Regulation of the TET1 protein in active DNA demethylation

Dissertation

Zur Erlangung des Grades

Doktor der Naturwissenschaften

Am Fachbereich Biologie

Der Johannes Gutenberg-Universität Mainz

Anna Luise Ernst

geb. am 28.03.1991 in Dortmund

Mainz, 2019

Dekan:

1. Berichterstatter:

2. Berichterstatterin:

Tag der mündlichen Prüfung: 06.12.2019

Table of contents

1 Summary

Active DNA demethylation plays an important role in various biological contexts, such as embryonic development and embryonic stem cell (ESC) differentiation. The Ten-Eleven Translocation (TET) family enzymes play a key role in this process, as they iteratively oxidize 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). 5fC and 5caC can subsequently be removed by Thymine DNA glycosylase (TDG) and unmethylated cytosine is restored by base excision repair (BER), completing the process of active DNA demethylation. Before the discovery of the role of TET enzymes in active DNA demethylation, the Growth arrest and DNA-damage-inducible protein 45 alpha (GADD45A) had already been implicated in active DNA demethylation, but its exact role remained unknown. Recent findings highlighted that GADD45A physically and functionally cooperates with the TET family member TET1 to mediate active DNA demethylation. However, the mechanism by which GADD45A enhances TET1 activity is yet poorly understood. In addition, how TET enzymes are targeted to genomic loci, especially to enhancers, remains largely unresolved. In fact, whereas TET binding sites are mostly found at promoters, the oxidation products of TET – 5hmC, 5fC and 5caC – are mostly enriched at enhancers. In this thesis, I explored two TET-related topics 1) TET1 regulation by GADD45A and 2) TET regulation by RAD21-dependent chromatin looping.

First, I found that the GADD45A interactor Mitogen-activated protein kinase kinase kinase 4 (MAP3K4) is most likely not involved in enhancing TET1 activity in a GADD45A-dependent manner. Then, I demonstrated that GADD45A can affect TET1 post-translational modifications: GADD45A decreases the ubiquitination and SUMOylation of full-length TET1, while it enhances the SUMOylation of the catalytic domain of TET1 (TET1CD). Finally, I established that GADD45A stabilizes TET1 protein levels, putatively by preventing Calpain-mediated decay. Nevertheless, the detailed mechanism behind this stabilization remains to be resolved in the future.

A proposed model by which TET enzymes are targeted to enhancers to oxidize 5mC is promoter-enhancer looping. Promoter-enhancer loops are established by the Cohesin complex and can be abolished by depleting one of its subunits, such as RAD21. To test this model, I investigated the distribution of 5fC in RAD21-depleted – and hence promoter-enhancer loop depleted – mouse ESCs (mESCs) and observed that the deposition of 5fC at enhancers is RAD21-loop independent. Although these results do not provide an ultimate explanation for the discrepancy between TET binding sites at promoters and the enrichment of 5hmC, 5fC and 5caC at enhancers, this analysis supports the conclusion that targeting of enhancer demethylation does not rely on RAD21-dependent chromatin loops.

2 Zusammenfassung

Aktive DNA-Demethylierung spielt eine wichtige Rolle in verschiedenen biologischen Zusammenhängen, wie beispielsweise der Embryonalentwicklung oder der embryonalen Stammzelldifferenzierung. Die Enzyme der Ten-Eleven Translocation (TET) Familie nehmen hierbei eine Schlüsselposition ein, indem sie 5-Methylcytosin (5mC) iterativ zu 5-Hydroxymethylcytosin (5hmC), 5-Formylcytosin (5fC), sowie 5-Carboxylcytosin (5caC) oxidieren. 5fC und 5caC können daraufhin von der Thymine DNA glycosylase (TDG) entfernt werden, und unmethyliertes Cytosin durch Enzyme der Basenexzisionsreparatur (BER) wiederhergestellt werden, welches den Prozess der aktiven DNA-Demethylierung vervollständigt. Bereits vor der Entdeckung der Rolle der TET Enzyme in der aktiven DNA-Demethylierung wurde das Protein Growth arrest and DNA-damage-inducible protein 45 alpha (GADD45A) mit der aktiven DNA-Demethylierung in Verbindung gebracht, jedoch blieb dessen Wirkmechanismus unklar. Neueste Erkenntnisse zeigen, dass GADD45A physisch und funktionell mit TET1 in der aktiven DNA-Demethylierung kooperiert, allerdings ist über die mechanistischen Details dieser Kooperation weiterhin wenig bekannt. Zudem bleibt ungelöst, wie TET Enzyme gezielt zu genomischen Loci rekrutiert werden. Tatsächlich befinden sich TET-Bindungsstellen überwiegend an Promotoren, während die TET-Oxidationsprodukte – 5hmC, 5fC, sowie 5caC – größtenteils an Enhancern angereichert sind. In der vorliegenden Arbeit untersuchte ich zwei TET-relevante Themen: 1) TET1 Regulation durch GADD45A und 2) TET Regulation durch RAD21-abhängiges Chromatin Looping.

Erstens stellte sich heraus, dass die mit GADD45A interagierende Mitogen-activated protein kinase kinase kinase 4 (MAP3K4) höchstwahrscheinlich nicht an dem TET1-aktivierenden Prozess durch GADD45A involviert ist. Außerdem demonstrierte ich, dass GADD45A posttranslationale Modifikationen von TET1 beeinflusst: GADD45A verringert die Ubiquitinierung sowie die SUMOylierung des Gesamtlängen-TET1-Proteins, während es die SUMOylierung der katalytischen TET1-Domäne verstärkt. Ferner stabilisiert GADD45A die TET1-Proteinspiegel, vermutlich durch die Verhinderung eines Calpain-vermittelten Proteinabbaus. Dennoch bleibt der vollständige Mechanismus dieser Stabilisation noch zu klären.

Ein vorgeschlagenes Modell, welches darlegt wie TET-Enzyme gezielt zu Enhancern zum Zwecke der 5mC Oxidierung rekrutiert werden, ist Promoter-Enhancer-Looping. Promoter-Enhancer Loops werden mittels des Cohesin Komplexes generiert und können durch Depletion von Untereinheiten, beispielsweise RAD21, aufgelöst werden. Um dieses Modell zu testen, erforschte ich die Verteilung von 5fC in RAD21 depletierten, und demzufolge Promoter-Enhancer-Loop depletierten Mausstammzellen und zeige, dass das Vorkommen von 5fC an Enhancern RAD21-Loop unabhängig ist. Wenngleich diese Ergebnisse keine endgültige Erläuterung der Diskrepanz zwischen TET-Bindungsstellen an Promotoren und der Anreicherung von 5hmC, 5fC, sowie 5caC an Enhancern liefern, so legen meine Ergebnisse nahe, dass DNA-Demethylierung, zumindest im Kontext RAD21-abhängiger Loops, unabhängig der dreidimensionalen Chromatinstruktur erfolgen kann.

3 Introduction

3.1 DNA methylation and demethylation

In 1898, W.G. Ruppel discovered a novel nucleic acid while attempting to isolate the bacterial toxin responsible for causing tuberculosis, which he named tuberculinic acid [1]. It took another 50 years until this modification of the DNA base cytosine (C), now known as 5-methylcytosine (5mC), was also discovered in mammalian DNA [2]. In eukaryotes, 5mC occurs in the context of cytosine-guanine dinucleotides (CpGs) and is present symmetrically on both strands of the DNA [3]. Today, we also know the enzymes responsible for setting these marks: DNA methyl-transferases (DNMTs), which can generate 5mC by the transfer of a methyl group from the cofactor S-adenosyl-L-methionine (SAM) to the 5' position of cytosine [4]. In mammals, DNMT3A and DNMT3B act as the *de novo* methyltransferases [5], whereas DNMT1 is the maintenance enzyme, which ensures that DNA methylation is maintained during DNA replication [6].

3.1.1 Molecular mechanisms of DNA methylation

The *de novo* methyltransferases DNMT3A and DNMT3B contain a highly conserved methyltransferase (MTase) domain, as well as two chromatin reading domains: ATRX-DNMT3-DNMT3L (ADD) and PWWP [3]. The ADD domain can bind to unmethylated histone H3 lysine4 (H3K4), which releases the auto-inhibition of the DNMT3 enzymes and allows for the *de novo* methylation of C. When H3K4 is trimethylated (H3K4me3), marking active and poised promoters, this prevents the binding of the ADD domain and leads to the auto-inhibition of DNMT3A/B, as the ADD domain then binds and inhibits the MTase domain [7,8]. The PWWP domain can bind to H3K36me3, a mark that is found in the gene bodies of actively transcribed genes, resulting in the methylation of these regions as well [9].

The maintenance methyltransferase DNMT1 cooperates with UHRF1, which recognizes hemi-methylated CpG dyads via its SET- and RING-associated (SRA) domain in order to ensure the inheritance of 5mC after DNA replication [10]. UHRF1 also harbors a tandem TUDOR-PHD domain, which binds to H3K9me2 and H3K9me3, two epigenetic marks associated with transcriptional repression [11]. UHRF1 recruits DNMT1 via its ubiquitin-like (UBL) domain, which can bind to the Replication foci targeting sequence (RFTS) of DNMT1 [12,13]. Without binding of UHRF1, the RFTS of DNMT1 folds into its MTase domain and inhibits the catalytic activity. The binding of UHRF1 promotes the binding of the RFTS to ubiquitinated H3 tails, which releases the auto-inhibition of DNMT1 to allow for the maintenance of symmetrical methylation of CpGs [14].

3.1.2 Molecular mechanisms of DNA demethylation

Originally, 5mC was associated with transcriptional repression, and now also has established roles in imprinting, X chromosome inactivation, embryonic development, tissue-specific gene expression, and cancer [15]. Due to the stable nature of the carbon-carbon bond, it was long thought that DNA methylation is an irreversible process that can only be removed in a passive manner, i.e. by dilution and *de novo* DNA synthesis in the absence of DNA methylation [16]. However, at the beginning of the new millennium, the first evidence emerged to support the

existence of an active removal pathway for 5mC: Mayer and colleagues showed that the mouse paternal genome is demethylated within 6-8 hours of fertilization, before the onset of DNA replication, thus proposing the existence of a DNA demethylation machinery [17].

DNA demethylation can occur in a passive manner, by inhibiting the maintenance of 5mC by DNMT1 during DNA replication. In the absence of functional DNMT1 or its cofactor UHRF1, DNA methylation will be successively lost over multiple rounds of DNA replication. This is how, for example, 5mC is erased in the maternal genome during mouse preimplantation development [10,18,19].

Various different mechanisms of active DNA demethylation have been proposed in mammals, which include: 1) DNA demethylation by nucleotide excision repair (NER) [20], 2) base-excision repair (BER) upon 5mC deamination by Activation-induced deaminase (AID) [21], and 3) 5mC oxidation mediated by the Ten-Eleven Translocation (TET) family enzymes followed by BER or passive dilution [22–26].

DNA demethylation by NER was initially proposed to occur via Growth arrest and DNA-damage-inducible protein 45 alpha (GADD45A). In this model, GADD45A is first targeted to sites of DNA demethylation and then recruits the DNA repair machinery through the endonuclease XPG. In this process, methylated cytosines are excised and repaired with unmethylated cytosines [20]. However, the details of this mechanism are a matter of debate [27] and rather, alternative functions for GADD45 proteins in DNA demethylation have been proposed, which will be discussed in chapter 3.2.

AID was shown to deaminate 5mC to thymidine *in vitro*, suggesting a possible role of DNA deaminases in active DNA demethylation [28]. Still, the direct *in vivo* evidence for a role of deaminases is missing and a subject of debate, as some studies have suggested a role of AID-mediated 5mC deamination in DNA demethylation [29–31], while others did not find any evidence for the involvement of AID in DNA demethylation and reprogramming [31–33]. One model proposes that AID could play an indirect role in DNA demethylation, by inducing lesions at nearby unmethylated C, leading to the replacement of 5mC in the vicinity [16,34].

The breakthrough in the field of active DNA demethylation came in 2009, when mammalian Ten-eleven translocation 1 (TET1), a protein originally identified as a fusion partner of the mixed lineage leukaemia (MLL) gene in acute myeloid leukaemia (AML) [35], was discovered to convert 5mC to 5-hydroxymethylcytosine (5hmC) [22,24]. Later, also TET2 and TET3 were shown to convert 5mC to 5hmC [36] and ultimately, all TET proteins were shown to further convert 5hmC into 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (Figure 3.1) [23,37]. Following TET-mediated oxidation of 5mC, unmethylated C can be restored via passive or active mechanisms: passively by DNA replication-dependent dilution [38–40] or actively via removal of 5fC and 5caC by Thymine DNA glycosylase (TDG) [23,25]. Analogous to 5mC, its oxidative derivatives – 5hmC, 5fC and 5caC – can also be removed in a passive manner via replication-dependent dilution, as

no maintenance machinery is known to exist for these bases. This process is known as active modification (by TET enzymes) followed by passive dilution [26].

Arguably, the most accepted mechanism of active DNA demethylation is the active modification by TET enzymes and the subsequent removal of 5fC and 5caC by TDG [41]. The removal of 5fC and 5caC by TDG generates an abasic (AP) site, which is restored to unmethylated C by BER (Figure 3.1). This process has been biochemically reconstituted by Weber and colleagues, who demonstrated that base excision by TDG is coupled to BER [42]. After TDG recognizes 5fC or 5caC, it cleaves the N-glycosylic bond between the DNA base and the sugar phosphate bone, generating an AP site. Subsequently, the DNA-(apurinic or apyrimidinic site) lyase (APEX1) cleaves the phosphodiester bond 5' to the AP site, creating a single-strand break. The single-strand break, containing a hydroxyl residue at the 3'-end and a deoxyribose phosphate at the 5' end, is then recognized by DNA polymerase β (Pol β), which fills the gap with an unmodified C. At the same time, Pol β removes the 5'-deoxyribose phosphate and recruits XRCC1 and DNA ligase 3 (LIG3) to seal the DNA ends [43].

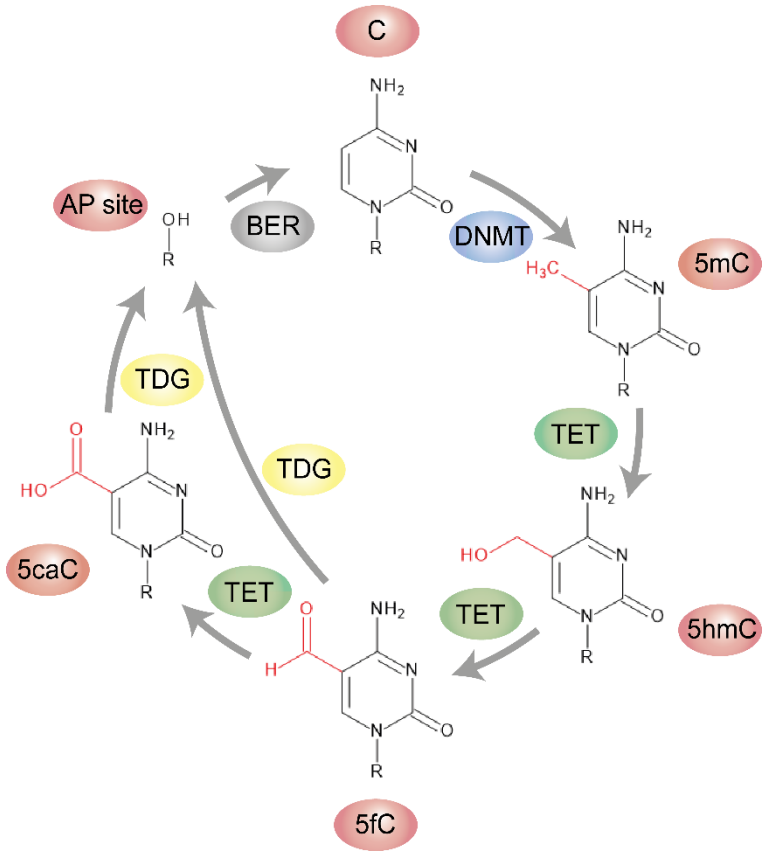

Figure 3.1 TET-TDG mediated active DNA demethylation
Cytosine (C) is methylated by DNMT1 (maintenance methyltransferase) or DNMT3A/B (*de novo* methyltransferases) at the 5' position. TET enzymes (TET1/2/3) oxidize the methyl group in an iterative manner to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). 5fC and 5caC are then excised by TDG, which generates an abasic (AP) site. Unmethylated C is finally restored by the base excision repair (BER) machinery. Modifications of C are highlighted in red. Figure based on [44].

3.1.3 Biological functions of DNA demethylation

DNA demethylation plays a role in numerous biological contexts, among them pre-implantation and primordial germ cell (PGC) development, embryonic stem cell (ESC) differentiation, neuronal functions, genomic stability and DNA damage repair, as well as cancer [41].

In mouse development, TET3 acts on the paternal genome of the fertilized zygote and oxidizes 5mC to 5hmC, to initiate active DNA demethylation [45]. This is followed by passive DNA demethylation of the paternal and maternal genome due to exclusion of DNMT1 from the nucleus during the subsequent cell divisions [19]. However, there are studies that argue with this model. First of all, it was proposed that also the maternal genome can undergo active DNA demethylation [46]. Second, it was demonstrated that the majority of global DNA demethylation occurs independently of replication [46,47]. Instead, the paternal genome could be demethylated independently of TET3, via an additional, yet unknown mechanism. In fact, it was shown that DNMT1 and DNMT3A are indeed active in the zygote and provide the substrate for TET3-mediated hydroxylation [47].

After implantation of the blastocyst, *de novo* methylation of the epiblast genome is rapidly restored by DNMT3A and DNMT3B [3]. Those stem cells, which eventually will form the later germline, undergo two additional rounds of DNA demethylation: one passive round and one active round, which is mediated by TET1 and TET2 [38,48,49]. Later, both male and female gametes are *de novo* methylated by DNMT3 enzymes to give rise to the gamete specific DNA methylation patterns. In general, TET proteins are not thought to be the drivers of active demethylation in the embryos and the germline, but rather the guardians of erroneous DNA methylation [3].

Mouse ESCs (mESCs), which are derived from the inner cell mass of the embryonic day 3.5 blastocyst, are a popular model system to study DNA methylation and demethylation dynamics, as they express both *Tet1* and *Tet2* [41]. Accordingly, they have detectable levels of 5hmC (~0.13 % of total C), 5fC (~0.002 % of total C) and 5caC (~0.0003 % of total C) [37]. *Tet1/2/3* triple-knockout mESCs have no detectable 5hmC levels. As expected, they display hypermethylation, especially at distal enhancers [50], and differentiation and developmental defects [51]. The finding that *Tet* triple-knockout mouse embryos display gastrulation defects further underlines the importance of the TET enzymes and DNA demethylation for proper early differentiation and embryonic development [52]. Importantly, the gastrulation defects are indeed due to impaired DNA demethylation, as in *Tet* triple-knockout embryos the *Lefty1* promoter and *Lefty2* enhancer are hypermethylated, which leads to impaired Nodal signaling, a core signaling pathway required for gastrulation. In this process, the catalytic activity of TET3 is essential for normal Lefty-Nodal signaling, as a catalytically inactive mutant of TET3 could not rescue the gastrulation defects.

In addition to stem cells, 5hmC is also highly abundant in neuronal tissues, such as the brain [24]. The neuronal defects observed in TET-deficient mice suggest that DNA demethylation plays an important role in neuronal functions. For example, TET1 deficiency leads to impaired hippocampal neurogenesis, as well as poor learning and memory formation [53]. Moreover, neuronal activity-

regulated genes including *Npas4* are downregulated, most likely due to promoter hypermethylation [54]. TET3 in primary cortical neurons of mice, on the other hand, was shown to be necessary for fear extinction learning by establishing 5hmC accumulation around extinction-related genes [55]. Recently, 5hmC generated by TET3 was also suggested to protect mice against ischemic injuries in the brain [56].

TET enzymes and DNA demethylation contribute to DNA repair and genomic stability in several ways. For example, at DNA damage sites, 5hmC colocalizes with the damage response proteins phosphorylated histone H2AX (γH2AX) and p53-binding protein 1 (53BP1). Moreover, TET enzymes are required for efficient repair after Aphidicolin-induced DNA damage, and a loss of TET enzymes leads to inaccurate chromosome segregation [57]. A loss of all TET enzymes in mESCs also leads to telomere abnormalities, increased sister chromatid exchange and chromosomal fusion in two additional studies [50,58]. *In vivo*, *Tet1* heterozygous knockout mice are more sensitive to X-ray exposure than wild-type mice, as they display a severe deterioration of the coat [59].

Lastly, defective DNA demethylation has been connected to cancer in various respects and especially the disruption of TET2's enzymatic activity can lead to myeloid tumorigenesis [60]. Numerous solid tumors display a reduction of 5hmC levels. The reduction of 5hmC originates either directly from TET mutations, or from mutations of TET regulators, such as Isocitrate dehydrogenase (IDH) 1/2 or Wilms tumor protein 1 (WT1) [41]. IDH1 and 2 catalyze the formation of α-ketoglutarate (α-KG), a cofactor of TET enzymes. Common mutations of IDH1 and 2 in AML and brain tumors can lead to a neomorphic enzyme activity that catalyzes the reduction of α-KG to 2-hydroxyglutarate, which ultimately inhibits TET enzymatic activity [61]. Wild-type WT1 can recruit TET2 to target genes, but AML-associated mutations or loss of WT1 lead to impaired hematopoietic differentiation and decreased 5hmC levels, probably due to failure of proper TET2 recruitment [62,63]. Paradoxically, also the widespread hypomethylation observed in cancer genomes was recently connected to TET loss of function. TET1 deficiency results in the relocalization of DNMT3A1 from heterochromatic to euchromatic regions that were previously occupied by TET1, thereby leading to hypomethylation in heterochromatic regions [64].

3.1.4 Molecular functions of TET proteins

TET proteins are iron(II)/α-KG-dependent dioxygenases, which iteratively oxidize 5mC to 5hmC, 5fC and 5caC (Figure 3.1). All three mammalian TET protein family members (TET1, 2 and 3) share a common C-terminal catalytic dioxygenase domain, consisting of a double-stranded β-helix domain (DSBH), which is split into two parts by a low complexity insert region and preceded by a cysteine-rich region (Figure 3.2) [41].

The analysis of the crystal structure of human TET2 bound to DNA gave important insights into the catalytic function of the C-terminal domain: the cysteine-rich and DSBH domains form a compact catalytic domain, brought together by two zinc fingers. Whereas the cysteine-rich domain stabilizes the DNA above the DSBH domain, the DSBH domain binds to iron(II) and α-KG and

oxidizes 5mC, 5hmC and 5fC, which can be inserted in the catalytic cavity by a base-flipping mechanism [65,66].

The N-terminal region and the low complexity region are mostly uncharacterized, except for a CXXC-type zinc finger at the N-terminus of TET1 and TET3 [41]. The CXXC domain can mediate binding to CpG-rich regions, as demonstrated by the crystal structure of the CXXC domain of *Xenopus* Tet3: It contains eight conserved cysteine residues, which coordinate two zinc ions. The zinc ions hold the mainly unstructured CXXC domain together, creating a crescent-shaped architecture to bind DNA [67]. However, the CXXC domains of TET1 and TET3 prefer different substrates: the TET1 CXXC domain can bind to unmodified C as well as 5mC and 5hmC in a CpG-rich context [68], whereas the TET3 CXXC domain requires an unmodified cytosine to bind to DNA, but not necessarily in a strict CpG context [67]. TET2 has evolutionarily lost its putative CXXC domain as a result of a chromosomal gene inversion, which created the gene *IDAX*. IDAX can bind to unmethylated CpGs and also interacts with the catalytic domain of TET2. Surprisingly, it does not recruit TET2 to CpGs in order to promote TET2-mediated oxidation of 5mC, 5hmC and 5fC, but instead it mediates caspase-dependent degradation of TET2. In fact, a similar mode of regulation was proposed for the CXXC domain of TET3 [69].

TET1 and TET3 have additional isoforms, termed TET1 short (TET1s) and TET3s, which lack the CXXC domain (Figure 3.2) [70,71]. The different TET isoforms likely play different roles in mouse development: while the full-length isoform of TET1 (TET1FL) is expressed in early embryos, ESCs and PGCs, TET1s is restricted to somatic cells. Interestingly, TET1s can still bind CpG islands despite lacking the CXXC domain, but its global chromatin affinity and DNA demethylation activity are reduced. In addition to the CXXC domain, also the first 131 amino acids (a.a.) of TET1FL can mediate global chromatin binding [70].

