | 1.2e-14 |
| | PRMT1 (high) | 21 | 2.00 | 2.20 | 1.07 | - | 0.01079 |
| 5.6D number | PRMT6 (low) | 3146 | 26 | 26.33 | 8.19 | 0.01 | $<$2.2e-16 |
| | PRMT6 (high) | 682 | 26 | 26.57 | 8.10 | 0.02 | $<$2.2e-16 |
| 5.6D area | PRMT6 (low) | 3146 | 1.91 | 2.15 | 0.99 | - | $<$2.2e-16 |

| Figure | Sample | n | Median | Mean | StDev | 95% CI | p-value |
|---|---|---|---|---|---|---|---|
| | PRMT6 (high) | 682 | 1.91 | 2.13 | 0.98 | - | $<$2.2e-16 |
| 5.7A | PRMT1 (high), MeCP2 wt (low) | 29 | 27 | 27.93 | 7.63 | 0.09 | - |
| | PRMT1 (high), MeCP2 wt (high) | 232 | 20 | 22.12 | 8.93 | 0.04 | - |
| | PRMT1 (low), MeCP2 wt (low) | 52 | 27.5 | 26.98 | 9.5 | 0.08 | - |
| | PRMT1 (low), MeCP2 wt (high) | 282 | 22 | 22.75 | 8.74 | 0.03 | - |
| | PRMT1 (high), MeCP2 3K (low) | 26 | 24.5 | 23.58 | 8.83 | 0.11 | 0.0806 |
| | PRMT1 (high), MeCP2 3K (high) | 144 | 21 | 22.18 | 9.08 | 0.05 | 0.7324 |
| | PRMT1 (low), MeCP2 3K (low) | 40 | 26 | 24.63 | 9.9 | 0.1 | 0.3401 |
| | PRMT1 (low), MeCP2 3K (high) | 178 | 20 | 21.31 | 8.33 | 0.04 | 0.07061 |
| | PRMT1 (high), MeCP2 3Q (low) | 36 | 22.5 | 22.75 | 6.58 | 0.07 | 0.004312 |
| | PRMT1 (high), MeCP2 3Q (high) | 226 | 19 | 20.83 | 7.3 | 0.03 | 0.25 |
| | PRMT1 (low), MeCP2 3Q (low) | 96 | 25 | 25.27 | 8.73 | 0.06 | 0.2347 |
| | PRMT1 (low), MeCP2 3Q (high) | 276 | 22 | 23 | 8.38 | 0.03 | 0.8832 |
| | PRMT1 (high), MeCP2 3L (low) | 34 | 24.5 | 25.71 | 8.97 | 0.1 | 0.2112 |
| | PRMT1 (high), MeCP2 3L (high) | 114 | 23 | 22.68 | 8.73 | 0.05 | 0.2948 |
| | PRMT1 (low), MeCP2 3L (low) | 110 | 26.5 | 25.55 | 9.27 | 0.06 | 0.4059 |
| | PRMT1 (low), MeCP2 3L (high) | 156 | 23 | 22.97 | 9.09 | 0.05 | 0.7617 |
| 5.7B | PRMT6 (high), MeCP2 wt (low) | 142 | 27.5 | 27.99 | 7.71 | 0.04 | - |
| | PRMT6 (high), MeCP2 wt (high) | 318 | 25 | 25.33 | 8.45 | 0.03 | - |
| | PRMT6 (low), MeCP2 wt (low) | 818 | 27 | 27.62 | 7.9 | 0.02 | - |
| | PRMT6 (low), MeCP2 wt (high) | 92 | 24 | 24.16 | 7.72 | 0.05 | - |
| | PRMT6 (high), MeCP2 3K (low) | 102 | 24 | 24.44 | 9.31 | 0.06 | 0.0005301 |
| | PRMT6 (high), MeCP2 3K (high) | 162 | 18 | 18.94 | 8.3 | 0.04 | 6.87e-15 |
| | PRMT6 (low), MeCP2 3K (low) | 528 | 25 | 25.44 | 8.88 | 0.02 | 3.08e-06 |
| | PRMT6 (low), MeCP2 3K (high) | 22 | 22 | 22.14 | 5.72 | 0.08 | 0.2956 |
| | PRMT6 (high), MeCP2 3Q (low) | 76 | 27 | 25.93 | 9.09 | 0.07 | 0.1689 |
| | PRMT6 (high), MeCP2 3Q (high) | 320 | 25 | 26.07 | 8.99 | 0.03 | 0.3974 |
| | PRMT6 (low), MeCP2 3Q (low) | 718 | 26 | 26.93 | 8.28 | 0.02 | 0.04095 |
| | PRMT6 (low), MeCP2 3Q (high) | 76 | 28 | 27.84 | 9.88 | 0.07 | 0.01197 |
| | PRMT6 (high), MeCP2 3L (low) | 68 | 25 | 24.84 | 9.77 | 0.07 | 0.04254 |
| | PRMT6 (high), MeCP2 3L (high) | 240 | 24 | 23.58 | 8.42 | 0.03 | 0.0241 |
| | PRMT6 (low), MeCP2 3L (low) | 470 | 26.5 | 26.44 | 8.77 | 0.03 | 0.01003 |
| | PRMT6 (low), MeCP2 3L (high) | 268 | 23 | 24.03 | 8.81 | 0.03 | 0.7235 |
| 5.7C | PRMT1 (high), MeCP2 R106K (low) | 50 | 24 | 24.26 | 8.66 | 0.08 | 0.03505 |
| | PRMT1 (high), MeCP2 R106K (high) | 248 | 24 | 24.32 | 8.38 | 0.03 | 0.002063 |
| | PRMT1 (low), MeCP2 R106K (low) | 260 | 25.5 | 25.19 | 7.97 | 0.03 | 0.1909 |
| | PRMT1 (low), MeCP2 R106K (high) | 440 | 24 | 24.77 | 8.74 | 0.03 | 0.002416 |
| | PRMT1 (high), MeCP2 R106Q (low) | 42 | 22 | 22.50 | 5.86 | 0.06 | 0.001865 |
| | PRMT1 (high), MeCP2 R106Q (high) | 64 | 23.5 | 23.55 | 7.73 | 0.06 | 0.09642 |
| | PRMT1 (low), MeCP2 R106Q (low) | 106 | 23.5 | 23.68 | 8.70 | 0.05 | 0.03174 |
| | PRMT1 (low), MeCP2 R106Q (high) | 60 | 25 | 25.48 | 7.95 | 0.06 | 0.01825 |
| | PRMT1 (high), MeCP2 R106L (low) | 86 | 24 | 24.60 | 6.95 | 0.05 | 0.03178 |
| | PRMT1 (high), MeCP2 R106L (high) | 152 | 24 | 25.10 | 7.80 | 0.04 | 0.0001461 |
| | PRMT1 (low), MeCP2 R106L (low) | 250 | 26.5 | 26.36 | 8.23 | 0.03 | 0.6033 |
| | PRMT1 (low), MeCP2 R106L (high) | 144 | 26 | 26.63 | 7.81 | 0.04 | 7.10e-06 |
| 5.7D | PRMT6 (high), MeCP2 R106K (low) | 82 | 21.5 | 22.89 | 10.05 | 0.07 | 3.54e-05 |
| | PRMT6 (high), MeCP2 R106K (high) | 168 | 22 | 21.22 | 9.11 | 0.04 | 5.71e-06 |
| | PRMT6 (low), MeCP2 R106K (low) | 542 | 23.5 | 23.90 | 9.32 | 0.03 | 7.25e-16 |
| | PRMT6 (low), MeCP2 R106K (high) | 68 | 23 | 23.25 | 9.31 | 0.07 | 0.4788 |
| | PRMT6 (high), MeCP2 R106Q (low) | 90 | 25.5 | 25.53 | 9.17 | 0.06 | 0.02262 |
| | PRMT6 (high), MeCP2 R106Q (high) | 18 | 28 | 27.17 | 11.09 | 0.17 | 0.3492 |
| | PRMT6 (low), MeCP2 R106Q (low) | 24 | 28 | 29.42 | 13.89 | 0.18 | 0.8737 |
| | PRMT6 (low), MeCP2 R106Q (high) | 16 | 26.5 | 25.50 | 8.02 | 0.13 | 0.4435 |

| Figure | Sample | n | Median | Mean | StDev | 95% CI | p-value |
|---|---|---|---|---|---|---|---|
| 5.7D | PRMT6 (high), MeCP2 R106L (low) | 384 | 24 | 23.70 | 8.80 | 0.03 | 4.07e-07 |
| | PRMT6 (high), MeCP2 R106L (high) | 174 | 24 | 23.89 | 7.44 | 0.04 | 0.09972 |
| | PRMT6 (low), MeCP2 R106L (low) | 610 | 24 | 23.66 | 8.36 | 0.02 | $<$2.2e-16 |
| | PRMT6 (low), MeCP2 R106L (high) | 30 | 24.5 | 24.37 | 8.45 | 0.10 | 0.7843 |

n: number of cells, StDev: standard deviation, CI: confidence interval; p-value: in comparison to wild type MeCP2.

**Table 5.6:** Post-translational modifications of MeCP2 identified by mass spectrometry analysis.

| | Amino acid* | Modification | Peptide sequence*** | Number of biological replicates (total) | Proteome Discoverer | Max Quant |
|---|---|---|---|---|---|---|
| NTD | R9/ K12** | met | AAAAATAAAAAAPSGGGGGGEEErLEEk | 5 | 0 | 5 |
| | R9** | dimet | AAAAATAAAAAAPSGGGGGGEEEr | 5 | 0 | 5 |
| | S13** | phos | LEEKsEDQDLQGLR | 6 | 2 | 6 |
| | K42 | ac | EGkHEPLQPSAHHSAEPAEAGK | 2 | 0 | 2 |
| | K61 | met / dimet | HEPLQPSAHHSAEPAEAGk | 2 / 5 | 1 / 4 | 2 / 4 |
| | S68 | phos | AETSESsGSAPAVPEASASPK | 2 | 0 | 2 |
| | S70 | phos | AETSESSGsAPAVPEASASPK | 1 | 1 | 0 |
| | S78 | phos | AETSESSGSAPAVPEAsASPK | 4 | 3 | 4 |
| | S80 | phos | AETSESSGSAPAVPEASAsPK | 7 | 7 | 7 |
| | K82 | met / dimet | AETSESSGSAPAVPEASASPk | 2 / 4 | 1 / 4 | 2 / 4 |
| MBD | R91 | met | DrGPMYDDPTLPEGWTR | 1 | 1 | 1 |
| | R106 | dimet | GPMYDDPTLPEGWTr | 4 | 1 | 4 |
| | T148 | phos | VGDtSLDPNDFDFTVTGR | 1 | 1 | 0 |
| | S149 | phos | VGDTsLDPNDFDFTVTGR | 1 | 1 | 1 |
| | T160 | phos | VGDTSLDPNDFDFTVtGRGSPSR | 3 | 0 | 3 |
| | R162 | met / dimet | VGDTSLDPNDFDFTVTGr(GSPSR) | 7 / 4 | 6 / 2 | 7 / 4 |
| ID | S164 | phos | VGDTSLDPNDFDFTVTGRGsPSR | 2 | 2 | 0 |
| | S166 | phos | VGDTSLDPNDFDFTVTGRGSPsR | 1 | 1 | 0 |
| | R167 | met | VGDTSLDPNDFDFTVTGRGSPSr | 5 | 2 | 5 |
| | S216 | phos | VLEKsPGK | 1 | 1 | 1 |
| | S229 | phos | MPFQAsPGGK | 5 | 2 | 5 |
| | T240 | phos | MPFQASPGGKGEGGGAtTSAQVMVIK | 1 | 1 | 0 |
| NID | K271 | ac | kPGSVVAAAAAEAK | 2 | 0 | 2 |
| | S274 | phos | KPGsVVAAAAAEAK | 5 | 3 | 5 |
| | K284 | met / dimet | KPGSVVAAAAAEAk | 1 / 4 | 1 / 3 | 1 / 4 |
| CTD | K321 | ac | EVVkPLLVSTLGEK | 5 | 4 | 1 |
| | K331 | met / dimet | EVVKPLLVSTLGEk | 2 / 4 | 1 / 4 | 2 / 4 |

NTD: N-terminal domain, MBD: Methyl-binding domain, ID: Intervening domain, NID: N-CoR interacting domain; CTD: C-terminal domain; met: methylation; dimet: dimethylation; phos: phosphorylation; ac: acetylation; *numbering according to mouse MeCP2 exon2 isoform (484 aa, uniprot Q9Z2D6-1) **only identified in mouse MeCP2 exon1 isoform (501 aa, uniprot Q9Z2D6-2) ***A peptide may bear more than one amino acid candidate for a post-translational modification. MaxQuant provides probability scores for the localization of each PTM site that can be retrieved from the data repository. Lower case indicates the modified amino acid. This table was taken from Schmidt *et al.*, 2022 [2].

**Table 5.7:** Plot statistics (Figure 5.20.)

| Figure | Sample | n | Median | Mean | StDev | 95% CI | p-value |
|---|---|---|---|---|---|---|---|
| 5.20A | PRMT1 (high), MeCP2 wt (low) | 29 | 1.91 | 2.16 | 1.00 | - | - |
| | PRMT1 (high), MeCP2 wt (high) | 232 | 2.10 | 2.36 | 1.16 | - | - |
| | PRMT1 (low), MeCP2 wt (low) | 52 | 2.03 | 2.24 | 1.08 | - | - |
| | PRMT1 (low), MeCP2 wt (high) | 282 | 2.10 | 2.37 | 1.14 | - | - |
| | PRMT1 (high), MeCP2 3K (low) | 26 | 1.91 | 2.21 | 1.13 | - | 0.7383 |
| | PRMT1 (high), MeCP2 3K (high) | 144 | 2.10 | 2.35 | 1.10 | - | 0.9675 |
| | PRMT1 (low), MeCP2 3K (low) | 40 | 1.91 | 2.19 | 1.03 | - | 0.2078 |
| | PRMT1 (low), MeCP2 3K (high) | 178 | 2.16 | 2.42 | 1.11 | - | 0.005726 |
| | PRMT1 (high), MeCP2 3Q (low) | 36 | 1.91 | 2.21 | 1.02 | - | 0.373 |
| | PRMT1 (high), MeCP2 3Q (high) | 226 | 2.28 | 2.51 | 1.17 | - | 7.70e-14 |
| | PRMT1 (low), MeCP2 3Q (low) | 96 | 2.03 | 2.27 | 1.07 | - | 0.4513 |
| | PRMT1 (low), MeCP2 3Q (high) | 276 | 2.22 | 2.48 | 1.18 | - | 8.05e-10 |
| | PRMT1 (high), MeCP2 3L (low) | 34 | 1.97 | 2.16 | 1.03 | - | 0.7695 |
| | PRMT1 (high), MeCP2 3L (high) | 114 | 2.10 | 2.38 | 1.13 | - | 0.4875 |
| | PRMT1 (low), MeCP2 3L (low) | 110 | 1.91 | 2.22 | 1.17 | - | 0.09646 |
| | PRMT1 (low), MeCP2 3L (high) | 156 | 2.03 | 2.30 | 1.29 | - | 5.46e-05 |
| 5.20B | PRMT6 (high), MeCP2 wt (low) | 142 | 1.91 | 2.11 | 0.96 | - | - |
| | PRMT6 (high), MeCP2 wt (high) | 318 | 1.97 | 2.23 | 1.04 | - | - |
| | PRMT6 (low), MeCP2 wt (low) | 818 | 1.91 | 2.11 | 0.96 | - | - |
| | PRMT6 (low), MeCP2 wt (high) | 92 | 2.03 | 2.29 | 1.07 | - | - |
| | PRMT6 (high), MeCP2 3K (low) | 102 | 2.10 | 2.38 | 1.24 | - | $<$2.2e-16 |
| | PRMT6 (high), MeCP2 3K (high) | 162 | 2.40 | 2.76 | 1.43 | - | $<$2.2e-16 |
| | PRMT6 (low), MeCP2 3K (low) | 528 | 2.03 | 2.28 | 1.08 | - | $<$2.2e-16 |
| | PRMT6 (low), MeCP2 3K (high) | 22 | 2.10 | 2.34 | 1.06 | - | 0.3049 |
| | PRMT6 (high), MeCP2 3Q (low) | 76 | 1.97 | 2.23 | 1.12 | - | 0.002288 |
| | PRMT6 (high), MeCP2 3Q (high) | 320 | 1.97 | 2.23 | 1.09 | - | 0.9272 |
| | PRMT6 (low), MeCP2 3Q (low) | 718 | 1.97 | 2.19 | 1.03 | - | 9.54e-15 |
| | PRMT6 (low), MeCP2 3Q (high) | 76 | 1.97 | 2.17 | 0.99 | - | 0.001012 |
| | PRMT6 (high), MeCP2 3L (low) | 68 | 2.03 | 2.29 | 1.19 | - | 9.98e-06 |
| | PRMT6 (high), MeCP2 3L (high) | 240 | 2.10 | 2.36 | 1.17 | - | 1.18e-11 |
| | PRMT6 (low), MeCP2 3L (low) | 470 | 1.97 | 2.19 | 1.00 | - | 6.59e-12 |
| | PRMT6 (low), MeCP2 3L (high) | 268 | 2.10 | 2.35 | 1.11 | - | 0.02225 |
| 5.20C | PRMT1 (high), MeCP2 R106K (low) | 50 | 1.97 | 2.15 | 0.94 | - | 0.7171 |
| | PRMT1 (high), MeCP2 R106K (high) | 248 | 2.03 | 2.24 | 1.03 | - | 4.82e-08 |
| | PRMT1 (low), MeCP2 R106K (low) | 260 | 1.97 | 2.21 | 0.99 | - | 0.7667 |
| | PRMT1 (low), MeCP2 R106K (high) | 440 | 2.03 | 2.24 | 1.05 | - | 1.23e-12 |
| | PRMT1 (high), MeCP2 R106Q (low) | 42 | 2.10 | 2.36 | 1.21 | - | 0.000767 |
| | PRMT1 (high), MeCP2 R106Q (high) | 64 | 1.97 | 2.18 | 1.01 | - | 5.72e-09 |
| | PRMT1 (low), MeCP2 R106Q (low) | 106 | 2.03 | 2.28 | 1.13 | - | 0.5803 |
| | PRMT1 (low), MeCP2 R106Q (high) | 60 | 1.91 | 2.15 | 1.05 | - | 1.16e-14 |
| | PRMT1 (high), MeCP2 R106L (low) | 86 | 1.97 | 2.16 | 0.99 | - | 0.8981 |
| | PRMT1 (high), MeCP2 R106L (high) | 152 | 1.91 | 2.07 | 0.92 | - | $<$2.2e-16 |
| | PRMT1 (low), MeCP2 R106L (low) | 250 | 1.91 | 2.13 | 0.99 | - | 0.0002114 |
| | PRMT1 (low), MeCP2 R106L (high) | 144 | 1.91 | 2.11 | 0.96 | - | $<$2.2e-16 |
| 5.20D | PRMT6 (high), MeCP2 R106K (low) | 82 | 2.10 | 2.37 | 1.17 | - | 8.70e-13 |
| | PRMT6 (high), MeCP2 R106K (high) | 168 | 2.16 | 2.48 | 1.33 | - | $<$2.2e-16 |
| | PRMT6 (low), MeCP2 R106K (low) | 542 | 2.10 | 2.35 | 1.16 | - | $<$2.2e-16 |
| | PRMT6 (low), MeCP2 R106K (high) | 68 | 2.03 | 2.31 | 1.18 | - | 0.9811 |
| | PRMT6 (high), MeCP2 R106Q (low) | 90 | 1.91 | 2.13 | 1.03 | - | 0.4898 |
| | PRMT6 (high), MeCP2 R106Q (high) | 18 | 1.97 | 2.12 | 0.78 | - | 0.7577 |
| | PRMT6 (low), MeCP2 R106Q (low) | 24 | 1.94 | 2.20 | 1.21 | - | 0.1254 |
| | PRMT6 (low), MeCP2 R106Q (high) | 16 | 1.79 | 2.15 | 1.01 | 0.01 | 0.2926 |

| Figure | Sample | n | Median | Mean | StDev | 95% CI | p-value |
|---|---|---|---|---|---|---|---|
| 5.20D | PRMT6 (high), MeCP2 R106L (low) | 384 | 1.97 | 2.20 | 1.13 | - | 0.005139 |
| | PRMT6 (high), MeCP2 R106L (high) | 174 | 1.91 | 2.16 | 0.99 | - | 0.002118 |
| | PRMT6 (low), MeCP2 R106L (low) | 610 | 1.97 | 2.25 | 1.10 | - | $<$ 2.2e-16 |
| | PRMT6 (low), MeCP2 R106L (high) | 30 | 2.00 | 2.23 | 1.04 | - | 0.2227 |

n: number of cells, StDev: standard deviation, CI: confidence interval; p-value: in comparison to wild type MeCP2.