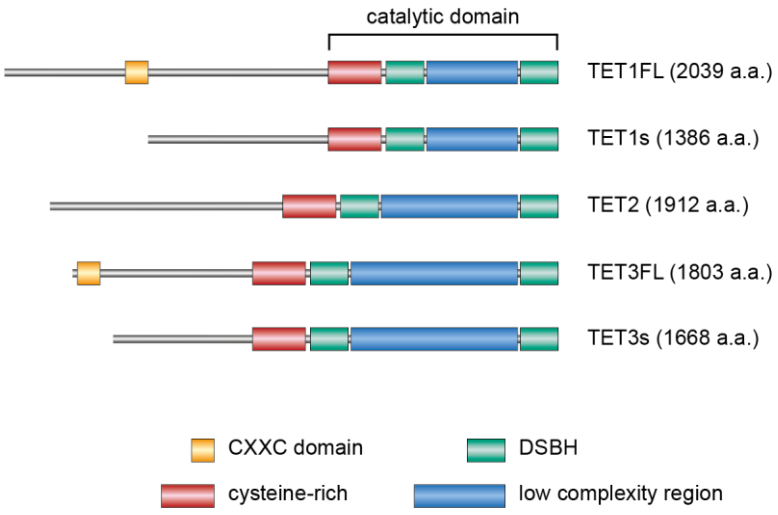

Figure 3.2 Structure of mouse TET proteins
TET family members share a common catalytic domain, consisting of a cysteine-rich region and a double-stranded β-helix domain (DSBH) domain, which is split in two parts by a low complexity insert region. Only full-length TET1 (TET1FL) and full-length TET3 (TET3FL) also have a CXXC domain, which can mediate CpG binding. a.a., amino acids, s, short. Figure based on [41].

Unexpectedly, TET enzymes are not exclusively binding to sites enriched for the oxidized 5mC products of TET, 5hmC, 5fC and 5caC. On the contrary, TET binding sites are more enriched at active promoters (marked by H3K4me3), bivalent promoters (marked by H3K4me3 and H3K27me3) and CpG islands, but not at distal regulatory elements [70,72–75]. In this respect, a study by Zhang and colleagues found no correlation of TET1 binding sites and TET1-mediated DNA demethylation sites, neither at distal DNase I hypersensitivity regions, which are indicative of regulatory elements, nor at promoters [70].

In addition to their catalytic functions, TET proteins also have non-catalytic functions. For example, TET1 contributes to transcriptional repression by interacting with the SIN3A-co-repressor complex [74]. TET1 can also form a complex with the histone acetyltransferase MOF, preventing the auto-acetylation of MOF. This increases the chromatin affinity of MOF, leading to higher H4K16 acetylation and ultimately allowing DNA repair and the maintenance of genome stability [59]. TET2 can interact with and recruit the Histone deacetylase 2 (HDAC2) to mediate repression of interleukin-6 transcription during inflammation [76]. Moreover, all TET proteins can regulate transcription via the recruitment of O-linked N-acetylglucosamine transferase (OGT) [73,77,78]. The downregulation of TET2 reduces the the addition of N-acetylglucosamine (O-GlcNAcetylation) of H2B serine 112, which in turn diminishes the transcription of target genes of OGT, H2B Ser 112 GlcNAc and TET2 [73]. In addition, both TET2 and TET3, in co-operation with OGT, promote binding of the H3K4 methyltransferase SETD1A to chromatin, to activate gene expression [78].

In vivo, catalytically inactive mutants of TET proteins can exert some biological functions. For example, a catalytically inactive mutant of TET1 remains able to regulate memory-related genes in mice [79] and a Tet3 mutant in *Xenopus laevis* can partially rescue developmental defects caused by a Tet3 depletion [67]. A recent study has shed more light on TET2's non-catalytic functions in hematopoiesis and revealed that the catalytic activity mainly regulates the myeloid lineage, while the non-catalytic functions are important for the lymphoid lineage [80].

3.1.5 Genomic distribution of the oxidative derivatives of 5mC

The oxidative derivatives of 5mC are present at highly variable levels, depending on the tissue and cell type. 5hmC is present at high levels in the central nervous system and embryonic tissues, but has low levels in most other tissues [41]. 5fC and 5caC levels are much less abundant than 5hmC, one reason being their active removal by TDG. In TDG-deficient mESCs, 5fC levels are about 6-fold increased, and 5caC levels are about 8-fold increased [81].

Several studies have shed light on the genomic distribution of 5hmC, 5fC and 5caC in mESCs. 5hmC is mostly located in gene-rich areas and increases in density from transcription start sites to transcription termination sites. 5hmC levels at promoters are generally low, whereas 5hmC is enriched at distal regulatory elements, especially at poised enhancers containing H3K4me1, and DNase I hypersensitivity sites. Interestingly, 5hmC levels are high near, but not directly at transcription factor binding sites [68,82–85]. 5fC and 5caC are found at major satellite repeats and poised enhancers in wild-type mESCs. Upon TDG depletion, 5fC and 5caC accumulate at

distal regulatory elements, TET binding sites, DNase I hypersensitivity regions, CCCTC-binding factor (CTCF) and p300 binding sites [81,86,87].

3.1.6 Regulation of TET proteins

3.1.6.1 Regulation of TET proteins by substrates and cofactors

The cofactors of TET proteins closely link TET enzymes to metabolism. α-KG is essential for TET enzymatic activity and is produced from isocitrate by the enzyme IDH as one step of the citric acid cycle. In the next steps, α-KG is converted to succinyl-coenzyme A by α-KG dehydrogenase, and further converted to succinate by the succinyl-coenzyme A synthetase. In mESCs, high levels of α-KG induce TET-dependent hypomethylation and support self-renewal while high levels of succinate promote differentiation [88]. Likewise, the modulation of the involved enzymes, such as overexpression of IDH1 or IDH2, can boost 5hmC levels in cells [89].

Vitamin C can interact with the catalytic domain of TET enzymes and thereby enhance their activity towards 5mC oxidation [90]. In mice, a gestational Vitamin C deficiency mimics a TET1 deficiency with regard to germline reprogramming, transcriptional profile, meiosis and fecundity. Specifically, Vitamin C deficiency leads to DNA hypermethylation at meiosis regulators and transposable elements in the embryonic germline, leading to failures in meiotic initiation and progression in fetal female germ cells [91]. Lastly, also iron(II) [92] and oxygen availability can influence TET activity, although the effects of oxygen on TET activity vary in cells and animals [41].

3.1.6.2 Post-translational modifications of TET proteins

TET proteins can be regulated by various post-translational modifications, which can affect chromatin binding as well as the catalytic activity. All three TET proteins can be monoubiquitinated by the E3 ubiquitin ligase complex cullin-ring finger ligase-4 (CRL4). Ubiquitination occurs on a highly conserved lysine (K) residue (human TET1 K1589, TET2 K1299 and TET3 K859) within the catalytic domain. The monoubiquitination is promoted by the binding of the DDB1- and CUL4-associated factor 1 (VprBP) to the catalytic domain of TET. VprBP is a binding partner of Damaged DNA binding protein 1 (DDB1), which itself is a member of the CRL4 complex. This monoubiquitination promotes the binding of TET to DNA and is essential for proper 5hmC production [93,94].

Furthermore, TET2 can be acetylated by the acetyltransferase p300 in the N-terminal region of TET2 at K110, and can be deacetylated by HDAC1 and HDAC2. Acetylation increases TET2 activity, possibly by preventing polyubiquitination and thereby protein degradation. Interestingly, TET2 acetylation also enhances DNMT1 binding, which also promotes TET2 protein stability. In addition, a DNMT1 knockdown reduces TET2 tightening to chromatin during oxidative stress [95]. This suggests that DNA methylation and DNA demethylation could be directly coupled.

All three TET family members are highly phosphorylated, mostly at the N-terminus and the low complexity insert region, whereas only a few phosphorylation sites have been identified in the catalytic domain. While the impact of phosphorylation on TET activity has only been scarcely

studied so far, it is known that phosphorylation can be suppressed via O-GlcNAcetylation of serines and threonines by the glycosyltransferase OGT [96]. O-GlcNAcetylation was shown to be required for both TET1-mediated 5hmC production and TET1-mediated target gene repression [97].

Furthermore, human TET1 can be modified by addition of poly [ADP-ribose] (PAR), a negatively charged polymer. PARylation of TET1 is catalyzed by poly [ADP-ribose] polymerase 1 (PARP1) and enhances TET1 activity [98].

3.1.6.3 Interaction partners of TET proteins

Several interaction partners of TET proteins have been identified, most of which help to target TET proteins to genomic loci. Several different transcription factors can target TET proteins to their binding sites. For example, NANOG physically interacts with TET1 and TET2 and recruits them to NANOG binding sites in mESCs, where TET1/2-mediated 5hmC production is required for proper somatic cell reprogramming [99]. In a similar manner, PR domain zinc finger protein 14 (PRDM14) physically interacts with TET1/2 and recruits them to its target loci in mESCs in order to induce DNA demethylation at pluripotency-associated genes, germline-specific genes and imprinted loci [100]. Recently, Early growth response protein 1 (EGR1) was discovered to recruit TET1 in order to induce DNA demethylation and the activation of downstream genes during postnatal mouse development [101].

LIN28A, a protein that can bind both DNA and RNA, can recruit TET1 to transcription start sites in mESCs to induce DNA demethylation, linking LIN28A and TET1 specifically to active transcription [102]. In addition, TET proteins also interact with repressive protein complexes in mESCs, for example with two subunits of the Polycomb Repressive Complex 2 (PRC2): Suppressor of zeste 12 protein homolog (SUZ12) and Enhancer of zeste homolog 2 (EZH2). These interactions are in line with the finding that 5hmC correlates with the histone mark H3K27me3, which is generated by PRC2. Accordingly, a SUZ12 knockdown reduces 5hmC levels at H3K27me3 positive regions, especially within bivalent (H3K4me3 and H3K27me3 positive) promoters. This finding combines both roles of TET1, its activating function via DNA demethylation, and its repressive function via PRC2. These two opposing functions of TET proteins could be important to maintain developmental genes in a poised state [103].

TET1 also physically interacts with other DNA demethylation factors, such as TDG [42] and GADD45A [104] (see chapter 3.2.2). An *in vitro* study demonstrated that TET1 can interact with TDG via an N-terminal domain containing the CXXC domain, and the C-terminal catalytic domain of TET1. The two enzymes closely work together in promoting 5mC oxidation and excision, suggesting a closely coupled DNA demethylation pathway [42].

3.2 The role of GADD45A in active DNA demethylation

3.2.1 Molecular functions of GADD45A

GADD45A belongs to a family of small (18 kDa), stress-inducible acid nuclear proteins with three members: GADD45-alpha (GADD45A), -beta (GADD45B), and –gamma (GADD45G). GADD45 proteins play a role in diverse processes such as cell cycle control, apoptosis and DNA repair [105]. In 2007, GADD45A was discovered to play a role in active DNA demethylation, which was proposed to occur via DNA repair through the recruitment of the endonuclease XPG [20]. Although other studies have confirmed a role of GADD45 proteins in active DNA demethylation, how and if GADD45 proteins mediate active DNA demethylation remains to be a matter of debate in the field, one reason being perhaps the lacking enzymatic activity of GADD45 proteins [27,31,33,106]. Since then, several mechanisms of GADD45-mediated DNA demethylation have been proposed: In zebrafish, Gadd45 proteins were shown to promote DNA demethylation via the recruitment of the deaminase AID, followed by thymine base excision by Methyl-CpG-binding domain protein 4 (Mbd4) [31]. However, this mechanism has been challenged [33]. More evidence for the involvement of GADD45 proteins in DNA demethylation emerged with the understanding of how GADD45 can be recruited to sites of DNA demethylation, and with the discovery of TET-TDG mediated active DNA demethylation (see chapters 3.1.2 and 3.2.2).

GADD45A can be directed to chromatin and mediate DNA demethylation by interacting with different chromatin factors. TAF12, a subunit of RNA Polymerase I and II-specific TBP-TAF complexes, can recruit GADD45A to the promoters of active genes [107]. The H3K4me3 reader Inhibitor of growth protein 1 (ING1) can recruit GADD45A to H3K4me3 positive regions via its PHD domain [108], and the long noncoding RNA TARID can recruit GADD45A to the *TCF21* promoter [109]. GADD45A-mediated locus-specific DNA demethylation was further corroborated by a recent study, which showed that *Gadd45a/b/g* triple-knockout mESCs display hypermethylation at ~ 7000 sites, but only minor global hypermethylation. The hypermethylated sites correlate with loci undergoing TET-TDG-dependent oxidation and are enriched for enhancers and coding exons [110]. Overall, these studies suggest that GADD45A has a role in gene-specific DNA demethylation rather than global DNA demethylation, as originally proposed [20].

In 2015, two studies consolidated GADD45A-mediated DNA demethylation and the TET-TDG pathway by discovering that GADD45A can interact with both TET1 [104] and TDG [111] (Figure 3.3). First of all, GADD45A physically interacts with TET1 to enhance 5hmC formation [104]. Second, GADD45A physically interacts with TDG to increase 5fC and 5caC removal in transfected HEK293T cells [111]. In addition, *Gadd45a/b* double- and *Gadd45a/b/g* triple-knockout mESCs displayed hypermethylation at loci that are also affected by TDG depletion [110,111]. Still, the exact mechanism by which GADD45A enhances TDG activity has not been resolved to date.

3.2.2 Interaction of GADD45A and TET1 in active DNA demethylation

Recent work in our research group linked the roles of GADD45A and TET1 in active DNA demethylation. As shown for TDG, GADD45A can also physically interact with TET1 to enhance TET1 activity. In particular, GADD45A enhances the formation of 5hmC by TET1. On the other hand, it reduces 5fC and 5caC levels, which is probably linked to its interaction with TDG (Figure 3.3). Importantly, GADD45A also interacts with the isolated catalytic domain of TET1 to enhance 5hmC production. In contrast, the enhanced removal of 5fC and 5caC mediated by GADD45A also requires the N-terminus of TET1. Moreover, GADD45A and full-length TET1 can synergistically activate a number of GADD45A target genes in HEK293T cells, which also requires the catalytic activity of TET1 [104]. Overall, this points to a dual function of GADD45A, namely to activate TET1 and also to enhance 5fC/5caC removal by TDG [104,111]. However, the mechanism by which GADD45A enhances TET1 activity and 5fC/5caC removal by TDG is still unresolved.

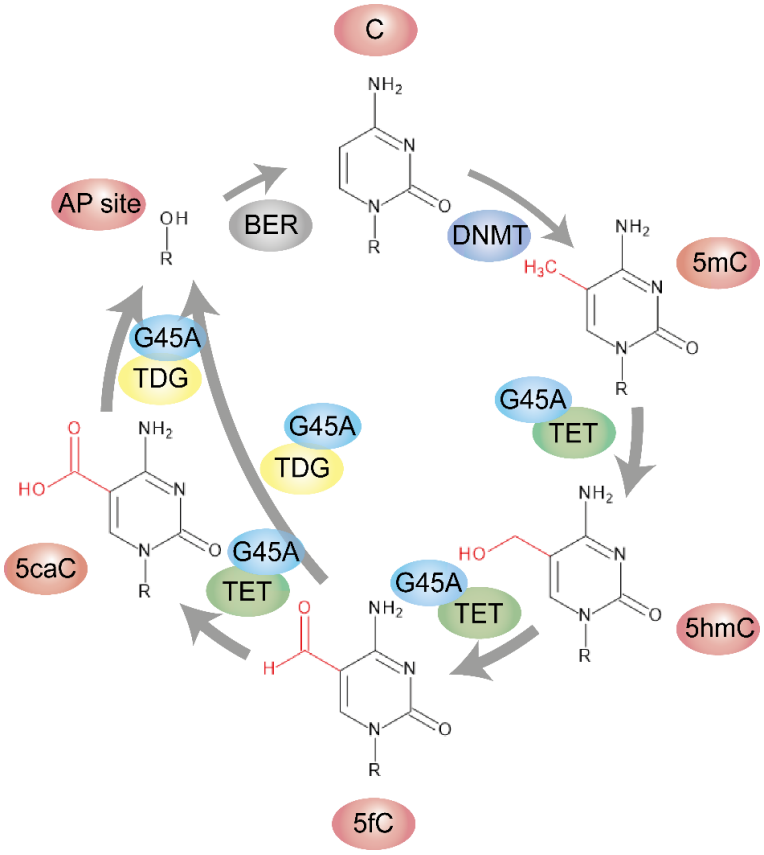

Figure 3.3 TET-TDG mediated active DNA demethylation and the role of GADD45A
Cytosine (C) is methylated by DNMT1 (maintenance methyltransferase) or DNMT3A/B (*de novo* methyltransferases) at the 5' position. TET enzymes (TET1/2/3) oxidize the methyl group in an iterative manner to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). 5fC and 5caC are then excised by TDG, which generates an abasic (AP) site. Unmethylated C is finally restored by the base excision repair (BER) machinery. Modifications of C are highlighted in red. GADD45A (G45A) has a dual function in this process, as it can 1) interact with TET1 to enhance TET1 oxidation activity and 2) interact with TDG to enhance the removal of 5fC and 5caC.

3.3 Mitogen-activated protein kinases (MAP kinases) as interaction partners of GADD45 proteins

3.3.1 Molecular functions of MAP kinases

As stress response proteins, GADD45 proteins have various interaction partners, among them members of the Mitogen-activated protein kinase (MAP kinase) family [112]. MAP kinase signaling pathways are highly conserved and ubiquitous mechanisms of eukaryotic cell regulation and respond to various stimuli such as hormones, growth factors, cytokines, agents acting through G protein-coupled receptors, Transforming growth factor β-related agents, pathogen-associated molecular patterns, and environmental stresses. MAP kinase signaling consists of a cascade comprised of three steps, which includes a MAP kinase kinase kinase, a MAP kinase kinase, and a MAP kinase (Figure 3.4). MAP kinases are proline directed kinases, which phosphorylate serines and threonines of protein targets only if they are followed by prolines. Substrate specificity is established through docking sites on the substrates [113,114].

3.3.2 Activation of MAP kinases by GADD45 proteins

In response to stress, GADD45 proteins can mediate the activation of Mitogen-activated protein kinase kinase kinase 4 (MAP3K4/MTK1/MEKK4) by directly binding to an N-terminal region of MAP3K4. The binding of GADD45 proteins interrupts the auto-inhibition of the C-terminal kinase domain of MAP3K4, caused by the interaction with the N-terminus of MAP3K4. Upon GADD45 binding, MAP3K4 can dimerize, allowing the autophosphorylation and activation of MAP3K4 [115]. An activation of MAP3K4 activates downstream kinases such as MAP2K3, MAP2K4 and MAP2K6, which in turn activate MAPK14 (p38α) and MAPK11 (p38β) [114]. In addition, GADD45A can directly bind to p38 (MAPK11,12,13,14) and mediate oncogenic H-Ras-induced activation of p38 in a MAP3K4-independent manner [116].

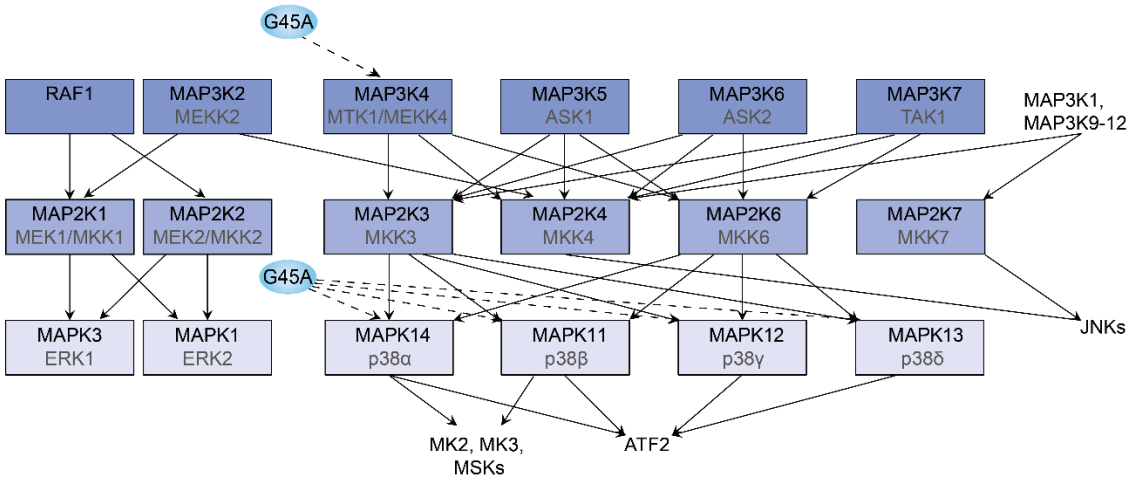

Figure 3.4 Overview of mammalian MAP kinase signaling and its interaction with GADD45A
MAP kinase signaling forms a cascade, where a MAP kinase kinase kinase (MAP3K) phosphorylates a MAP kinase kinase (MAP2K), which in turn phosphorylates a MAP kinase (MAPK). MAP kinases can activate additional kinases. Solid lines indicate an activation via direct phosphorylation by an upstream kinase, while dashed lines indicate an activation via GADD45A (G45A).

3.4 The 3D genome architecture

Beyond the regulation of active DNA demethylation via TET interaction partners such as GADD45A, also the three-dimensional (3D) genome architecture has been implicated in shaping the DNA (de)methylome [41]. Human diploid cells contain roughly 6 billion base pairs of DNA. This corresponds to about 2 meters of DNA per cell [117]. However, human cells only have a diameter of about 10-100 µM. The answer to how the entire DNA fits into the cell nucleus is: the eukaryotic DNA is packaged into chromatin, a complex of DNA, associated proteins and RNA. The first level of chromatin packaging is the nucleosome, which was discovered in the 1970s [118]. Nucleosomes are further organized into the 11 nm chromatin fiber, which folds at the submegabase scale into higher-order domains called topologically associating domains (TADs) (Figure 3.5 A) [119]. TADs have a high frequency of self-interaction and are separated from other TADs by boundary elements, such as the zinc finger protein CTCF and the Cohesin complex. Whereas TADs are conserved across tissues, sub-TADs (finer scale separation) and meta-TADs (larger scale separation) differ across cell types [120]. The sub-TADs, also known as intra-TADs, constrain promoter-enhancer interactions [121]. At the megabase scale, chromatin separates further into an active, so-called compartment A, and an inactive compartment B [122].

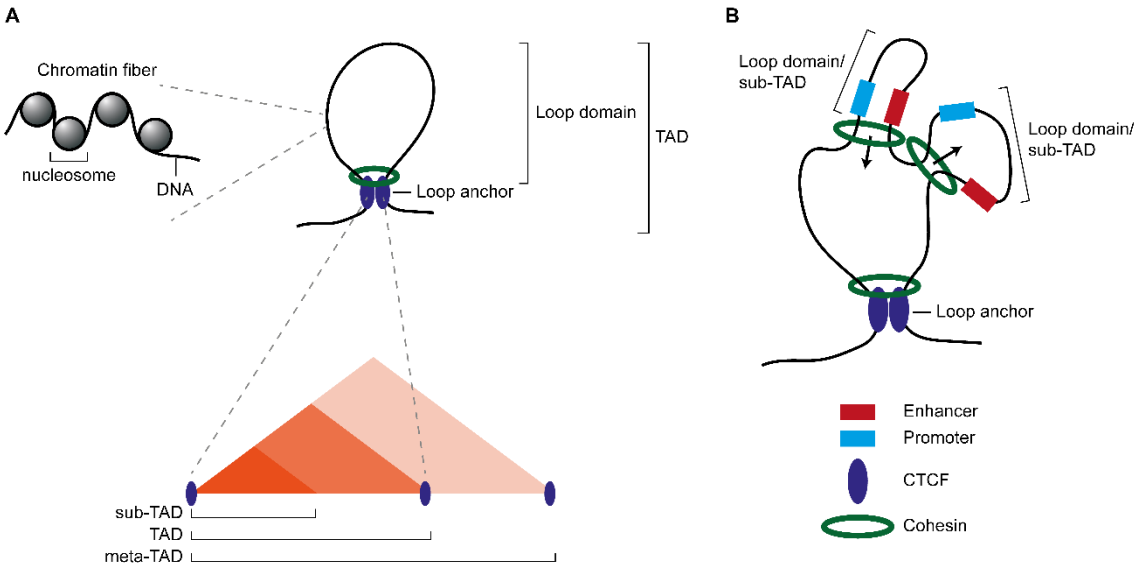

Figure 3.5 Hierarchical model of chromatin organization
A) The eukaryotic DNA is wrapped around histones, forming nucleosomes, which are organized into the 11 nm chromatin fiber. At the submegabase scale, chromatin folds into TADs, regions that display a high frequency of self-interaction. TADs are separated by domain boundaries. These domain boundaries are created by CTCF and the Cohesin complex, which form so-called loop anchors. TADs can be further separated into sub-TADs, which are regions with a higher interaction frequency (dark red), and can also form meta-TADs, regions with an overall lower interaction frequency than the sub-TADs or TADs (light red). The lower panel represents interaction signals as obtained from Hi-C data. B) The loop extrusion model proposes a dynamic process of chromatin structuring, in which the Cohesin complex moves along the chromatin until it reaches CTCF binding sites. In this extrusion process, additional loop domains, also called sub-TADs, are formed and enhancers can be brought into close proximity to promoters. The arrows indicate movement of the Cohesin complex along chromatin.

TADs have been mainly defined by the so-called Hi-C method, a method based on chromosome conformation capture that was first described in 2009 by Lieberman-Aiden and colleagues. In this method, DNA fragments that are originally in close spatial proximity are ligated. Subsequently, interactions between DNA fragments are identified via next generation sequencing [122]. Today, the existence of TADs has been supported by alternative methods such as high resolution microscopy [123]. Yet, whereas the original Hi-C method gives an overview over the 3D structure in a population of cells and found that TADs are highly conserved, high resolution microscopy and also the more recent single-cell Hi-C identified a large number of cell-specific TADs [123,124].

3.4.1 Establishing TADs and promoter-enhancer contacts

3.4.1.1 The role of the Cohesin complex in chromatin looping

Cohesin is a multi-unit, ring-shaped protein complex that is highly conserved in eukaryotes (Figure 3.6) [125]. Originally, it was discovered as an essential factor in cell division, during which it promotes sister chromatid cohesion, thus ensuring proper chromatid segregation during anaphase [126]. Cohesin's ring-shaped structure is formed by the coiled-coil domains of the Structural maintenance of chromosomes subunits 1 and 3 (SMC1 and SMC3), which form a heterodimer. SMC1 and SMC3 contain ATPase domains, which are connected by a protein called RAD21. This tripartite ring has a diameter of about 40 nm, which can enclose one or two chromatin strands [127]. The fourth subunit, SA1 or SA2, contacts RAD21 and additional Cohesin subunits transiently associate with the complex, such as PDS5 and WAPL. SA1/2, PDS5 and WAPL are necessary to remove Cohesin complexes from chromatin [125].

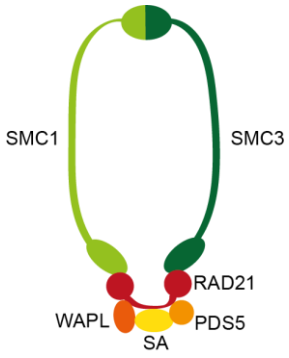

Figure 3.6 The Cohesin Core Complex
The ring-shaped Cohesin complex consists of the Structural maintenance of chromosomes subunits 1 and 3 (SMC1 and SMC3), which form a heterodimer. SMC1 and SMC3 contain ATPase domains, which are connected by RAD21, closing the ring structure. The fourth subunit, SA1 or SA2, contacts RAD21 and plays a role in the first part of Cohesin release. Additional Cohesin subunits associate transiently with the complex, such as PDS5 and WAPL, which are essential for Cohesin removal.

In 2008, Wendt and colleagues discovered a new function of Cohesin: shaping the 3D chromatin structure. Their study showed that the binding sites of Cohesin greatly overlap with the binding sites of CTCF, a factor already known to be important for transcriptional insulation [128]. Today, the loop extrusion model is the most accepted model that describes how Cohesin and CTCF cooperate to shape the 3D genome. Cohesin is loaded onto chromatin by the NIPBL-MAU2

complex and moves along the chromatin until it is blocked by CTCF bound to convergent CTCF binding sites [129]. The extrusion process can occur actively via the motor activity of the ATPase domains of SMC1 and SMC3 [130], via constant Cohesin loading and diffusion [131], or via RNA Polymerase II during transcription [129]. Through this process, Cohesin and CTCF form loop domains. During the extrusion process, regulatory regions such as enhancers and promoters are brought into close proximity (Figure 3.5 B). Conversely, enhancers and promoters lying within different TADs can be isolated from one another by boundary elements, such as the loop anchors consisting of CTCF and Cohesin [120].

Evidence for Cohesin's importance in loop domain formation comes from several studies all published in 2017, in which either the Cohesin subunit RAD21, the Cohesin loading factor NIPBL, or the release factors WAPL and PDS5 were depleted. Rao and colleagues acutely depleted RAD21 in HCT116 cells using the auxin-inducible degron system. Already 40 min of RAD21 loss depleted the loop domains, whereas auxin withdrawal could restore the majority of loops within 60 min [132]. Moreover, the Cre-driven conditional deletion of *Rad21* in mouse oocytes also depleted loops in the one-cell embryo [133]. In a study by Schwarzer and colleagues, *Nipbl* was deleted in mice, specifically in non-dividing hepatocytes, using a tamoxifen-inducible Cre driver. The depletion of NIPBL led to a strong reduction of chromatin-bound Cohesin; more notably, TADs disappeared globally [134]. Finally, also the release factors WAPL and PDS5 were shown to play a role in proper loop formation: by depleting WAPL in HAP1 cells, Haarhuis and colleagues found that WAPL restricts chromatin loop extension [135]. This finding was also confirmed in mouse zygotes by Gassler and colleagues [133]. Accordingly, PDS5 depletion by RNAi in HeLa cells decreased TAD numbers but increased TAD sizes [136]. Interestingly, the above-mentioned studies also show that Cohesin restricts compartmentalization, i.e. the separation of chromatin into an active compartment A, and an inactive compartment B. In addition, Cohesin loss causes superenhancers to co-localize, thereby allowing aberrant transcription of nearby genes [132]. Overall, Cohesin has a complex role in the 3D genome organization; to allow distant genomic regions to interact, but also to insulate genomic regions from one another.

3.4.1.2 The role of additional factors in promoting promoter-enhancer contacts

Although the Cohesin complex can bring together enhancers and promoters, and CTCF can act as a boundary to restrict aberrant promoter-enhancer interactions outside of TADs, promoter-enhancer contacts are not directly mediated by Cohesin and CTCF. Instead, additional factors have been implicated in the promotion of promoter-enhancer contacts. These include Mediator, a transcriptional co-activator complex of RNA polymerase II transcription [137], its interactor Estrogen-related receptor b (ESRRB) [138], the transcription factors Yin Yang 1 (YY1) [139], ZNF143 [140] and OCT4 [141], long noncoding RNAs [142], and the CHD4-ADNP-HP1 (ChAHP) complex [143].

Mediator can form a complex with Cohesin, and loops are formed between sites co-occupied by both Mediator and Cohesin. In contrast to CTCF and Cohesin co-bound sites, which tend to be more conserved, Mediator-Cohesin co-bound sites were found to be cell-type specific [137]. YY1

is a zinc-finger transcription factor, which is ubiquitously expressed and can promote structural chromatin interactions by forming dimers, in a similar manner as CTCF. To monitor promoter-enhancer contacts after YY1 depletion, Weintraub and colleagues used the degradation tag (dTAG) system to deplete YY1 in mESCs and performed HiChIP, a protein-centric chromatin conformation method, for H3K27ac. YY1 depletion led to a decreased interaction frequency of enhancers and promoters, which was accompanied by significant expression changes of 60 % of the genes connected by YY1 promoter-enhancer loops [139]. Another zinc-finger protein, ZNF143 was also shown to bind to promoter-enhancer loops in human cancer cell lines and also mouse sperm [140,144]. Some ZNF143 binding sites only co-localized with Cohesin, but not CTCF binding sites. Interestingly, these were preferentially short-range interactions, implicating that ZNF143 might play a role in the regulation of sub-TADs [144]. A recent preprint discovered another transcription factor that can target Cohesin to CTCF-unbound sites: OCT4. The authors propose a model by which the pioneer transcription factor OCT4 can create open chromatin regions, which allows Cohesin to bind [141]. Cohesin was also found to be enriched at a new class of long noncoding RNA loci, called trait-relevant intergenic long noncoding RNAs. As these long noncoding RNA loci are also enriched for enhancer-like chromatin signatures, this study suggests an involvement of long noncoding RNAs in shaping promoter-enhancer interactions [142]. Recently, the ChAHP complex, a complex comprised of the zinc finger transcription factor ADNP, the chromatin remodeler CHD4 and the heterochromatin binding protein HP1, was demonstrated to counteract chromatin looping by competing for CTCF binding sites [143].

However, how Cohesin, CTCF and the various regulators act together in promoting promoter-enhancer contacts has not been studied in detail. Furthermore, yet uncharacterized or cell-type specific factors might play additional roles. In fact, a study by Rubin and colleagues proposed that there might be Cohesin-independent promoter-enhancer loops, which are formed during differentiation. During epidermal differentiation, one class of promoter-enhancer contacts established contacts, while the second class already existed in undifferentiated cells. Notably, while the second class was associated with Cohesin, the first class was Cohesin independent. Instead, the newly established promoter-enhancer contacts were bound by the differentiation induced transcription factors KLF4 and ZNF750. Since these two factors also bind to other regions, they do not globally regulate promoter-enhancer looping, but rather are cell-type specific regulators of promoter-enhancer contacts [145].

In contrast, Matthews and Waxman computationally predicted and experimentally validated that CTCF and Cohesin also form anchor loop structures within TADs, which they refer to as intra-TAD loops. These intra-TAD loops constrain promoter-enhancer contacts and show conservation among various mouse and human tissues [121]. A recent preprint has further underlined the importance of Cohesin in mediating promoter-enhancer contacts. Depletion of the release factor WAPL resulted in decreased contact frequency between promoters and regulatory elements [141]. In conclusion, tissue invariant, i.e. CTCF/Cohesin as well as tissue specific mechanisms, e.g. transcription factors, likely cooperate to promote and also to constrain promoter-enhancer contacts.

3.4.2 The role of promoter-enhancer looping in gene regulation

It has long been proposed that enhancers regulate promoters via looping [146]. However, only in recent years more light has been shed on the relationship between chromatin structure and gene regulation. While disturbances of the local chromatin structure can have an immense impact on the regulation of individual genes, disturbances of global players, such as CTCF and Cohesin, have a surprisingly small effect [132,147].

An example for the regulation of an individual locus is the thoroughly studied β-globin locus, which contains several β-globin genes. In erythroid cells, a distal enhancer, referred to as the locus control region, physically contacts different β-globin promoters at different developmental stages to stimulate transcription of the appropriate β-globin. By artificially targeting the promoter of the repressed fetal β-globin gene to the distal enhancer, transcription of the fetal β-globin gene could be greatly enhanced in human erythroblasts [148].

Still, existing promoter-enhancer contacts do not necessarily lead to active transcription. Developmental studies in flies and mice have demonstrated that many promoter-enhancer contacts persist during different developmental stages, although the transcriptional programs are quite different [144,149]. In fact, many promoter-enhancer contacts are even associated with paused RNA polymerase [149]. Thus, chromatin structure seems to be important for gene regulation, but is not the only mode of regulation.

In this context, manipulation of conserved regulators of chromatin looping does not greatly affect gene expression. Depletion of the Cohesin subunit RAD21 leads to a loss of loops in HCT116 cells, but only to minor gene expression changes. Out of 12,222 genes that were expressed in untreated cells, only 66 genes showed a ≥ 2-fold change [132]. Two recent preprints found similarly modest, although slightly more widespread, reductions in gene transcription in mESCs [141,150]. In addition, Rhodes and colleagues discovered a new role for Cohesin in counteracting long-range polycomb chromatin domain interactions. Out of 365 downregulated genes after RAD21 depletion, 251 genes were bound by the PRC1 member RING1B [150].

Manipulating Cohesin loading via NIPBL deletion in mouse hepatocytes led to differential gene expression of about a thousand genes, with more genes downregulated than upregulated. Exogenic regions, i.e. intergenic or antisense intragenic regions, were more often upregulated than downregulated [134]. When Cohesin release was inhibited by WAPL deletion in HAP1 cells, a similar number of genes was misregulated [135]. Although these numbers are higher than the number of differentially expressed genes after RAD21 depletion, the genes misregulated after NIPBL and WAPL depletion still only correspond to a subset of genes. In addition, even in regions where loop length increased after WAPL depletion, genes were still identified for which gene expression did not change [135].

Loss of CTCF in mESCs disrupted TAD insulation, giving rise to ectopic contacts across TAD boundaries. However, effects on gene transcription were also modest, with 370 differentially

expressed genes after 1 day of CTCF depletion. A subset of genes was upregulated, possibly due to loss of CTCF's insulator function, and resulting ectopic activation by enhancers [147].

In summary, disrupting 3D chromatin structuring factors does affect gene expression, but to a very limited extent. A reason could be that multiple factors shape the 3D chromatin in concert to ensure a fine-tuned gene regulation. Thus, the factors still intact in the described experimental conditions might compensate for the loss of one or only a few of these factors.

3.4.3 The interplay between chromatin structure and DNA (de)methylation

3.4.3.1 Chromatin structure as a regulator for active DNA (de)methylation

The direct impact of disturbing promoter-enhancer looping on gene expression is rather modest. Hence, the question arises what other functions promoter-enhancer looping might have. The oxidation products of TET – 5hmC, 5fC and 5caC – are largely found at distal elements, whereas TET enzymes themselves are usually found at promoters. One model to explain this conundrum is that TET enzymes are recruited to enhancers via promoter-enhancer looping [41]. A mechanistic proof of this model is still missing, but several independent studies have formed a basis of support for this model.

A study by Song and colleagues found that 5hmC and 5fC are enriched especially at enhancers predicted to link to promoters, suggesting that the loop structure might be important to mediate 5mC oxidation [86]. In addition, two studies by Sun and colleagues showed that 5hmC and 5fC are located at CTCF-bound regions, i.e. regions that are important for constraining promoter-enhancer contacts. At these regions, 5hmC and 5fC are distributed in a symmetrical and regularly spaced oscillating manner. The distance between neighboring 5hmC and 5fC peaks correlates with the unit length of the DNA wrapped around one nucleosome, ~150 bp. Thus, the nucleosome structure is likely important for TET regulation, perhaps by affecting the accessibility of 5mC and 5hmC, the substrates for TET enzymes [83,84].

A recent study discovered GADD45A and ING1 as novel factors involved in organismal aging via DNA demethylation of C/EBPβ-dependent superenhancers. This local DNA demethylation was dependent on long-range chromatin looping. While the TET1 interactor GADD45A was found to occupy mostly superenhancers, the second key protein, ING1, occupied promoters. As not only GADD45A, but also ING1 can interact with TET1, long-range chromatin looping was proposed to bring all of these factors together in order to induce DNA demethylation [151].

Lastly, at distal regulatory regions, DNA binding factors have been implicated in shaping the methylome. Of note, CTCF binding itself was suggested to initiate local DNA demethylation [152].

3.4.3.2 DNA (de)methylation as a regulator for chromatin structure

In a converse fashion, DNA (de)methylation can also regulate CTCF/Cohesin binding. CTCF itself is methylation sensitive and upon hypermethylation of its binding sites, CTCF binding is reduced [153]. In addition, DNA methylation was recently suggested to regulate CTCF/Cohesin-mediated chromatin interactions via the Methyl-CpG-binding protein 2 (MeCP2). A loss of DNMT3B reduced RAD21-mediated chromatin interactions in human ESCs. Xu and Corces proposed that this is

due to a loss of hemimethylation, likely because of an altered physical interaction with MeCP2 [154]. MeCP2 can, as described by its name, bind to methylated C occurring in a CpG context, but also binds to 5hmC [155]. In addition, it can alter the chromatin state, e.g. by affecting histone H1 binding [156] or by recruiting the chromatin remodeler ATRX in order to allow CTCF binding [157]. Moreover, MeCP2 interacts with SMC1 and SMC3, two members of the Cohesin complex, and MeCP2's interactor ATRX was shown to be required for the chromatin occupancy of Cohesin and CTCF [158].

Whether DNA (de)methylation regulates chromatin structure, or if the chromatin structure determines the distribution of 5mC, 5hmC, 5fC and 5caC is unclear. It is possible that both control each other in a feedback mechanism to ensure the proper regulation of the genome, as was proposed by a recent study by Wiehle and colleagues [159]. On the one hand, they established CTCF as an insulator against the spreading of DNA methylation and proposed that CTCF sites could act as "bookmarks" for DNA demethylation. On the other hand, they showed that TET1/2 deficiency leads to a gain of DNA methylation and increased nucleosome occupancy, and thereby to an altered chromatin structure. Uniting both regulatory pathways, the increased nucleosome occupancy itself can affect CTCF binding, which in turn changes DNA methylation patterns.

3.5 Aim of the thesis

Research of the past 10 years has revealed many milestones regarding the understanding of active DNA demethylation in vertebrates. Today we comprehend the details of the biochemical mechanism of TET-mediated oxidation of 5mC and its oxidative derivatives in considerable detail. The regulation of TET enzymes by post-translational modifications and interacting proteins has also been widely studied; however, how post-translational modifications and interactors directly affect the enzymatic activity of TET remains uncertain. The small protein GADD45A is required for TET1 induced 5mC oxidation and DNA demethylation [104]. Yet, the mechanism by which GADD45A enhances TET1 activity and 5fC/5caC removal is still unresolved. In addition, a conundrum in the field of DNA demethylation has not been solved thus far: whereas TET binding sites are mostly found at promoters, the oxidation products of TET, 5hmC, 5fC and 5caC are mostly enriched at enhancers. My thesis addresses two TET-related topics: 1) TET1 regulation by GADD45A and 2) TET regulation by RAD21-dependent chromatin looping.

In order to unravel the mechanism of GADD45A-mediated TET1 regulation, I explored several different possibilities. I considered that GADD45A could affect post-translational modifications of TET1 and monitored GADD45A-mediated changes of TET1-phosphorylation, -ubiquitination and -SUMOylation and their impact on TET1 activity. In addition, I focused on a few potential GADD45A-dependent interactors of TET1 and analyzed their roles in regulating TET1.

To address the discrepancy between the location of TET binding sites and 5hmC, 5fC and 5caC sites, I first generated a system where chromatin loops could be globally depleted. I used mESCs as a model system and CRISPR/Cas9 to establish a conditional RAD21-knockout system. To investigate whether the loss of promoter-enhancer loops leads to a loss of 5fC at enhancers, I depleted the loops by depleting RAD21 and monitored 5fC levels via DNA immunoprecipitation followed by sequencing.

4 Results

4.1 Investigating mechanisms of GADD45A-mediated TET1 regulation

Our laboratory recently showed that GADD45A is required for TET1 induced 5mC oxidation and DNA demethylation [104]. However, at the start of my PhD project, the mechanism by which GADD45A enhances TET1 activity and 5fC/5caC removal was still unresolved. In the following chapters, I explored various possible mechanisms by which GADD45A could enhance TET1 activity: by the interaction with MAP kinases (4.1.2), by enhancing ubiquitination (4.1.3) or SUMOylation (4.1.4) and by the recruitment of interactor proteins (4.1.5 and 4.1.6).

4.1.1 Activity of a GADD45A-TET1 fusion protein

In order to study the functional interaction between GADD45A and TET1 in more detail, I generated a fusion construct of GADD45A and the catalytic domain of TET1 (TET1CD), with an N-terminal FLAG-HA tag and a flexible glycine-serine (GS)-linker between GADD45A and TET1CD (Figure 4.1 A, top). I transfected HEK293T cells with two GADD45A-TET1CD fusion constructs (originating from two different bacterial clones), as well as three control TET1CD constructs, in three different concentrations, and monitored the protein expression by Western blotting (Figure 4.1 A, bottom). The GADD45A-TET1CD fusion proteins could be detected both with an HA and a GADD45A-antibody, and they were similarly or better expressed than the TET1CD proteins.

To test the enzymatic activity of the GADD45A-TET1CD fusion proteins, genomic DNA was purified from HEK293T cells overexpressing the constructs and subjected to quantitative mass spectrometry. When TET1CD was transfected at the highest concentration (2 µg/well in a 12-well dish in a total volume of 600 µl), 5hmC levels were increased 18-fold compared to empty vector transfection (Figure 4.1 C, compare “TET1CD” to “Control”, black bars). However, 5hmC levels were only 6-fold enriched using the GADD45A-TET1CD fusion protein (Figure 4.1 C, compare “GADD45A-TET1CD” to “Control”, black bars). This effect was more pronounced for 5fC and 5caC levels, as they could barely be detected after overexpression of the GADD45A-TET1CD fusion protein, whereas they were easily detectable after TET1CD transfection (Figure 4.1 D and E). Thus, instead of the expected enhanced activity, the GADD45A-TET1CD fusion construct had a reduced activity towards 5mC oxidation. One reason could be that the 10-amino acid linker (GGGGSGGGGS) between GADD45A and TET1 is not sufficiently long and flexible, and the fusion protein had a reduced activity because of steric hindrance. Consequently, I did not use the GADD45A-TET1 fusion constructs in further experiments to study the functional interaction between these two proteins.

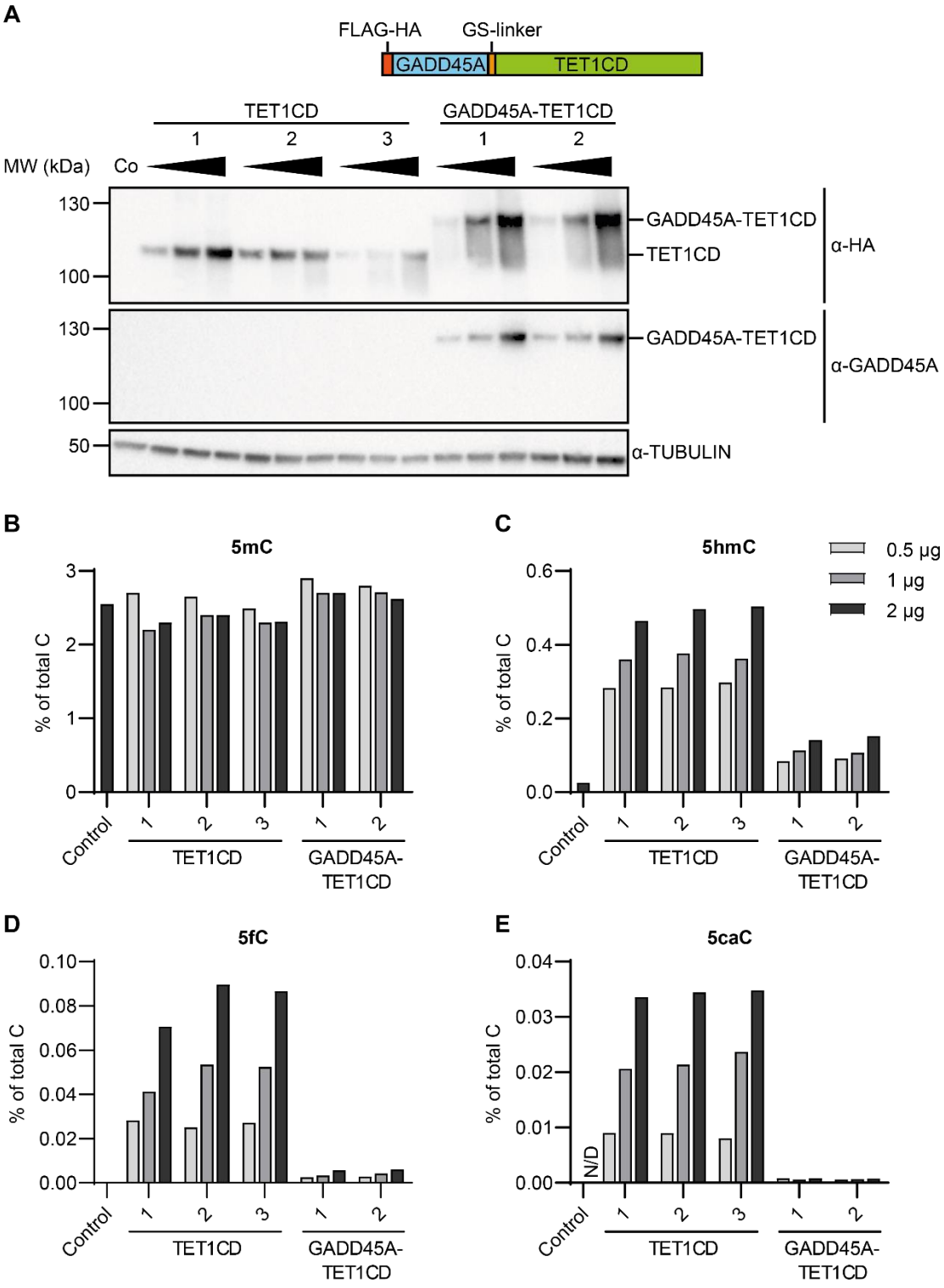

Figure 4.1 The GADD45A-TET1CD fusion protein has a reduced catalytic activity
HEK293T cells were transfected for 48 h with empty vector (Control), TET1CD, or GADD45A-TET1CD fusion constructs. (A) Top: Scheme of the GADD45A-TET1CD fusion construct. Bottom: Western blot showing expression levels of TET1CD and GADD45A-TET1CD. Plasmids were transfected in 3 different concentrations, 0.5 µg, 1 µg and 2 µg (per well of a 12-well dish in a total volume of 600 µl) represented by the triangle above. MW, molecular weight; Co, control. (B-E) Quantitative mass spectrometry measurements of genomic cytosine modifications. Mass spectrometry data by Dr. Michael Musheev.