## Supplementary figures

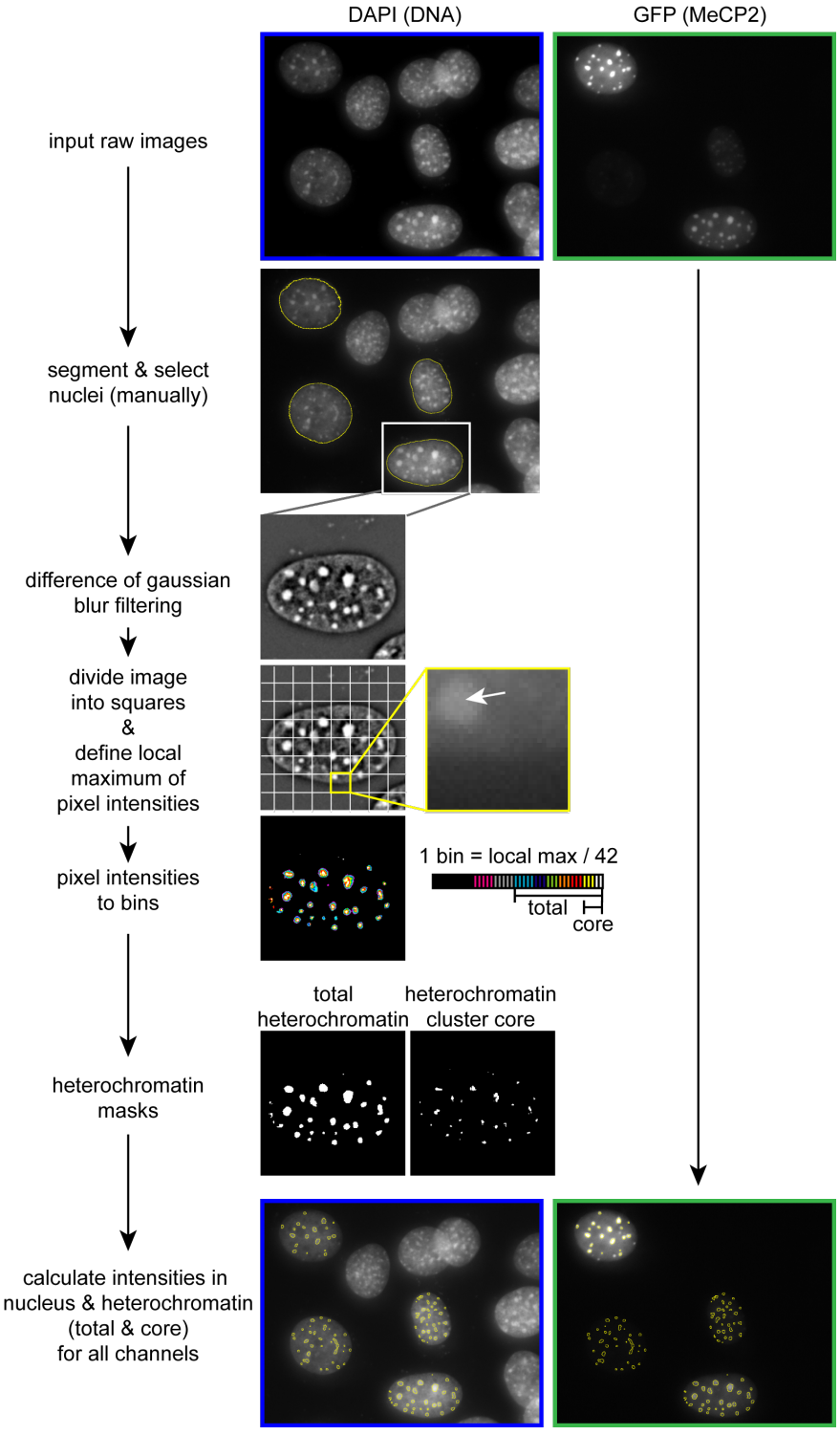



**Figure 5.9: Image analysis pipeline using imageJ.** Images were taken on a wide-field fluorescence microscope (Axiovert 200, Zeiss). Nuclei and heterochromatin segmentation were carried out based on the DAPI channel and subsequently used for fluorescent intensity measurements on the GFP channel. Nuclei segmentation was performed manually and images were filtered using a difference of gaussian blur

**Figure 5.9 (*previous page*):** filter for heterochromatin segmentation. Each image was divided into squares, and the local maximum of each square was determined, set as the maximum for pixel intensity binning and heterochromatin clusters as well as the heterochromatin cluster cores were defined by taking a specific number of bins for the mask. The heterochromatin segmentation procedure was carried out using a self-made ImageJ macro published in [1]. The figure was taken from Schmidt *et al.*, 2022 [2].

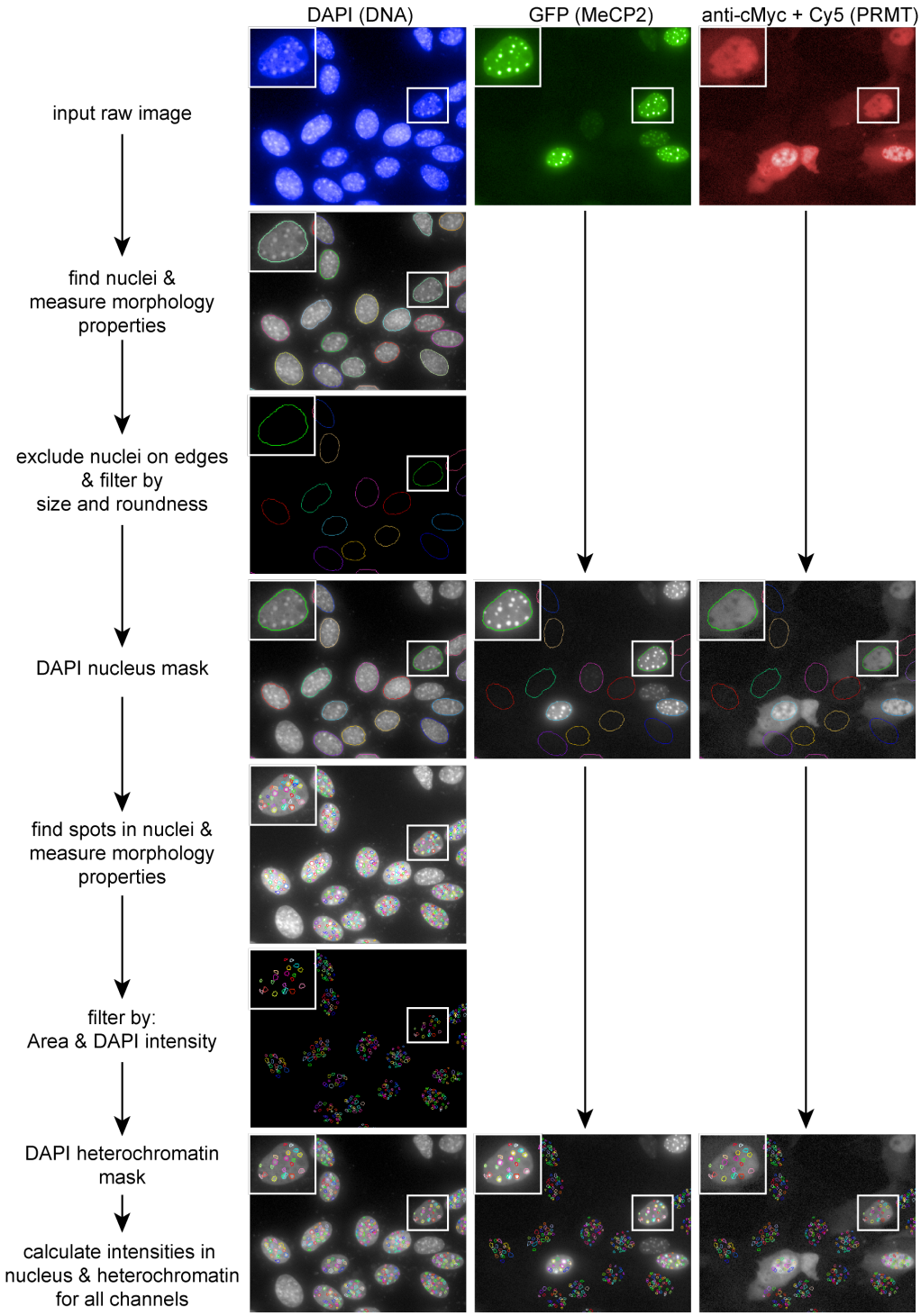



**Figure 5.10: High content screening microscopy analysis pipeline.** The analysis of high content screening microscopy images was done using the PerkinElmer Harmony software. Based on the DAPI channel, nuclei were segmented, those on edges were excluded and nuclei were filtered by size and roundness. For heterochromatin cluster segmentation in the DAPI channel, spots were found inside the nuclei and filtered by area and total DAPI intensity. The nuclei and heterochromatin mask were used to calculate fluorescent intensities in all channels. The figure was taken from Schmidt *et al.*, 2022 [2].

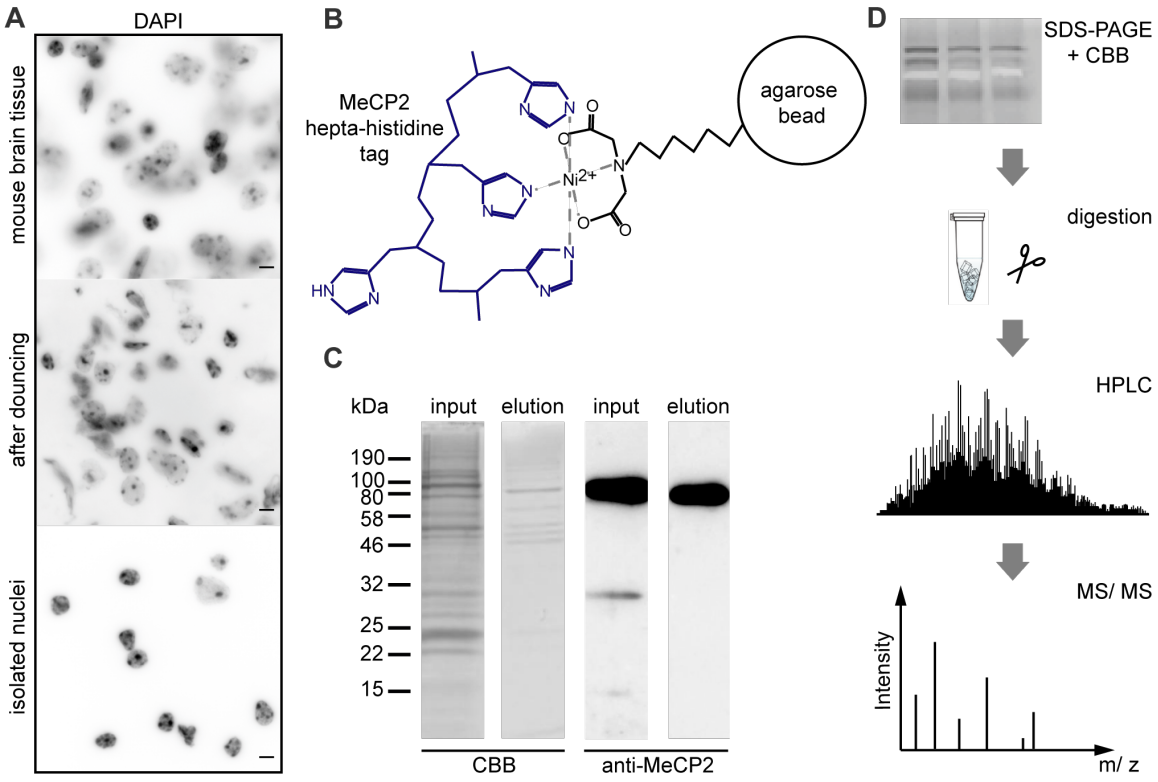



**Figure 5.11: MeCP2 enrichment from mouse brain tissue and mass spectrometry analysis.** (A) Images of DNA stained with DAPI from mouse brain tissue, tissue after douncing and isolated nuclei. Scale bar 5 µm. (B) Scheme of the binding of MeCP2 hepta-histidine tag to Ni-IDA agarose beads. It is not clear whether two or three histidines are involved in the interaction [373]. The protein backbone is shown simplified as described before [373]. (C) Coomassie (CBB) stained SDS-PAGE and Western blot with antibodies specific to MeCP2 showing the input and elution fraction of the MeCP2 enrichment procedure. (D) Workflow of sample preparation and subsequent HPLC-MS/MS analysis involving SDS-PAGE and Coomassie staining, gel band extraction, trypsin digestion, and HPLC-coupled mass spectrometry. The figure was modified from Schmidt *et al.*, 2022 [2].

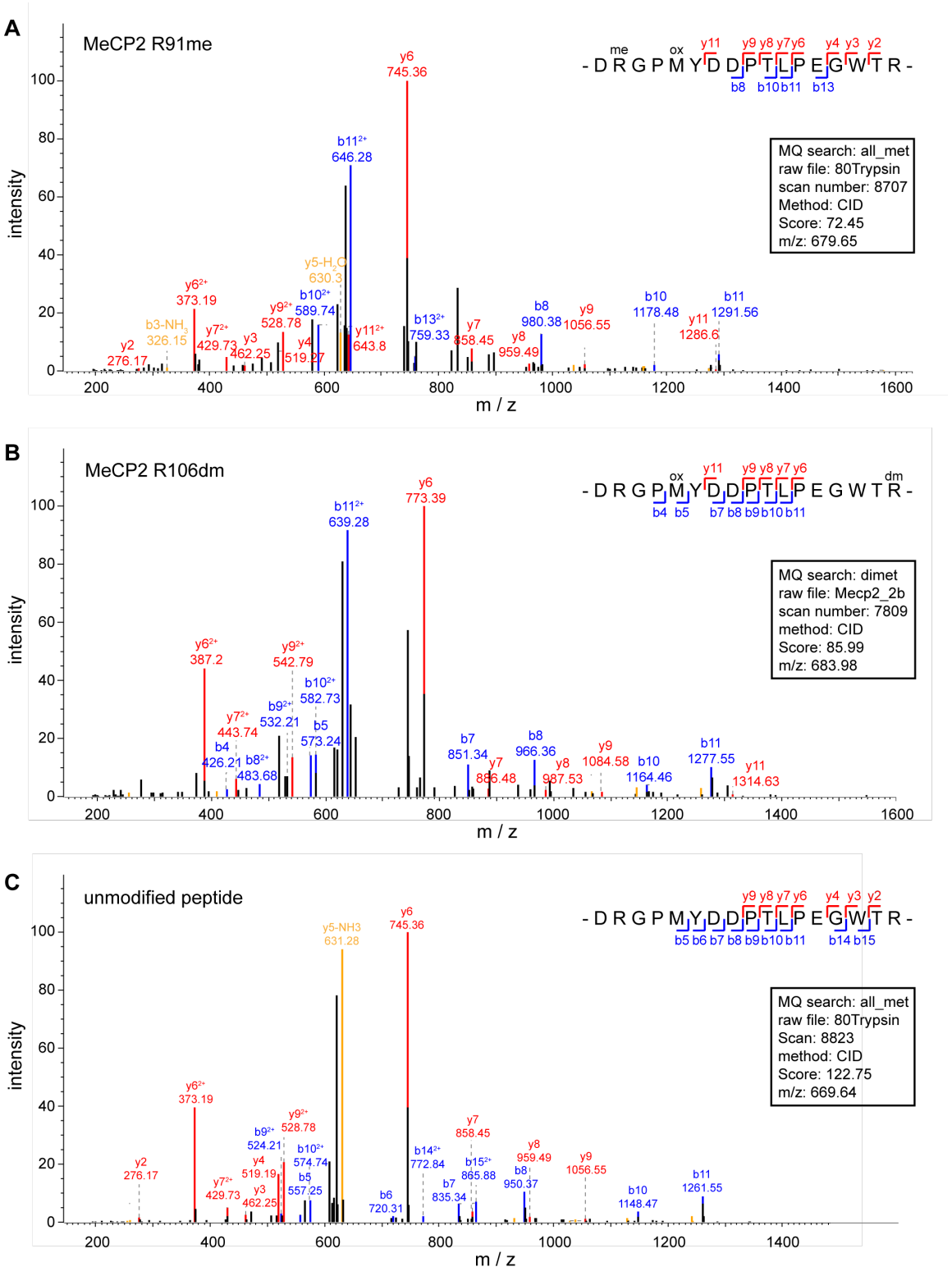



**Figure 5.12: Manual validation of the arginine methylation sites R91 and R106 identified by mass spectrometry.** Exemplary MS/MS spectra for the identification of MeCP2 R91 methylation (me) (A), MeCP2 R106 dimethylation (dm) (B), and an exemplary spectrum of the same peptide without modifications (C). The spectra were exported from the MaxQuant software [265] visualization tool, and fragments of interest were labeled. The identified y- and b-ions are indicated on the peptide sequence and the details of the spectra are given in the black box. The figure was taken from Schmidt *et al.*, 2022 [2].

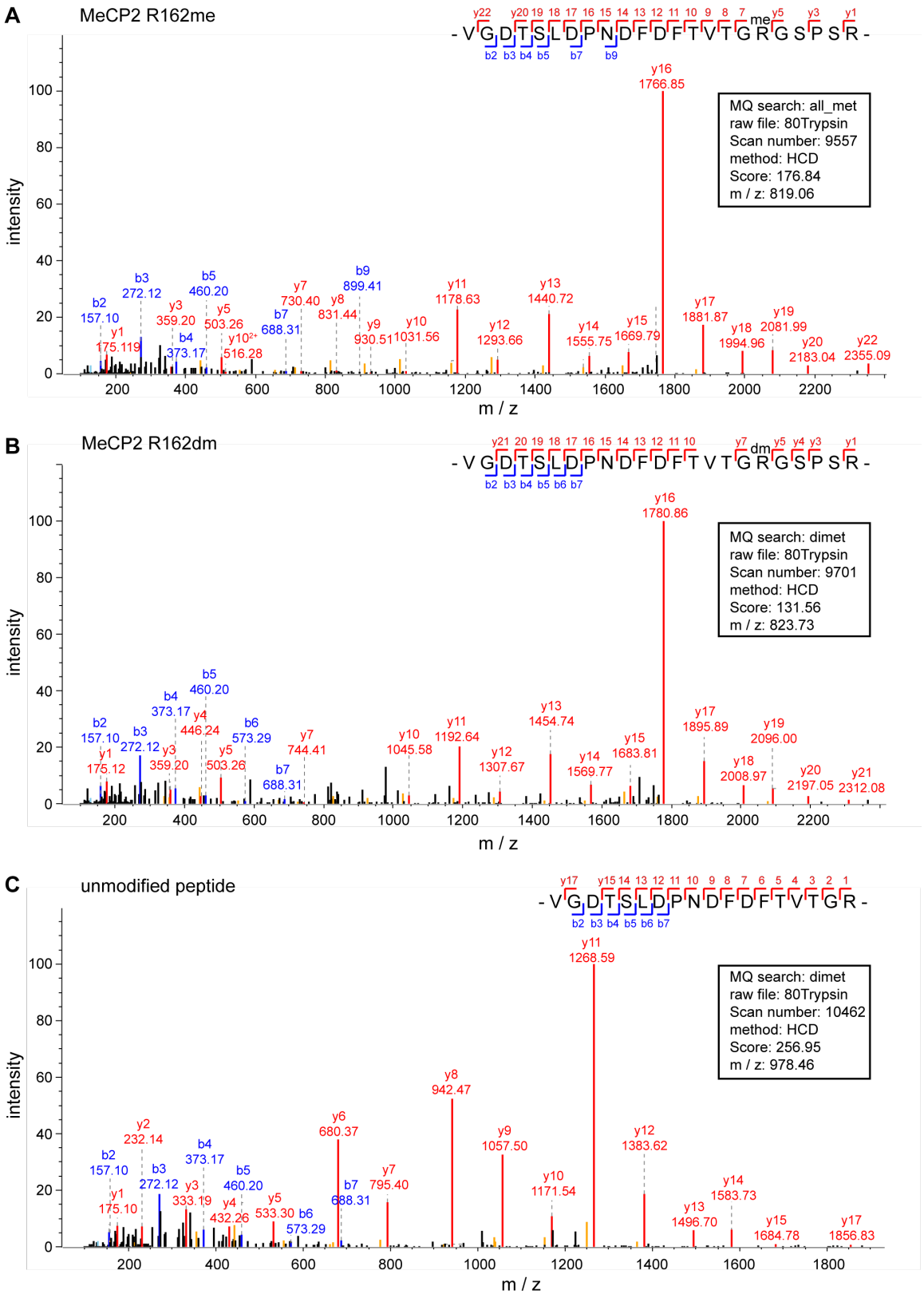



**Figure 5.13: Manual validation of the arginine methylation site R162 identified by mass spectrometry.** Exemplary MS/MS spectra for the identification of MeCP2 R162 methylation (me) (A), MeCP2 R162 dimethylation (dm) (B), and an exemplary spectrum of a peptide without the missed-cleavage and without modifications (C). The spectra were exported from the MaxQuant software [265] visualization tool, and fragments of interest were labeled. The identified y- and b-ions are indicated on

**Figure 5.13 (*previous page*):** the peptide sequence and the details of the spectra are given in the black box. The figure was taken from Schmidt *et al.*, 2022 [2].

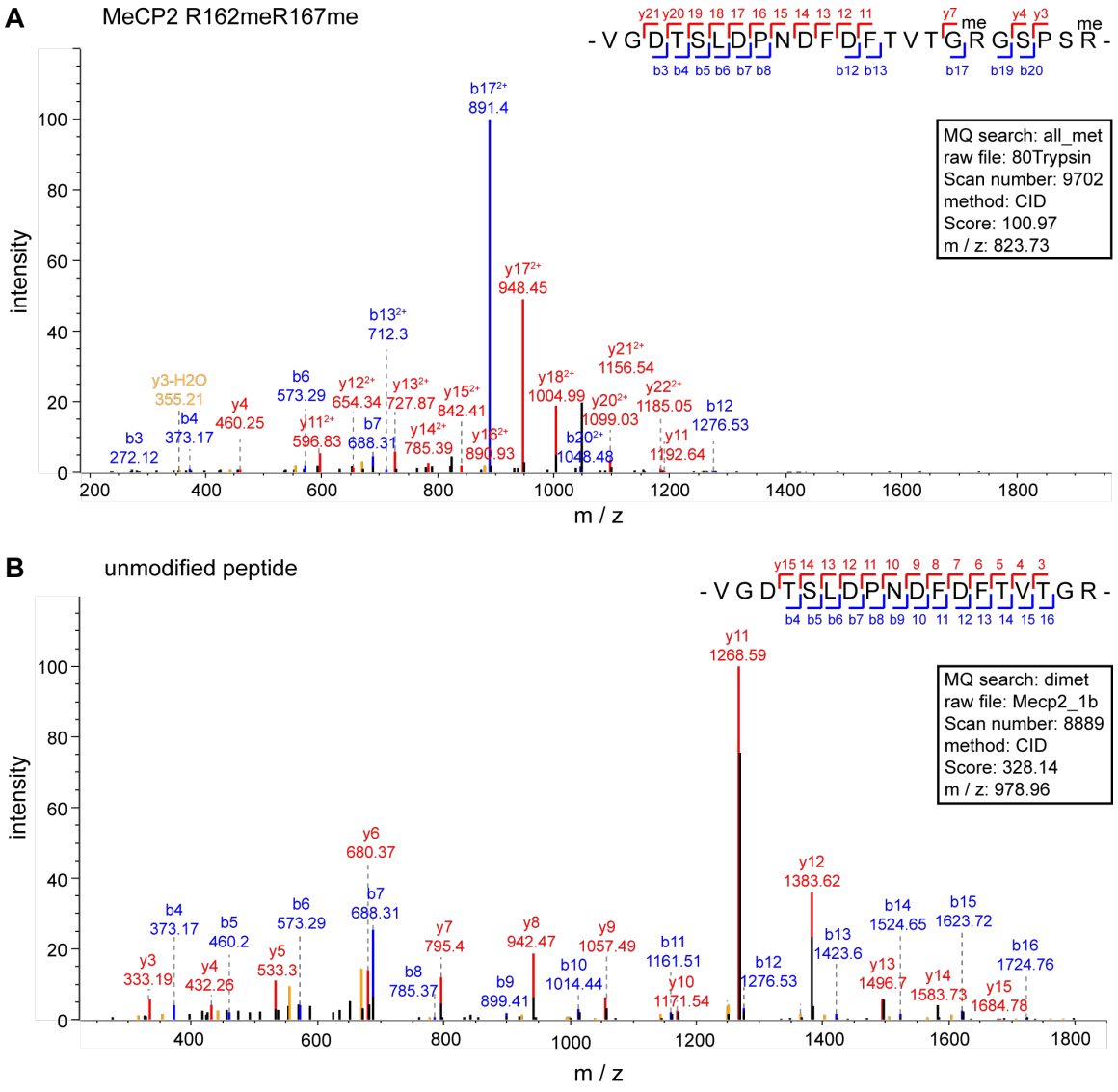



**Figure 5.14: Manual validation of the arginine methylation site R167 identified by mass spectrometry.** Exemplary MS/MS spectra for the identification of MeCP2 R167 methylation (me) (A) and an exemplary spectrum of a peptide without the missed cleavage and without modifications (B). The localization of R167 methylation cannot be determined from the spectrum shown in A, as it could also be a dimethylation on R162. The spectra were exported from the MaxQuant software [265] visualization tool, and fragments of interest were labeled. The identified y- and b-ions are indicated on the peptide sequence and the details of the spectra are given in the black box. The figure was taken from Schmidt *et al.*, 2022 [2].