4.1.2 GADD45A and MAP kinase mediated TET1 regulation

Unpublished data from our laboratory suggested that the MAP kinase MAP3K4, a known interactor of GADD45A [160], might play a role in enhancing TET1 activity in a GADD45A-dependent manner (Dr. Andrea Schäfer, personal communication). The binding of GADD45A to MAP3K4 mediates the autophosphorylation of MAP3K4, thereby activating the kinase [115]. Thus, TET1 could be a potential phosphorylation target of MAP3K4 or its downstream kinases and I investigated how GADD45A, MAP kinases and TET1 functionally interact.

4.1.2.1 GADD45A and MAP3K4 decrease TET1 oxidation products

To study the effect of GADD45A and MAP3K4 on TET1 enzymatic activity, these proteins were transiently overexpressed in different combinations in HEK293T cells. The cell line HEK293T was chosen as it is easy to manipulate and cost-efficient. However, it has very low endogenous TET1 expression levels, and consequently, very low 5hmC, 5fC and 5caC levels. Therefore, also TET1 has to be overexpressed to increase the levels of oxidized 5mC derivatives.

The overexpression of TET1, GADD45A and MAP3K4 was confirmed by Western blot analysis (data not shown). 5mC and its oxidative derivatives – 5hmC, 5fC and 5caC – were measured by quantitative mass spectrometry (Figure 4.2 A-D). Whereas 5mC levels were not affected by the overexpression of the different proteins on a genomic scale, TET1 overexpression led to a 6-fold and 50-fold increase in 5hmC and 5fC, respectively, and raised 5caC levels to a detectable level (Figure 4.2 B-D, compare dark grey bars to light grey bars). Co-expression of GADD45A and TET1 significantly increased 5hmC levels (compared to TET1 overexpression alone, $p<0.005$), while 5caC levels were reduced ($p<0.0005$), consistent with previous reports [104,111]. MAP3K4, TET1 and GADD45A co-expression reduced 5hmC levels compared to TET1 overexpression alone (Figure 4.2 B, compare green bar to dark grey bar). Of note, the observed differences in 5hmC levels did not arise from different TET1 protein expression levels (data not shown). This observed reduction is surprising, as it is opposite to the increase in 5hmC levels caused by GADD45A and TET1 co-expression. Moreover, it mimics the depletion of GADD45A, which also leads to a decrease in 5hmC levels [104].

5fC and 5caC levels were also reduced when MAP3K4 was co-expressed with TET1 and GADD45A ($p<0.0005$) (Figure 4.2 C and D, compare green bars to dark grey bars). Also, the co-expression of MAP3K4 and TET1, without GADD45A, decreased 5caC levels ($p<0.0005$) (compare blue bars to dark grey bars). The decrease in global 5hmC, 5fC and 5caC levels by the co-expression of GADD45A and MAP3K4 can be interpreted in contrasting ways: First, TET1 enzymatic activity could be inhibited by MAP3K4. However, it is also possible that GADD45A and MAP3K4 might enhance TET1-mediated 5mC oxidation, as the reduction of 5hmC, 5fC and 5caC might be the result of an increased TET1 enzymatic activity coupled to an accelerated removal of 5fC and 5caC by TDG. In addition, it also has to be considered that active MAP3K4 might solely affect the removal of 5fC and 5caC by TDG, and not TET1 activity directly.

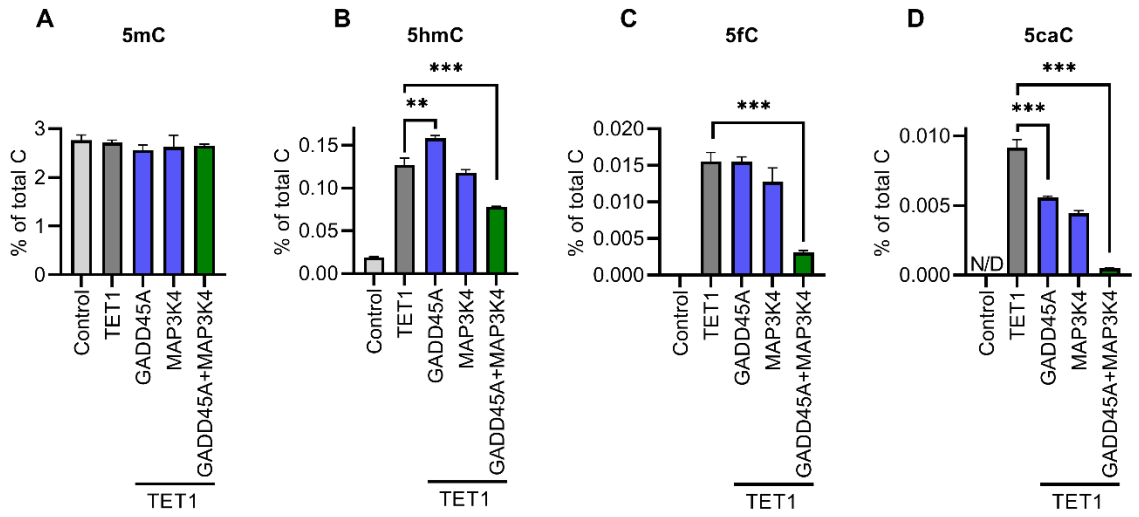

Figure 4.2 GADD45A and MAP3K4 decrease TET1 oxidation products
Quantitative mass spectrometry measurements of genomic levels of 5mC (A), 5hmC (B), 5fC (C) and 5caC (D) in HEK293T cells, which were transiently transfected for 48 h with plasmids encoding the indicated proteins or empty vector (Control). N/D, non detectable. Data shown as mean ± SD, n=3, ** = p-value < 0.005, *** = p-value < 0.0005, according to student's t-test. Mass spectrometry data by Dr. Michael Musheev.

4.1.2.2 siRNA knockdown of selected MAP kinases does not rescue the MAP3K4-induced reduction of 5hmC, 5fC and 5caC levels

MAP3K4 is part of a complex network of MAP kinases, with many kinases being redundant (see Figure 3.4) [113]. In order to unravel whether MAP3K4 is directly responsible for the reduction of 5hmC, 5fC and 5caC levels (see Figure 4.2), or if downstream kinases of MAP3K4 or other MAP kinases are involved, I performed siRNA knockdowns of candidate MAP kinases followed by quantitative mass spectrometry to monitor oxidized 5mC levels. I also overexpressed TET1, GADD45A and MAP3K4 to monitor if the siRNA knockdowns rescued the above-mentioned reduction of oxidized 5mC derivatives.

Successful knockdown of the MAP kinases was confirmed via reverse transcription followed by quantitative PCR (RT-qPCR), showing siRNA knockdown efficiencies between 40 and 80 %. For one kinase, MAP3K6, the siRNA knockdown was not successful (data not shown). The overexpression of TET1, GADD45A and MAP3K4 was confirmed by Western blot analysis and of note, their expression levels were not affected by the siRNA knockdowns (data not shown).

None of the transfected siRNAs rescued the reduction of 5hmC, 5fC and 5caC levels by MAP3K4 (Figure 4.3 B-D, compare green bars to blue bars). This suggests that none of the tested MAP kinases is required for the reduction of 5hmC, 5fC and 5caC levels by MAP3K4 and GADD45A. Alternatively, higher knockdown efficiencies might be required to affect 5hmC, 5fC and 5caC levels.

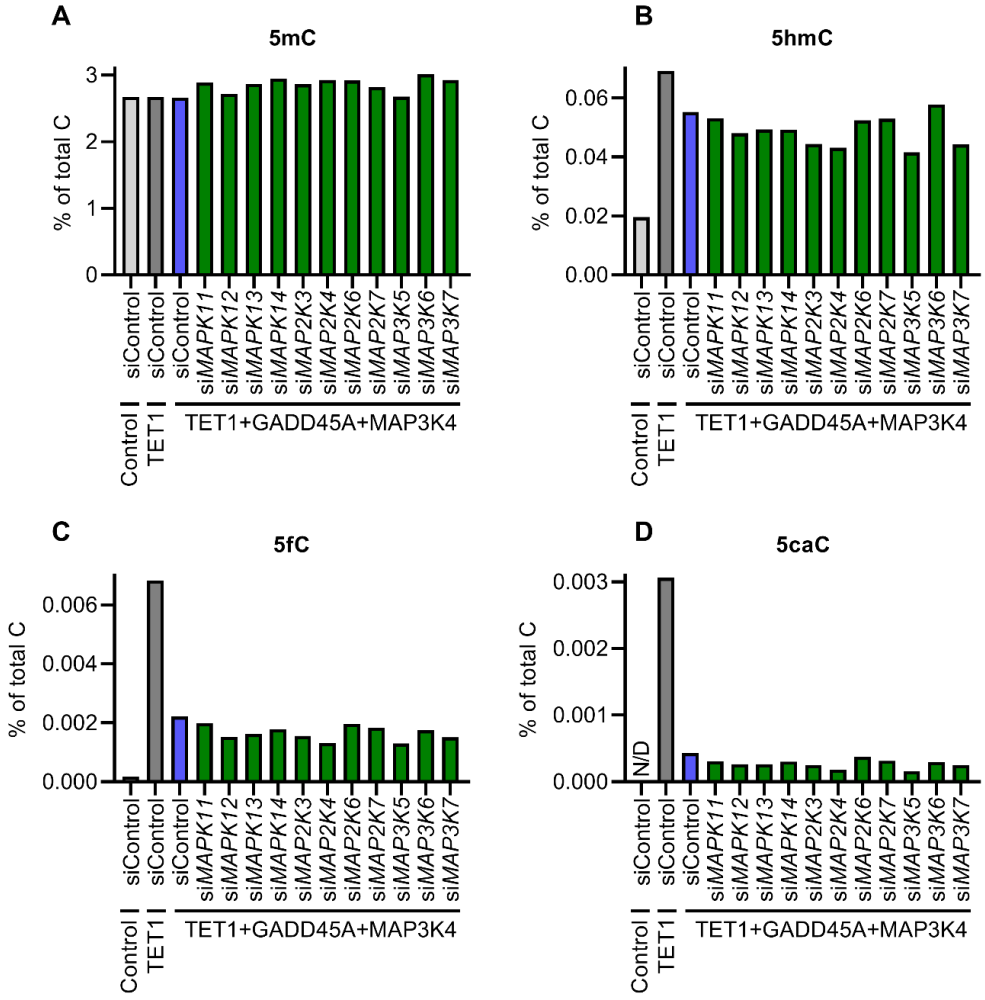

Figure 4.3 siRNA knockdown of selected MAP kinases does not rescue the MAP3K4-induced reduction of 5hmC, 5fC and 5caC levels
Quantitative mass spectrometry measurements of the levels of 5mC (A), 5hmC (B), 5fC (C) and 5caC (D) in HEK293T cells. Cells were transfected with control or specific siRNAs 24 h prior to DNA transfection with empty vector (Control), or a combination of TET1, GADD45A or MAP3K4 and harvested 72 h *post* siRNA transfection. N/D, non detectable. Mass spectrometry data by Dr. Michael Musheev.

As many MAP kinases are redundant, the knockdown of only one kinase might not have any effects on target proteins. Thus, I also performed combined knockdowns of selected MAP kinases. Depletion of the MAP kinases was again monitored via RT-qPCR, showing siRNA knockdown efficiencies between 40 and 60 %, with the exception of MAPK12, MAP2K6 and MAP3K6, for which the knockdown was not successful (data not shown). The overexpression of TET1, GADD45A and MAP3K4 was confirmed by Western blot analysis and similar to the single MAP kinase siRNA knockdowns, the combined knockdowns did not affect the expression levels of TET1, GADD45A and MAP3K4 (data not shown).

Considering 5mC oxidation, the combined knockdowns of MAP kinases also did not rescue the reduction of 5hmC, 5fC and 5caC levels by MAP3K4 and GADD45A (Figure 4.4 B-D, compare green bars to blue bars). This suggests that the tested MAP kinases are most likely not involved

in reducing 5hmC, 5fC and 5caC levels in a MAP3K4-dependent manner. Still, since the siRNA-mediated knockdowns did not lead to a complete depletion of the mRNA levels, it does not fully exclude a possible role of the tested kinases in TET1 regulation. Also, additional downstream MAP kinases could be involved, which were not tested in these experiments.

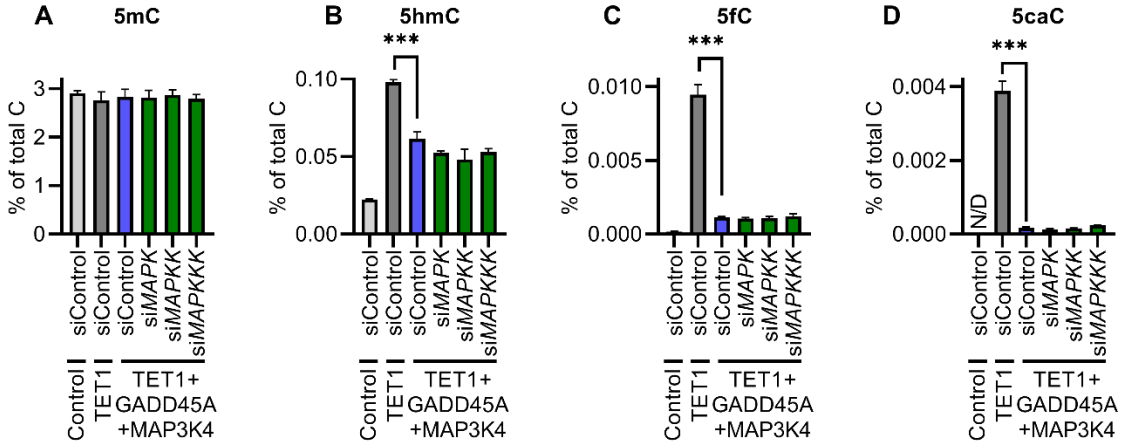

Figure 4.4 Combined siRNA knockdown of MAP kinases does not rescue the MAP3K4-induced reduction of 5hmC, 5fC and 5caC levels
Quantitative mass spectrometry measurements of the levels of 5mC (A), 5hmC (B), 5fC (C) and 5caC (D) in HEK293T cells. Cells were transfected with control or a combination of specific siRNAs targeting MAP kinases 24 h prior to DNA transfection with plasmids encoding the indicated proteins, and harvested 72 h *post* siRNA transfection. N/D, non detectable. Data shown as mean ± SD, n=3, *** = p-value < 0.0005, according to student's t-test. Mass spectrometry data by Dr. Michael Musheev.

4.1.2.3 The p38 inhibitor SB202190 reduces TET1 activity

An alternative approach to reducing the MAP kinase activity is using specific MAP kinase inhibitors. I used two different inhibitors, SB202190 and U0126: SB202190 inhibits MAPK11 (p38B) and MAPK14 (p38A) [161], and U0126 inhibits MAP2K1 (MEK1) and MAP2K2 (MEK2) [162]. Similar to previous siRNA knockdown experiments (see 4.1.2.2), TET1, MAP3K4 and GADD45A were overexpressed in HEK293T cells, this time with the addition of the inhibitor treatments.

The U0126 inhibitor treatment did not lead to any change in 5mC levels or 5hmC levels (Figure 4.5 A and B, compare dark blue bars to grey bars). Since the success of the U0126 inhibitor treatment was not tested, e.g. by determining the phosphorylation levels of downstream MAP2K1/MAP2K2 target proteins, it cannot be excluded that a higher concentration of the inhibitor would have been necessary to cause effects. However, other studies have shown that a concentration of 10 µM of the U0126 inhibitor depletes the phosphorylation of downstream MAP2K1/MAP2K2 proteins [163,164], which is a lower concentration than the 25 µM used in this study. Assuming that the U0126 inhibitor treatment was successful, this suggests that MAP2K1 and MAP2K2 are not involved in MAP3K4-mediated regulation of TET1.

The SB202190 inhibitor also did not affect 5mC levels. When only TET1 was overexpressed, SB202190 treatment slightly decreased 5hmC levels, although not significantly (Figure 4.5 B, compare light blue bar to grey bar). When GADD45A was co-expressed, however, 5hmC was significantly reduced ($p<0.0005$) (compare light blue bar to grey bar). Thus, SB202190 treatment blocked the increase of 5hmC levels by GADD45A. When MAP3K4 and TET1, or MAP3K4, GADD45A and TET1 were co-expressed, there was also a reduction of 5hmC by the SB202190-treatment (compare light blue bars to grey bars).

If MAP3K4 acted on TET1 via its downstream kinases MAPK11 (p38B) and MAPK14 (p38A), the use of the SB202190 inhibitor should block the effect of MAP3K4, i.e. increase 5hmC levels to the levels produced by TET1 overexpression alone. This, however, was not observed here. In contrast, the results suggest that MAPK14 and MAPK11 might enhance TET1 activity independently of MAP3K4. Therefore, GADD45A could regulate TET1 via MAP3K4 and MAPK11/14 via two independent pathways and with opposing effects.

MAP kinases can affect the activity of their target proteins via phosphorylation [113]. Thus, to further address MAP3K4-mediated TET1 regulation, experiments had been performed in our laboratory by Dr. Andrea Schäfer to determine phosphorylation sites of TET1 via quantitative proteomics, in collaboration with the research group of Dr. Petra Beli (IMB Mainz). GADD45A and constitutively active MAP3K4 enhanced the phosphorylation of three common serine residues of TET1, namely serine 322, serine 871 and serine 1851. Furthermore, the siRNA knockdown of MAP3K4 impaired GADD45A-enhanced TET1 phosphorylation at two of these residues, serine 322 and serine 871 (Dr. Andrea Schäfer, personal communication).

To investigate which downstream MAP kinases are involved in these phosphorylation changes, I decided to also use a quantitative proteomics approach: Stable Isotope Labeling with Amino acids in Cell culture (SILAC) was used to quantitatively compare TET1 phosphorylation sites from inhibitor-treated cells vs. DMSO-treated controls. The SB202190 inhibitor reduced phosphorylation of two TET1 phosphorylation sites, serine 116 and serine 871, to a SILAC ratio of 0.5, compared to control treatment (Figure 4.5. C, light blue bars). Interestingly, phosphorylation of serine 871 was enhanced by the overexpression of GADD45A or constitutively active MAP3K4 (Dr. Andrea Schäfer, personal communication), indicating that this phosphorylation site might play a role in GADD45A-MAP3K4-mediated regulation of TET1.

The U0126 inhibitor reduced the phosphorylation of serine 45 to a SILAC ratio of 0.6, and enhanced serine 1303 phosphorylation to a ratio of 1.5. However, the latter was only enhanced in one of two replicates (Figure 4.5. C, dark blue bars) and these phosphorylation sites were not affected by GADD45A depletion or overexpression (Dr. Andrea Schäfer, personal communication).

To further elucidate the role of phosphorylation in regulating TET1 activity, TET1 mutants lacking one, two or three phosphorylation sites were generated by mutating serines to alanines. These

included mutations of serines 116 and 871; however, the TET1 mutants did not affect TET1 enzymatic activity (Dr. Andrea Schäfer, personal communication).

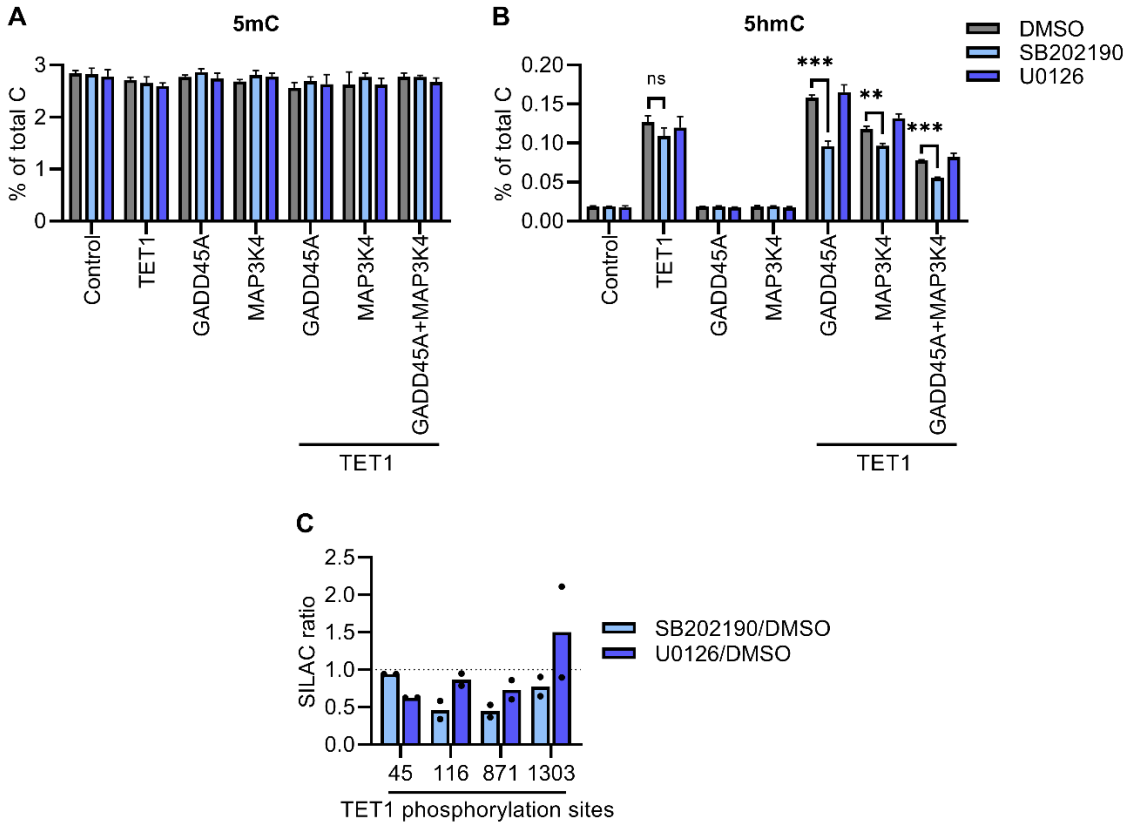

Figure 4.5 The p38-inhibitor SB202190 reduces TET1 activity
(A and B) Quantitative mass spectrometry measurements of the levels of 5mC (A), and 5hmC (B) in HEK293T cells, which were transiently transfected for 48 h with plasmids encoding the indicated proteins or empty vector (Control). At the time of transfection, cells were treated with DMSO, 25 µM SB202190 (MAPK11/MAPK14-inhibitor) or 25 µM U0126 (MAP2K1/MAP2K2-inhibitor). Data shown as mean ± SD, n=3, ns = non significant, ** = p-value < 0.005, *** = p-value < 0.0005, according to student's t-test. Mass spectrometry data by Dr. Michael Musheev. (C) TET1 phosphorylation site changes upon SB202190 or U0126 treatment in relation to control (DMSO) treatment, identified and quantified by SILAC and phosphoproteomic analysis. Experiment was performed in duplicates, bars display the mean of the two experiments and individual data points are shown. Phosphoproteomic analysis by Andrea Voigt and Dr. Petra Beli.

4.1.2.4 GADD45A and MAP3K4 reduce TET1 activity

Since the cooperation of GADD45A and MAP3K4 in the regulation of TET1 activity can be interpreted in two ways: 1) impairing TET1 activity and 2) enhancing TET1 activity, an additional method was required to resolve the dynamics of 5mC oxidation.

I performed metabolic labeling of HEK293T cells followed by mass spectrometry to differentiate between newly produced and previously formed 5mC oxidation products. To label 5mC and its derivatives in cells, "labeled" L-Methionine-(methyl-$^{13}C,D_3$) can be added to the cell culture medium. This is then metabolized to labeled methionine-derived SAM within cells, which in turn is the substrate for DNMTs, therefore generating labeled 5mC [4].

HEK293T cells were "labeled" for 14 days with L-Methionine-(methyl-$^{13}C,D_3$), achieving a labeling efficiency of 97 % (data not shown). Then, cells were transfected with TET1, GADD45A or MAP3K4. At the time of transfection, the label was withdrawn, i.e., the "labeled" medium was replaced with medium containing "nonlabeled" L-Methionine. Cells were harvested after different time points and the ratios of "nonlabeled", i.e. newly produced 5mC, 5hmC, 5fC, and 5caC vs. "labeled", i.e. old 5mC, 5hmC, 5fC and 5caC, were determined via quantitative mass spectrometry.

In general, the ratios of nonlabeled vs. labeled 5mC, 5hmC, 5fC and 5caC increased over time, as 5mC and its oxidative derivatives were newly generated (Figure 4.6). The ratios of nonlabeled vs. labeled 5hmC did not change upon the co-expression of GADD45A and MAP3K4, compared to TET1 overexpression alone (Figure 4.6 B, compare green bars to dark grey bars). GADD45A without MAP3K4, however, slightly increased the ratios of nonlabeled/labeled 5hmC, although not significantly (Figure 4.6 B, compare blue bars to dark grey bars). If GADD45A and MAP3K4 indeed enhanced TET1 enzymatic activity, as would have been one possible conclusion from the experiments depicted in Figure 4.2, one would expect an increase of nonlabeled 5hmC and a strong decrease in labeled 5hmC, resulting in an increased ratio of nonlabeled/labeled 5hmC. Since this was not the case, it indicates that active MAP3K4 does not increase the turnover of 5hmC, but rather reduces TET1 activity.

Unexpectedly, the ratios of nonlabeled vs. labeled 5fC decreased when TET1 was overexpressed (Figure 4.6 C, compare dark grey bars to light grey bars). One possible explanation is that TET1 preferentially oxidizes the parental, i.e. labeled 5mC, to 5hmC and further to 5fC. In addition, in order to produce nonlabeled 5fC, nonlabeled 5mC and 5hmC have to be generated first, which would explain the decrease of the nonlabeled vs. labeled ratio. When GADD45A and MAP3K4 were co-expressed, the nonlabeled vs. labeled ratios of 5fC significantly increased compared to TET1 overexpression alone ($p<0.05$) (Figure 4.6 C, compare blue and green bars to dark grey bars). Still, the increase of the ratios was only modest. For 5caC, the results were similar, although here the increase of the ratio of nonlabeled vs. labeled 5caC by the co-expression of GADD45A and MAP3K4 ($p<0.0005$) was more pronounced than for 5fC (Figure 4.6 D).

The results for 5fC and 5caC suggest that GADD45A alone, and GADD45A in combination with MAP3K4 indeed increase the turnover of these modifications. However, since GADD45A and MAP3K4 do not increase the turnover of 5hmC, GADD45A and MAP3K4 in combination do not generally enhance TET1 activity. The increased 5fC and 5caC ratios might rather be due to an enhanced removal by TDG, perhaps because TDG is recruited preferentially to the parental, i.e. labeled 5fC and 5caC by GADD45A and TET1. Considering all the data, I conclude that GADD45A and MAP3K4 together reduce the activity of TET1. Consequently, MAP3K4 does not play a role in enhancing TET1 activity in a GADD45A-dependent manner.

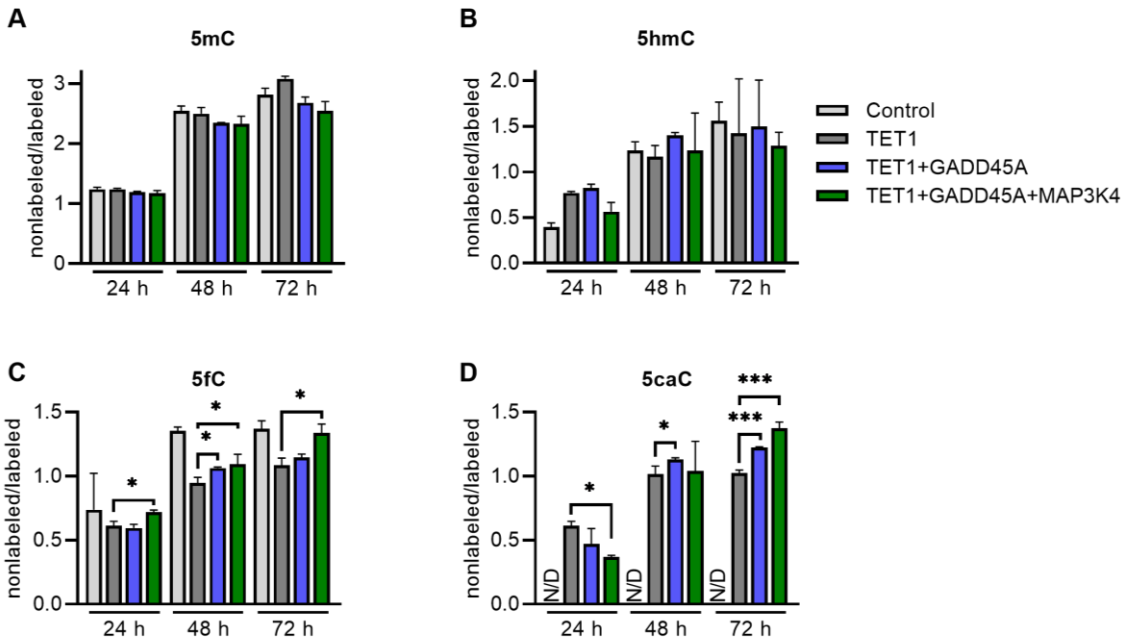

Figure 4.6 GADD45A and MAP3K4 do not change the ratio of nonlabeled/labeled 5hmC
(A-D) Quantitative mass spectrometry measurements of genomic DNA in HEK293T cells, which were transiently transfected for 24 h, 48 h, or 72 h with plasmids encoding the indicated proteins. Cells were incubated with "labeled" L-Methionine-(methyl-$^{13}C,D_3$) for 14 days and medium was replaced with "nonlabeled" L-Methionine at the time of transfection. Displayed are the ratios of "nonlabeled" 5mC vs. "labeled" $^{13}C,D_3$-5mC (A), "nonlabeled" 5hmC vs. "labeled" $^{13}C,D_2$-5hmC (B), "nonlabeled" 5fC vs. "labeled" $^{13}C,D_1$-5fC (C), and "nonlabeled" 5caC vs. "labeled" ^{13}C-5caC (D). Data shown as mean ± SD, n=3, * = p-value < 0.05, *** = p-value < 0.0005, according to student's t-test. N/D, non detectable. Mass spectrometry data by Dr. Michael Musheev.

4.1.3 The role of GADD45A in TET1 ubiquitination

4.1.3.1 GADD45A decreases TET1 ubiquitination

As TET1 is known to be ubiquitinated and this promotes DNA binding [94], GADD45A could also regulate TET1 activity via affecting TET1 ubiquitination. In order to investigate if GADD45A affects the ubiquitination of TET1, I performed Co-Immunoprecipitations (Co-IPs) of Ubiquitin and TET1 in the presence or absence of GADD45A. The Co-IPs were done with overexpressed proteins in HEK293T cells, using an antibody against HA-tagged Ubiquitin to co-immunoprecipitate FLAG-tagged TET1.

Ubiquitinated TET1 could be successfully co-immunoprecipitated with the HA (Ubiquitin)-antibody, as a smear in the IP fraction was detected for FLAG-TET1 (Figure 4.7 A, “TET1+UB”). Interestingly, GADD45A co-expression reduced the ubiquitination of TET1, while GFP co-expression, serving as a control, had no effect on the ubiquitination (compare lane “TET1+GADD45A+UB” to lane “TET+GFP+UB”).

Using another experimental approach to validate this finding, GFP-tagged TET1 was affinity purified with GFP-trap beads, allowing high salt washes under denaturing conditions to reduce unspecific background. Here, Ubiquitin was not ectopically expressed, in contrast to the experiment described above. In line with the previous results, GADD45A reduced the ubiquitination of TET1 in the IP fraction (Figure 4.7 B, compare lane “GFP-TET1+GADD45A” to lane “GFP-TET1”). As monoubiquitination of TET1 K1589 enhances TET1 activity [94] and GADD45A also enhances TET1 activity, the GADD45A-mediated decrease of TET1 ubiquitination likely has a different role. Perhaps, GADD45A has a stabilizing role on TET1, by preventing its ubiquitination and subsequent proteasomal degradation.

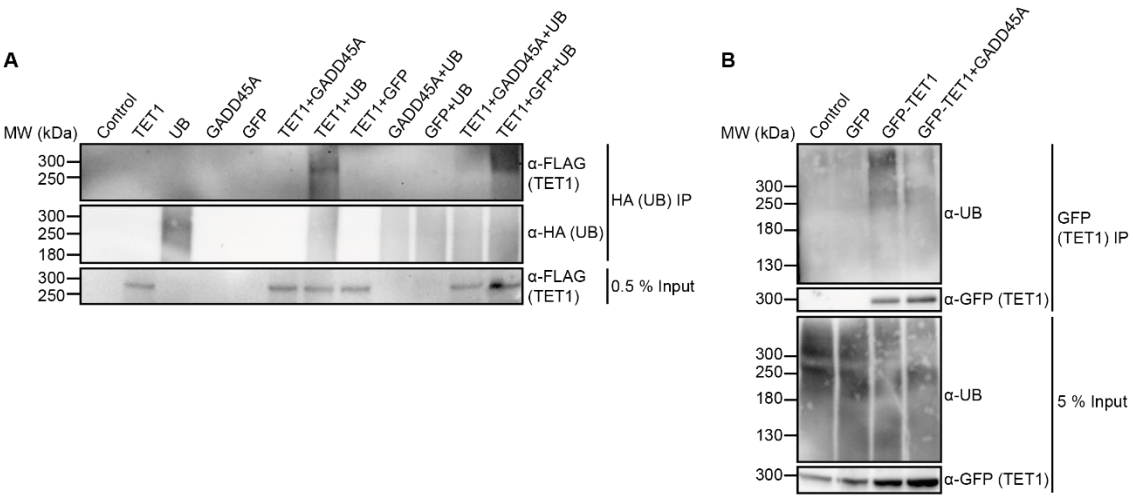

Figure 4.7 GADD45A decreases TET1 ubiquitination
(A) HA-affinity purification of whole cell extracts from HEK293T cells that were transiently transfected for 48 h with HA-Ubiquitin (UB), FLAG-TET1, GADD45A, GFP or a combination thereof. Proteins were separated via SDS-PAGE, subjected to Western blotting and detected using antibodies indicated on the right. (B) GFP-affinity purification of whole cell extracts from HEK293T cells that were transiently transfected for 48 h with GFP-TET1, GADD45A or GFP using GFP-Trap Agarose. MW, molecular weight.

To identify GADD45A-dependent TET1 ubiquitination sites, affinity purification of TET1 followed by quantitative mass spectrometry was performed in collaboration with Matthias Ostermaier and Dr. Petra Beli. However, several experimental optimization approaches with different digestion enzymes and using both mouse or human TET1 only yielded a 50 % coverage of the TET1 protein sequence. Only one GADD45A-dependent ubiquitination site, K328 in mouse TET1, was discovered when GADD45A was depleted (data not shown). Unfortunately, this site could not be reproduced in another pulldown. Mainly because of the low sequence coverage, we decided to not further explore TET1 ubiquitination via quantitative mass spectrometry.

4.1.3.2 GADD45A increases TET1 protein stability

To test if GADD45A stabilizes TET1 by preventing TET1 ubiquitination and subsequent degradation, I performed cycloheximide (CHX) chase assays of ectopic and endogenous TET1 in human HEK293T and mESCs, respectively, in the presence and absence of GADD45A. CHX inhibits protein biosynthesis by blocking translational elongation and can therefore be used to monitor the stability of a protein over time [165].

In HEK293T cells, 6 h of CHX treatment reduced ectopic TET1 protein levels to 20 % (Figure 4.8 A and B, “Control”). When GADD45A was co-expressed, no reduction of TET1 protein levels was observed after 6 h (Figure 4.8 A and B, “GADD45A”), suggesting that GADD45A indeed stabilizes TET1 protein levels.

To exclude any unspecific effects due to the overexpression of TET1 in HEK293T cells, I repeated the CHX assay in a knock-in CRISPR/Cas9 mESC line, which contains an HA-tag at the endogenous TET1 locus, enabling the detection of TET1 protein with an HA antibody [166]. Indeed, GADD45A overexpression in mESCs reduced the depletion of endogenous TET1 protein levels to 42 % after 12 h, compared to 4 % remaining protein levels without GADD45A co-expression (Figure 4.8 C and D, compare “GADD45A”).

To test whether GADD45A prevents the proteasomal degradation of TET1, the proteasome was inhibited with MG132. Surprisingly, MG132 treatment did not prevent degradation of TET1 in the CHX-chase assay (data not shown). Since CALPAIN1, a calcium-dependent protease, was reported to target TET1 for degradation [46], I tested if the Calpain inhibitor Calpeptin would stabilize TET1 protein levels similarly to GADD45A. Indeed, when Calpeptin was added to the HEK293T cells, TET1 protein levels were only reduced to 70 % after 6 h, compared to 20 % without Calpeptin (Figure 4.8 A and B, compare “CHX+Calpeptin” to “CHX”). Importantly, GADD45A did not further stabilize TET1 protein levels in the presence of Calpeptin. This was only the case for HEK293T cells, however, as in mESCs this effect of Calpeptin was not observed (data not shown). In conclusion, GADD45A stabilizes TET1 protein levels independently of the proteasome. This could be one mechanism of how GADD45A enhances TET1 activity.

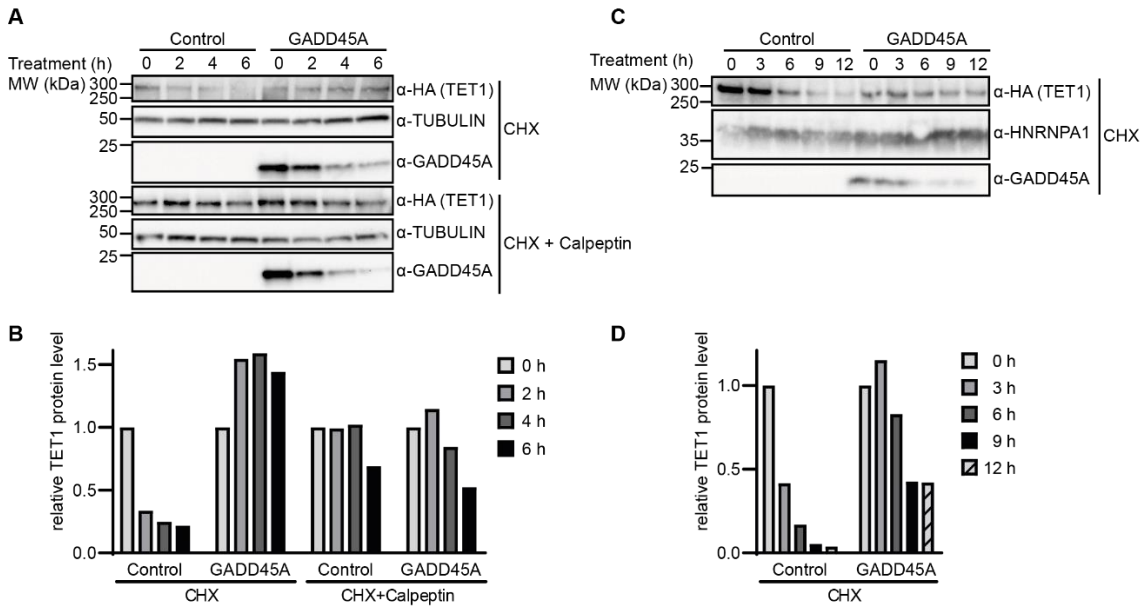

Figure 4.8 GADD45A increases TET1 protein stability

(A) CHX chase assay in HEK293T cells that were transiently transfected for 48 h with HA-TET1, GADD45A or empty vector (Control) and treated with 100 µg/ml CHX or 100 µg/ml CHX + 20 µM Calpeptin for the indicated times. Whole cell protein lysates were harvested, separated via SDS-PAGE, subjected to Western blotting and detected using antibodies indicated on the right. MW, molecular weight. (B) Quantification of TET1 protein levels in (A), normalized to TUBULIN as loading control. (C) CHX chase assay in HA-TET1 mESCs that were transiently transfected for 48 h with GADD45A or empty vector (Control) and treated with 50 µg/ml CHX for the indicated times. Nuclear protein lysates were harvested, separated via SDS-PAGE, subjected to Western blotting and detected using antibodies indicated on the right. Representative images of 2 independent experiments are shown. (D) Quantification of TET1 protein levels in (C), normalized to HNRNPA1 as loading control.

4.1.4 The role of GADD45A in TET1 SUMOylation

4.1.4.1 GADD45A affects TET1 SUMOylation

As SUMOylation is a well-known post-translational modification, which regulates protein activity [167], GADD45A-dependent SUMOylation of TET1 could be another mechanism of how GADD45A regulates TET1.

To explore if TET1 is SUMOylated, Co-IPs of HA-tagged SUMO and FLAG-tagged TET1 were performed in the presence or absence of GADD45A. Indeed, when HA-SUMO2 and FLAG-TET1 were overexpressed in HEK293T cells, the HA-antibody could immunoprecipitate FLAG-TET1, indicating that TET1 is indeed SUMOylated (Figure 4.9 A, “TET1+SUMO2”). GADD45A reduced the SUMOylation of TET1 (Figure 4.9 A, compare lane “TET1+GADD45A+SUMO2” to lane “TET1+GFP+SUMO2”).

To corroborate these findings with endogenous SUMO, GFP-tagged TET1 was affinity purified with GFP-trap beads, and endogenous SUMO was detected by Western blotting. The GFP-affinity purification confirmed the results obtained with ectopic SUMO (data not shown). To test whether the SUMOylation lies within the catalytic domain of TET1, I overexpressed a GFP-tagged TET1CD, and performed a GFP-pulldown.

Whereas GADD45A reduced the SUMOylation of full-length TET1, GADD45A increased the SUMOylation of TET1CD in the IP fraction (Figure 4.9 B). This difference seems peculiar at first, but it could be possible that decreased SUMOylation of the full-length protein masks the increased SUMOylation of the catalytic domain. As GADD45A can also enhance the activity of the isolated catalytic domain of TET1 [104], the functional interplay between GADD45A and the catalytic domain of TET1 will be further explored in the following chapter.

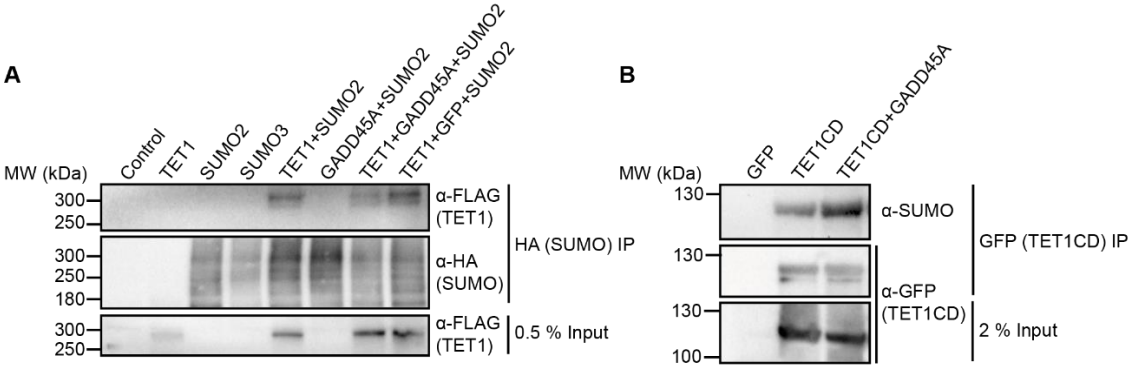

Figure 4.9 GADD45A affects TET1 SUMOylation
(A) HA-affinity purification of whole cell extracts from HEK293T cells that were transiently transfected for 48 h with FLAG-TET1, HA-SUMO2/3, GADD45A, GFP, empty vector (Control) or a combination thereof. Proteins were separated via SDS-PAGE, subjected to Western blotting and detected using antibodies indicated on the right. (B) GFP-affinity purification of whole cell extracts from HEK293T cells that were transiently transfected for 48 h with GFP-TET1CD, GADD45A or GFP using GFP-Trap Agarose. MW, molecular weight.

4.1.4.2 Predicted TET1 SUMOylation-deficient single mutants do not affect the functional cooperativity with GADD45A

To further explore the effect of SUMOylation on TET1 activity, SUMOylation sites in the catalytic domain of TET1 were predicted using the SUMOplot™ tool. This tool predicts SUMOylation sites according to a match with the SUMO-consensus sequence BKxD/E, where B is a hydrophobic residue, K is the lysine conjugated to SUMO, x is any amino acid, and D/E stands for aspartic acid or glutamic acid. Four potential SUMOylation sites were found in the catalytic domain of human TET1 that were conserved in mouse and rat: K1813, K1842, K1863, and K2067. The conservation gives more confidence that the predicted SUMOylation sites indeed play a role in TET1 protein regulation.

Using site-directed mutagenesis, I generated lysine to alanine and lysine to arginine mutants of the predicted and conserved SUMOylation sites in the catalytic domain of TET1, and tested the activity of the mutants, first by quantitative mass spectrometry. However, none of the mutations had an impact on global 5mC oxidation in HEK293T cells upon TET1CD overexpression, as the mutants produced similar levels of 5hmC, 5fC and 5caC as wild-type TET1CD (Figure 4.10 B-D, grey bars). Moreover, the ability of GADD45A to enhance TET1CD-mediated 5hmC, 5fC and 5caC production was not affected in the mutants (Figure 4.10 B-D, compare blue bars to grey bars). The only exception was a lacking increase of 5caC by GADD45A in the TET1 K2067R mutant, although there was still a slight trend.

As a more direct assay to monitor the activity of the TET1CD mutants, I employed an *in vitro* TET1 oxidation assay, where I tested the effect of a K2067R mutation and the effect of GADD45A overexpression on 5mC oxidation. I FLAG-affinity purified wild-type FLAG-TET1CD or mutant FLAG-TET1CD from HEK293F suspension cells that had been transfected with expression constructs for these proteins. Subsequently, the *in vitro* TET1 oxidation assay was performed using an isotopically labeled methylated amplicon as a substrate, which allows the discrimination from contaminating genomic DNA that could have been immunoprecipitated with the proteins. TET1 oxidation products originating from the labeled methylated amplicon – 5hmC+3, 5fC+3 and 5caC+3 – were then measured by quantitative mass spectrometry (where +3 indicates the increased molecular weight of the isotopes).

In the *in vitro* TET1 oxidation assay, the addition of wild-type TET1CD decreased 5mC levels of the labeled methylated amplicon (originally 48 % 5mC+3) to 28 % 5mC+3 after 24 min (Figure 4.10 E, “wt+GFP”). At the same time, 5hmC+3, 5fC+3 and 5caC+3 were produced (Figure 4.10 F-H). When GADD45A was added to the reaction, production of 5hmC+3, 5fC+3 and 5caC+3 slightly increased (compare green curves to black curves).

Similar to what was observed for global 5mC oxidation (Figure 4.10 B-D), the K2067R mutant did not affect TET1CD activity in the *in vitro* TET1 oxidation assay (compare grey curves to black curves), and more importantly, the slight increase in 5mC oxidation by GADD45A was still observed (Figure 4.10 F-H compare blue curves to grey curves).

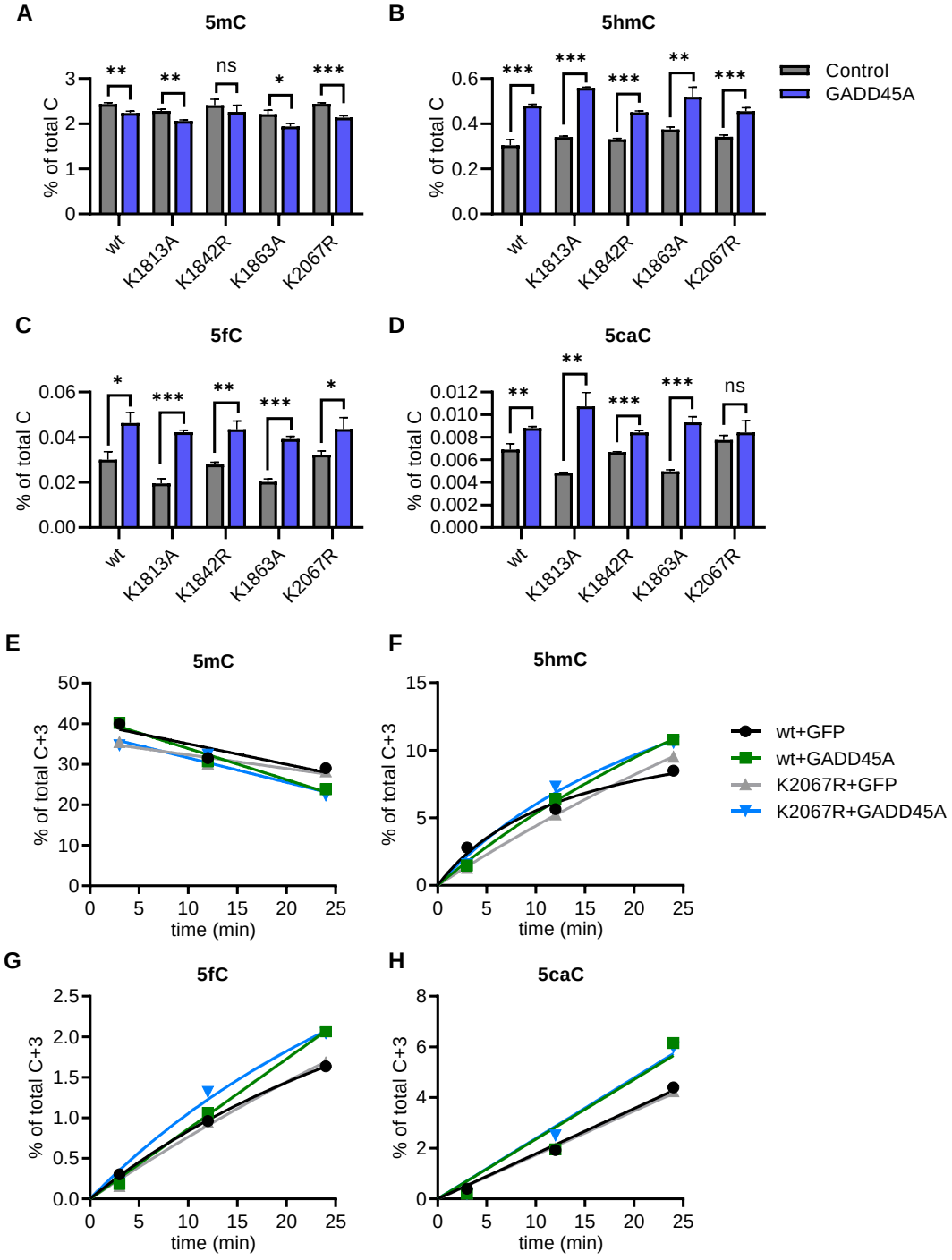

Figure 4.10 TET1 SUMOylation-deficient mutants do not affect the functional cooperativity with GADD45A

(A-D) Quantitative mass spectrometry measurements of the levels of 5mC (A), 5hmC (B), 5fC (C) and 5caC (D) in HEK293T cells. Cells were transfected with wild-type (wt) or mutant (K1813A, K1842R, K1863A, K2067R) TET1CD, in addition to empty vector (Control) or GADD45A, and harvested after 48 h. Data shown as mean ± SD, n=3, ns = non significant, * = p-value < 0.05, ** = p-value < 0.005, *** = p-value < 0.0005, according to student's t-test. (E-H) *In vitro* TET1 oxidation assay. HEK293F cells were transfected with wild-type (wt) or mutant (K2067R) FLAG-TET1CD, in addition to GFP (Control) or GADD45A, and harvested after 48 h. FLAG-TET1CD was affinity purified and the *in vitro* TET1 oxidation assay was performed using an isotopically labeled methylated amplicon as a substrate. DNA was purified and subjected to quantitative mass spectrometry. Data shown as mean, n=2. Mass spectrometry data by Dr. Michael Musheev.