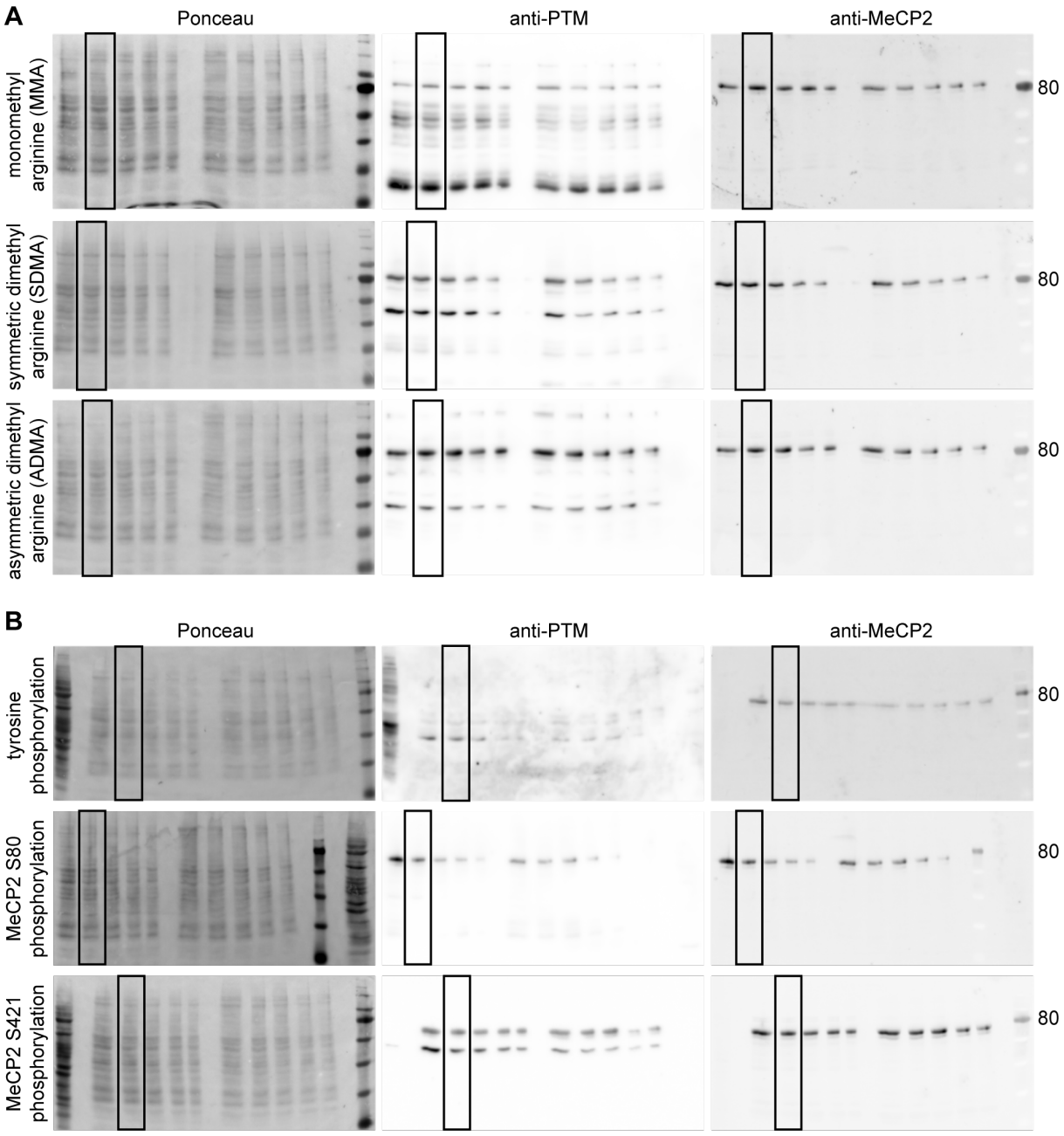



**Figure 5.15: Western blots for detection of MeCP2 post-translational modifications on mouse brain nuclei.** Western blot analysis of mouse brain nuclei extracts from $10^6$ nuclei per lane tested for (A) monomethyl arginine (MMA), symmetric dimethylarginine (SDMA), asymmetric dimethyl arginine (ADMA), (B) tyrosine phosphorylation, MeCP2 S80 phosphorylation, MeCP2 S421 phosphorylation and reprobed with an antibody specific for MeCP2. Shown are the full membranes stained for total protein with Ponceau S stain, incubated with PTM-specific antibodies, and with MeCP2-specific antibodies. Boxes mark the lanes of interest shown in figure 5.1. The figure was modified from Schmidt *et al.*, 2022 [2].

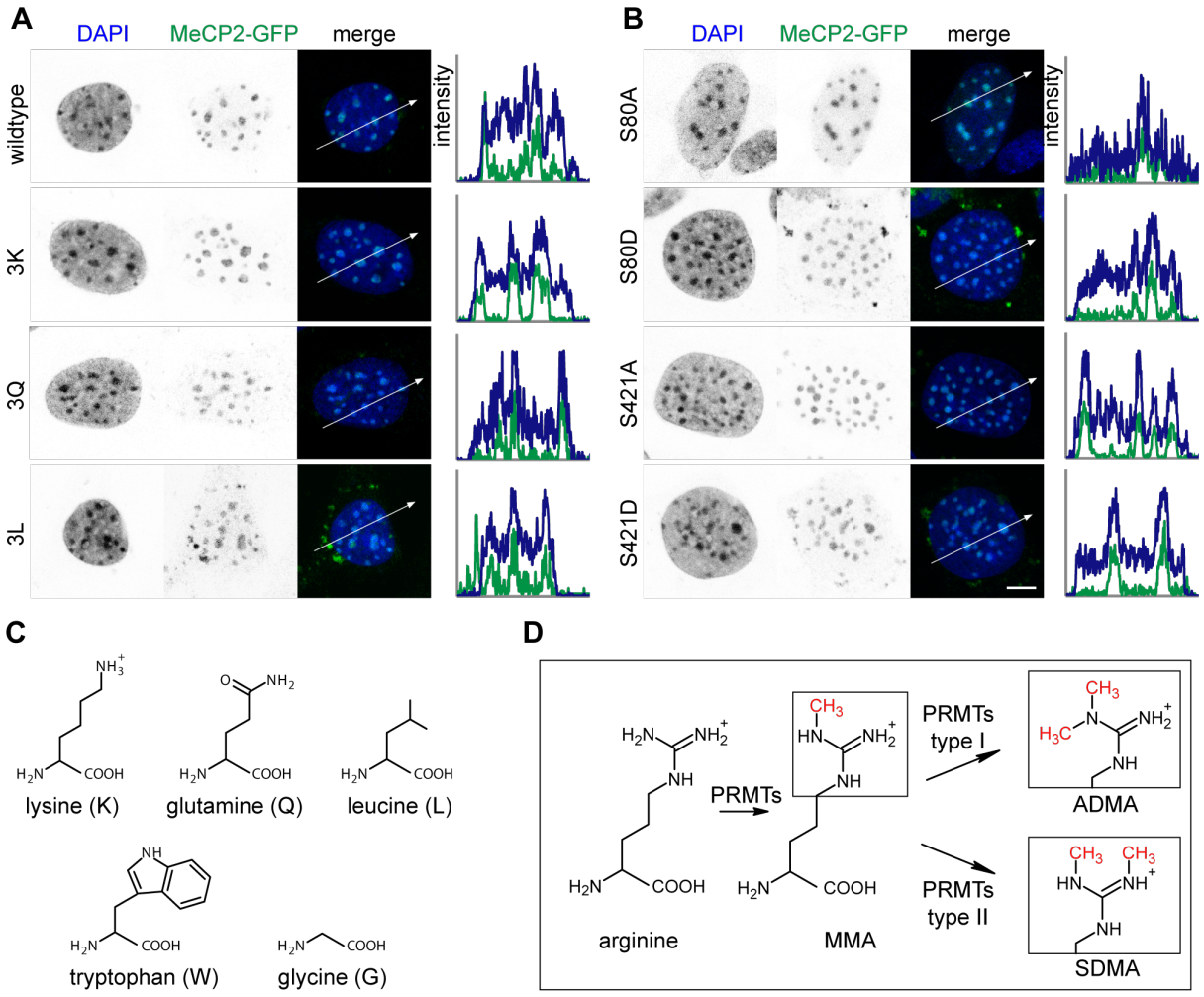



**Figure 5.16: MeCP2-GFP mutated for modified sites.** Subcellular localization of MeCP2-GFP mutated for arginine methylated (A) and phosphorylated (B) sites in MTF Mecp2 -/y cells. (C) Structure of amino acids used as substitutes for methylated arginine sites. (D) DNA staining of a C2C12 mouse myoblast cell. Scale bars 5 µm. The figure was modified from Schmidt *et al.*, 2022 [2].

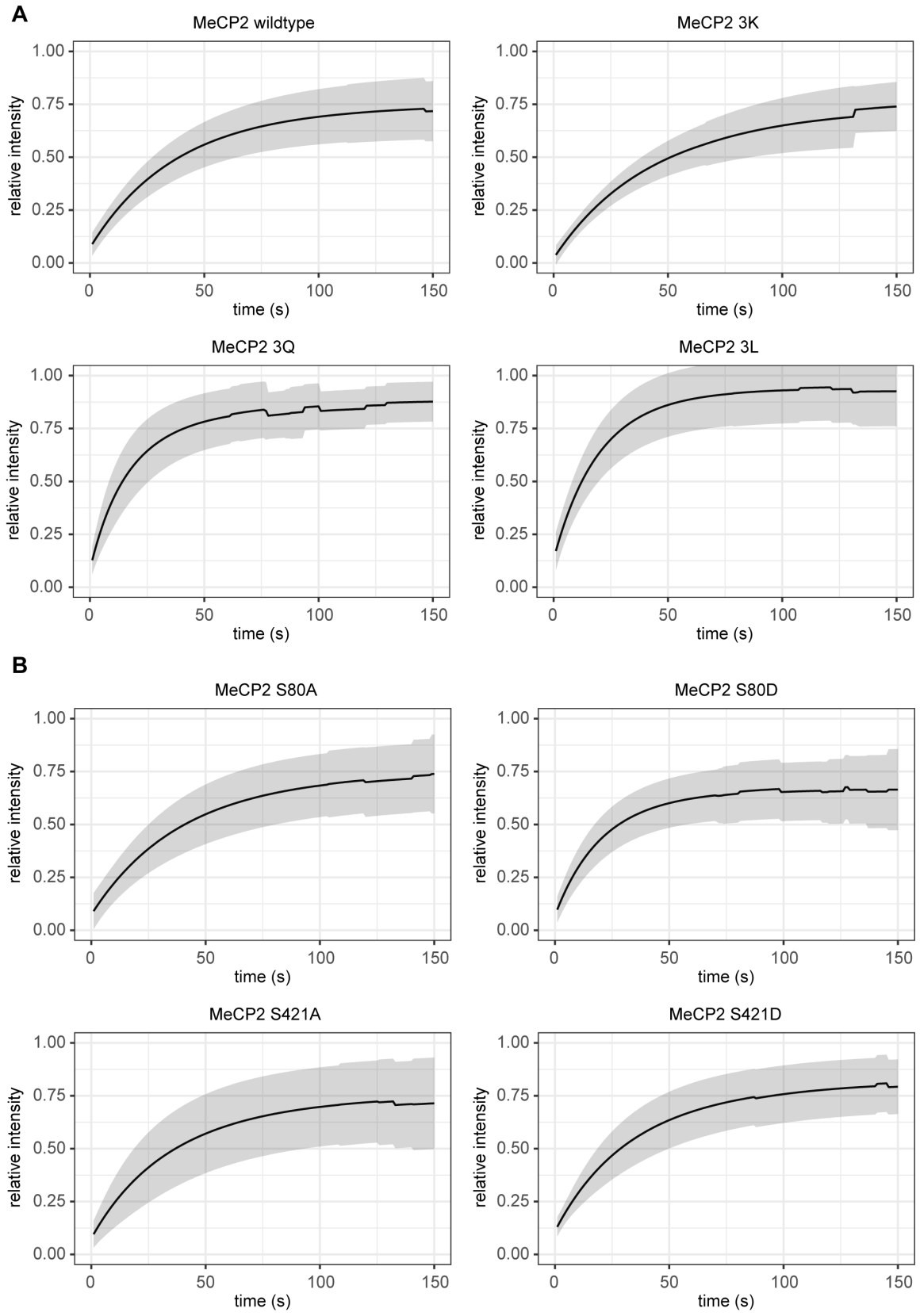



**Figure 5.17: Fluorescence recovery after photobleaching curves of 3x arginine methylation (A) and single serine phosphorylation mutants (B) for each individual sample with respective standard deviation.** Plotted was the mean of all individual fitted curves per timepoint with the standard deviation for each individual timepoint (gray shading). The figure was taken from Schmidt *et al.*, 2022 [2].

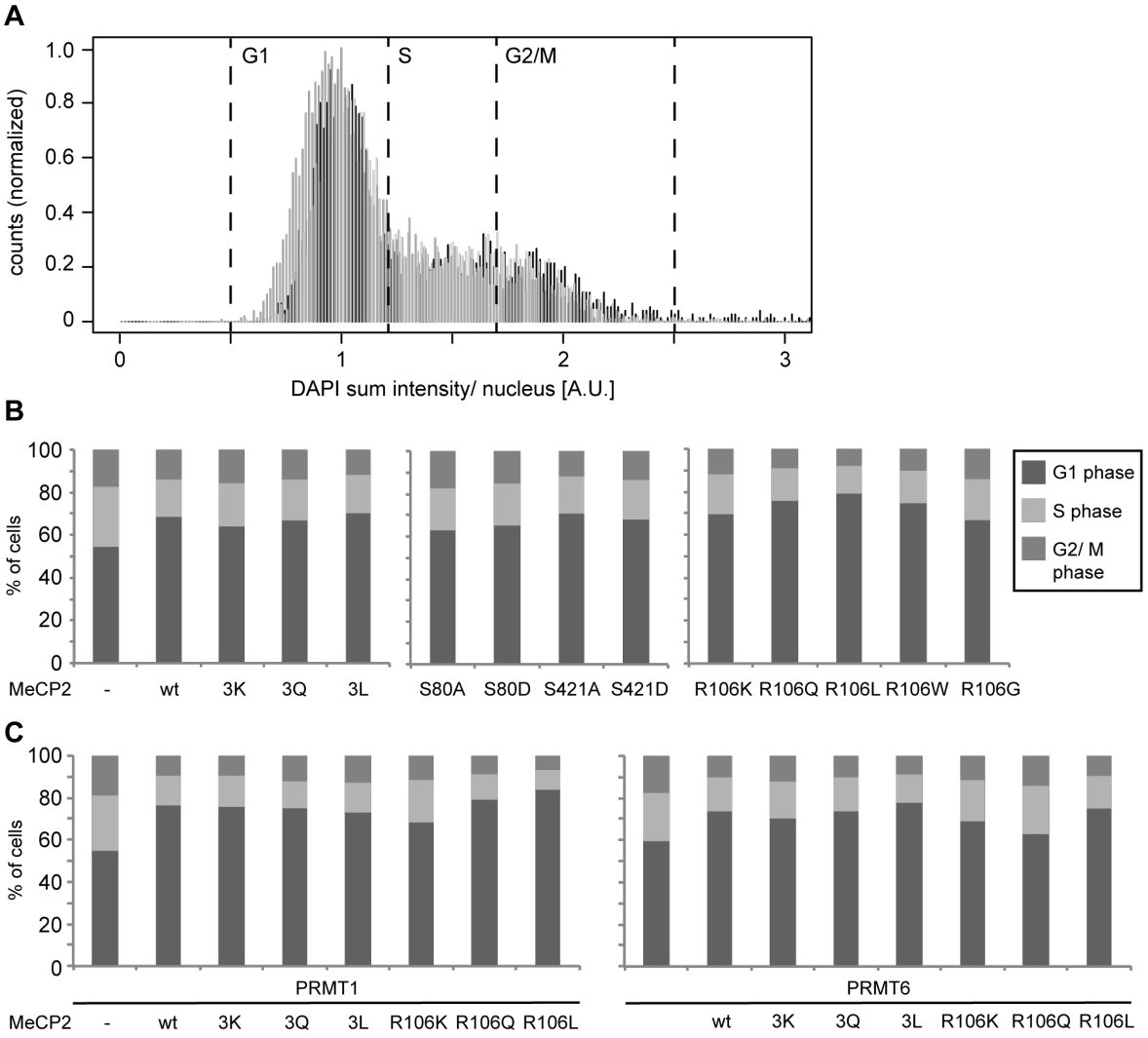



**Figure 5.18: Cell cycle distribution analysis of mouse myoblast cells.** (A) Exemplary cell cycle distribution plotted as the count of cells per DAPI sum intensity in the nucleus. The different intervals for cell cycle phases G1, S, and G2/ M phase are indicated. (B) Bar diagrams depicting the percentage of cells per cell cycle phase as indicated in (A). Mock transfected cells (MeCP2 -), wild type (wt) MeCP2, and MeCP2 mutant transfected cells are shown in (B), and transfections of the PRMTs 1 and 6 alone (MeCP2 -) and together with MeCP2 wild type and MeCP2 mutants are shown in (C). The figure was taken from Schmidt *et al.*, 2022 [2].

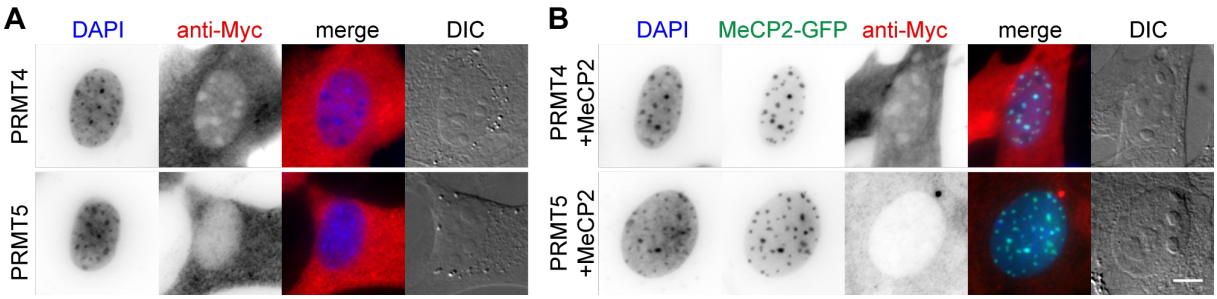



**Figure 5.19: Subcellular localization of protein arginine methyltransferases (PRMTs) 4 and 5.** Subcellular localization of PRMT4 and 5 in C2C12 mouse myoblast cells in the absence (A) and presence (B) of MeCP2. Scale bar 5 µm. The figure was taken from Schmidt *et al.*, 2022 [2].

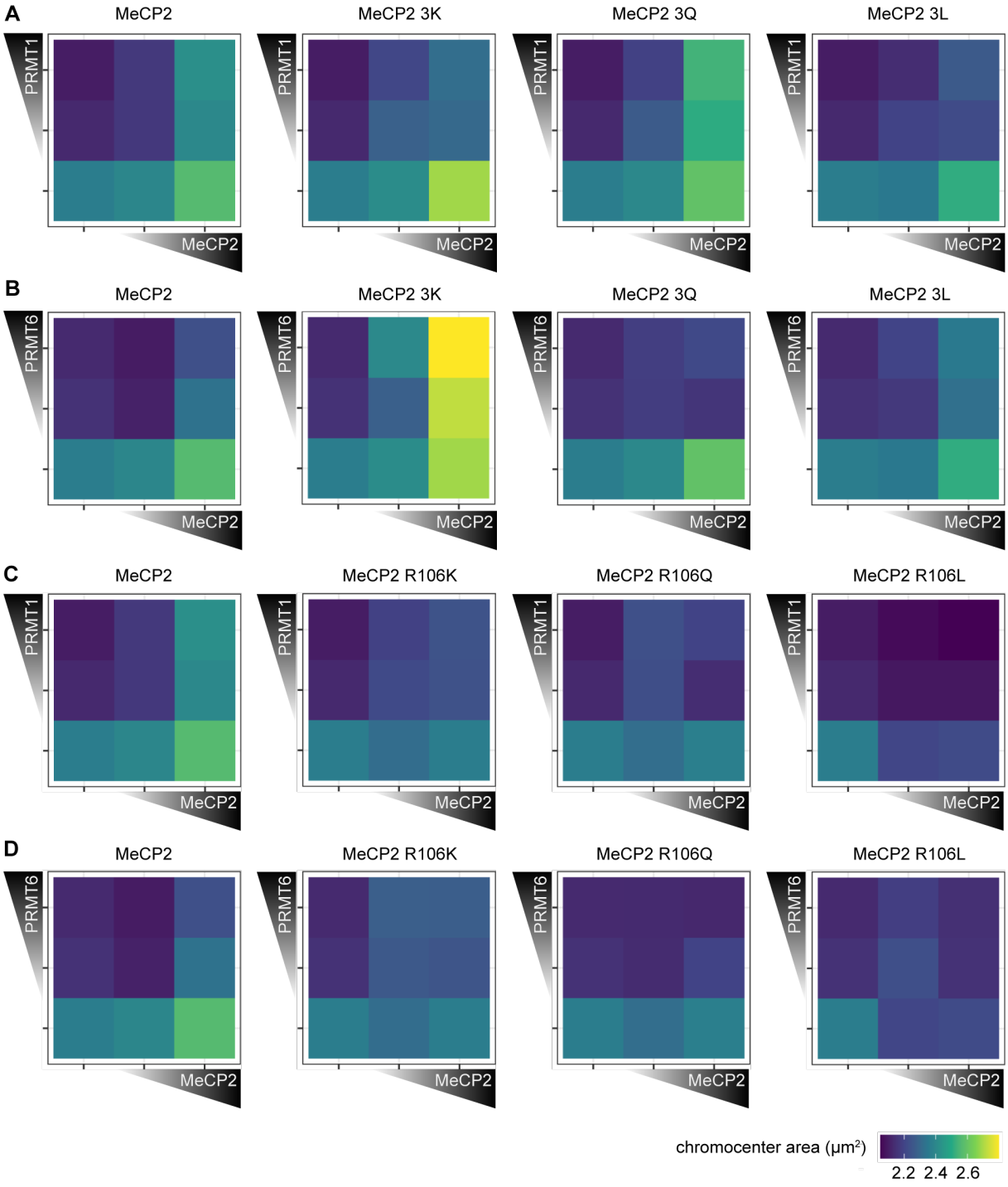



**Figure 5.20: Heterochromatin clustering represented by heterochromatin cluster areas of MeCP2 mutants in the presence of protein arginine methyltransferases (PRMTs) 1 and 6.** Heatmaps show the heterochromatin cluster areas obtained by high-content screening microscopy of C2C12 cells cotransfected with MeCP2 3x mutants and PRMT1 (A) or PRMT6 (B) and with MeCP2 R106 mutants and PRMT1 (B) and PRMT6 (D). Cells were binned for low and high fluorescence intensity of both MeCP2 (GFP channel) and PRMT (Cy5). Heterochromatin cluster areas are shown as means of at least 26 cells from at least two biological replicates. The figure was taken from Schmidt *et al.*, 2022 [2].

# 6 Conclusion & Outlook

Constitutive heterochromatin is comprised of repetitive DNA sequences and is consistently silenced between cell types. In mouse cells, it is localized at pericentromeric regions which cluster during the interphase to form chromocenters. About half of the mammalian genome is composed of mostly silenced repeat elements and transposons [374] that form a basic foundation for regulating networks and genome formatting and are hence crucial for genome function. Moreover, the importance of gene silencing is underscored by many studies describing various disease phenotypes upon derepression of heterochromatic regions (reviewed in [59]). Thus, it is of great importance to study heterochromatin organization and function to understand how its compaction is achieved and maintained in the cell.

A crucial step to understanding heterochromatin organization is to identify the factors involved in the process. In this study, we analyzed the proteomic composition of pericentromeric heterochromatin in mouse tissues. In comparison to previous reports spanning methods based on hybridization, immunoprecipitation, or requiring cell line engineering (see table 1.1), we adapted a protocol for unbiased isolation of chromocenters in the native state from tissues (see Prusov & Zatsepina [94, 95]). More importantly, we aimed to quantitatively analyze the pericentromeric heterochromatin proteome *in vivo* in mouse brain and liver, as these tissues showed differences in heterochromatin organization upon DNA staining. The comparison of different tissues allows the identification of conserved heterochromatin proteins between tissues on the one hand and tissue-specific differences in protein abundance on the other hand. In addition to the core histones, we found several histone variants enriched in the chromocenters. Histone macroH2A.1 and H3.3 were described to occur at heterochromatic regions before [21, 295], but the finding of H2AX bound to chromocenters in mouse tissues indicates a possible function of the histone variant in pericentromeric heterochromatin in addition to its role in DNA repair. Furthermore, the chromatin proteins MeCP2 and ATRX were upregulated in the brain and histone H1.4 was upregulated in the liver. All of them were shown to be involved in heterochromatin organization and are thus interesting candidates for tissue-specific chromocenter organizers. Histone H1 subtypes differ in their amino acid composition, and their chromatin binding affinities and were proposed to be involved in differential packing of specific regions of chromatin [15, 375]. Of note, it was proposed that histone H1 and MeCP2 can substitute each other, as they compete for a common chromatin binding site [253] and MeCP2 deficient cells showed an increase of H1 [138, 325]. The opposing tendencies in levels between brain and liver we obtained from the quantitative mass spectrometry analysis strengthen the evidence for a correlation of the levels of H1 and MeCP2 and their competition for chromatin binding sites. Another interesting point is the interaction of MeCP2 and ATRX [172, 315], as we identified both upregulated in the brain. They regulate gene expression by remodeling chromatin structure [315] and ATRX might contribute to MeCP2-mediated heterochromatin organization during terminal differentiation [316]. Still, as we analyzed chromocenters from whole tissues, we cannot rule out that the abundance of some proteins was underestimated, especially in cases of proteins with cell type-specific chromocenter accumulation. Our findings show that heterochromatin organization is maintained by the

interplay of many tightly regulated factors and only unraveling all these interactions will allow a complete understanding of the complex. Our study can serve as a basis for future studies of chromocenter organization, as we identified many protein candidates that might be involved in chromocenter formation and function. Nevertheless, more experiments are required to confirm the proteins with a functional role in chromocenter organization. In addition, future studies should focus on the analysis of additional mouse tissues to confirm the results of this study and to get a more comprehensive list of conserved chromocenter proteins.

As described above, the epigenetic reader MeCP2 is enriched in mouse brain chromocenters and was previously shown to induce dose-dependent heterochromatin clustering upon overexpression in cells [55]. MeCP2 protein levels were shown to increase during differentiation [55, 233, 313, 314]. In addition, mutations in the *MECP2* gene are the main cause of Rett syndrome, a neurological disorder, [116] and deviations in MeCP2 protein levels caused disease phenotypes [244–247, 328]. These results prove that MeCP2 is an important component of heterochromatin, and its level is tightly regulated *in vivo*. Therefore, we focused on the quantification of MeCP2 in different systems in the second part of this study. Estimation of the MeCP2 amount in the mouse brain and in (transfected) mouse myoblast cells allowed us to compare the protein levels from both systems. Using fluorescence microscopy, we estimated the absolute protein concentration in the nucleus and heterochromatin compartments. The MeCP2 protein level in untransfected C2C12 mouse myoblast cells was almost not detectable so the cells can be used as a functional *Mecp2*-null system. The MeCP2 amount in low-level transfected myoblast cells was within the same range as the MeCP2 level in the mouse brain, indicating that the cellular system can be used to study MeCP2 function in the range of physiological protein levels. In addition, the concentration values can be applied to *in vitro* experiments enabling comparison to the *in vivo* physiological situation.

MeCP2 is an intrinsically disordered protein with a high amount of positively charged amino acids [124] and was shown to be post-translationally modified [3, 332]. Of note, many of the described modifications were identified in big proteomic screens on cell lines, and only a few of them were validated or further characterized [3, 332]. Thus, we performed a proteomics analysis of post-translational modifications of MeCP2 in mouse brains. As overexpression of a plasmid coding for MeCP2 in cellular systems can mimic the *in vivo* dose-dependent heterochromatin clustering observed during differentiation, we used this system to study MeCP2-dependent heterochromatin organization in cells. By mass spectrometry analysis we showed that MeCP2 carries various post-translational modifications, among them phosphorylations on S80 and S421, which lead to minor changes in either heterochromatin binding kinetics or clustering in the cellular system. In addition, we found that MeCP2 is monomethylated, symmetrically and asymmetrically dimethylated on several arginines and that these modifications alter heterochromatin organization. Interestingly, we identified the Rett syndrome mutation site R106 as a dimethylation site. Thus, we showed that MeCP2 carries various post-translational modifications in mouse brains and that MeCP2-mediated heterochromatin organization is modulated by arginine methylation and to a lesser extent also by serine phosphorylation. The results prove, that PTMs play an important role in the regulation of protein function and further studies of modifications of other

chromatin proteins will be necessary to unravel the mysteries of heterochromatin organization in cells.

Heterochromatin is characterized by high levels of cytosine methylation, methyl-binding proteins, repressive histone marks like H3K9me3, and proteins binding to these marks. Nevertheless, the mechanism of how heterochromatin clustering and chromocenter formation take place is still the focus of discussion. An increase in transcription of major satellite repeats observed upon chromocenter formation during early mouse development suggests a role of major satellite transcripts in heterochromatin organization [75]. In addition, DNA and chromatin binding proteins are involved in chromocenter organization. HP1, which binds to H3K9me via its chromodomain, was shown to be dispensable for chromocenter formation [376] and did not induce chromocenter clustering [55]. On the contrary, MeCP2 was shown to induce chromatin reorganization during terminal development and dose-dependent chromocenter clustering upon overexpression of *Mecp2-GFP* in cells [55]. Still, *Mecp2* knockout mice maintain their heterochromatin organization in chromocenters, indicating that MeCP2 is not required for chromocenter formation. Instead, the functional redundancy of MeCP2 might explain this observation, as other MBD proteins like MBD2 were shown to have similar properties [55]. Also, the involvement of sequence-specific DNA binding proteins like HMGA1 in chromocenter organization was reported [377, 378]. Recently, liquid-liquid phase separation was proposed as a mechanism that might be responsible for heterochromatin cluster formation, as chromocenters display liquid-like properties [257–259, 274]. Indeed, several heterochromatin proteins were shown to undergo or promote phase separation under differing conditions, e.g. MeCP2, HP1 and histone H1 [1, 255–258, 274, 379, 380]. Many chromocenter associated proteins show characteristics of phase separating proteins such as intrinsically disordered regions and the capacity to form multivalent interactions [26, 381, 382]. Phase separation depends on factors like protein concentration, salt concentration, crowding agents and the presence of (methylated) DNA [1]. Due to the many factors influencing the process, adjusting the conditions for *in vitro* phase separation experiments is crucial to allow comparison to physiological conditions. Thus, we recently used our results from MeCP2 protein concentration determination to prove MeCP2 phase separation under physiological protein levels [1]. Liquid-liquid phase separation may also play a role in neurodegenerative diseases [383, 384]. Interestingly, studies on inverted nuclei from rod photoreceptor cells of nocturnal animals indicate that heterochromatin interactions are required for nuclear compartmentalization, whereas euchromatin interactions are dispensable, and an association of heterochromatin with lamina is required for the formation of the conventional nuclear architecture [72].

Summarizing, we contributed to the unraveling of the mechanisms of heterochromatin organization and more specifically the role of the chromatin protein MeCP2 in this process. We identified candidate heterochromatin proteins that might contribute to tissue specific heterochromatin organization. Furthermore, we established a calibration system for MeCP2 levels, allowing the comparison of *in cellulo* and *in vitro* experiments with physiological protein levels *in vivo*. Finally, we identified and validated modifications of MeCP2 in the brain and show that these can modulate its ability to bind as well as reorganize heterochromatin, which may play a role in the pathology of the Rett syndrome.

Altogether, future studies should focus not only on the identification of proteins involved in heterochromatin organization, but as well on their levels *in vivo*, their post-translational modifications, and their phase separation properties. The regulation of chromatin organization mediated by PTMs on histone tails reinforces the importance of modifications for the regulation of protein function. There might be common factors required for proper heterochromatin organization as well as redundant factors which can be substituted by others in a cell type specific manner. For this reason, future studies will also need to elucidate differences in heterochromatin composition between tissues and cell types. In this study, we present experimental concepts to address these questions and our results set a basis for further studies. The better we comprehend heterochromatin organization, the greater our chances of understanding and preventing diseases associated with heterochromatin misfunction.

# 7 References


1. Zhang, H., Romero, H., Schmidt, A., Gagova, K., Qin, W., Bertulat, B., Lehmkuhl, A., Milden, M., Eck, M., Meckel, T., Leonhardt, H. & Cardoso, M. C. MeCP2-induced heterochromatin organization is driven by oligomerization-based liquid-liquid phase separation and restricted by DNA methylation. *Nucleus* **13,** 1–34 (2022).
2. Schmidt, A., Frei, J., Poetsch, A., Chittka, A., Zhang, H., Aßmann, C., Lehmkuhl, A., Bauer, U.-M., Nuber, U. A. & Cardoso, M. C. MeCP2 heterochromatin organization is modulated by arginine methylation and serine phosphorylation. *Frontiers in Cell and Developmental Biology* **10,** 941493 (2022).
3. Schmidt, A., Zhang, H. & Cardoso, M. MeCP2 and chromatin compartmentalization. *Cells* **9,** 878 (2020).
4. Finn, E. H. & Misteli, T. Molecular basis and biological function of variability in spatial genome organization. *Science* **365,** eaaw9498 (2019).
5. Kornberg, R. D. Chromatin structure: a repeating unit of histones and DNA. *Science* **184,** 868–871 (1974).
6. Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E. & Moudrianakis, E. N. The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. *Proceedings of the National Academy of Sciences of the United States of America* **88,** 10148–10152 (1991).
7. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389,** 251–260 (1997).
8. Talbert, P. B. & Henikoff, S. Histone variants–ancient wrap artists of the epigenome. *Nature Reviews. Molecular Cell Biology* **11,** 264–275 (2010).
9. Singh, R., Bassett, E., Chakravarti, A. & Parthun, M. R. Replication-dependent histone isoforms: a new source of complexity in chromatin structure and function. *Nucleic Acids Research* **46,** 8665–8678 (2018).
10. Marzluff, W. F., Gongidi, P., Woods, K. R., Jin, J. & Maltais, L. J. The human and mouse replication-dependent histone genes. *Genomics* **80,** 487–498 (2002).
11. Martire, S. & Banaszynski, L. A. The roles of histone variants in fine-tuning chromatin organization and function. *Nature Reviews. Molecular Cell Biology* **21,** 522–541 (2020).
12. Arents, G. & Moudrianakis, E. N. The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. *Proceedings of the National Academy of Sciences of the United States of America* **92,** 11170–11174 (1995).
13. Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293,** 1074–1080 (2001).
14. Taylor, B. C. & Young, N. L. Combinations of histone post-translational modifications. *The Biochemical Journal* **478,** 511–532 (2021).
15. Happel, N. & Doenecke, D. Histone H1 and its isoforms: contribution to chromatin structure and function. *Gene* **431,** 1–12 (2009).
16. Jin, C. & Felsenfeld, G. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes & Development* **21,** 1519–1529 (2007).
17. Hake, S. B., Garcia, B. A., Kauer, M., Baker, S. P., Shabanowitz, J., Hunt, D. F. & Allis, C. D. Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* **102,** 6344–6349 (2005).
18. Sitbon, D., Boyarchuk, E., Dingli, F., Loew, D. & Almouzni, G. Histone variant H3.3 residue S31 is essential for Xenopus gastrulation regardless of the deposition pathway. *Nature Communications* **11,** 1256 (2020).

19. Martire, S., Gogate, A. A., Whitmill, A., Tafessu, A., Nguyen, J., Teng, Y.-C., Tastemel, M. & Banaszynski, L. A. Phosphorylation of histone H3.3 at serine 31 promotes p300 activity and enhancer acetylation. *Nature Genetics* **51,** 941–946 (2019).

20. Armache, A., Yang, S., Robbins, L. E., Durmaz, C., Daman, A. W., Jeong, J. Q., Martínez de Paz, A., Ravishankar, A., Arslan, T., Lin, S., Panchenko, T., Garcia, B. A., Hake, S. B., Li, H., Allis, C. D. & Josefowicz, S. Z. Phosphorylation of the ancestral histone variant H3.3 amplifies stimulation-induced transcription. *BioRxiv* (2019).

21. Goldberg, A. D., Banaszynski, L. A., Noh, K.-M., Lewis, P. W., Elsaesser, S. J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X., Wen, D., Chapgier, A., DeKelver, R. C., Miller, J. C., Lee, Y.-L., Boydston, E. A., Holmes, M. C., Gregory, P. D., Greally, J. M., Rafii, S., Yang, C., Scambler, P. J., Garrick, D., Gibbons, R. J., Higgs, D. R., Cristea, I. M., Urnov, F. D., Zheng, D. & Allis, C. D. Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* **140,** 678–691 (2010).

22. Wong, L. H., McGhie, J. D., Sim, M., Anderson, M. A., Ahn, S., Hannan, R. D., George, A. J., Morgan, K. A., Mann, J. R. & Choo, K. H. A. ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Research* **20,** 351–360 (2010).

23. Heitz, E. *Das Heterochromatin der Moose.* 762–818 (1928).

24. Jost, K. L., Bertulat, B. & Cardoso, M. C. Heterochromatin and gene positioning: inside, outside, any side? *Chromosoma* **121,** 555–563 (2012).

25. Baccarini, P. Sulle cinesi vegetative del "Cynomorium coccineum L.". *N Giorn Bot Ital N Ser* **15,** 189–203 (1908).

26. Brändle, F., Frühbauer, B. & Jagannathan, M. Principles and functions of pericentromeric satellite DNA clustering into chromocenters. *Seminars in Cell & Developmental Biology* **128,** 26–39 (2022).

27. Nishibuchi, G. & Déjardin, J. The molecular basis of the organization of repetitive DNA-containing constitutive heterochromatin in mammals. *Chromosome Research* **25,** 77–87 (2017).

28. Ekwall, K. Genome-wide analysis of HDAC function. *Trends in Genetics* **21,** 608–615 (2005).

29. Zhang, Y., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Bird, A. & Reinberg, D. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes & Development* **13,** 1924–1935 (1999).

30. Humphrey, G. W., Wang, Y., Russanova, V. R., Hirai, T., Qin, J., Nakatani, Y. & Howard, B. H. Stable histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and Mta-L1. *The Journal of Biological Chemistry* **276,** 6817–6824 (2001).

31. Kelly, R. D. W. & Cowley, S. M. The physiological roles of histone deacetylase (HDAC) 1 and 2: complex co-stars with multiple leading parts. *Biochemical Society Transactions* **41,** 741–749 (2013).

32. K, N., Allis, C. D. & Grewal, S. I. Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science* **293,** 1150–1155 (2001).

33. Turner, B. M. Cellular memory and the histone code. *Cell* **111,** 285–291 (2002).

34. Zhang, Y. & Reinberg, D. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes & Development* **15,** 2343–2360 (2001).

35. Brown, S. W. Heterochromatin. *Science* **151,** 417–425 (1966).

36. Trojer, P. & Reinberg, D. Facultative heterochromatin: is there a distinctive molecular signature? *Molecular Cell* **28,** 1–13 (2007).

37. Saksouk, N., Simboeck, E. & Déjardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics & Chromatin* **8,** 3 (2015).

38. Pardue, M. L. & Gall, J. G. Chromosomal localization of mouse satellite DNA. *Science* **168,** 1356–1358 (1970).

39. Wong, A. K. & Rattner, J. B. Sequence organization and cytological localization of the minor satellite of mouse. *Nucleic Acids Research* **16,** 11645–11661 (1988).

40. Joseph, A., Mitchell, A. R. & Miller, O. J. The organization of the mouse satellite DNA at centromeres. *Experimental Cell Research* **183,** 494–500 (1989).

41. Thakur, J., Packiaraj, J. & Henikoff, S. Sequence, chromatin and evolution of satellite DNA. *International Journal of Molecular Sciences* **22,** 4309 (2021).

42. Hörz, W. & Altenburger, W. Nudeotide sequence of mouse satellite DNA. *Nucleic Acids Research* **9,** 683–696 (1981).

43. Komissarov, A. S., Gavrilova, E. V., Demin, S. J., Ishov, A. M. & Podgornaya, O. I. Tandemly repeated DNA families in the mouse genome. *BMC Genomics* **12,** 531 (2011).

44. Kuznetsova, I. S., Prusov, A. N., Enukashvily, N. I. & Podgornaya, O. I. New types of mouse centromeric satellite DNAs. *Chromosome Research* **13,** 9–25 (2005).

45. Kuznetsova, I., Podgornaya, O. & Ferguson-Smith, M. A. High-resolution organization of mouse centromeric and pericentromeric DNA. *Cytogenetic and Genome Research* **112,** 248–255 (2006).

46. Bartholdi, M. F. Nuclear distribution of centromeres during the cell cycle of human diploid fibroblasts. *Journal of Cell Science* **99,** 255–263 (1991).

47. Guenatri, M., Bailly, D., Maison, C. & Almouzni, G. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *The Journal of Cell Biology* **166,** 493–505 (2004).

48. Ostromyshenskii, D. I., Chernyaeva, E. N., Kuznetsova, I. S. & Podgornaya, O. I. Mouse chromocenters DNA content: sequencing and in silico analysis. *BMC Genomics* **19,** 151 (2018).

49. Probst, A. V. & Almouzni, G. Pericentric heterochromatin: dynamic organization during early development in mammals. *Differentiation; Research in Biological Diversity* **76,** 15–23 (2008).

50. Hsu, T. C. & Arrighi, F. E. Distribution of constitutive heterochromatin in mamallian chromosomes. *Chromosoma* **34,** 243–253 (1971).

51. Manuelidis, L. Different central nervous system cell types display distinct and nonrandom arrangements of satellite DNA sequences. *Proceedings of the National Academy of Sciences of the United States of America* **81,** 3123–3127 (1984).

52. Martou, G. & De Boni, U. Nuclear topology of murine, cerebellar Purkinje neurons: changes as a function of development. *Experimental Cell Research* **256,** 131–139 (2000).

53. Mayer, R., Brero, A., von Hase, J., Schroeder, T., Cremer, T. & Dietzel, S. Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. *BMC Cell Biology* **6,** 44 (2005).

54. Solovei, I., Grandi, N., Knoth, R., Volk, B. & Cremer, T. Positional changes of pericentromeric heterochromatin and nucleoli in postmitotic Purkinje cells during murine cerebellum development. *Cytogenetic and Genome Research* **105,** 302–310 (2004).

55. Brero, A., Easwaran, H. P., Nowak, D., Grunewald, I., Cremer, T., Leonhardt, H. & Cardoso, M. C. Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation. *The Journal of Cell Biology* **169,** 733–743 (2005).

56. Singleton, M. K., Gonzales, M. L., Leung, K. N., Yasui, D. H., Schroeder, D. I., Dunaway, K. & LaSalle, J. M. MeCP2 is required for global heterochromatic and nucleolar changes during activity-dependent neuronal maturation. *Neurobiology of Disease* **43,** 190–200 (2011).

57. Bertulat, B., De Bonis, M. L., Della Ragione, F., Lehmkuhl, A., Milden, M., Storm, C., Jost, K. L., Scala, S., Hendrich, B., D'Esposito, M. & Cardoso, M. C. MeCP2 dependent heterochromatin reorganization during neural differentiation of a novel Mecp2-deficient embryonic stem cell reporter line. *Plos One* **7,** e47848 (2012).

58. Janssen, A., Colmenares, S. U. & Karpen, G. H. Heterochromatin: guardian of the genome. *Annual Review of Cell and Developmental Biology* **34,** 265–288 (2018).

59. Hahn, M., Dambacher, S. & Schotta, G. Heterochromatin dysregulation in human diseases. *Journal of Applied Physiology* **109,** 232–242 (2010).

60. Bernard, P., Maure, J. F., Partridge, J. F., Genier, S., Javerzat, J. P. & Allshire, R. C. Requirement of heterochromatin for cohesion at centromeres. *Science* **294,** 2539–2542 (2001).

61. Carone, D. M. & Lawrence, J. B. Heterochromatin instability in cancer: from the Barr body to satellites and the nuclear periphery. *Seminars in Cancer Biology* **23,** 99–108 (2013).

62. Agarwal, N., Becker, A., Jost, K. L., Haase, S., Thakur, B. K., Brero, A., Hardt, T., Kudo, S., Leonhardt, H. & Cardoso, M. C. MeCP2 Rett mutations affect large scale chromatin organization. *Human Molecular Genetics* **20,** 4187–4195 (2011).

63. Casas-Delucchi, C. S., Becker, A., Bolius, J. J. & Cardoso, M. C. Targeted manipulation of heterochromatin rescues MeCP2 Rett mutants and re-establishes higher order chromatin organization. *Nucleic Acids Research* **40,** e176 (2012).

64. Schildkraut, C. L. & Maio, J. J. Studies on the intranuclear distribution and properties of mouse satellite DNA. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* **161,** 76–93 (1968).