In this study, only single lysine mutants were generated and double or multiple lysine to arginine mutants might be necessary to completely abolish SUMOylation of the catalytic domain of TET1. Moreover, additional lysines, which were not selected for mutation, might play a role in GADD45A-mediated TET1 regulation. Still, I could not find any strong evidence that GADD45A enhances TET1 activity via SUMOylation.

4.1.4.3 GADD45A enhances the SUMOylation of TDG

GADD45A can not only interact with and enhance TET1 activity in active DNA demethylation. In fact, Li and colleagues showed that GADD45A also physically interacts with TDG and thereby enhances 5fC and 5caC removal [111]. Since it was also shown that TDG can be SUMOylated and this increases its substrate turnover [168], GADD45A could enhance TDG activity, and thereby 5fC and 5caC removal, by promoting SUMOylation of TDG. In this mechanism, GADD45A could act as the SUMO donor, as GADD45A itself can be SUMOylated [169].

To explore if the SUMOylation of TDG is affected by GADD45A, I performed Co-IPs of HA-tagged SUMO and FLAG-tagged TDG with or without GADD45A overexpression. When SUMO and TDG were co-expressed, an additional protein band appeared for TDG in the input as well as in the IP fraction, which represents SUMOylated TDG (Figure 4.11, compare lanes “SUMO” to “Control”). GADD45A strongly enhanced the SUMOylation of TDG, and even a third band appeared in the IP fraction, probably representing the conjugation of two SUMO molecules to TDG (Figure 4.11, compare lanes “SUMO+GADD45A” to “SUMO+GFP”). Therefore, GADD45A enhances the SUMOylation of TDG and this could be the mechanism by which GADD45A enhances 5fC and 5caC removal.

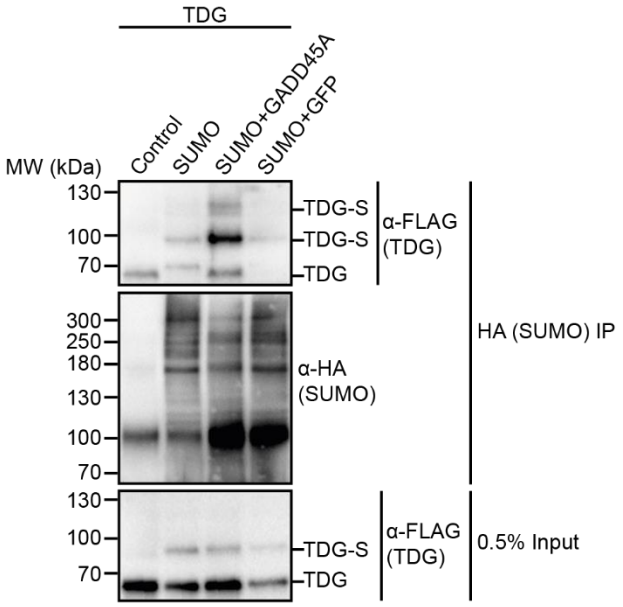

Figure 4.11 GADD45A enhances the SUMOylation of TDG
HA-affinity purification of whole cell extracts from HEK293T cells that were transiently transfected for 48 h with FLAG-TDG, HA-SUMO2, GADD45A, GFP or combination thereof. Proteins were separated via SDS-PAGE and subjected to Western blotting and detected using antibodies indicated on the right. TDG-S, SUMOylated TDG. MW, molecular weight.

4.1.5 GADD45A and the TET1 interactome

In order to identify possible GADD45A-dependent TET1 interactors involved in ubiquitination or SUMOylation, TET1 was affinity purified in the presence or absence of GADD45A and subjected to quantitative mass spectrometry (in collaboration with Dr. Petra Beli). Unfortunately, four independent replicates only had a small overlap (only 18 common TET1 interactors), and GADD45A-dependet interactors were even more variable between replicates. Some of these few top candidates, which were also found in previous pulldowns performed by Dr. Andrea Schäfer (personal communication), were proteins involved in post-translational modifications, e.g. E3 ligases (Table 4.1). The candidates included members of the superfamily of tripartite motif-containing (TRIM) E3 ligases. These proteins all contain a RING domain, which enables them to conjugate Ubiquitin, SUMO, or Ubiquitin-like protein ISG15 (ISG15) to target proteins [170]. The interaction of GADD45A, TET1 and members of the TRIM family will be explored in the following chapter.

Protein	Function
TRIM24	E3 Ubiquitin Ligase
TRIM25	E3 Ubiquitin/ISG15 Ligase
TRIM28	E3 Ubiquitin/SUMO Ligase
TRIM33	E3 Ubiquitin Ligase
HUWE1	E3 Ubiquitin Ligase
STUB1	E3 Ubiquitin Ligase

Table 4.1 Selection of candidate TET1 interactors identified via quantitative mass spectrometry

4.1.6 Interplay between TET1, GADD45A and the E3 Ligases TRIM25 and TRIM28

4.1.6.1 TRIM25 and TRIM28 knockdown decreases TET1 activity

In chapters 4.1.3 and 4.1.4, I showed that GADD45A affects the ubiquitination and SUMOylation of TET1. Since GADD45A lacks any known enzymatic activity, additional co-factors are needed to affect these post-translational modifications of TET1. In this chapter, I will explore whether potential TET1 interactors with an E3 ligase activity, which were identified via quantitative mass spectrometry (Table 4.1), play a role in GADD45A-dependent regulation of TET1.

First, the effect of the potential interactors on global 5mC oxidation was tested in loss of function experiments in HEK293T cells, followed by quantitative mass spectrometry. Depletion of selected TRIM family members was monitored via RT-qPCR, showing siRNA knockdown efficiencies between 40 and 70 % (data not shown). The overexpression of TET1 and GADD45A was confirmed by Western blot analysis and expression levels were not affected by the knockdown of the TRIM family members (data not shown).

Depletion of GADD45A via siRNA on a background of TET1 overexpression significantly decreased 5hmC levels compared to siControl-knockdown ($p<0.05$) (Figure 4.12 B, grey bars). The siRNA knockdowns of TRIM25 and TRIM28 decreased 5hmC levels even further ($p<0.05$ and $p<0.005$, respectively). In contrast, the depletion of TRIM24 and TRIM33, which together with TRIM28 form the Tif1 protein family, did not affect 5hmC levels.

When GADD45A was co-expressed with TET1 and no specific siRNA was transfected (mock or siControl), 5hmC levels significantly increased ($p<0.05$), as previously observed. However, when siRNAs against TRIM25 and TRIM28 were transfected, this increase by GADD45A was not observed (Figure 4.12 B, compare blue bars to grey bars).

Regarding 5fC and 5caC, the effect of the GADD45A, TRIM25 and TRIM28 knockdowns were similar to the effects observed towards 5hmC levels: depletion of each of the three factors significantly decreased 5fC and 5caC levels, when compared to the non-specific siControl siRNA (Figure 4.12 C and D, grey bars). As for 5hmC, siRNA knockdowns of TRIM24 and TRIM33 had no effect.

GADD45A co-expression with TET1 reduced the levels of 5fC and 5caC under control conditions (mock and siControl) (Figure 4.12 C and D, compare blue bars to grey bars). However, a combined knockdown of TRIM25 or TRIM28 together with overexpression of GADD45A did not reduce levels further than either knockdown of TRIM25 or TRIM28, or overexpression of GADD45A alone (compare blue bars to grey bars). In contrast, TRIM24 and TRIM33 knockdown in combination with GADD45A overexpression reduced 5fC and 5caC levels, similarly as control treatments (mock and siControl). Of note, siRNA knockdowns of TRIM25 and TRIM28 also reduced global 5hmC, 5fC and 5caC levels produced by the catalytic domain of TET, in the same way as a GADD45A siRNA knockdown (data not shown). These results suggest that TRIM25 and TRIM28 might be involved in enhancing TET1 activity in a GADD45A-dependent manner.

In HEK293T cells, TET1 needs to be ectopically expressed in order to reach reliable and detectable 5fC and 5caC levels, which might introduce a bias due to variation in TET1 protein levels. Although the protein levels were always monitored via Western blot (data not shown) and the effects observed for the depletion of TRIM25 and TRIM28 cannot be simply explained by a reduction of TET1 protein levels, a system without TET1 overexpression is more reliable. Thus, I also performed knockdowns of the top candidates in mESCs. mESCs have higher endogenous TET1 expression levels compared to HEK293T and consequently, also higher 5hmC, 5fC and 5caC levels.

In mESCs, GADD45A depletion did not affect 5hmC levels, but like in HEK293T cells, slightly reduced 5fC and 5caC levels, although not significantly (Figure 4.12 F-H). TRIM25 and TRIM28 depletion also only mildly decreased 5fC and 5caC levels (non significant), but the trend was similar as in HEK293T cells.

Despite the observed discrepancies in HEK293T cells and mESCs, TRIM25 and TRIM28 might be promising candidates involved in GADD45A-mediated TET1 regulation and will be further explored in the following chapter.

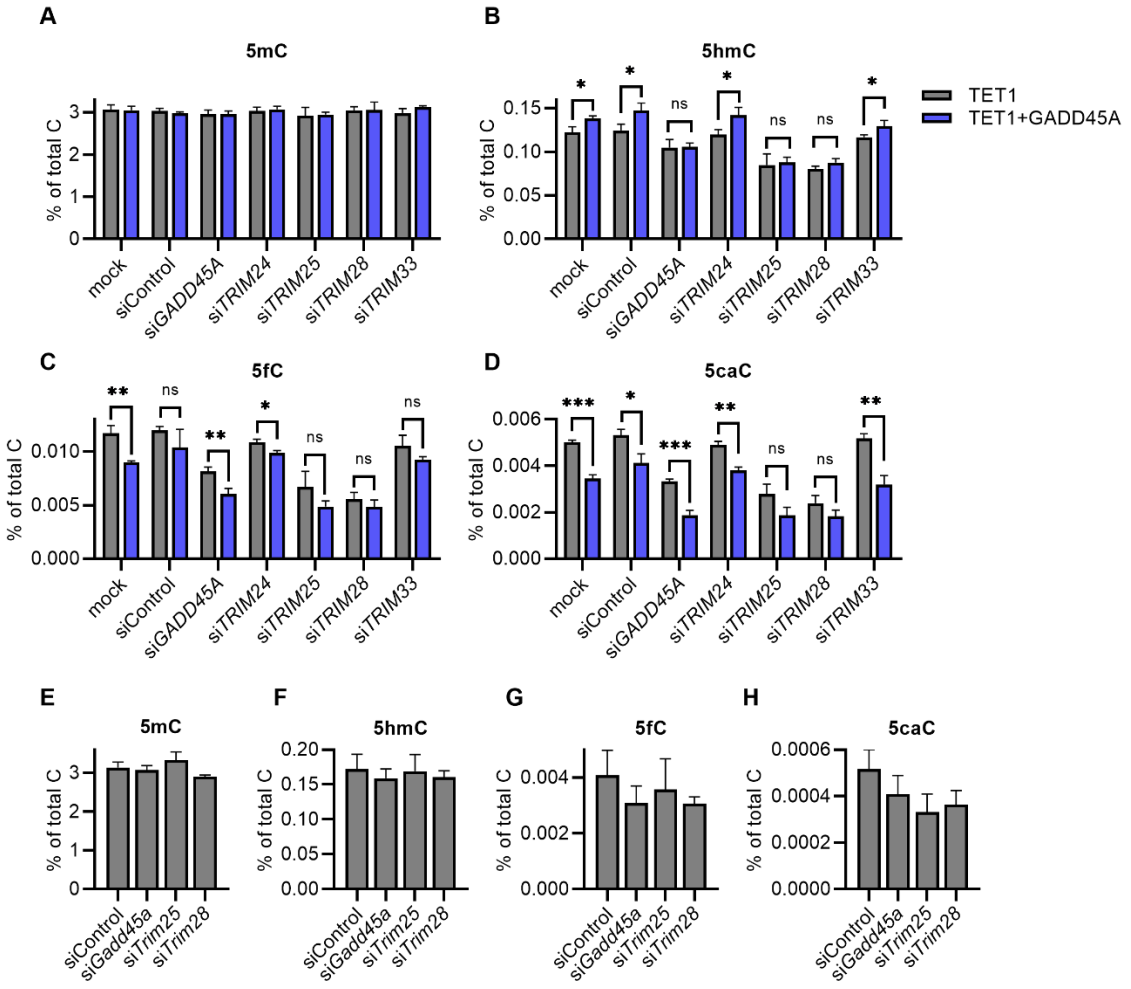

Figure 4.12 TRIM25 and TRIM28 knockdown decreases TET1 activity
(A-D) Quantitative mass spectrometry measurements of the levels of 5mC (A), 5hmC (B), 5fC (C) and 5caC (D) in HEK293T cells. Cells were transfected with control or specific siRNAs 24 h prior to DNA transfection with TET1 alone or TET1+GADD45A, and harvested after 72 h. In "mock" treated samples, cells were only treated with transfection reagent but no siRNA. (E-H) Quantitative mass spectrometry measurements of the levels of 5mC (E), 5hmC (F), 5fC (G) and 5caC (H) in mESCs. Cells were transfected with control or specific siRNAs and harvested after 48 h. All data are shown as mean ± SD, n=3, ns = non significant, * = p-value < 0.05, ** = p-value < 0.005, *** = p-value < 0.0005, according to student's t-test. Mass spectrometry data by Dr. Michael Musheev.

4.1.6.2 TRIM25 and TRIM28 knockdown does not affect TET1 activity *in vitro*

To corroborate the roles of TRIM25 and TRIM28 in regulating TET1 activity, I performed a more direct assay investigating the functional interaction between TRIM25/TRIM28 and TET1, an *in vitro* TET1 oxidation assay (see chapter 4.1.4.2). The FLAG-tagged catalytic domain of TET1 was transiently overexpressed in HEK293T cells and TRIM25 or TRIM28 were depleted using siRNAs.

FLAG-TET was purified from cell extracts and the *in vitro* TET1 oxidation assay was performed using an isotopically labeled methylated amplicon.

Addition of wild-type TET1CD sligthly decreased 5mC+3 levels of the isotopically labeled methylated amplicon purified from siControl-treated cells, whereas 5hmC+3 production increased over time, as expected (Figure 4.13 A and B, black curves). 5hmC production after 3 and 12 min was not affected by either, TRIM25 or TRIM28 knockdown (Figure 4.13 B, compare blue and green curves to black curve). After 24 min, 5hmC production seemed decreased by TRIM25 knockdown; however, Western blot data (not shown) indicated that this was probably due to decreased TET1CD protein levels in the samples. Hence, while the depletion of TRIM25 and TRIM28 decreased 5hmC levels on a global scale, the depletion of the two factors did not affect TET1CD oxidation activity in this *in vitro* TET1 oxidation assay.

These results suggest that TRIM25 and TRIM28 might indirectly, rather than directly, regulate TET1. For example, they could be involved in modifying and targeting TET1 to genomic loci, analogous to the ubiquitination by CRL4 that enhances TET1 binding to DNA [94]. Such effects might require a chromatin context, whereas these effects might not be observed in an *in vitro* assay using an amplicon as a TET1 substrate. Still, also off-target effects of the siRNAs cannot be excluded. In addition, overexpression of TRIM25 or TRIM28 did not affect TET1 activity, neither on a global scale nor in the *in vitro* TET1 oxidation assay (data not shown). Furthermore, an interaction between TRIM25 or TRIM28 and TET1 could not be confirmed via Co-IPs (data not shown). This makes it unlikely that TRIM25 or TRIM28 are directly involved in the regulation, and specifically the GADD45A-mediated regulation, of TET1.

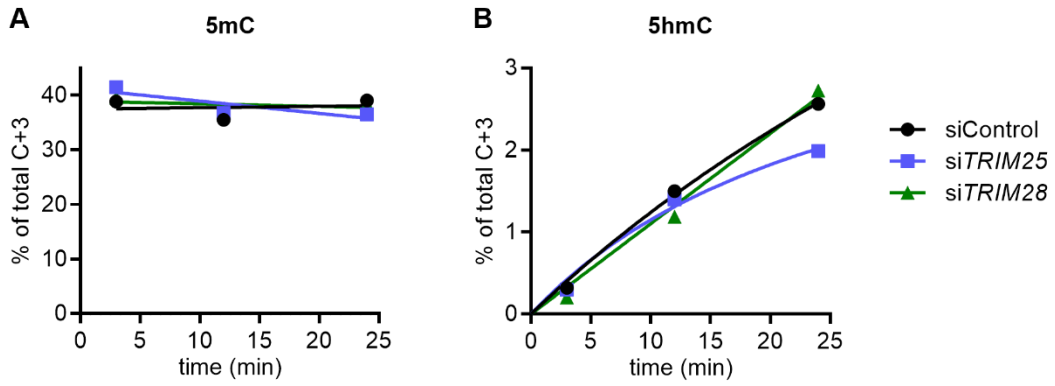

Figure 4.13 TRIM25 and TRIM28 knockdown does not affect TET1 activity *in vitro*
In vitro TET1 oxidation assay. HEK293T cells were transfected with control or specific siRNAs 24 h prior to DNA transfection with the catalytic domain of FLAG-TET1 and harvested after 72 h. FLAG-TETCD was affinity purified and the *in vitro* TET1 oxidation assay was performed using an isotopically labeled methylated amplicon as a substrate. DNA was purified and subjected to quantitative mass spectrometry to quantify 5mC (A) and 5hmC (B). Data shown as mean, n=2. Mass spectrometry data by Dr. Michael Musheev.

4.2 Linking TET activity to distal regulatory elements

A yet unresolved question in the demethylation field is why TET binding sites are mostly found at promoters, whereas the oxidation products of TET – 5hmC, 5fC and 5caC – are largely found at distal elements. One explanation could be that TET proteins are targeted to distal elements via promoter-enhancer looping [41]. In the following chapters, I will present how I tested this hypothesis by depleting promoter-enhancer loops and measuring if this leads to a loss of oxidized 5mC derivatives at enhancers. Promoter-enhancer loops can be abolished via depletion of RAD21, a core component of the Cohesin complex [132]. However, since RAD21 is an essential protein for cell division, as it is involved in sister chromatid cohesion [126], a system in which RAD21 can be depleted in a controlled manner was needed. Therefore, I used the Auxin-inducible degron (AID) system, which allows fast depletion of proteins, shortening the half-life of AID-tagged proteins to only 10-20 min [171]. As a model system, I chose mESCs, because they have high endogenous TET1 and TET2 expression levels, and thus, high 5hmC, 5fC and 5caC levels [41].

4.2.1 Auxin-Inducible Degron tagging of *Rad21* in mESCs

In order to establish a conditional RAD21 depletion system in mESCs, I first replaced the stop codons of both *Rad21* alleles with C-terminal AID-tags. Second, I stably introduced a transgene encoding the Tir1 F-box (OsTIR1) protein from Oryza sativa (rice). OsTIR1 can bind to AID in the presence of the plant hormone auxin and triggers a proteasome-dependent degradation of the AID-tagged protein [171].

RAD21-degron cell lines were generated by transfecting E14tg2a mESCs with three different plasmids: one plasmid containing the guide RNA (gRNA) (targeting the C-terminus of the *Rad21* gene) and the Cas9 endonuclease sequence, and two plasmids containing an *AID-3xHA*-tag and either a *hygromycin* or a *neomycin* resistance gene, flanked by *Rad21* homology arms. This allowed for antibiotic selection of cells containing an *AID-3xHA*-tag at both *Rad21* alleles (Figure 4.14 A, top). 144 single clones were picked and analyzed for homozygous *AID-3xHA* integration at the *Rad21* locus, using genotyping PCRs (data not shown). The genotyping PCRs revealed 37 double positive clones for an insertion of the *hygromycin* and the *neomycin* resistance genes at the *Rad21* locus. These clones were expanded, protein lysates were prepared, and lysates were subjected to Western blotting. A successful *AID-3xHA* insertion at the *Rad21* locus is expected to translate to an approximately 10 kDa larger protein compared to the wild-type RAD21 protein. The Western blot analysis revealed two clones (A1 and F5), which had homozygous *AID-3xHA* insertions, as indicated by a single 130 kDa protein band being detected using an HA-antibody, and by the detection of the same sized band using a RAD21-antibody (Figure 4.14 B, lanes 1 and 14). For four other clones, the HA-antibody also detected a band at this size (A3, F7, C7_2 and D10_2); however, the RAD21-antibody detected a second lower band. This suggests that the *AID-3xHA*-tag was only inserted at one of the two *Rad21* alleles, where the lower band corresponds to the endogenous, untagged RAD21 protein, and the higher band corresponds to the AID-3xHA-tagged RAD21 protein.

In a second transfection and selection round, *OsTir1* was integrated at the *Tigre* (tightly regulated) locus [172] in the two positive *Rad21-AID-3xHA* cell lines (Figure 4.14 A, bottom). The *Tigre* locus was chosen as it allows for a robust transgene expression in mESCs, and transgenes inserted at this locus were observed to have higher expression levels than if inserted at the *Rosa26* locus or randomly [147,173]. Thus, the two clones A1 and F5 were once more transfected with two plasmids: one containing a gRNA (targeting the *Tigre* locus to allow for insertion of the *OsTir1* transgene) and the Cas9 endonuclease sequence, and one containing a V5-tagged *OsTir1* expression cassette and a *puromycin* resistance gene flanked by arms homologous to the *Tigre* locus. Following puromycin selection, a total of 192 single clones were picked and analyzed for successful *OsTir1* integration at the *Tigre* locus, using genotyping PCRs (data not shown). The genotyping PCRs revealed 29 *OsTir1*-positive clones, which were further expanded to test for a depletion of RAD21 after auxin addition.

Acute RAD21 depletion was induced by treating cells with 500 µM of the auxin analog Indole-3-acetic acid sodium salt (referred to as “auxin” henceforth) for 6 h. Afterwards, protein lysates were subjected to Western blot analysis, using antibodies against HA and RAD21 to monitor the depletion of RAD21, and V5 to analyze the expression of OsTir1. This analysis revealed 10 conditional RAD21 depletion cell lines (data not shown). Of these, four clones with the most efficient RAD21 depletion upon auxin addition were selected and more thoroughly tested in time-course experiments.