65. Rae, M. M. & Franke, W. W. The interphase distribution of satellite DNA-containing heterochromatin in mouse nuclei. *Chromosoma* **39,** 443–456 (1972).

66. Oh, J., Yeom, S., Park, J. & Lee, J.-S. The regional sequestration of heterochromatin structural proteins is critical to form and maintain silent chromatin. *Epigenetics & Chromatin* **15,** 5 (2022).

67. Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M. B., Talhout, W., Eussen, B. H., de Klein, A., Wessels, L., de Laat, W. & van Steensel, B. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453,** 948–951 (2008).

68. Vertii, A., Ou, J., Yu, J., Yan, A., Pagès, H., Liu, H., Zhu, L. J. & Kaufman, P. D. Two contrasting classes of nucleolus-associated domains in mouse fibroblast heterochromatin. *Genome Research* **29,** 1235–1249 (2019).

69. Penagos-Puig, A. & Furlan-Magaril, M. Heterochromatin as an important driver of genome organization. *Frontiers in cell and developmental biology* **8,** 579137 (2020).

70. Solovei, I., Kreysing, M., Lanctôt, C., Kösem, S., Peichl, L., Cremer, T., Guck, J. & Joffe, B. Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. *Cell* **137,** 356–368 (2009).

71. Solovei, I., Wang, A. S., Thanisch, K., Schmidt, C. S., Krebs, S., Zwerger, M., Cohen, T. V., Devys, D., Foisner, R., Peichl, L., Herrmann, H., Blum, H., Engelkamp, D., Stewart, C. L., Leonhardt, H. & Joffe, B. LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell* **152,** 584–598 (2013).

72. Falk, M., Feodorova, Y., Naumova, N., Imakaev, M., Lajoie, B. R., Leonhardt, H., Joffe, B., Dekker, J., Fudenberg, G., Solovei, I. & Mirny, L. A. Heterochromatin drives compartmentalization of inverted and conventional nuclei. *Nature* **570,** 395–399 (2019).

73. Lu, J. & Gilbert, D. M. Proliferation-dependent and cell cycle regulated transcription of mouse pericentric heterochromatin. *The Journal of Cell Biology* **179,** 411–421 (2007).

74. Rudert, F., Bronner, S., Garnier, J. M. & Dollé, P. Transcripts from opposite strands of gamma satellite DNA are differentially expressed during mouse development. *Mammalian Genome* **6,** 76–83 (1995).

75. Probst, A. V., Okamoto, I., Casanova, M., El Marjou, F., Le Baccon, P. & Almouzni, G. A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Developmental Cell* **19,** 625–638 (2010).

76. Martens, J. H. A., O'Sullivan, R. J., Braunschweig, U., Opravil, S., Radolf, M., Steinlein, P. & Jenuwein, T. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *The EMBO Journal* **24,** 800–812 (2005).

77. Terranova, R., Sauer, S., Merkenschlager, M. & Fisher, A. G. The reorganisation of constitutive heterochromatin in differentiating muscle requires HDAC activity. *Experimental Cell Research* **310,** 344–356 (2005).

78. Enukashvily, N. I., Malashicheva, A. B. & Waisertreiger, I. S.-R. Satellite DNA spatial localization and transcriptional activity in mouse embryonic E-14 and IOUD2 stem cells. *Cytogenetic and Genome Research* **124,** 277–287 (2009).

79. Enukashvily, N. I., Donev, R., Waisertreiger, I. S.-R. & Podgornaya, O. I. Human chromosome 1 satellite 3 DNA is decondensed, demethylated and transcribed in senescent cells and in A431 epithelial carcinoma cells. *Cytogenetic and Genome Research* **118,** 42–54 (2007).

80. Rizzi, N., Denegri, M., Chiodi, I., Corioni, M., Valgardsdottir, R., Cobianchi, F., Riva, S. & Biamonti, G. Transcriptional activation of a constitutive heterochromatic domain of the human genome in response to heat shock. *Molecular Biology of the Cell* **15,** 543–551 (2004).

81. Valgardsdottir, R., Chiodi, I., Giordano, M., Cobianchi, F., Riva, S. & Biamonti, G. Structural and functional characterization of noncoding repetitive RNAs transcribed in stressed human cells. *Molecular Biology of the Cell* **16,** 2597–2604 (2005).

82. Jolly, C., Metz, A., Govin, J., Vigneron, M., Turner, B. M., Khochbin, S. & Vourc'h, C. Stress-induced transcription of satellite III repeats. *The Journal of Cell Biology* **164,** 25–33 (2004).

83. Valgardsdottir, R., Chiodi, I., Giordano, M., Rossi, A., Bazzini, S., Ghigna, C., Riva, S. & Biamonti, G. Transcription of Satellite III non-coding RNAs is a general stress response in human cells. *Nucleic Acids Research* **36,** 423–434 (2008).

84. Eymery, A., Callanan, M. & Vourc'h, C. The secret message of heterochromatin: new insights into the mechanisms and function of centromeric and pericentric repeat sequence transcription. *The International Journal of Developmental Biology* **53,** 259–268 (2009).

85. Alexiadis, V., Ballestas, M. E., Sanchez, C., Winokur, S., Vedanarayanan, V., Warren, M. & Ehrlich, M. RNAPol-ChIP analysis of transcription from FSHD-linked tandem repeats and satellite DNA. *Biochimica et Biophysica Acta* **1769,** 29–40 (2007).

86. Zhu, Q., Pao, G. M., Huynh, A. M., Suh, H., Tonnu, N., Nederlof, P. M., Gage, F. H. & Verma, I. M. BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* **477,** 179–184 (2011).

87. Ting, D. T., Lipson, D., Paul, S., Brannigan, B. W., Akhavanfard, S., Coffman, E. J., Contino, G., Deshpande, V., Iafrate, A. J., Letovsky, S., Rivera, M. N., Bardeesy, N., Maheswaran, S. & Haber, D. A. Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* **331,** 593–596 (2011).

88. Gauchier, M., van Mierlo, G., Vermeulen, M. & Dejardin, J. Purification and enrichment of specific chromatin loci. *Nature Methods* **17,** 380–389 (2020).

89. Van Mierlo, G. & Vermeulen, M. Chromatin Proteomics to Study Epigenetics - Challenges and Opportunities. *Molecular & Cellular Proteomics* **20,** 100056 (2021).

90. Torrente, M. P., Zee, B. M., Young, N. L., Baliban, R. C., LeRoy, G., Floudas, C. A., Hake, S. B. & Garcia, B. A. Proteomic interrogation of human chromatin. *Plos One* **6,** e24747 (2011).

91. Van Mierlo, G., Dirks, R. A. M., De Clerck, L., Brinkman, A. B., Huth, M., Kloet, S. L., Saksouk, N., Kroeze, L. I., Willems, S., Farlik, M., Bock, C., Jansen, J. H., Deforce, D., Vermeulen, M., Déjardin, J., Dhaenens, M. & Marks, H. Integrative Proteomic Profiling Reveals PRC2-Dependent Epigenetic Crosstalk Maintains Ground-State Pluripotency. *Cell Stem Cell* **24,** 123–137.e8 (2019).

92. Federation, A. J., Nandakumar, V., Searle, B. C., Stergachis, A., Wang, H., Pino, L. K., Merrihew, G., Ting, Y. S., Howard, N., Kutyavin, T., MacCoss, M. J. & Stamatoyannopoulos, J. A. Highly parallel quantification and compartment localization of transcription factors and nuclear proteins. *Cell reports* **30,** 2463–2471.e5 (2020).

93. Becker, J. S., McCarthy, R. L., Sidoli, S., Donahue, G., Kaeding, K. E., He, Z., Lin, S., Garcia, B. A. & Zaret, K. S. Genomic and proteomic resolution of heterochromatin and its restriction of alternate fate genes. *Molecular Cell* **68,** 1023–1037.e15 (2017).

94. Prusov, A. N. & Zatsepina, O. V. Isolation of the chromocenter fraction from mouse liver nuclei. *Biochemistry (Moscow)* **67,** 423–431 (2002).

95. Zatsepina, O. V., Zharskaya, O. O. & Prusov, A. N. *Isolation of the constitutive heterochromatin from mouse liver nuclei* in *The nucleus* 169–180 (2008).

96. Déjardin, J. & Kingston, R. E. Purification of proteins associated with specific genomic Loci. *Cell* **136,** 175–186 (2009).

97. Marzec, P., Armenise, C., Pérot, G., Roumelioti, F.-M., Basyuk, E., Gagos, S., Chibon, F. & Déjardin, J. Nuclear-receptor-mediated telomere insertion leads to genome instability in ALT cancers. *Cell* **160,** 913–927 (2015).

98. Saksouk, N., Barth, T. K., Ziegler-Birling, C., Olova, N., Nowak, A., Rey, E., Mateos-Langerak, J., Urbach, S., Reik, W., Torres-Padilla, M.-E., Imhof, A., Déjardin, J. & Simboeck, E. Redundant mechanisms to form silent chromatin at pericentromeric regions rely on BEND3 and DNA methylation. *Molecular Cell* **56,** 580–594 (2014).

99. Gauchier, M., Kan, S., Barral, A., Sauzet, S., Agirre, E., Bonnell, E., Saksouk, N., Barth, T. K., Ide, S., Urbach, S., Wellinger, R. J., Luco, R. F., Imhof, A. & Déjardin, J. SETDB1-dependent heterochromatin stimulates alternative lengthening of telomeres. *Science Advances* **5,** eaav3673 (2019).

100. Dai, Y., Kennedy-Darling, J., Shortreed, M. R., Scalf, M., Gasch, A. P. & Smith, L. M. Multiplexed Sequence-Specific Capture of Chromatin and Mass Spectrometric Discovery of Associated Proteins. *Analytical Chemistry* **89,** 7841–7846 (2017).

101. Guillen-Ahlers, H., Rao, P. K., Levenstein, M. E., Kennedy-Darling, J., Perumalla, D. S., Jadhav, A. Y. L., Glenn, J. P., Ludwig-Kubinski, A., Drigalenko, E., Montoya, M. J., Göring, H. H., Anderson, C. D., Scalf, M., Gildersleeve, H. I. S., Cole, R., Greene, A. M., Oduro, A. K., Lazarova, K., Cesnik, A. J., Barfknecht, J., Cirillo, L. A., Gasch, A. P., Shortreed, M. R., Smith, L. M. & Olivier, M. HyCCAPP as a tool to characterize promoter DNA-protein interactions in Saccharomyces cerevisiae. *Genomics* **107,** 267–273 (2016).

102. Kennedy-Darling, J., Guillen-Ahlers, H., Shortreed, M. R., Scalf, M., Frey, B. L., Kendziorski, C., Olivier, M., Gasch, A. P. & Smith, L. M. Discovery of Chromatin-Associated Proteins via Sequence-Specific Capture and Mass Spectrometric Protein Identification in Saccharomyces cerevisiae. *Journal of Proteome Research* **13,** 3810–3825 (2014).

103. Grolimund, L., Aeby, E., Hamelin, R., Armand, F., Chiappe, D., Moniatte, M. & Lingner, J. A quantitative telomeric chromatin isolation protocol identifies different telomeric states. *Nature Communications* **4,** 2848 (2013).

104. Majerská, J., Redon, S. & Lingner, J. Quantitative telomeric chromatin isolation protocol for human cells. *Methods* **114,** 28–38 (2017).

105. Fujita, T., Asano, Y., Ohtsuka, J., Takada, Y., Saito, K., Ohki, R. & Fujii, H. Identification of telomere-associated molecules by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP). *Scientific Reports* **3,** 3171 (2013).

106. Liu, X., Zhang, Y., Chen, Y., Li, M., Zhou, F., Li, K., Cao, H., Ni, M., Liu, Y., Gu, Z., Dickerson, K. E., Xie, S., Hon, G. C., Xuan, Z., Zhang, M. Q., Shao, Z. & Xu, J. In Situ Capture of Chromatin Interactions by Biotinylated dCas9. *Cell* **170,** 1028–1043.e19 (2017).

107. Tsui, C., Inouye, C., Levy, M., Lu, A., Florens, L., Washburn, M. P. & Tjian, R. dCas9-targeted locus-specific protein isolation method identifies histone gene regulators. *Proceedings of the National Academy of Sciences of the United States of America* **115,** E2734–E2741 (2018).

108. Garcia-Exposito, L., Bournique, E., Bergoglio, V., Bose, A., Barroso-Gonzalez, J., Zhang, S., Roncaioli, J. L., Lee, M., Wallace, C. T., Watkins, S. C., Opresko, P. L., Hoffmann, J.-S. & O'Sullivan, R. J. Proteomic profiling reveals a specific role for translesion DNA polymerase $\eta$ in the alternative lengthening of telomeres. *Cell reports* **17,** 1858–1871 (2016).

109. Schmidtmann, E., Anton, T., Rombaut, P., Herzog, F. & Leonhardt, H. Determination of local chromatin composition by CasID. *Nucleus* **7,** 476–484 (2016).

110. Qiu, W., Xu, Z., Zhang, M., Zhang, D., Fan, H., Li, T., Wang, Q., Liu, P., Zhu, Z., Du, D., Tan, M., Wen, B. & Liu, Y. Determination of local chromatin interactions using a combined CRISPR and peroxidase APEX2 system. *Nucleic Acids Research* **47,** e52 (2019).

111. Gao, X. D., Tu, L.-C., Mir, A., Rodriguez, T., Ding, Y., Leszyk, J., Dekker, J., Shaffer, S. A., Zhu, L. J., Wolfe, S. A. & Sontheimer, E. J. C-BERST: defining subnuclear proteomic landscapes at genomic elements with dCas9-APEX2. *Nature Methods* **15,** 433–436 (2018).

112. Rozanova, S., Barkovits, K., Nikolov, M., Schmidt, C., Urlaub, H. & Marcus, K. Quantitative mass spectrometry-based proteomics: an overview. *Quantitative Methods in Proteomics*, 85–116 (2021).

113. Ludwig, A. K., Zhang, P. & Cardoso, M. C. Modifiers and readers of DNA modifications and their impact on genome structure, expression, and stability in disease. *Frontiers in genetics* **7,** 115 (2016).

114. Rausch, C., Hastert, F. D. & Cardoso, M. C. DNA modification readers and writers and their interplay. *Journal of molecular biology* **432,** 1731–1746 (2020).

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

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

117. Kriaucionis, S. & Bird, A. The major form of MeCP2 has a novel N-terminus generated by alternative splicing. *Nucleic Acids Research* **32,** 1818–1823 (2004).

118. Mnatzakanian, G. N., Lohi, H., Munteanu, I., Alfred, S. E., Yamada, T., MacLeod, P. J. M., Jones, J. R., Scherer, S. W., Schanen, N. C., Friez, M. J., Vincent, J. B. & Minassian, B. A. A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. *Nature Genetics* **36,** 339–341 (2004).

119. Dragich, J. M., Kim, Y.-H., Arnold, A. P. & Schanen, N. C. Differential distribution of the MeCP2 splice variants in the postnatal mouse brain. *The Journal of Comparative Neurology* **501,** 526–542 (2007).

120. Olson, C. O., Zachariah, R. M., Ezeonwuka, C. D., Liyanage, V. R. B. & Rastegar, M. Brain region-specific expression of MeCP2 isoforms correlates with DNA methylation within Mecp2 regulatory elements. *Plos One* **9,** e90645 (2014).

121. Fichou, Y., Nectoux, J., Bahi-Buisson, N., Rosas-Vargas, H., Girard, B., Chelly, J. & Bienvenu, T. The first missense mutation causing Rett syndrome specifically affecting the MeCP2_e1 isoform. *Neurogenetics* **10,** 127–133 (2009).

122. Saunders, C. J., Minassian, B. E., Chow, E. W. C., Zhao, W. & Vincent, J. B. Novel exon 1 mutations in MECP2 implicate isoform MeCP2_e1 in classical Rett syndrome. *American Journal of Medical Genetics. Part A* **149A,** 1019–1023 (2009).

123. Sheikh, T. I., de Paz, A. M., Akhtar, S., Ausió, J. & Vincent, J. B. MeCP2_E1 N-terminal modifications affect its degradation rate and are disrupted by the Ala2Val Rett mutation. *Human Molecular Genetics* **26,** 4132–4141 (2017).

124. Adams, V. H., McBryant, S. J., Wade, P. A., Woodcock, C. L. & Hansen, J. C. Intrinsic disorder and autonomous domain function in the multifunctional nuclear protein, MeCP2. *The Journal of Biological Chemistry* **282,** 15057–15064 (2007).

125. Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N. & Bird, A. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393,** 386–389 (1998).

126. Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J. & Wolffe, A. P. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genetics* **19,** 187–191 (1998).

127. Kokura, K., Kaul, S. C., Wadhwa, R., Nomura, T., Khan, M. M., Shinagawa, T., Yasukawa, T., Colmenares, C. & Ishii, S. The Ski protein family is required for MeCP2-mediated transcriptional repression. *The Journal of Biological Chemistry* **276,** 34115–34121 (2001).

128. Lunyak, V. V., Burgess, R., Prefontaine, G. G., Nelson, C., Sze, S.-H., Chenoweth, J., Schwartz, P., Pevzner, P. A., Glass, C., Mandel, G. & Rosenfeld, M. G. Corepressor-dependent silencing of chromosomal regions encoding neuronal genes. *Science* **298,** 1747–1752 (2002).

129. Lyst, M. J., Ekiert, R., Ebert, D. H., Merusi, C., Nowak, J., Selfridge, J., Guy, J., Kastan, N. R., Robinson, N. D., de Lima Alves, F., Rappsilber, J., Greenberg, M. E. & Bird, A. Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. *Nature Neuroscience* **16,** 898–902 (2013).

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

131. Meehan, R. R., Lewis, J. D. & Bird, A. P. Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Research* **20,** 5085–5092 (1992).

132. Ghosh, R. P., Nikitina, T., Horowitz-Scherer, R. A., Gierasch, L. M., Uversky, V. N., Hite, K., Hansen, J. C. & Woodcock, C. L. Unique physical properties and interactions of the domains of methylated DNA binding protein 2. *Biochemistry* **49,** 4395–4410 (2010).

133. Nikitina, T., Shi, X., Ghosh, R. P., Horowitz-Scherer, R. A., Hansen, J. C. & Woodcock, C. L. Multiple modes of interaction between the methylated DNA binding protein MeCP2 and chromatin. *Molecular and Cellular Biology* **27,** 864–877 (2007).

134. Lyst, M. J., Connelly, J., Merusi, C. & Bird, A. Sequence-specific DNA binding by AT-hook motifs in Me CP 2. *FEBS Letters* **590,** 2927–2933 (2016).

135. Wakefield, R. I., Smith, B. O., Nan, X., Free, A., Soteriou, A., Uhrin, D., Bird, A. P. & Barlow, P. N. The solution structure of the domain from MeCP2 that binds to methylated DNA. *Journal of Molecular Biology* **291,** 1055–1065 (1999).

136. Ho, K. L., McNae, I. W., Schmiedeberg, L., Klose, R. J., Bird, A. P. & Walkinshaw, M. D. MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Molecular Cell* **29,** 525–531 (2008).

137. Hansen, J. C., Wexler, B. B., Rogers, D. J., Hite, K. C., Panchenko, T., Ajith, S. & Black, B. E. DNA binding restricts the intrinsic conformational flexibility of methyl CpG binding protein 2 (MeCP2). *The Journal of Biological Chemistry* **286,** 18938–18948 (2011).

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

139. Chahrour, M., Jung, S. Y., Shaw, C., Zhou, X., Wong, S. T. C., Qin, J. & Zoghbi, H. Y. MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* **320,** 1224–1229 (2008).

140. Yasui, D. H., Peddada, S., Bieda, M. C., Vallero, R. O., Hogart, A., Nagarajan, R. P., Thatcher, K. N., Farnham, P. J. & Lasalle, J. M. Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. *Proceedings of the National Academy of Sciences of the United States of America* **104,** 19416–19421 (2007).

141. Lagger, S., Connelly, J. C., Schweikert, G., Webb, S., Selfridge, J., Ramsahoye, B. H., Yu, M., He, C., Sanguinetti, G., Sowers, L. C., Walkinshaw, M. D. & Bird, A. MeCP2 recognizes cytosine methylated tri-nucleotide and di-nucleotide sequences to tune transcription in the mammalian brain. *PLoS Genetics* **13,** e1006793 (2017).

142. Mo, A., Mukamel, E. A., Davis, F. P., Luo, C., Henry, G. L., Picard, S., Urich, M. A., Nery, J. R., Sejnowski, T. J., Lister, R., *et al.* Epigenomic signatures of neuronal diversity in the mammalian brain. *Neuron* **86,** 1369–1384 (2015).

143. Renthal, W., Boxer, L. D., Hrvatin, S., Li, E., Silberfeld, A., Nagy, M. A., Griffith, E. C., Vierbuchen, T. & Greenberg, M. E. Characterization of human mosaic Rett syndrome brain tissue by single-nucleus RNA sequencing. *Nature neuroscience* **21,** 1670–1679 (2018).

144. Mellén, M., Ayata, P., Dewell, S., Kriaucionis, S. & Heintz, N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* **151,** 1417–1430 (2012).

145. Spruijt, C. G., Gnerlich, F., Smits, A. H., Pfaffeneder, T., Jansen, P. W. T. C., Bauer, C., Münzel, M., Wagner, M., Müller, M., Khan, F., Eberl, H. C., Mensinga, A., Brinkman, A. B., Lephikov, K., Müller, U., Walter, J., Boelens, R., van Ingen, H., Leonhardt, H., Carell, T. & Vermeulen, M. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell* **152,** 1146–1159 (2013).

146. Pfaffeneder, T., Hackner, B., Truss, M., Münzel, M., Müller, M., Deiml, C. A., Hagemeier, C. & Carell, T. The discovery of 5-formylcytosine in embryonic stem cell DNA. *Angewandte Chemie International Edition* **50,** 7008–7012 (2011).

147. Szulwach, K. E., Li, X., Li, Y., Song, C.-X., Wu, H., Dai, Q., Irier, H., Upadhyay, A. K., Gearing, M., Levey, A. I., Vasanthakumar, A., Godley, L. A., Chang, Q., Cheng, X., He, C. & Jin, P. 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nature Neuroscience* **14,** 1607–1616 (2011).

148. Frauer, C., Rottach, A., Meilinger, D., Bultmann, S., Fellinger, K., Hasenöder, S., Wang, M., Qin, W., Söding, J., Spada, F. & Leonhardt, H. Different binding properties and function of CXXC zinc finger domains in Dnmt1 and Tet1. *Plos One* **6,** e16627 (2011).

149. Chen, L., Chen, K., Lavery, L. A., Baker, S. A., Shaw, C. A., Li, W. & Zoghbi, H. Y. MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome. *Proceedings of the National Academy of Sciences of the United States of America* **112,** 5509–5514 (2015).

150. Kinde, B., Gabel, H. W., Gilbert, C. S., Griffith, E. C. & Greenberg, M. E. Reading the unique DNA methylation landscape of the brain: Non-CpG methylation, hydroxymethylation, and MeCP2. *Proceedings of the National Academy of Sciences of the United States of America* **112,** 6800–6806 (2015).

151. Ludwig, A. K., Zhang, P., Hastert, F. D., Meyer, S., Rausch, C., Herce, H. D., Müller, U., Lehmkuhl, A., Hellmann, I., Trummer, C., Storm, C., Leonhardt, H. & Cardoso, M. C. Binding of MBD proteins to DNA blocks Tet1 function thereby modulating transcriptional noise. *Nucleic Acids Research* **45,** 2438–2457 (2017).

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

153. Yu, F., Zingler, N., Schumann, G. & Strätling, W. H. Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. *Nucleic acids research* **29,** 4493–4501 (2001).

154. Zhang, P., Ludwig, A. K., Hastert, F. D., Rausch, C., Lehmkuhl, A., Hellmann, I., Smets, M., Leonhardt, H. & Cardoso, M. C. L1 retrotransposition is activated by Ten-eleven-translocation protein 1 and repressed by methyl-CpG binding proteins. *Nucleus* **8,** 548–562 (2017).

155. Ren, J., Wen, L., Gao, X., Jin, C., Xue, Y. & Yao, X. DOG 1.0: illustrator of protein domain structures. *Cell Research* **19,** 271–273 (2009).

156. Fuks, F., Hurd, P. J., Wolf, D., Nan, X., Bird, A. P. & Kouzarides, T. The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *The Journal of Biological Chemistry* **278,** 4035–4040 (2003).

157. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410,** 116–120 (2001).

158. Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C. & Kouzarides, T. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410,** 120–124 (2001).

159. Agarwal, N., Hardt, T., Brero, A., Nowak, D., Rothbauer, U., Becker, A., Leonhardt, H. & Cardoso, M. C. MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation. *Nucleic Acids Research* **35,** 5402–5408 (2007).

160. Yamamoto, K. & Sonoda, M. Self-interaction of heterochromatin protein 1 is required for direct binding to histone methyltransferase, SUV39H1. *Biochemical and Biophysical Research Communications* **301,** 287–292 (2003).

161. Kimura, H. & Shiota, K. Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1. *The Journal of Biological Chemistry* **278,** 4806–4812 (2003).

162. Fuks, F., Burgers, W. A., Brehm, A., Hughes-Davies, L. & Kouzarides, T. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nature genetics* **24,** 88–91 (2000).

163. Rountree, M. R., Bachman, K. E. & Baylin, S. B. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nature genetics* **25,** 269–277 (2000).

164. Becker, A., Allmann, L., Hofstätter, M., Casà, V., Weber, P., Lehmkuhl, A., Herce, H. D. & Cardoso, M. C. Direct homo- and hetero-interactions of MeCP2 and MBD2. *Plos One* **8,** e53730 (2013).

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

166. Bedford, M. T., Chan, D. C. & Leder, P. FBP WW domains and the Abl SH3 domain bind to a specific class of proline-rich ligands. *The EMBO Journal* **16,** 2376–2383 (1997).

167. Buschdorf, J. P. & Strätling, W. H. A WW domain binding region in methyl-CpG-binding protein MeCP2: impact on Rett syndrome. *Journal of Molecular Medicine* **82,** 135–143 (2004).

168. Young, J. I., Hong, E. P., Castle, J. C., Crespo-Barreto, J., Bowman, A. B., Rose, M. F., Kang, D., Richman, R., Johnson, J. M., Berget, S. & Zoghbi, H. Y. Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proceedings of the National Academy of Sciences of the United States of America* **102,** 17551–17558 (2005).

169. Suzuki, M., Yamada, T., Kihara-Negishi, F., Sakurai, T. & Oikawa, T. Direct association between PU.1 and MeCP2 that recruits mSin3A-HDAC complex for PU.1-mediated transcriptional repression. *Oncogene* **22,** 8688–8698 (2003).

170. Matsumura, S., Persson, L. M., Wong, L. & Wilson, A. C. The latency-associated nuclear antigen interacts with MeCP2 and nucleosomes through separate domains. *Journal of Virology* **84,** 2318–2330 (2010).

171. Krithivas, A., Fujimuro, M., Weidner, M., Young, D. B. & Hayward, S. D. Protein interactions targeting the latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus to cell chromosomes. *Journal of virology* **76,** 11596–11604 (2002).

172. Kernohan, K. D., Jiang, Y., Tremblay, D. C., Bonvissuto, A. C., Eubanks, J. H., Mann, M. R. W. & Bérubé, N. G. ATRX partners with cohesin and MeCP2 and contributes to developmental silencing of imprinted genes in the brain. *Developmental Cell* **18,** 191–202 (2010).

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

174. Harikrishnan, K. N., Chow, M. Z., Baker, E. K., Pal, S., Bassal, S., Brasacchio, D., Wang, L., Craig, J. M., Jones, P. L., Sif, S. & El-Osta, A. Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. *Nature Genetics* **37,** 254–264 (2005).

175. Leoh, L. S., van Heertum, B., De Rijck, J., Filippova, M., Rios-Colon, L., Basu, A., Martinez, S. R., Tungteakkhun, S. S., Filippov, V., Christ, F., De Leon, M., Debyser, Z. & Casiano, C. A. The stress oncoprotein LEDGF/p75 interacts with the methyl CpG binding protein MeCP2 and influences its transcriptional activity. *Molecular Cancer Research* **10,** 378–391 (2012).

176. Gonzales, M. L., Adams, S., Dunaway, K. W. & LaSalle, J. M. Phosphorylation of distinct sites in MeCP2 modifies cofactor associations and the dynamics of transcriptional regulation. *Molecular and Cellular Biology* **32,** 2894–2903 (2012).

177. Long, S. W., Ooi, J. Y. Y., Yau, P. M. & Jones, P. L. A brain-derived MeCP2 complex supports a role for MeCP2 in RNA processing. *Bioscience reports* **31,** 333–343 (2011).

178. Jeffery, L. & Nakielny, S. Components of the DNA methylation system of chromatin control are RNA-binding proteins. *The Journal of Biological Chemistry* **279,** 49479–49487 (2004).

179. Bracaglia, G., Conca, B., Bergo, A., Rusconi, L., Zhou, Z., Greenberg, M. E., Landsberger, N., Soddu, S. & Kilstrup-Nielsen, C. Methyl-CpG-binding protein 2 is phosphorylated by homeodomain-interacting protein kinase 2 and contributes to apoptosis. *EMBO Reports* **10,** 1327–1333 (2009).

180. Lombardi, L. M., Zaghlula, M., Sztainberg, Y., Baker, S. A., Klisch, T. J., Tang, A. A., Huang, E. J. & Zoghbi, H. Y. An RNA interference screen identifies druggable regulators of MeCP2 stability. *Science Translational Medicine* **9,** eaaf7588 (2017).

181. Becker, A., Zhang, P., Allmann, L., Meilinger, D., Bertulat, B., Eck, D., Hofstaetter, M., Bartolomei, G., Hottiger, M. O., Schreiber, V., Leonhardt, H. & Cardoso, M. C. Poly(ADP-ribosyl)ation of Methyl CpG Binding Domain Protein 2 Regulates Chromatin Structure. *The Journal of Biological Chemistry* **291,** 4873–4881 (2016).

182. Mari, F., Azimonti, S., Bertani, I., Bolognese, F., Colombo, E., Caselli, R., Scala, E., Longo, I., Grosso, S., Pescucci, C., Ariani, F., Hayek, G., Balestri, P., Bergo, A., Badaracco, G., Zappella, M., Broccoli, V., Renieri, A., Kilstrup-Nielsen, C. & Landsberger, N. CDKL5 belongs to the same molecular pathway of MeCP2 and it is responsible for the early-onset seizure variant of Rett syndrome. *Human Molecular Genetics* **14,** 1935–1946 (2005).

183. Lin, C., Franco, B. & Rosner, M. R. CDKL5/Stk9 kinase inactivation is associated with neuronal developmental disorders. *Human Molecular Genetics* **14,** 3775–3786 (2005).

184. Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V. & Skrzypek, E. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Research* **43,** D512–20 (2015).

185. Humphrey, S. J., Yang, G., Yang, P., Fazakerley, D. J., Stöckli, J., Yang, J. Y. & James, D. E. Dynamic adipocyte phosphoproteome reveals that Akt directly regulates mTORC2. *Cell Metabolism* **17,** 1009–1020 (2013).

186. Shiromizu, T., Adachi, J., Watanabe, S., Murakami, T., Kuga, T., Muraoka, S. & Tomonaga, T. Identification of missing proteins in the neXtProt database and unregistered phosphopeptides in the PhosphoSitePlus database as part of the Chromosome-centric Human Proteome Project. *Journal of Proteome Research* **12,** 2414–2421 (2013).

187. Bian, Y., Song, C., Cheng, K., Dong, M., Wang, F., Huang, J., Sun, D., Wang, L., Ye, M. & Zou, H. An enzyme assisted RP-RPLC approach for in-depth analysis of human liver phosphoproteome. *Journal of Proteomics* **96,** 253–262 (2014).

188. Sharma, K., D'Souza, R. C. J., Tyanova, S., Schaab, C., Wiśniewski, J. R., Cox, J. & Mann, M. Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell reports* **8,** 1583–1594 (2014).

189. Huttlin, E. L., Jedrychowski, M. P., Elias, J. E., Goswami, T., Rad, R., Beausoleil, S. A., Villén, J., Haas, W., Sowa, M. E. & Gygi, S. P. A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **143,** 1174–1189 (2010).

190. Mertins, P., Mani, D. R., Ruggles, K. V., Gillette, M. A., Clauser, K. R., Wang, P., Wang, X., Qiao, J. W., Cao, S., Petralia, F., Kawaler, E., Mundt, F., Krug, K., Tu, Z., Lei, J. T., Gatza, M. L., Wilkerson, M., Perou, C. M., Yellapantula, V., Huang, K.-l., Lin, C., McLellan, M. D., Yan, P., Davies, S. R., Townsend, R. R., Skates, S. J., Wang, J., Zhang, B., Kinsinger, C. R., Mesri, M., Rodriguez, H., Ding, L., Paulovich, A. G., Fenyö, D., Ellis, M. J., Carr, S. A. & CPTAC, N. Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* **534,** 55–62 (2016).

191. Dephoure, N., Zhou, C., Villén, J., Beausoleil, S. A., Bakalarski, C. E., Elledge, S. J. & Gygi, S. P. A quantitative atlas of mitotic phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* **105,** 10762–10767 (2008).

192. Zanivan, S., Gnad, F., Wickström, S. A., Geiger, T., Macek, B., Cox, J., Fässler, R. & Mann, M. Solid tumor proteome and phosphoproteome analysis by high resolution mass spectrometry. *Journal of Proteome Research* **7,** 5314–5326 (2008).

193. Tweedie-Cullen, R. Y., Reck, J. M. & Mansuy, I. M. Comprehensive mapping of post-translational modifications on synaptic, nuclear, and histone proteins in the adult mouse brain. *Journal of Proteome Research* **8,** 4966–4982 (2009).

194. Zhou, Z., Hong, E. J., Cohen, S., Zhao, W.-N., Ho, H.-Y. H., Schmidt, L., Chen, W. G., Lin, Y., Savner, E., Griffith, E. C., Hu, L., Steen, J. A. J., Weitz, C. J. & Greenberg, M. E. Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription, dendritic growth, and spine maturation. *Neuron* **52,** 255–269 (2006).

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

196. Ebert, D. H., Gabel, H. W., Robinson, N. D., Kastan, N. R., Hu, L. S., Cohen, S., Navarro, A. J., Lyst, M. J., Ekiert, R., Bird, A. P. & Greenberg, M. E. Activity-dependent phosphorylation of MeCP2 threonine 308 regulates interaction with NCoR. *Nature* **499,** 341–345 (2013).

197. Mertins, P., Yang, F., Liu, T., Mani, D. R., Petyuk, V. A., Gillette, M. A., Clauser, K. R., Qiao, J. W., Gritsenko, M. A., Moore, R. J., Levine, D. A., Townsend, R., Erdmann-Gilmore, P., Snider, J. E., Davies, S. R., Ruggles, K. V., Fenyo, D., Kitchens, R. T., Li, S., Olvera, N., Dao, F., Rodriguez, H., Chan, D. W., Liebler, D., White, F., Rodland, K. D., Mills, G. B., Smith, R. D., Paulovich, A. G., Ellis, M. & Carr, S. A. Ischemia in tumors induces early and sustained phosphorylation changes in stress kinase pathways but does not affect global protein levels. *Molecular & Cellular Proteomics* **13,** 1690–1704 (2014).

198. Geoghegan, V., Guo, A., Trudgian, D., Thomas, B. & Acuto, O. Comprehensive identification of arginine methylation in primary T cells reveals regulatory roles in cell signalling. *Nature Communications* **6,** 6758 (2015).

199. Kettenbach, A. N., Schweppe, D. K., Faherty, B. K., Pechenick, D., Pletnev, A. A. & Gerber, S. A. Quantitative phosphoproteomics identifies substrates and functional modules of Aurora and Polo-like kinase activities in mitotic cells. *Science Signaling* **4,** rs5 (2011).

200. Jung, S. Y., Li, Y., Wang, Y., Chen, Y., Zhao, Y. & Qin, J. Complications in the assignment of 14 and 28 Da mass shift detected by mass spectrometry as in vivo methylation from endogenous proteins. *Analytical Chemistry* **80,** 1721–1729 (2008).

201. Bergo, A., Strollo, M., Gai, M., Barbiero, I., Stefanelli, G., Sertic, S., Cobolli Gigli, C., Di Cunto, F., Kilstrup-Nielsen, C. & Landsberger, N. Methyl-CpG binding protein 2 (MeCP2) localizes at the centrosome and is required for proper mitotic spindle organization. *The Journal of Biological Chemistry* **290,** 3223–3237 (2015).

202. D'Annessa, I., Gandaglia, A., Brivio, E., Stefanelli, G., Frasca, A., Landsberger, N. & Di Marino, D. Tyr120Asp mutation alters domain flexibility and dynamics of MeCP2 DNA binding domain leading to impaired DNA interaction: Atomistic characterization of a Rett syndrome causing mutation. *Biochimica et biophysica acta. General subjects* **1862,** 1180–1189 (2018).

203. Wagner, S. A., Beli, P., Weinert, B. T., Nielsen, M. L., Cox, J., Mann, M. & Choudhary, C. A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. *Molecular & Cellular Proteomics* **10,** M111.013284 (2011).

204. Olsen, J. V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M. L., Jensen, L. J., Gnad, F., Cox, J., Jensen, T. S., Nigg, E. A., Brunak, S. & Mann, M. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Science Signaling* **3,** ra3 (2010).

205. Guo, A., Gu, H., Zhou, J., Mulhern, D., Wang, Y., Lee, K. A., Yang, V., Aguiar, M., Kornhauser, J., Jia, X., Ren, J., Beausoleil, S. A., Silva, J. C., Vemulapalli, V., Bedford, M. T. & Comb, M. J. Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. *Molecular & Cellular Proteomics* **13,** 372–387 (2014).

206. Larsen, S. C., Sylvestersen, K. B., Mund, A., Lyon, D., Mullari, M., Madsen, M. V., Daniel, J. A., Jensen, L. J. & Nielsen, M. L. Proteome-wide analysis of arginine monomethylation reveals widespread occurrence in human cells. *Science Signaling* **9,** rs9 (2016).

207. Jungmichel, S., Rosenthal, F., Altmeyer, M., Lukas, J., Hottiger, M. O. & Nielsen, M. L. Proteome-wide identification of poly (ADP-Ribosyl) ation targets in different genotoxic stress responses. *Molecular cell* **52,** 272–285 (2013).

208. Stefanelli, G., Gandaglia, A., Costa, M., Cheema, M. S., Di Marino, D., Barbiero, I., Kilstrup-Nielsen, C., Ausió, J. & Landsberger, N. Brain phosphorylation of MeCP2 at serine 164 is developmentally regulated and globally alters its chromatin association. *Scientific Reports* **6,** 28295 (2016).

209. Yi, T., Zhai, B., Yu, Y., Kiyotsugu, Y., Raschle, T., Etzkorn, M., Seo, H.-C., Nagiec, M., Luna, R. E., Reinherz, E. L., Blenis, J., Gygi, S. P. & Wagner, G. Quantitative phosphoproteomic analysis reveals system-wide signaling pathways downstream of SDF-1/CXCR4 in breast cancer stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **111,** E2182–E2190 (2014).

210. Carrier, M., Joint, M., Lutzing, R., Page, A. & Rochette-Egly, C. Phosphoproteome and Transcriptome of RA-Responsive and RA-Resistant Breast Cancer Cell Lines. *Plos One* **11,** e0157290 (2016).

211. Lundby, A., Lage, K., Weinert, B. T., Bekker-Jensen, D. B., Secher, A., Skovgaard, T., Kelstrup, C. D., Dmytriyev, A., Choudhary, C., Lundby, C. & Olsen, J. V. Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns. *Cell reports* **2,** 419–431 (2012).

212. Akimov, V., Barrio-Hernandez, I., Hansen, S. V. F., Hallenborg, P., Pedersen, A.-K., Bekker-Jensen, D. B., Puglia, M., Christensen, S. D. K., Vanselow, J. T., Nielsen, M. M., Kratchmarova, I., Kelstrup, C. D., Olsen, J. V. & Blagoev, B. UbiSite approach for comprehensive mapping of lysine and N-terminal ubiquitination sites. *Nature Structural & Molecular Biology* **25,** 631–640 (2018).

213. Cheng, J., Huang, M., Zhu, Y., Xin, Y.-J., Zhao, Y.-K., Huang, J., Yu, J.-X., Zhou, W.-H. & Qiu, Z. SUMOylation of MeCP2 is essential for transcriptional repression and hippocampal synapse development. *Journal of Neurochemistry* **128,** 798–806 (2014).

214. Chen, R.-Q., Yang, Q.-K., Lu, B.-W., Yi, W., Cantin, G., Chen, Y.-L., Fearns, C., Yates, J. R. & Lee, J.-D. CDC25B mediates rapamycin-induced oncogenic responses in cancer cells. *Cancer Research* **69,** 2663–2668 (2009).

215. Wu, Z., Cheng, Z., Sun, M., Wan, X., Liu, P., He, T., Tan, M. & Zhao, Y. A chemical proteomics approach for global analysis of lysine monomethylome profiling. *Molecular & Cellular Proteomics* **14,** 329–339 (2015).

216. Liu, F., Ni, J.-J., Huang, J.-J., Kou, Z.-W. & Sun, F.-Y. VEGF overexpression enhances the accumulation of phospho-S292 MeCP2 in reactive astrocytes in the adult rat striatum following cerebral ischemia. *Brain Research* **1599,** 32–43 (2015).

217. Parker, B. L., Yang, G., Humphrey, S. J., Chaudhuri, R., Ma, X., Peterman, S. & James, D. E. Targeted phosphoproteomics of insulin signaling using data-independent acquisition mass spectrometry. *Science Signaling* **8,** rs6 (2015).

218. Beli, P., Lukashchuk, N., Wagner, S. A., Weinert, B. T., Olsen, J. V., Baskcomb, L., Mann, M., Jackson, S. P. & Choudhary, C. Proteomic investigations reveal a role for RNA processing factor THRAP3 in the DNA damage response. *Molecular Cell* **46,** 212–225 (2012).

219. Weinert, B. T., Schölz, C., Wagner, S. A., Iesmantavicius, V., Su, D., Daniel, J. A. & Choudhary, C. Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation. *Cell reports* **4,** 842–851 (2013).

220. Dhayalan, A., Kudithipudi, S., Rathert, P. & Jeltsch, A. Specificity analysis-based identification of new methylation targets of the SET7/9 protein lysine methyltransferase. *Chemistry & Biology* **18,** 111–120 (2011).

221. Wu, Q., Cheng, Z., Zhu, J., Xu, W., Peng, X., Chen, C., Li, W., Wang, F., Cao, L., Yi, X., Wu, Z., Li, J. & Fan, P. Suberoylanilide hydroxamic acid treatment reveals crosstalks among proteome, ubiquitylome and acetylome in non-small cell lung cancer A549 cell line. *Scientific Reports* **5,** 9520 (2015).

222. Yang, F., Stenoien, D. L., Strittmatter, E. F., Wang, J., Ding, L., Lipton, M. S., Monroe, M. E., Nicora, C. D., Gristenko, M. A., Tang, K., Fang, R., Adkins, J. N., Camp, D. G., Chen, D. J. & Smith, R. D. Phosphoproteome profiling of human skin fibroblast cells in response to low- and high-dose irradiation. *Journal of Proteome Research* **5,** 1252–1260 (2006).

223. Grimsrud, P. A., Carson, J. J., Hebert, A. S., Hubler, S. L., Niemi, N. M., Bailey, D. J., Jochem, A., Stapleton, D. S., Keller, M. P., Westphall, M. S., Yandell, B. S., Attie, A. D., Coon, J. J. & Pagliarini, D. J. A quantitative map of the liver mitochondrial phosphoproteome reveals posttranslational control of ketogenesis. *Cell Metabolism* **16,** 672–683 (2012).

224. Deng, J. V., Rodriguiz, R. M., Hutchinson, A. N., Kim, I.-H., Wetsel, W. C. & West, A. E. MeCP2 in the nucleus accumbens contributes to neural and behavioral responses to psychostimulants. *Nature Neuroscience* **13,** 1128–1136 (2010).

225. Li, H., Zhong, X., Chau, K. F., Williams, E. C. & Chang, Q. Loss of activity-induced phosphorylation of MeCP2 enhances synaptogenesis, LTP and spatial memory. *Nature Neuroscience* **14,** 1001–1008 (2011).

226. Wang, Z., Udeshi, N. D., O'Malley, M., Shabanowitz, J., Hunt, D. F. & Hart, G. W. Enrichment and site mapping of O-linked N-acetylglucosamine by a combination of chemical/enzymatic tagging, photochemical cleavage, and electron transfer dissociation mass spectrometry. *Molecular & Cellular Proteomics* **9,** 153–160 (2010).

227. Alfaro, J. F., Gong, C.-X., Monroe, M. E., Aldrich, J. T., Clauss, T. R. W., Purvine, S. O., Wang, Z., Camp, D. G., Shabanowitz, J., Stanley, P., Hart, G. W., Hunt, D. F., Yang, F. & Smith, R. D. Tandem mass spectrometry identifies many mouse brain O-GlcNAcylated proteins including EGF domain-specific O-GlcNAc transferase targets. *Proceedings of the National Academy of Sciences of the United States of America* **109,** 7280–7285 (2012).

228. Trinidad, J. C., Barkan, D. T., Gulledge, B. F., Thalhammer, A., Sali, A., Schoepfer, R. & Burlingame, A. L. Global identification and characterization of both O-GlcNAcylation and phosphorylation at the murine synapse. *Molecular & Cellular Proteomics* **11,** 215–229 (2012).

229. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V. & Mann, M. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325,** 834–840 (2009).

230. Schweppe, D. K., Rigas, J. R. & Gerber, S. A. Quantitative phosphoproteomic profiling of human non-small cell lung cancer tumors. *Journal of Proteomics* **91,** 286–296 (2013).