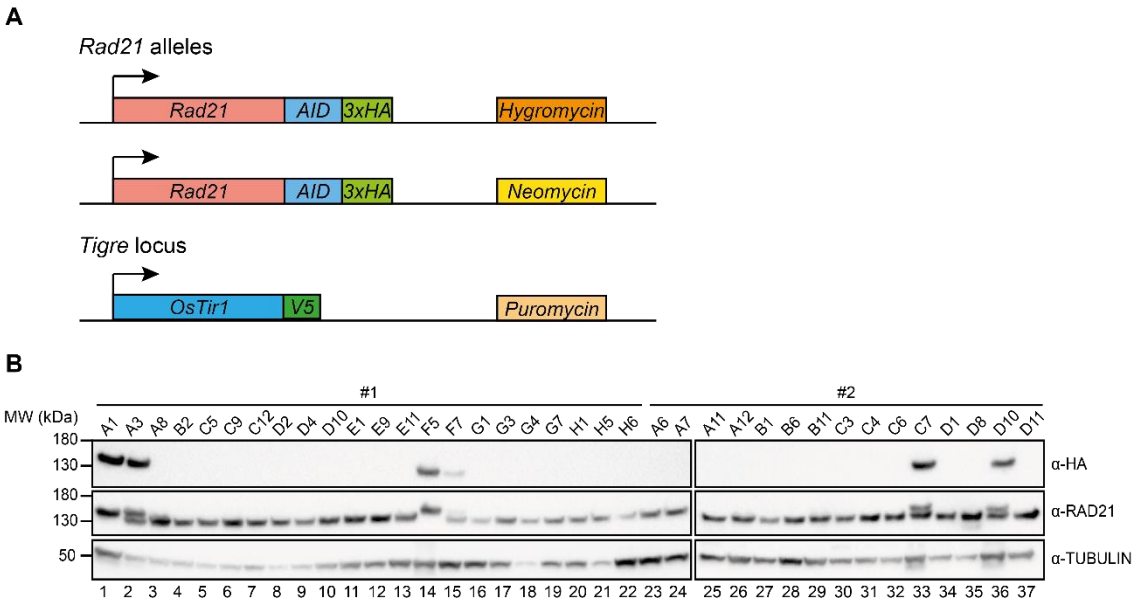

Figure 4.14 Auxin-Inducible Degron tagging of *Rad21* in mESCs
(A) Scheme showing the transgenic loci of the *Rad21-AID-3xHA* cell lines. (B) Western blot of whole cell protein lysates of the indicated mESC-clones, which originated from two different 96-well plates (#1 and #2), using antibodies indicated on the right. MW, molecular weight.

4.2.2 Acute Protein Depletion of RAD21 in mESCs

The dynamics of acute RAD21 depletion were monitored using auxin treatment intervals ranging from 0.5 - 24 h followed by Western blotting and immunofluorescence staining, using four different *Rad21-AID-3xHA* cell lines (A1_B4, A1_B6, F5_D4 and F5_F6). A short auxin treatment of 0.5 h already reduced RAD21 levels by more than 50 %, and 1 h was enough to almost completely reduce RAD21 levels (Figure 4.15 A, “α-HA”, and B, “RAD21”). After 2 h of treatment, cell morphology started to change compared to untreated cells, as the cell colonies became smaller and started detaching from the culture dish (data not shown). After 4 h of auxin treatment, higher numbers of mitotic cells were found compared to untreated cells (Figure 4.15 B, “DAPI”, indicated by arrows). This could be the result of mitotic arrest, due to premature sister-chromatid separation after RAD21 depletion [174]. These experiments show that interpretation of data acquired after more than 4 h of auxin treatment will be compromised by interference with cell cycle and chromosome condensation. Of note, the same results were observed for all four tested cell lines (data shown only for cell line F5_F6).

Next, quantitative mass spectrometry was performed to measure the levels of the oxidative derivatives of 5mC in the *Rad21-AID-Tir1* mESC-lines. Cell lines F5_D4 and F5_F6 (originating from clone F5 from the first transfection round) had higher levels of 5hmC, 5fC and 5caC than cell lines A1_B4 and A1_B6 (originating from clone A1 from the first transfection round) (Figure 4.15 C). To monitor if RAD21 depletion affects global levels of 5mC and its oxidative derivatives, genomic DNA was also isolated after auxin-mediated RAD21 depletion and similarly measured by quantitative mass spectrometry. After 24 h of RAD21 depletion, 5fC levels were almost reduced by half, suggesting that RAD21 depletion might affect 5fC globally (Figure 4.15 D). Yet, this might also be related to the mitotic arrest mentioned earlier. In fact, the enzyme responsible for the removal of 5fC, TDG, is controlled in a cell cycle dependent manner and is degraded in S-phase, but highly expressed in the G2-phase [175]. To analyze cell cycle differences in more detail, flow cytometry of propidium iodide stained cells could be performed for future experiments.

Cell lines F5_D4 and F5_F6 had comparable 5mC, 5hmC, 5fC and 5caC levels as wild-type mESCs (compare Figure 4.12 and Figure 4.14, also see [37]). Therefore, one of these cell lines containing higher levels of 5hmC, 5fC and 5caC – cell line F5_F6 – was selected for further downstream analyses.

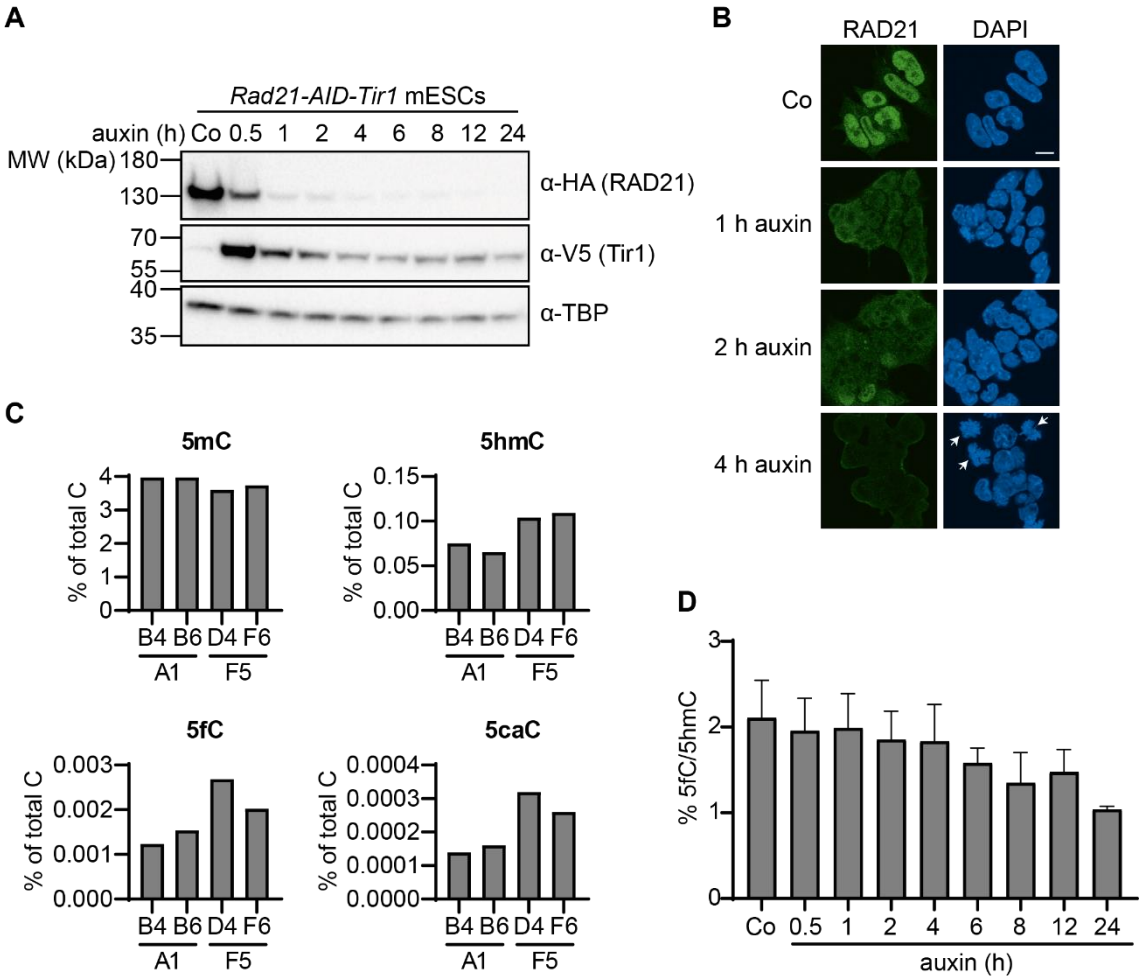

Figure 4.15 Acute Depletion of RAD21 using the AID System in mESCs
(A) Western blot of nuclear protein lysates from *Rad21-AID-Tir1*-mESC-line F5_F6, treated with 500 µM auxin for the indicated time, using antibodies indicated on the right. MW, molecular weight. (B) Immunofluorescence staining of cell line F5_F6 using a RAD21-antibody, DAPI was used to stain nuclei. The scale bar corresponds to 10 µM, arrows indicate mitotic cells. Experiment in (A) and (B) was performed with four different *Rad21-AID-Tir1*-mESC-lines (A1_B4, A1_B6, F5_D4 and F5_F6), representative data for cell line F5_F6 is shown. (C) Quantitative mass spectrometry measurements of the levels of 5mC, 5hmC, 5fC and 5caC in the indicated *Rad21-AID-Tir1*-mESC-lines. (D) Quantitative mass spectrometry measurements of 5fC (normalized to 5hmC) in the four different *Rad21-AID-Tir1*-mESC-lines, treated with 500 µM auxin for the indicated hours. Data shown as mean ± SD, n=4. Mass spectrometry data by Dr. Michael Musheev. Co, control.

4.2.2.1 Prolonged RAD21 depletion leads to a G2/M arrest and reduced 5fC accumulation

To investigate if loop depletion affects the oxidation of 5mC at enhancers, I performed 5hmC and 5fC DNA immunoprecipitation (DIP) followed by qPCR. Investigated regions and respective qPCR primers were bioinformatically selected and designed (in collaboration with Dr. Emil Karaulanov, IMB Mainz) according to the following criteria: The candidate regions for qPCR amplification overlap with annotated enhancers [176], are 5hmC and 5fC positive [81,177,178], and loop to a promoter with a TET1 binding site [166,179]. The promoters have low 5hmC and 5fC levels [81,177,178] and enhancers do not contain a TET1 binding site [166,179].

During the course of this study, Rhodes and colleagues published a RAD21-degron mESC-line similar to ours and showed that chromatin loops are depleted after 6 h of auxin treatment [150]. I decided to use the same time point for my experiments to ensure that also in my system, loops were completely depleted. This meant accepting the risk that some of the results I would obtain might result from the unintended impacts of RAD21 depletion on the cell cycle. A relatively short time of auxin-mediated RAD21 depletion would likely not affect already established 5hmC and 5fC, unless these modifications are quickly removed. Hence, the effects of loop depletion on TET activity would be expected to be larger for newly established 5hmC and 5fC. To increase the amount of newly established 5hmC and 5fC after RAD21 depletion, TET activity can be enhanced by Vitamin C addition after RAD21 depletion. Vitamin C can interact with the catalytic domain of TET enzymes and thereby enhance their activity towards 5mC oxidation [90].

Genomic DNA was extracted and subjected to 5fC or 5hmC DIP followed by qPCR to quantify enrichment of TET oxidation products at candidate enhancer regions under four different experimental conditions: First, cells were treated with auxin for 6 h to induce depletion of loops; untreated cells served as control. Second, cells were treated with Vitamin C 2 h after auxin addition (for a total of 4 h) in order to temporarily enhance TET activity, untreated cells served as control (Figure 4.16 A).

5hmC and 5fC DIP successfully enriched target regions. 5hmC was enriched 15-30-fold at the candidate enhancer regions, compared to a control region, while 5hmC enrichment at the corresponding promoters was much lower (up to 7-fold) (Figure 4.16 B, dark green bars). 5fC DIP enriched *Cgnl1,* a 5fC positive control region [81] 13-fold compared to a 5fC negative control region. The 5fC enrichment at the candidate enhancer regions was up to 4-fold compared to the negative control, and the enrichment at the corresponding promoters was up to 3-fold (Figure 4.16 C, dark green bars).

Vitamin C treatment increased 5hmC enrichment moderately at enhancer regions, but not at promoter regions (Figure 4.16 B, compare dark blue to dark green bars). Similarly, upon Vitamin C treatment, 5fC enrichment at enhancer regions was increased up to 14-fold compared to the control region, whereas the Vitamin C effect was more moderate at the promoter regions (Figure 4.16 C, compare dark blue to dark green bars). On the other hand, auxin-mediated RAD21 depletion had no effect on 5hmC and 5fC enrichment (Figure 4.16 B and C, compare light green to dark green bars).

The enrichment of 5hmC and 5fC in response to combined auxin and Vitamin C treatments differed: The 5hmC DIP revealed no differences of 5hmC enrichment between auxin treated and untreated cells upon Vitamin C treatment (Figure 4.16 B, compare light blue to dark blue bars). Vitamin C-induced 5fC enrichment, however, was significantly impaired at four candidate enhancer regions, whereas the same impairment was only observed at one promoter region (Figure 4.16 C, compare light blue to dark blue bars).

This suggests that the depletion of RAD21, which leads to a subsequent depletion of chromatin loops, impairs the acute, Vitamin C-induced accumulation of 5fC at enhancer regions. However, a similar trend, although not significant, was also observed for the non-enhancer regions. This indicates that the reduced 5fC accumulation may not be enhancer-specific and therefore RAD21-loop independent.

To rule out unspecific effects caused by possible cell cycle perturbations, cell cycle profiles for the four different conditions were generated by flow cytometry of propidium iodide stained cells. Untreated cells displayed a cell cycle profile typical for mESCs with a small proportion of cells in G1-phase and a high proportion of cells in S-phase (Figure 4.16 D) [180]. The addition of Vitamin C did not alter the cell cycle profile, whereas the addition of auxin strongly did: after 6 h auxin treatment, no G1 peak could be detected, the S-phase was smaller than in the untreated cells, and most cells were found in G2/M, suggesting that auxin treatment led to a G2/M arrest.

Considering the lack of 5fC accumulation at enhancers and the unimpaired 5hmC accumulation after combined auxin and Vitamin C treatment, it is possible that not the oxidation activity of TET enzymes was impaired, but rather that removal of 5fC by TDG was enhanced. In fact, TDG is degraded in S-phase, but highly expressed in the G2-phase [175]. This would explain why 5fC levels are lower in cells that were treated with both auxin and Vitamin C, compared to Vitamin C alone, as they arrested in G2/M-phase, where 5fC is excised by TDG. This explanation, however, is not supported by the unchanged enrichment of 5fC upon auxin treatment alone. Moreover, following Vitamin C treatment, the accumulation of newly generated 5hmC is much smaller than that of 5fC. Therefore, possible differences of 5hmC accumulation mediated by RAD21 depletion might be below detection limit.

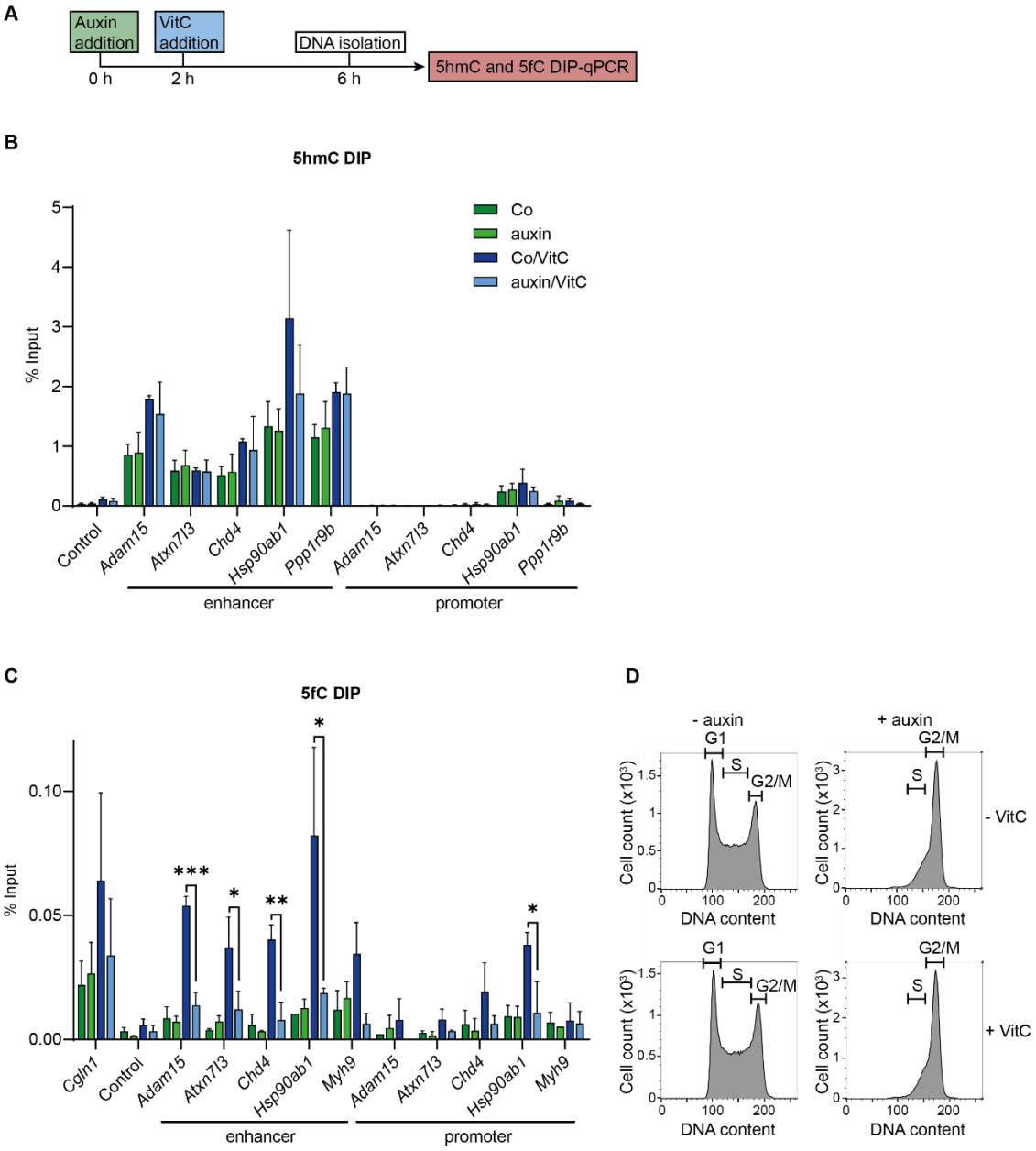

Figure 4.16 Prolonged RAD21 depletion leads to a G2/M arrest and reduced 5fC accumulation
(A) Experimental workflow and timeline depicting the combined auxin and Vitamin C treatment. *Rad21-AID-Tir1*-mESCs were treated with 500 µM auxin, treated 2 h later with 40 µg/ml Vitamin C (VitC), and harvested after a total of 6 h. (B) qPCR of immunoprecipitated genomic DNA samples using a 5hmC-antibody. Co, control. (C) qPCR of immunoprecipitated genomic DNA samples using a 5fC-antibody. Data shown as mean ± SD, n=3, * = p-value < 0.05, ** = p-value < 0.005, *** = p-value < 0.0005, according to student's t-test. (D) Cell cycle profiles of propidium iodide stained cells, measured by flow cytometry. DNA content corresponds to propidium iodide intensity.

4.2.2.2 G1/S synchronization avoids cell cycle perturbations by RAD21 depletion

To avoid cell cycle differences between RAD21-depleted and control cells, I controlled cell cycle progression by synchronizing the *Rad21-AID-Tir1*-mESCs at the G1/S boundary by double thymidine arrest. Since preliminary results from our laboratory suggested that TET enzymes are most active in S-Phase (Lorenz Hexemer and Dr. Andrea Schäfer, personal communication), cells were released by thymidine removal, thereby allowing for new production of 5hmC, 5fC and 5caC mediated by the activity of the TET enzymes. One hour before release from the thymidine arrest, auxin was added to deplete RAD21 and at the time of release, Vitamin C was added to the cells. Genomic DNA was isolated, and global 5mC, 5hmC, 5fC and 5caC levels were measured by quantitative mass spectrometry (Figure 4.17 A).

The genomic levels of all monitored cytosine modifications – 5mC, 5hmC, 5fC and 5caC – increased after double thymidine arrest and decreased again after release (Figure 4.17 B-E, dark green bars). The addition of auxin alone had no effect on the levels of these modifications (compare light green bars to dark green bars). Vitamin C increased 5hmC levels 2-fold, 5fC levels 6-fold, and 5caC levels 8-fold (compare dark blue bars to dark green bars). When auxin was added in combination with Vitamin C, only small changes could be detected. For the time points 2 h and 3 h after release, 5hmC and 5fC levels were slightly reduced compared to non-auxin treated samples (Figure 4.17 C and D, compare light blue bars to light green bars).

To monitor cell cycle perturbations, cell cycle profiles of the four different conditions were generated by flow cytometry of propidium iodide stained cells. Compared to untreated cells, most cells at time of release (“0 h”) were arrested at the G1/S boundary (Figure 4.17 F). In auxin non-treated cells, the release enabled the cells to migrate through S-phase until most cells reached G2/M 4 h after release. At this time point, a small proportion of cells had already reached G1-phase to start a new cell cycle. The addition of Vitamin C did not affect the dynamics of the cell cycle profiles. In contrast, the addition of auxin induced a G2/M arrest, starting to be visible 3 h after release, as was also observed in non-synchronized cells (Figure 4.16 D). However, 3 h after release, the cell cycle differences between auxin-treated and non-treated samples were still subtle. At the same time, the effects of auxin depletion in combination with Vitamin C treatment on global 5hmC, 5fC and 5caC levels were the largest.

Therefore, I selected the 3 h interval as a good time point for DIP followed by sequencing (DIP-seq). DIP-seq would allow investigating the site-specific effects of RAD21 depletion on 5mC oxidation globally and corroborating findings made earlier for selected enhancers using DIP followed by qPCR. In order to achieve greater effects of RAD21 depletion on 5hmC, 5fC and 5caC levels, auxin treatment during the thymidine block was further prolonged. By testing different time points, adding auxin 10 h before release of the thymidine block was determined to be the best time point (data not shown).

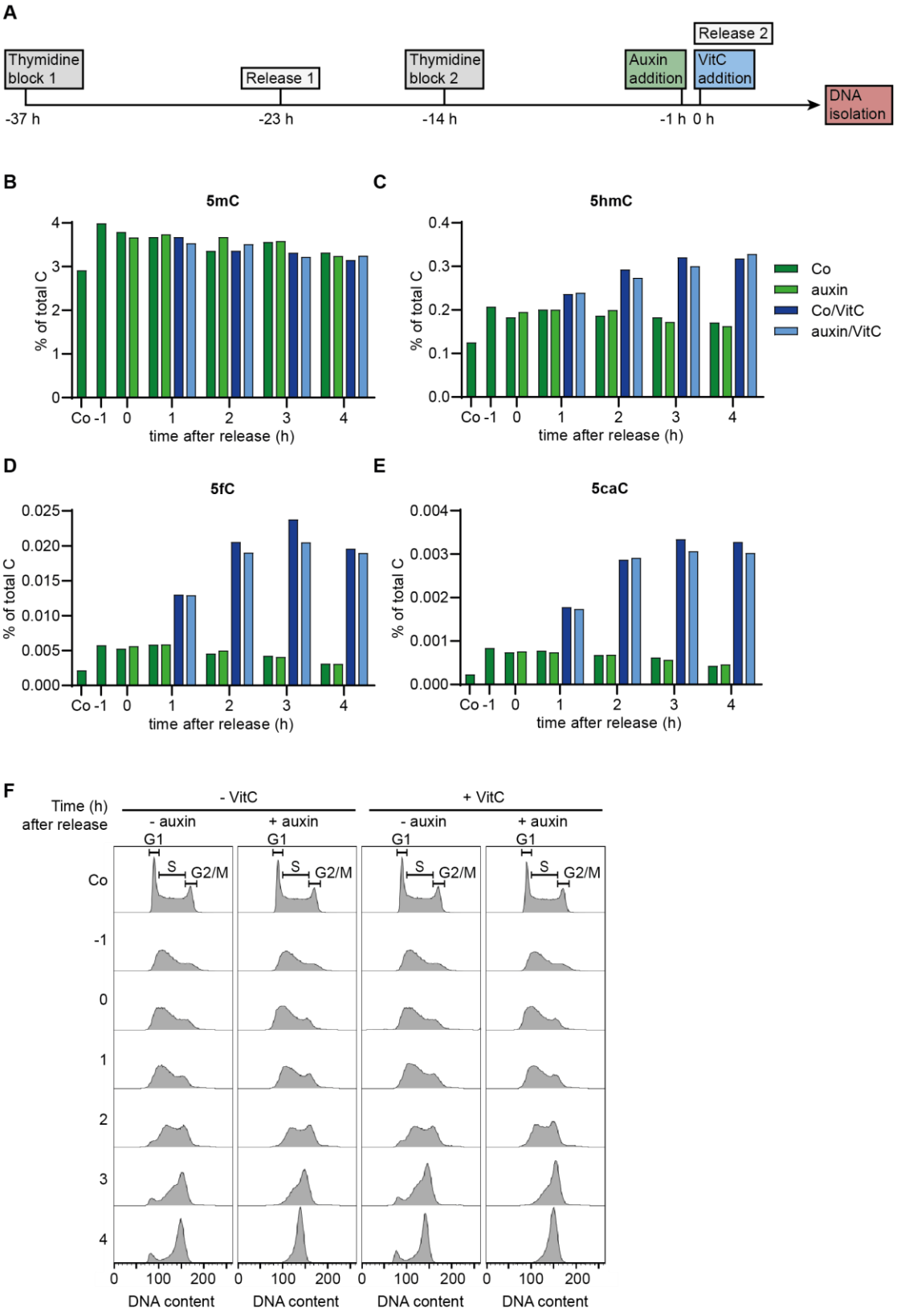

Figure 4.17 G1/S synchronization avoids cell cycle perturbations by RAD21 depletion
(A) Experimental workflow and timeline. *Rad21-AID-Tir1*-mESCs were synchronized in the G1/S-phase by double thymidine block. Cells were treated with 500 µM auxin 1 h before release, and 40 µg/ml Vitamin C (VitC) was added at the time of release. Cells were harvested after the indicated times. (B-E) Quantitative mass spectrometry measurements of the levels of 5mC (B), 5hmC (C), 5fC (D) and 5caC (E). Mass spectrometry data by Dr. Michael Musheev. Co, control. (F) Cell cycle profiles of propidium iodide stained cells, measured by flow cytometry. DNA content corresponds to propidium iodide intensity.