231. Chen, W. G., Chang, Q., Lin, Y., Meissner, A., West, A. E., Griffith, E. C., Jaenisch, R. & Greenberg, M. E. Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science* **302,** 885–889 (2003).

232. Martinowich, K., Hattori, D., Wu, H., Fouse, S., He, F., Hu, Y., Fan, G. & Sun, Y. E. DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science* **302,** 890–893 (2003).

233. Song, C., Feodorova, Y., Guy, J., Peichl, L., Jost, K. L., Kimura, H., Cardoso, M. C., Bird, A., Leonhardt, H., Joffe, B. & Solovei, I. DNA methylation reader MECP2: cell type- and differentiation stage-specific protein distribution. *Epigenetics & Chromatin* **7,** 17 (2014).

234. Cohen, S., Gabel, H. W., Hemberg, M., Hutchinson, A. N., Sadacca, L. A., Ebert, D. H., Harmin, D. A., Greenberg, R. S., Verdine, V. K., Zhou, Z., Wetsel, W. C., West, A. E. & Greenberg, M. E. Genome-wide activity-dependent MeCP2 phosphorylation regulates nervous system development and function. *Neuron* **72,** 72–85 (2011).

235. Inui, K., Akagi, M., Ono, J., Tsukamoto, H., Shimono, K., Mano, T., Imai, K., Yamada, M., Muramatsu, T., Sakai, N. & Okada, S. Mutational analysis of MECP2 in Japanese patients with atypical Rett syndrome. *Brain & Development* **23,** 212–215 (2001).

236. Hagberg, B., Aicardi, J., Dias, K. & Ramos, O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Annals of Neurology* **14,** 471–479 (1983).

237. Hagberg, B. Clinical manifestations and stages of Rett syndrome. *Mental retardation and developmental disabilities research reviews* **8,** 61–65 (2002).

238. Goffin, D., Allen, M., Zhang, L., Amorim, M., Wang, I.-T. J., Reyes, A.-R. S., Mercado-Berton, A., Ong, C., Cohen, S., Hu, L., Blendy, J. A., Carlson, G. C., Siegel, S. J., Greenberg, M. E. & Zhou, Z. Rett syndrome mutation MeCP2 T158A disrupts DNA binding, protein stability and ERP responses. *Nature Neuroscience* **15,** 274–283 (2011).

239. Brown, K., Selfridge, J., Lagger, S., Connelly, J., De Sousa, D., Kerr, A., Webb, S., Guy, J., Merusi, C., Koerner, M. V. & Bird, A. The molecular basis of variable phenotypic severity among common missense mutations causing Rett syndrome. *Human Molecular Genetics* **25,** 558–570 (2016).

240. Ballestar, E., Yusufzai, T. M. & Wolffe, A. P. 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,** 7100–7106 (2000).

241. Georgel, P. T., Horowitz-Scherer, R. A., Adkins, N., Woodcock, C. L., Wade, P. A. & Hansen, J. C. Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. *The Journal of Biological Chemistry* **278,** 32181–32188 (2003).

242. Baker, S. A., Chen, L., Wilkins, A. D., Yu, P., Lichtarge, O. & Zoghbi, H. Y. An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. *Cell* **152,** 984–996 (2013).

243. Yusufzai, T. M. & Wolffe, A. P. Functional consequences of Rett syndrome mutations on human MeCP2. *Nucleic Acids Research* **28,** 4172–4179 (2000).

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

245. Chen, R. Z., Akbarian, S., Tudor, M. & Jaenisch, R. Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nature Genetics* **27,** 327–331 (2001).

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

247. Van Esch, H., Bauters, M., Ignatius, J., Jansen, M., Raynaud, M., Hollanders, K., Lugtenberg, D., Bienvenu, T., Jensen, L. R., Gecz, J., Moraine, C., Marynen, P., Fryns, J.-P. & Froyen, G. Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. *American Journal of Human Genetics* **77,** 442–453 (2005).

248. Luikenhuis, S., Giacometti, E., Beard, C. F. & Jaenisch, R. 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,** 6033–6038 (2004).

249. Garg, S. K., Lioy, D. T., Cheval, H., McGann, J. C., Bissonnette, J. M., Murtha, M. J., Foust, K. D., Kaspar, B. K., Bird, A. & Mandel, G. Systemic delivery of MeCP2 rescues behavioral and cellular deficits in female mouse models of Rett syndrome. *The Journal of Neuroscience* **33,** 13612–13620 (2013).

250. Giacometti, E., Luikenhuis, S., Beard, C. & Jaenisch, R. Partial rescue of MeCP2 deficiency by postnatal activation of MeCP2. *Proceedings of the National Academy of Sciences of the United States of America* **104,** 1931–1936 (2007).

251. Kerr, B., Soto C, J., Saez, M., Abrams, A., Walz, K. & Young, J. I. Transgenic complementation of MeCP2 deficiency: phenotypic rescue of Mecp2-null mice by isoform-specific transgenes. *European Journal of Human Genetics* **20,** 69–76 (2012).

252. Robinson, L., Guy, J., McKay, L., Brockett, E., Spike, R. C., Selfridge, J., De Sousa, D., Merusi, C., Riedel, G., Bird, A. & Cobb, S. R. Morphological and functional reversal of phenotypes in a mouse model of Rett syndrome. *Brain: A Journal of Neurology* **135,** 2699–2710 (2012).

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

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

255. Wang, L., Hu, M., Zuo, M.-Q., Zhao, J., Wu, D., Huang, L., Wen, Y., Li, Y., Chen, P., Bao, X., Dong, M.-Q., Li, G. & Li, P. Rett syndrome-causing mutations compromise MeCP2-mediated liquid-liquid phase separation of chromatin. *Cell Research* **30,** 393–407 (2020).

256. Fan, C., Zhang, H., Fu, L., Li, Y., Du, Y., Qiu, Z. & Lu, F. Rett mutations attenuate phase separation of MeCP2. *Cell discovery* **6,** 38 (2020).

257. Larson, A. G., Elnatan, D., Keenen, M. M., Trnka, M. J., Johnston, J. B., Burlingame, A. L., Agard, D. A., Redding, S. & Narlikar, G. J. Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature* **547,** 236–240 (2017).

258. Strom, A. R., Emelyanov, A. V., Mir, M., Fyodorov, D. V., Darzacq, X. & Karpen, G. H. Phase separation drives heterochromatin domain formation. *Nature* **547,** 241–245 (2017).

259. Larson, A. G. & Narlikar, G. J. The role of phase separation in heterochromatin formation, function, and regulation. *Biochemistry* **57,** 2540–2548 (2018).

260. Brown, K. E., Guest, S. S., Smale, S. T., Hahm, K., Merkenschlager, M. & Fisher, A. G. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **91,** 845–854 (1997).

261. Wessel, D. & Flügge, U. I. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Analytical Biochemistry* **138,** 141–143 (1984).

262. Kanashova, T., Popp, O., Orasche, J., Karg, E., Harndorf, H., Stengel, B., Sklorz, M., Streibel, T., Zimmermann, R. & Dittmar, G. Differential proteomic analysis of mouse macrophages exposed to adsorbate-loaded heavy fuel oil derived combustion particles using an automated sample-preparation workflow. *Analytical and Bioanalytical Chemistry* **407,** 5965–5976 (2015).

263. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. R. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nature Protocols* **4,** 484–494 (2009).

264. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* **26,** 1367–1372 (2008).

265. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **11,** 2301–2319 (2016).

266. Consortium, U. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research* **47,** D506–D515 (2019).

267. Geiger, T., Wisniewski, J. R., Cox, J., Zanivan, S., Kruger, M., Ishihama, Y. & Mann, M. Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. *Nature Protocols* **6,** 147–157 (2011).

268. R Core Team. *R: A language and environment for statistical computing* R Foundation for Statistical Computing (Vienna, Austria, 2013).

269. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10,** 48 (2009).

270. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9,** 671–675 (2012).

271. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. & Cardona, A. Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9,** 676–682 (2012).

272. Kalimuthu, S. N. & Chetty, R. Gene of the month: SMARCB1. *Journal of Clinical Pathology* **69,** 484–489 (2016).

273. Klose, R. J. & Bird, A. P. MeCP2 behaves as an elongated monomer that does not stably associate with the Sin3a chromatin remodeling complex. *The Journal of Biological Chemistry* **279,** 46490–46496 (2004).

274. Gibson, B. A., Doolittle, L. K., Schneider, M. W. G., Jensen, L. E., Gamarra, N., Henry, L., Gerlich, D. W., Redding, S. & Rosen, M. K. Organization of chromatin by intrinsic and regulated phase separation. *Cell* **179,** 470–484.e21 (2019).

275. Kochanova, N. Y., Schauer, T., Mathias, G. P., Lukacs, A., Schmidt, A., Flatley, A., Schepers, A., Thomae, A. W. & Imhof, A. A multi-layered structure of the interphase chromocenter revealed by proximity-based biotinylation. *Nucleic Acids Research* **48,** 4161–4178 (2020).

276. Minc, E., Allory, Y., Worman, H. J., Courvalin, J. C. & Buendia, B. Localization and phosphorylation of HP1 proteins during the cell cycle in mammalian cells. *Chromosoma* **108,** 220–234 (1999).

277. Hayakawa, T., Haraguchi, T., Masumoto, H. & Hiraoka, Y. Cell cycle behavior of human HP1 subtypes: distinct molecular domains of HP1 are required for their centromeric localization during interphase and metaphase. *Journal of Cell Science* **116,** 3327–3338 (2003).

278. Gilbert, N., Boyle, S., Sutherland, H., de Las Heras, J., Allan, J., Jenuwein, T. & Bickmore, W. A. Formation of facultative heterochromatin in the absence of HP1. *The EMBO Journal* **22,** 5540–5550 (2003).

279. Ritou, E., Bai, M. & Georgatos, S. D. Variant-specific patterns and humoral regulation of HP1 proteins in human cells and tissues. *Journal of Cell Science* **120,** 3425–3435 (2007).

280. Eberhart, A., Feodorova, Y., Song, C., Wanner, G., Kiseleva, E., Furukawa, T., Kimura, H., Schotta, G., Leonhardt, H., Joffe, B. & Solovei, I. Epigenetics of eu- and heterochromatin in inverted and conventional nuclei from mouse retina. *Chromosome Research* **21,** 535–554 (2013).

281. Kwon, S. H. & Workman, J. L. The changing faces of HP1: From heterochromatin formation and gene silencing to euchromatic gene expression: HP1 acts as a positive regulator of transcription. *Bioessays: News and Reviews in Molecular, Cellular and Developmental Biology* **33,** 280–289 (2011).

282. Boumendil, C., Hari, P., Olsen, K. C. F., Acosta, J. C. & Bickmore, W. A. Nuclear pore density controls heterochromatin reorganization during senescence. *Genes & Development* **33,** 144–149 (2019).

283. Iglesias, N., Paulo, J. A., Tatarakis, A., Wang, X., Edwards, A. L., Bhanu, N. V., Garcia, B. A., Haas, W., Gygi, S. P. & Moazed, D. Native chromatin proteomics reveals a role for specific nucleoporins in heterochromatin organization and maintenance. *Molecular Cell* **77,** 51–66.e8 (2020).

284. Pascual-Garcia, P. & Capelson, M. The nuclear pore complex and the genome: organizing and regulatory principles. *Current Opinion in Genetics & Development* **67,** 142–150 (2021).

285. Towbin, B. D., Gonzalez-Sandoval, A. & Gasser, S. M. Mechanisms of heterochromatin subnuclear localization. *Trends in Biochemical Sciences* **38,** 356–363 (2013).

286. Ji, X., Dadon, D. B., Abraham, B. J., Lee, T. I., Jaenisch, R., Bradner, J. E. & Young, R. A. Chromatin proteomic profiling reveals novel proteins associated with histone-marked genomic regions. *Proceedings of the National Academy of Sciences of the United States of America* **112,** 3841–3846 (2015).

287. Lee, S. K. & Wang, W. Roles of topoisomerases in heterochromatin, aging, and diseases. *Genes* **10,** 884 (2019).

288. Durand-Dubief, M., Persson, J., Norman, U., Hartsuiker, E. & Ekwall, K. Topoisomerase I regulates open chromatin and controls gene expression in vivo. *The EMBO Journal* **29,** 2126–2134 (2010).

289. Sperling, A. S., Jeong, K. S., Kitada, T. & Grunstein, M. Topoisomerase II binds nucleosome-free DNA and acts redundantly with topoisomerase I to enhance recruitment of RNA Pol II in budding yeast. *Proceedings of the National Academy of Sciences of the United States of America* **108,** 12693–12698 (2011).

290. Dinh, T. T., Gao, L., Liu, X., Li, D., Li, S., Zhao, Y., O'Leary, M., Le, B., Schmitz, R. J., Manavella, P. A., Li, S., Weigel, D., Pontes, O., Ecker, J. R. & Chen, X. DNA topoisomerase 1$\alpha$ promotes transcriptional silencing of transposable elements through DNA methylation and histone lysine 9 dimethylation in Arabidopsis. *PLoS Genetics* **10,** e1004446 (2014).

291. Mengoli, V., Bucciarelli, E., Lattao, R., Piergentili, R., Gatti, M. & Bonaccorsi, S. The analysis of mutant alleles of different strength reveals multiple functions of topoisomerase 2 in regulation of Drosophila chromosome structure. *PLoS Genetics* **10,** e1004739 (2014).

292. Bell, J. C. & Straight, A. F. Condensing chromosome condensation. *Nature Cell Biology* **17,** 964–965 (2015).

293. Blattes, R., Monod, C., Susbielle, G., Cuvier, O., Wu, J.-h., Hsieh, T.-s., Laemmli, U. K. & Käs, E. Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide. *The EMBO Journal* **25,** 2397–2408 (2006).

294. Buschbeck, M., Uribesalgo, I., Wibowo, I., Rué, P., Martin, D., Gutierrez, A., Morey, L., Guigó, R., López-Schier, H. & Di Croce, L. The histone variant macroH2A is an epigenetic regulator of key developmental genes. *Nature Structural & Molecular Biology* **16,** 1074–1079 (2009).

295. Douet, J., Corujo, D., Malinverni, R., Renauld, J., Sansoni, V., Posavec Marjanović, M., Cantariño, N., Valero, V., Mongelard, F., Bouvet, P., Imhof, A., Thiry, M. & Buschbeck, M. MacroH2A histone variants maintain nuclear organization and heterochromatin architecture. *Journal of Cell Science* **130,** 1570–1582 (2017).

296. Kozlowski, M., Corujo, D., Hothorn, M., Guberovic, I., Mandemaker, I. K., Blessing, C., Sporn, J., Gutierrez-Triana, A., Smith, R., Portmann, T., Treier, M., Scheffzek, K., Huet, S., Timinszky, G., Buschbeck, M. & Ladurner, A. G. MacroH2A histone variants limit chromatin plasticity through two distinct mechanisms. *EMBO Reports* **19,** e44445 (2018).

297. Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *The Journal of Biological Chemistry* **273,** 5858–5868 (1998).

298. Kuo, L. J. & Yang, L.-X. Gamma-H2AX - a novel biomarker for DNA double-strand breaks. *In Vivo* **22,** 305–309 (2008).

299. Soldi, M. & Bonaldi, T. The proteomic investigation of chromatin functional domains reveals novel synergisms among distinct heterochromatin components. *Molecular & Cellular Proteomics* **12,** 764–780 (2013).

300. Bassing, C. H., Chua, K. F., Sekiguchi, J., Suh, H., Whitlow, S. R., Fleming, J. C., Monroe, B. C., Ciccone, D. N., Yan, C., Vlasakova, K., Livingston, D. M., Ferguson, D. O., Scully, R. & Alt, F. W. Increased ionizing radiation sensitivity and genomic instability in the absence of histone H2AX. *Proceedings of the National Academy of Sciences of the United States of America* **99,** 8173–8178 (2002).

301. Fernandez-Capetillo, O., Mahadevaiah, S. K., Celeste, A., Romanienko, P. J., Camerini-Otero, R. D., Bonner, W. M., Manova, K., Burgoyne, P. & Nussenzweig, A. H2AX is required for chromatin remodeling and inactivation of sex chromosomes in male mouse meiosis. *Developmental Cell* **4,** 497–508 (2003).

302. Lewis, P. W., Elsaesser, S. J., Noh, K.-M., Stadler, S. C. & Allis, C. D. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proceedings of the National Academy of Sciences of the United States of America* **107,** 14075–14080 (2010).

303. Drané, P., Ouararhni, K., Depaux, A., Shuaib, M. & Hamiche, A. The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes & Development* **24,** 1253–1265 (2010).

304. Wong, L. H., McGhie, J. D., Sim, M., Anderson, M. A., Ahn, S., Hannan, R. D., George, A. J., Morgan, K. A., Mann, J. R. & Choo, K. H. A. ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. *Genome Research* **20,** 351–360 (2010).

305. Szenker, E., Ray-Gallet, D. & Almouzni, G. The double face of the histone variant H3.3. *Cell Research* **21,** 421–434 (2011).

306. McDowell, T. L., Gibbons, R. J., Sutherland, H., O'Rourke, D. M., Bickmore, W. A., Pombo, A., Turley, H., Gatter, K., Picketts, D. J., Buckle, V. J., Chapman, L., Rhodes, D. & Higgs, D. R. Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* **96,** 13983–13988 (1999).

307. Law, M. J., Lower, K. M., Voon, H. P. J., Hughes, J. R., Garrick, D., Viprakasit, V., Mitson, M., De Gobbi, M., Marra, M., Morris, A., Abbott, A., Wilder, S. P., Taylor, S., Santos, G. M., Cross, J., Ayyub, H., Jones, S., Ragoussis, J., Rhodes, D., Dunham, I., Higgs, D. R. & Gibbons, R. J. ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. *Cell* **143,** 367–378 (2010).

308. Sadic, D., Schmidt, K., Groh, S., Kondofersky, I., Ellwart, J., Fuchs, C., Theis, F. J. & Schotta, G. Atrx promotes heterochromatin formation at retrotransposons. *EMBO Reports* **16,** 836–850 (2015).

309. He, Q., Kim, H., Huang, R., Lu, W., Tang, M., Shi, F., Yang, D., Zhang, X., Huang, J., Liu, D. & Songyang, Z. The Daxx/Atrx Complex Protects Tandem Repetitive Elements during DNA Hypomethylation by Promoting H3K9 Trimethylation. *Cell Stem Cell* **17,** 273–286 (2015).

310. Gibbons, R. J., McDowell, T. L., Raman, S., O'Rourke, D. M., Garrick, D., Ayyub, H. & Higgs, D. R. Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. *Nature Genetics* **24,** 368–371 (2000).

311. Clynes, D., Higgs, D. R. & Gibbons, R. J. The chromatin remodeller ATRX: a repeat offender in human disease. *Trends in Biochemical Sciences* **38,** 461–466 (2013).

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

313. Shahbazian, M. D., Antalffy, B., Armstrong, D. L. & Zoghbi, H. Y. Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. *Human Molecular Genetics* **11,** 115–124 (2002).

314. Jung, B. P., Jugloff, D. G. M., Zhang, G., Logan, R., Brown, S. & Eubanks, J. H. The expression of methyl CpG binding factor MeCP2 correlates with cellular differentiation in the developing rat brain and in cultured cells. *Journal of Neurobiology* **55,** 86–96 (2003).

315. Kernohan, K. D., Vernimmen, D., Gloor, G. B. & Bérubé, N. G. Analysis of neonatal brain lacking ATRX or MeCP2 reveals changes in nucleosome density, CTCF binding and chromatin looping. *Nucleic Acids Research* **42,** 8356–8368 (2014).

316. Marano, D., Fioriniello, S., Fiorillo, F., Gibbons, R. J., D'Esposito, M. & Della Ragione, F. ATRX Contributes to MeCP2-Mediated Pericentric Heterochromatin Organization during Neural Differentiation. *International Journal of Molecular Sciences* **20,** 5371 (2019).

317. Lennox, R. W. & Cohen, L. H. The histone H1 complements of dividing and nondividing cells of the mouse. *The Journal of Biological Chemistry* **258,** 262–268 (1983).

318. Meergans, T., Albig, W. & Doenecke, D. Varied expression patterns of human H1 histone genes in different cell lines. *DNA and Cell Biology* **16,** 1041–1049 (1997).

319. Franke, K., Drabent, B. & Doenecke, D. Expression of murine H1 histone genes during postnatal development. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1398,** 232–242 (1998).

320. Wisniewski, J. R., Zougman, A., Krüger, S. & Mann, M. Mass spectrometric mapping of linker histone H1 variants reveals multiple acetylations, methylations, and phosphorylation as well as differences between cell culture and tissue. *Molecular & Cellular Proteomics* **6,** 72–87 (2007).

321. Fan, Y., Sirotkin, A., Russell, R. & Ayala, J. Individual Somatic H1 Subtypes Are Dispensable for Mouse Development Even in Mice Lacking the H10 Replacement Subtype. *Molecular and cellular biology*, 7933–7943 (2001).

322. Parseghian, M. H., Newcomb, R. L., Winokur, S. T. & Hamkalo, B. A. The distribution of somatic H1 subtypes is non-random on active vs. inactive chromatin: distribution in human fetal fibroblasts. *Chromosome Research* **8,** 405–424 (2000).

323. Fan, Y., Nikitina, T., Morin-Kensicki, E. M., Zhao, J., Magnuson, T. R., Woodcock, C. L. & Skoultchi, A. I. H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo. *Molecular and Cellular Biology* **23,** 4559–4572 (2003).

324. Fan, Y., Nikitina, T., Zhao, J., Fleury, T. J., Bhattacharyya, R., Bouhassira, E. E., Stein, A., Woodcock, C. L. & Skoultchi, A. I. Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. *Cell* **123,** 1199–1212 (2005).

325. Pearson, E. C., Bates, D. L., Prospero, T. D. & Thomas, J. O. Neuronal nuclei and glial nuclei from mammalian cerebral cortex. Nucleosome repeat lengths, DNA contents and H1 contents. *European Journal of Biochemistry / FEBS* **144,** 353–360 (1984).

326. Jost, K. L., Rottach, A., Milden, M., Bertulat, B., Becker, A., Wolf, P., Sandoval, J., Petazzi, P., Huertas, D., Esteller, M., Kremmer, E., Leonhardt, H. & Cardoso, M. C. Generation and characterization of rat and mouse monoclonal antibodies specific for MeCP2 and their use in X-inactivation studies. *Plos One* **6,** e26499 (2011).

327. Yaffe, D. & Saxel, O. Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature* **270,** 725–727 (1977).

328. Samaco, R. C., Fryer, J. D., Ren, J., Fyffe, S., Chao, H.-T., Sun, Y., Greer, J. J., Zoghbi, H. Y. & Neul, J. L. A partial loss of function allele of methyl-CpG-binding protein 2 predicts a human neurodevelopmental syndrome. *Human Molecular Genetics* **17,** 1718–1727 (2008).

329. Rothbauer, U., Zolghadr, K., Muyldermans, S., Schepers, A., Cardoso, M. C. & Leonhardt, H. A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Molecular & Cellular Proteomics* **7,** 282–289 (2008).

330. Vaughn, J. L., Goodwin, R. H., Tompkins, G. J. & McCawley, P. The establishment of two cell lines from the insectspodoptera frugiperda (lepidoptera; noctuidae). *In vitro* **13,** 213–217 (1977).

331. Inuzuka, L. M., Guerra-Peixe, M., Macedo-Souza, L. I., Pedreira, C. C., Gurgel-Giannetti, J., Monteiro, F. P., Ramos, L., Costa, L. A., Crippa, A. C. d. S., Lourenco, C. M., Pachito, D. V., Sukys-Claudino, L., Gaspar, L. S., Antoniuk, S. A., Dutra, L. P. d. S., Diniz, S. S. L., Pires, R. B., Garzon, E. & Kok, F. MECP2-related conditions in males: A systematic literature review and 8 additional cases. *European Journal of Paediatric Neurology* **34,** 7–13 (2021).

332. Bellini, E., Pavesi, G., Barbiero, I., Bergo, A., Chandola, C., Nawaz, M. S., Rusconi, L., Stefanelli, G., Strollo, M., Valente, M. M., Kilstrup-Nielsen, C. & Landsberger, N. MeCP2 post-translational modifications: a mechanism to control its involvement in synaptic plasticity and homeostasis? *Frontiers in Cellular Neuroscience* **8,** 236 (2014).

333. Dyballa, N. & Metzger, S. Fast and sensitive colloidal coomassie G-250 staining for proteins in polyacrylamide gels. *Journal of Visualized Experiments*, e1431 (2009).

334. Cerletti, M., Paggi, R. A., Guevara, C. R., Poetsch, A. & De Castro, R. E. Global role of the membrane protease LonB in Archaea: Potential protease targets revealed by quantitative proteome analysis of a lonB mutant in Haloferax volcanii. *Journal of Proteomics* **121,** 1–14 (2015).

335. Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry* **5,** 976–989 (1994).

336. Heinz, K. S., Casas-Delucchi, C. S., Török, T., Cmarko, D., Rapp, A., Raska, I. & Cardoso, M. C. Peripheral re-localization of constitutive heterochromatin advances its replication timing and impairs maintenance of silencing marks. *Nucleic Acids Research* **46,** 6112–6128 (2018).

337. Krishnaraj, R., Ho, G. & Christodoulou, J. RettBASE: Rett syndrome database update. *Human Mutation* **38,** 922–931 (2017).

338. Yang, Y., Kucukkal, T. G., Li, J., Alexov, E. & Cao, W. Binding Analysis of Methyl-CpG Binding Domain of MeCP2 and Rett Syndrome Mutations. *ACS Chemical Biology* **11,** 2706–2715 (2016).

339. Martin, R. M., Ter-Avetisyan, G., Herce, H. D., Ludwig, A. K., Lättig-Tünnemann, G. & Cardoso, M. C. Principles of protein targeting to the nucleolus. *Nucleus* **6,** 314–325 (2015).

340. Nicholson, T. B., Chen, T. & Richard, S. The physiological and pathophysiological role of PRMT1-mediated protein arginine methylation. *Pharmacological Research* **60,** 466–474 (2009).

341. Bedford, M. T. Arginine methylation at a glance. *Journal of Cell Science* **120,** 4243–4246 (2007).

342. Kleinschmidt, M. A., Streubel, G., Samans, B., Krause, M. & Bauer, U.-M. The protein arginine methyltransferases CARM1 and PRMT1 cooperate in gene regulation. *Nucleic Acids Research* **36,** 3202–3213 (2008).

343. Herrmann, F., Pably, P., Eckerich, C., Bedford, M. T. & Fackelmayer, F. O. Human protein arginine methyltransferases in vivo–distinct properties of eight canonical members of the PRMT family. *Journal of Cell Science* **122,** 667–677 (2009).

344. Stopa, N., Krebs, J. E. & Shechter, D. The PRMT5 arginine methyltransferase: many roles in development, cancer and beyond. *Cellular and Molecular Life Sciences* **72,** 2041–2059 (2015).

345. Frankel, A., Yadav, N., Lee, J., Branscombe, T. L., Clarke, S. & Bedford, M. T. The novel human protein arginine N-methyltransferase PRMT6 is a nuclear enzyme displaying unique substrate specificity. *The Journal of Biological Chemistry* **277,** 3537–3543 (2002).

346. Bedford, M. T. & Clarke, S. G. Protein arginine methylation in mammals: who, what, and why. *Molecular Cell* **33,** 1–13 (2009).

347. Wesche, J., Kühn, S., Kessler, B. M., Salton, M. & Wolf, A. Protein arginine methylation: a prominent modification and its demethylation. *Cellular and Molecular Life Sciences* **74,** 3305–3315 (2017).

348. Le Romancer, M., Treilleux, I., Leconte, N., Robin-Lespinasse, Y., Sentis, S., Bouchekioua-Bouzaghou, K., Goddard, S., Gobert-Gosse, S. & Corbo, L. Regulation of estrogen rapid signaling through arginine methylation by PRMT1. *Molecular Cell* **31,** 212–221 (2008).

349. Sylvestersen, K. B., Horn, H., Jungmichel, S., Jensen, L. J. & Nielsen, M. L. Proteomic analysis of arginine methylation sites in human cells reveals dynamic regulation during transcriptional arrest. *Molecular & Cellular Proteomics* **13,** 2072–2088 (2014).

350. Tsai, W.-C., Gayatri, S., Reineke, L. C., Sbardella, G., Bedford, M. T. & Lloyd, R. E. Arginine demethylation of G3BP1 promotes stress granule assembly. *The Journal of Biological Chemistry* **291,** 22671–22685 (2016).

351. Chang, B., Chen, Y., Zhao, Y. & Bruick, R. K. JMJD6 is a histone arginine demethylase. *Science* **318,** 444–447 (2007).

352. Walport, L. J., Hopkinson, R. J., Chowdhury, R., Schiller, R., Ge, W., Kawamura, A. & Schofield, C. J. Arginine demethylation is catalysed by a subset of JmjC histone lysine demethylases. *Nature Communications* **7,** 11974 (2016).

353. An, W., Kim, J. & Roeder, R. G. Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53. *Cell* **117,** 735–748 (2004).

354. Stein, C., Riedl, S., Rüthnick, D., Nötzold, R. R. & Bauer, U.-M. The arginine methyltransferase PRMT6 regulates cell proliferation and senescence through transcriptional repression of tumor suppressor genes. *Nucleic Acids Research* **40,** 9522–9533 (2012).

355. Stein, C., Nötzold, R. R., Riedl, S., Bouchard, C. & Bauer, U.-M. The Arginine Methyltransferase PRMT6 Cooperates with Polycomb Proteins in Regulating HOXA Gene Expression. *Plos One* **11,** e0148892 (2016).

356. Hyllus, D., Stein, C., Schnabel, K., Schiltz, E., Imhof, A., Dou, Y., Hsieh, J. & Bauer, U.-M. PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation. *Genes & Development* **21,** 3369–3380 (2007).

357. Guccione, E., Bassi, C., Casadio, F., Martinato, F., Cesaroni, M., Schuchlautz, H., Lüscher, B. & Amati, B. Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. *Nature* **449,** 933–937 (2007).

358. Bouchard, C., Sahu, P., Meixner, M., Nötzold, R. R., Rust, M. B., Kremmer, E., Feederle, R., Hart-Smith, G., Finkernagel, F., Bartkuhn, M., Savai Pullamsetti, S., Nist, A., Stiewe, T., Philipsen, S. & Bauer, U.-M. Genomic location of PRMT6-dependent H3R2 methylation is linked to the transcriptional outcome of associated genes. *Cell reports* **24,** 3339–3352 (2018).

359. Yoshimatsu, M., Toyokawa, G., Hayami, S., Unoki, M., Tsunoda, T., Field, H. I., Kelly, J. D., Neal, D. E., Maehara, Y., Ponder, B. A. J., Nakamura, Y. & Hamamoto, R. Dysregulation of PRMT1 and PRMT6, Type I arginine methyltransferases, is involved in various types of human cancers. *International Journal of Cancer* **128,** 562–573 (2011).

360. Veland, N., Hardikar, S., Zhong, Y., Gayatri, S., Dan, J., Strahl, B. D., Rothbart, S. B., Bedford, M. T. & Chen, T. The arginine methyltransferase PRMT6 regulates DNA methylation and contributes to global DNA hypomethylation in cancer. *Cell reports* **21,** 3390–3397 (2017).

361. Tang, J., Frankel, A., Cook, R. J., Kim, S., Paik, W. K., Williams, K. R., Clarke, S. & Herschman, H. R. PRMT1 is the predominant type I protein arginine methyltransferase in mammalian cells. *The Journal of Biological Chemistry* **275,** 7723–7730 (2000).

362. Vecsler, M., Simon, A. J., Amariglio, N., Rechavi, G. & Gak, E. MeCP2 deficiency downregulates specific nuclear proteins that could be partially recovered by valproic acid in vitro. *Epigenetics* **5,** 61–67 (2010).

363. Evich, M., Stroeva, E., Zheng, Y. G. & Germann, M. W. Effect of methylation on the side-chain pKa value of arginine. *Protein Science* **25,** 479–486 (2016).

364. Lorton, B. M. & Shechter, D. Cellular consequences of arginine methylation. *Cellular and Molecular Life Sciences* **76,** 2933–2956 (2019).

365. McBride, A. E. & Silver, P. A. State of the arg: protein methylation at arginine comes of age. *Cell* **106,** 5–8 (2001).

366. Tan, C. P. & Nakielny, S. Control of the DNA methylation system component MBD2 by protein arginine methylation. *Molecular and Cellular Biology* **26,** 7224–7235 (2006).

367. Cuddapah, V. A., Pillai, R. B., Shekar, K. V., Lane, J. B., Motil, K. J., Skinner, S. A., Tarquinio, D. C., Glaze, D. G., McGwin, G., Kaufmann, W. E., Percy, A. K., Neul, J. L. & Olsen, M. L. Methyl-CpG-binding protein 2 (MECP2) mutation type is associated with disease severity in Rett syndrome. *Journal of Medical Genetics* **51,** 152–158 (2014).

368. Bienvenu, T., Carrié, A., de Roux, N., Vinet, M. C., Jonveaux, P., Couvert, P., Villard, L., Arzimanoglou, A., Beldjord, C., Fontes, M., Tardieu, M. & Chelly, J. MECP2 mutations account for most cases of typical forms of Rett syndrome. *Human Molecular Genetics* **9,** 1377–1384 (2000).

369. Fukuda, T., Yamashita, Y., Nagamitsu, S., Miyamoto, K., Jin, J.-J., Ohmori, I., Ohtsuka, Y., Kuwajima, K., Endo, S., Iwai, T., Yamagata, H., Tabara, Y., Miki, T., Matsuishi, T. & Kondo, I. Methyl-CpG binding protein 2 gene (MECP2) variations in Japanese patients with Rett syndrome: pathological mutations and polymorphisms. *Brain & Development* **27,** 211–217 (2005).

370. Zahorakova, D., Lelkova, P., Gregor, V., Magner, M., Zeman, J. & Martasek, P. MECP2 mutations in Czech patients with Rett syndrome and Rett-like phenotypes: novel mutations, genotype-phenotype correlations and validation of high-resolution melting analysis for mutation scanning. *Journal of Human Genetics* **61,** 617–625 (2016).

371. Kucukkal, T. G., Yang, Y., Uvarov, O., Cao, W. & Alexov, E. Impact of rett syndrome mutations on mecp2 MBD stability. *Biochemistry* **54,** 6357–6368 (2015).

372. Pedretti, A., Granito, C., Mazzolari, A. & Vistoli, G. Structural Effects of Some Relevant Missense Mutations on the MECP2-DNA Binding: A MD Study Analyzed by Rescore+, a Versatile Rescoring Tool of the VEGA ZZ Program. *Molecular informatics* **35,** 424–433 (2016).

373. Block, H., Maertens, B., Spriestersbach, A., Brinker, N., Kubicek, J., Fabis, R., Labahn, J. & Schäfer, F. Immobilized-metal affinity chromatography (IMAC): a review. *Methods in enzymology* **463,** 439–473 (2009).

374. Gregory, T. R. Synergy between sequence and size in large-scale genomics. *Nature Reviews. Genetics* **6,** 699–708 (2005).

375. Tchouatcha-Tchouassom, J. C., Julliard, J. H. & Roux, B. Isolation and characterisation of five histone H1 subtypes from adult rat liver. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1009,** 121–128 (1989).

376. Mateos-Langerak, J., Brink, M. C., Luijsterburg, M. S., van der Kraan, I., van Driel, R. & Verschure, P. J. Pericentromeric heterochromatin domains are maintained without accumulation of HP1. *Molecular Biology of the Cell* **18,** 1464–1471 (2007).

377. Jagannathan, M., Cummings, R. & Yamashita, Y. M. A conserved function for pericentromeric satellite DNA. *eLife* **7,** e34122 (2018).

378. Jagannathan, M., Cummings, R. & Yamashita, Y. M. The modular mechanism of chromocenter formation in Drosophila. *eLife* **8,** e43938 (2019).

379. Sanulli, S., Trnka, M. J., Dharmarajan, V., Tibble, R. W., Pascal, B. D., Burlingame, A. L., Griffin, P. R., Gross, J. D. & Narlikar, G. J. Disorganization of the histone core promotes organization of heterochromatin into phase-separated droplets. *BioRxiv* (2018).

380. Shakya, A., Park, S., Rana, N. & King, J. T. Liquid-Liquid Phase Separation of Histone Proteins in Cells: Role in Chromatin Organization. *Biophysical Journal* **118,** 753–764 (2020).

381. Li, P., Banjade, S., Cheng, H.-C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J. V., King, D. S., Banani, S. F., Russo, P. S., Jiang, Q.-X., Nixon, B. T. & Rosen, M. K. Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483,** 336–340 (2012).

382. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nature Reviews. Molecular Cell Biology* **18,** 285–298 (2017).

383. Wegmann, S., Eftekharzadeh, B., Tepper, K., Zoltowska, K. M., Bennett, R. E., Dujardin, S., Laskowski, P. R., MacKenzie, D., Kamath, T., Commins, C., Vanderburg, C., Roe, A. D., Fan, Z., Molliex, A. M., Hernandez-Vega, A., Muller, D., Hyman, A. A., Mandelkow, E., Taylor, J. P. & Hyman, B. T. Tau protein liquid-liquid phase separation can initiate tau aggregation. *The EMBO Journal* **37,** e98049 (2018).

384. Zbinden, A., Pérez-Berlanga, M., De Rossi, P. & Polymenidou, M. Phase separation and neurodegenerative diseases: A disturbance in the force. *Developmental Cell* **55,** 45–68 (2020).

# 8 Annex

## 8.1 Abbreviations

| | |
|---|---|
| **ADMA** | asymmetric dimethyl arginine |
| **ChIP** | chromatin immunoprecipitation |
| **CID** | collision-induced dissociation |
| **CTD** | C-terminal domain |
| **DAPI** | 4',6-diamidino-2-phenylindole |
| **FACS** | fluorescence activated cell sorting |
| **FRAP** | fluorescence recovery after photobleaching |
| **GAR** | glycine and arginine rich |
| **GFP** | green fluorescent protein |
| **HCD** | high-energy collisional dissociation |
| **HP1** | heterochromatin protein 1 |
| **ID** | intervening domain |
| **MBD** | methyl binding domain |
| **MeCP2** | Methyl-CpG binding protein 2 |
| **MMA** | monomethyl arginine |
| **MS** | mass spectrometry |
| **NID** | N-CoR interacting domain |
| **NTD** | N-terminal domain |
| **PCR** | polymerase chain reaction |
| **phos** | phosphorylation |
| **PRMT** | protein arginine methyltransferase |
| **PTM** | post-translational modification |
| **ROI** | region of interest |
| **SDMA** | symmetric dimethyl arginine |
| **SDS-PAGE** | sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| **TRD** | transcriptional repression domain |
| **(U)HPLC** | ultra high performance liquid chromatography |
| **WB** | Western blot |

## 8.2 List of contributions

**The proteomic composition of constitutive heterochromatin in mouse tissues**

Annika Schmidt performed all experiments, data analysis and writing including figure and table preparation except for:

- Figure 3.3, 3.4, 3.5, 3.11, 3.12, 3.13: The mass spectrometry data used for analysis was acquired by Stephanie Meyer (chromocenter isolation) and Oliver Popp, Gunnar Dittmar (mass spectrometry sample preparation and measurements); Jiaxuan Chen, Falk Butter performed the MaxQuant search; Jiaxuan Chen and Oliver Popp provided help during mass spectrometry data analysis
- Figure 3.7, 3.10: Stephanie Meyer performed the TOPII and histone H2AX immunofluorescence staining; Katalina Gagova performed the TOPI immunofluorescence staining

**MeCP2 heterochromatin organization is modulated by arginine methylation and serine phosphorylation**

Annika Schmidt performed all experiments, data analysis and writing including figure and table preparation except for:

- Jana Frei contributed isolated mouse brains for Western blot analysis and generated the MeCP2 mutant plasmids
- Ansgar Poetsch contributed to mass spectrometry data acquisition and data analysis
- Uta-Maria Bauer and Chris Aßmann provided the PRMT plasmids
- Ulrike A. Nuber, Ansgar Poetsch, Alexandra Chittka and Uta-Maria Bauer commented on the text

## 8.3 Acknowledgments

First, I would like to thank Prof. Dr. M. Cristina Cardoso for giving me the opportunity to do my PhD in her lab. I am grateful for everything I learned and for many interesting scientific discussions. Thank you for always believing in my projects, whenever I could not.
I would also like to thank Prof. Dr. Ulrike A. Nuber for being my second supervisor and for helpful discussions during the meetings.
Many thanks go to all the cooperation partners who worked with me over time. Especially the people from the plant biochemistry group of Dr. Ansgar Poetsch in Bochum, who provided me with all the help I needed to prepare and measure my samples there. Thank you for the discussions and for your support with the mass spectrometry data analysis.
I would also like to thank Dr. Falk Butter and Dr. Jiaxuan Chen from the proteomics facility at IMB Mainz for measuring our samples and for their support during data analysis. Thanks also go to Prof. Dr. Frederik Lermyte for measuring our samples and good discussions, although the project did not make it into this thesis. Thank you to Prof. Dr. Uta-Maria Bauer and Chris Aßmann for providing the PRMT plasmids. Thanks to Stephanie Meyer for her work on the chromocenter project. Thanks to Dr. Oliver Popp and Prof. Dr. Gunnar Dittmar for the mass spectrometry analysis of the chromocenter fractions.
Thank you to Jana, for cloning, culturing cells, and never giving up on differentiating HT22 cells... I really enjoyed the time we worked together on the project.
Most importantly, thank you to all the past and present members of the Cardoso lab for a great time! I enjoyed all the discussions, working in the lab, and all the time we spent together. I would not want to miss a single evening out with you wonderful crazy people ;-)
Thank you, Anne, for always being there and making everyone (also and especially the cells) feel comfortable in the lab. Thank you, Manu, for listening to all my experimental problems and for the amazing food. Thanks to Diana for numerous Maxipreps. Thank you, Anne, Diana, Manu, and everyone else helping with the mice.
Many thanks to the (old) methylation group, Cathia, Florian, and Peng, for discussions and experimental advice, especially at the beginning when I joined the lab. Many thanks to Alex for spending his precious time on solving all possible problems with microscopes, computers, and analysis. Thanks to Hector for changing the ImageJ macro hundreds of times. Thanks to Hui for performing the revision experiments in record time. Thank you to Katalina for joining me in high-throughput Western blotting for antibody validation. Thank you, Stefania, for cheering me up on amazing afternoon sun walks in the park during writing time.
Thanks to Cathia, Christian, and Maruthi for proofreading the thesis.
Last but not least I thank my family, Christian, and my friends for supporting me on my way, without you this would not have been possible. Danke für eure Unterstützung, euer Verständnis und einfach alles, Ihr seid die Besten!

## 8.4 Declaration - Ehrenwörtliche Erklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit entsprechend den Regeln guter wissenschaftlicher Praxis selbstständig und ohne unzulässige Hilfe Dritter angefertigt habe.

Sämtliche aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie sämtliche von Anderen direkt oder indirekt übernommenen Daten, Techniken und Materialien sind als solche kenntlich gemacht. Die Arbeit wurde bisher bei keiner anderen Hochschule zu Prüfungszwecken eingereicht. Die eingereichte elektronische Version stimmt mit der schriftlichen Version überein.

Darmstadt, den 22.09.2022

_____

Annika Schmidt

## 8.5 Curriculum Vitae

**Personal details**

| | |
|---|---|
| Name | Annika Schmidt |

**Academic education**

| | |
|---|---|
| 2017 - 2022 | **PhD thesis at Technische Universität Darmstadt**<br>Department of Biology<br>Cell biology and epigenetics, Prof. Dr. M. Cristina Cardoso<br>"Heterochromatin composition, organization and MeCP2 post-translational modifications" |
| 2014 - 2016 | **Master of Science in Biochemistry at Ruhr-Universität Bochum**<br>Focal point: Protein structure and function<br>Thesis: Medical Proteome Center, Functional proteomics, Prof. Dr. Katrin Marcus<br>"Establishment of a mass spectrometry-based method for the identification of body fluids in forensic samples" |
| 2011 - 2014 | **Bachelor of Science in Biochemistry at Ruhr-Universität Bochum**<br>Thesis: Molecular Neurobiochemistry, Prof. Dr. Rolf Heumann<br>"Cloning of cDNAs to improve protein transduction - intracellular non-cleavable version" |

**Publications**

| | |
|---|---|
| 2022 | **Annika Schmidt**, Jana Frei*, Ansgar Poetsch*, Alexandra Chittka, Hui Zhang, Chris Aßmann, Anne Lehmkuhl, Uta-Maria Bauer, Ulrike A. Nuber, M. Cristina Cardoso<br>**MeCP2 heterochromatin organization is modulated by arginine methylation and serine phosphorylation**<br>Frontiers in Cell and Developmental Biology 2022, 10: 941493; doi: 10.3389/fcell.2022.941493 |
| 2022 | Hui Zhang*, Hector Romero*, **Annika Schmidt**, Katalina Gagova, Weihua Qin, Bianca Bertulat, Anne Lehmkuhl, Manuela Milden, Malte Eck, Tobias Meckel, Heinrich Leonhardt & M. Cristina Cardoso<br>**MeCP2-induced heterochromatin organization is driven by oligomerization-based liquid–liquid phase separation and restricted by DNA methylation**<br>Nucleus 2022, 13 (1): 1-34; doi: 10.1080/19491034.2021.2024691 |
| 2020 | **Annika Schmidt***, Hui Zhang* & M. Cristina Cardoso<br>**MeCP2 and chromatin compartmentalization**<br>Cells 2020, 9 (4): 878; doi: 10.3390/cells9040878 |

*equally contributing authors