4.2.3 Genome-wide identification of RAD21-dependent 5fC sites

4.2.3.1 5fC DIP-Seq after acute RAD21 depletion

To investigate whether loss of RAD21 leads to a loss of the oxidation products of 5mC at enhancers, I chose to focus on only one of the oxidation products – 5fC – for the following reasons: First, the depletion of RAD21 impaired the Vitamin C-induced accumulation of 5fC at enhancer regions after 6 h of auxin-mediated RAD21 depletion, whereas the accumulation of 5hmC was not significantly impaired at tested enhancers (Figure 4.16 B and C). Second, Vitamin C increased 5fC levels more than it increased 5hmC levels (Figure 4.16 B and C, Figure 4.17 B-E).

First, *Rad21-AID-Tir1*-mESCs were synchronized at the G1/S boundary by double thymidine arrest before acute RAD21 depletion and Vitamin C treatment. RAD21 was depleted for a total of 13 h, 10 h during the thymidine block, and 3 h after release. Next, Vitamin C was added for 11 h to enhance TET activity when RAD21 was depleted (8 h during the thymidine block and 3 h after release). Finally, genomic DNA was extracted and subjected to 5fC DIP-seq (see Figure 4.18 A for experimental outline).

The efficient depletion of RAD21 was confirmed by Western blot (Figure 4.18 B), and cell cycle analysis showed only small differences between auxin treated and untreated samples (Figure 4.18 C). To confirm an efficient enrichment of 5fC-positive regions by the 5fC-antibody, the enrichment of candidate enhancer regions was tested by qPCR. In addition, qPCRs were performed with samples immunoprecipitated by IgG, as a further negative control. In untreated samples, the 5fC antibody enriched the enhancer regions more than 80-fold, compared to IgG (Figure 4.18 D, compare dark green bars to striped dark green bars). In addition, there was a more than 10-fold enrichment of 5fC-immunoprecipitated enhancer regions compared to the negative control region, whereas promoter regions were not enriched (dark green bars). Vitamin C increased the enrichment of enhancer regions up to 2-fold, but not of promoter or control regions (compare dark blue bars to dark green bars). Yet, the addition of auxin did not affect the enrichment of the enhancers, neither without Vitamin C, nor with Vitamin C treatment. As only four enhancers were considered, however, we decided to proceed with the sequencing of the samples, as they might not be representative for the majority of enhancers in the cell.

Paired-end sequencing of the samples yielded an average of 2.4×10^7 reads per sample. However, mapping of the paired-end reads resulted in many discordant and singleton reads in the DIP samples, but not in the Input samples. We think that this peculiarity originates from non-homologous single-strand DNA hybridization during the DIP procedure and Library prep. In fact, other studies performed the adaptor-ligation with paired-end adaptors before the DIP procedure, which would decrease the number of discordant and singleton reads due to non-homologous single-strand DNA hybridization [181,182]. To overcome this problem, single reads (read 1 from each pair) were mapped to the mm10 mouse genome. Single read-mapping was also used in other 5fC DIP-seq studies in mESCs, and should therefore also give conclusive data about 5fC enrichment [81].

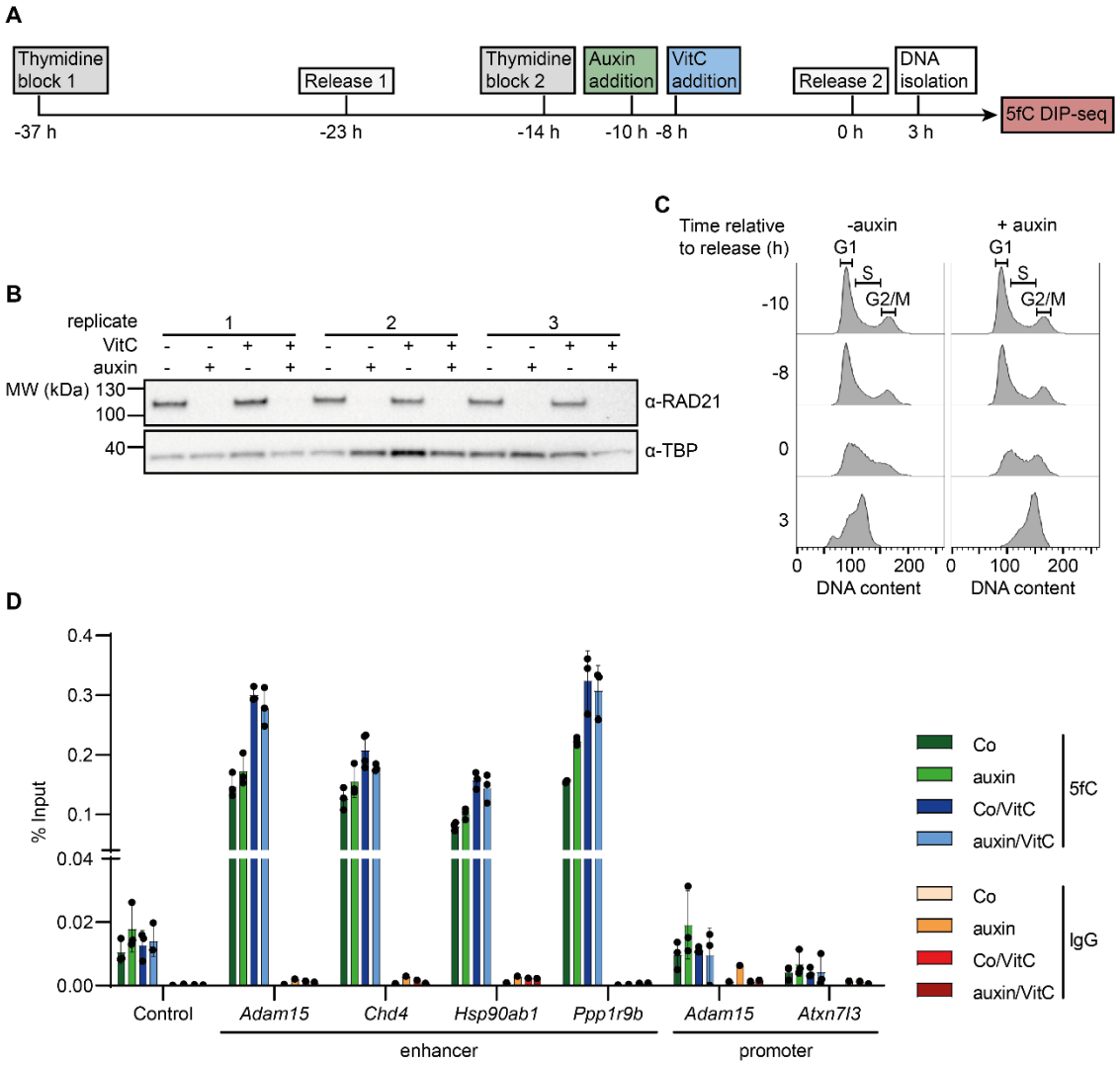

Figure 4.18 Generation of samples for 5fC DIP-Seq
(A) Experimental workflow and timeline. *Rad21-AID-Tir1*-mESCs were synchronized in the G1/S-phase by double thymidine block. Cells were treated with 500 µM auxin 10 h before release, and 40 µg/ml Vitamin C 8 h before release. (B) Western blot of nuclear protein lysates showing the successful depletion of RAD21 after auxin addition. Protein samples were harvested from the same dishes as DNA samples for DIP. MW, molecular weight. (C) Cell cycle profile of propidium iodide stained cells, measured by flow cytometry. DNA content corresponds to propidium iodide intensity. (D) qPCR of immunoprecipitated genomic DNA samples using a 5fC-antibody or IgG. Data shown as mean ± SD, n=3. Co, control.

Our dataset yielded 203,404 common 5fC peaks (present in at least two DIP samples). This 5fC peak set showed a good overlap with 5fC peaks identified by Shen and colleagues, as 80 % of their identified 5fC peaks (approximately 20,000 peaks) were also found in our dataset (Figure 4.19 A) [81]. Our identified 5fC peaks were enriched for TET1 binding sites [166], mESC enhancers [176], and Satellite and Simple Repeats, while LINE L1 elements were depleted (Figure 4.19 F, compare dark grey to light grey bars).

Unexpectedly, differential analysis revealed no deregulated 5fC accumulation in the RAD21-depleted (“auxin”) samples compared to the control samples, using DESeq2 [183] with a relaxed

FDR cut-off of 10 % (Figure 4.19 B). This result was further corroborated also with the tools EdgeR [184], DiffBind [185] and MEDIPS [186]. Moreover, RAD21 neither significantly affected any 5fC peaks in the cells treated with Vitamin C (Figure 4.19 C). Vitamin C treatment alone, on the other hand, affected thousands of peaks. DESeq2 differential enrichment analysis revealed that 3,732 5fC peaks were increased and 1,322 reduced at 10 % FDR (Figure 4.19 D). The Vitamin C-induced peaks were enriched at TET1 binding sites and mESC enhancers, while the Vitamin C-suppressed peaks were enriched at Satellite and Simple Repeats (Figure 4.19 F). A similar outcome was observed when analyzing Vitamin C effects in the presence of auxin, with even more induced 5fC peaks (5,985), but a similar number of suppressed 5fC peaks (1,283) as in the cells not treated with auxin (Figure 4.19 E).

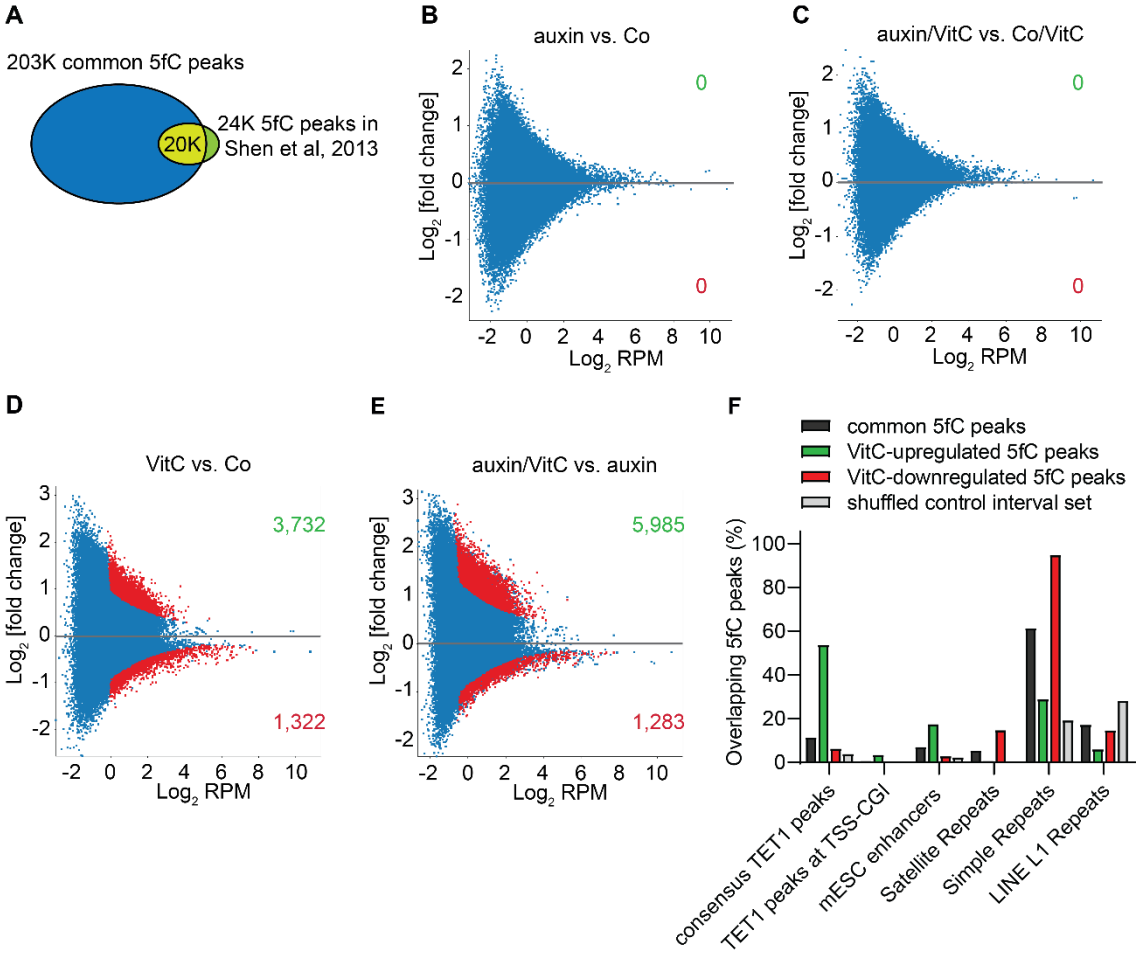

Figure 4.19 Depletion of RAD21 does not affect 5fC deposition in mESCs
(A) Overlap of the 203,404 common 5fC peaks identified by 5fC DIP-seq with the 5fC peaks identified by Shen et al [81]. (B-C) MA-plots from differential enrichment analyses of auxin-treated vs. untreated cells in the absence (B) or presence (C) of Vitamin C using DESeq2. No differential peaks were found with a relaxed 10 % FDR cut-off. Co, control. (D-E) MA-plots showing the effects of Vitamin C treatment in the absence (D) or presence (E) auf auxin, using DESeq2 with a 10 % FDR cut-off. Significantly changed peaks are highlighted red: 3,732 5fC peaks are increased and 1,322 reduced in the absence of auxin, and 5,985 increased/1,283 reduced in the presence of auxin. RPM, reads per million. (F) Feature overlap of the 203,404 common 5fC peaks, Vitamin C-induced (3,732) and Vitamin C-suppressed (1,322) 5fC peaks, as well as a shuffled interval set of the common 5fC peaks as a randomized control. Bioinformatic analysis and MA-plots by Dr. Emil Karaulanov.

I thus conclude that 5fC deposition at enhancers is RAD21-loop independent, at least under the experimental conditions used in this study (Figure 4.20). Although we did not investigate if all loops were indeed depleted, a recent study in mESCs, which uses a similar RAD21-degron mESC-line, showed that loops are depleted already 6 h after RAD21 depletion [150]. Of course, it is possible that by using another experimental set-up, e.g. using different time points, RAD21-dependent changes of 5fC peaks might be revealed. Nevertheless, it is likely that TET activity is targeted to distal elements via other means, and not by promoter-enhancer looping.

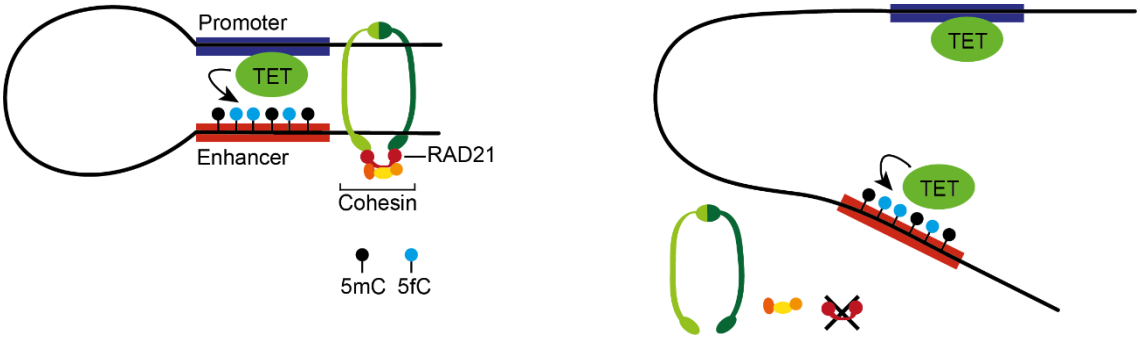

Figure 4.20 5fC deposition at enhancers is RAD21-loop independent
Model showing that 5fC deposition at enhancers is not affected by RAD21 depletion.

4.2.4 Regulation of TET activity by PAR

4.2.4.1 5hmC levels at selected enhancers are not altered by PAR depletion

Recently, the enzyme Poly [ADP-ribose] polymerase 1 (PARP1) emerged as a player in shaping the 3D chromatin structure. PARP1 can add poly [ADP-ribose] (PAR), a negatively charged polymer, to target proteins and also itself, in a process termed PARylation [187]. Its role in shaping chromatin structure includes, amongst others, the interaction with and PARylation of CTCF, which mediates chromatin fiber interactions [188].

To test if instead of RAD21, PARylation might play a role in regulating TET activity at enhancers, I inhibited PAR synthesis in mESCs using two PARP-inhibitors: Olaparib and PJ34. In order to determine if the inhibitor treatment was successful, a Western blot using a PAR-antibody was performed. Olaparib treatment caused a weak reduction of the PAR-signal in the cytoplasmic and nuclear fractions in a concentration dependent manner, while PJ34 at 10 µM completely depleted the PAR signal (Figure 4.21 A).

The effects of PAR depletion on global levels of 5mC, 5hmC, 5fC and 5caC were monitored via quantitative mass spectrometry. While PARP inhibition did not affect 5mC and 5fC levels, 10 µM PJ34 significantly increased 5hmC and 5caC levels, compared to the control DMSO treatment ($p<0.05$) (Figure 4.21 B-E). Still, the increase was rather modest: 1.1-fold for 5hmC and 1.2-fold for 5caC. Conversely, 1 µM Olaparib treatment significantly decreased 5hmC levels, however, 10 µM Olaparib treatment had no effect. Since the reduction of PAR levels according to the Western blot were actually weakest with 1 µM Olaparib treatment, this is most likely an unspecific effect.

Although no strong great global changes of the oxidative products of 5mC could be detected, site-specific changes at enhancers could still be possible. Hence, I performed 5hmC DIP-qPCR upon PARP inhibition. In control (DMSO)-treated samples, candidate enhancers were enriched about 20-fold compared to the negative control upon 5hmC DIP (Figure 4.21 F, “DMSO”, grey bars). The treatment with the PARP-inhibitors did not affect 5hmC enrichment at any of the candidate enhancer regions tested (compare “Olaparib”, green bars, and “PJ34”, blue bars, with “DMSO”, grey bars). The enrichment of 5hmC levels at promoters was weaker than at the enhancers, as observed before (see Figure 4.16), and again, no effect by PARP-inhibition was observed.

In summary, PARP inhibition did not lead to strong global changes of 5hmC, 5fC and 5caC. In addition, the enrichment of 5hmC at candidate enhancer regions was not altered by PARP inhibition. Thus, PARylation likely does not play a role in targeting TET to distal genomic elements.

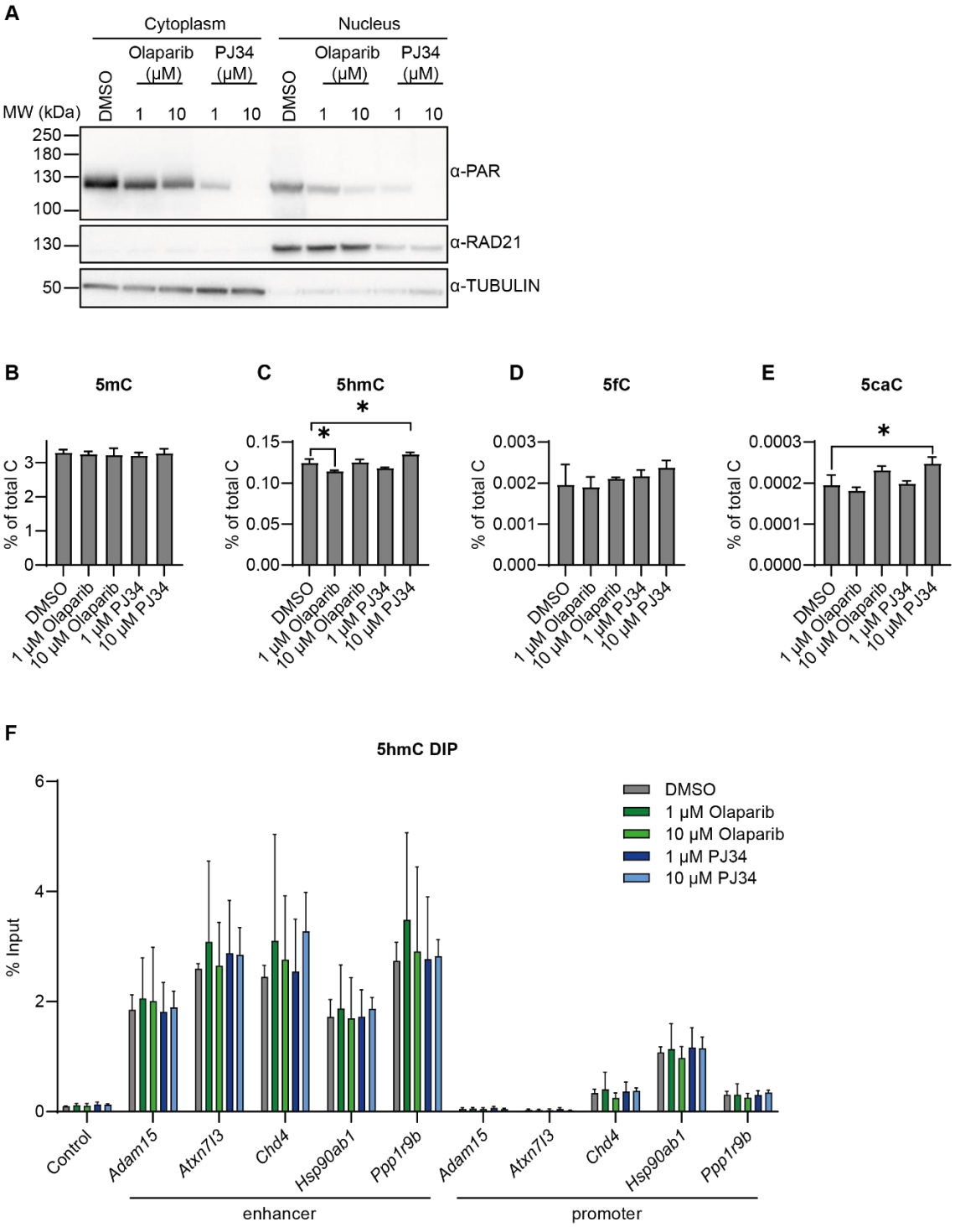

Figure 4.21 5hmC deposition at enhancers is not affected by PAR depletion
mESCs were treated with the indicated PARP inhibitors for 48 h. (A) Western blot of nuclear and cytoplasmic protein lysates, showing the depletion of PAR by the indicated inhibitors. RAD21 and TUBULIN antibodies were used as nuclear and cytoplasmic markers, respectively. MW, molecular weight. (B-E) Quantitative mass spectrometry measurements of the levels of 5mC (B), 5hmC (C), 5fC (D) and 5caC (E). (F) qPCR of immunoprecipitated genomic DNA using a 5hmC-antibody. Data shown as mean ± SD, n=3, * = p-value < 0.05, according to student's t-test. Mass spectrometry data by Dr. Michael Musheev.

5 Discussion

5.1 Mechanisms of GADD45A-mediated TET1 regulation

GADD45A and TET1 are both well-known players in the relatively young field of active DNA demethylation [20,22]. GADD45A physically interacts with TET1 and thereby enhances TET1 activity [104]. However, the mechanism behind this cooperation is still poorly understood. GADD45A does not directly demethylate DNA, as it lacks any known enzymatic function. It rather promotes active DNA demethylation by acting as an adaptor protein and by bringing together different factors involved in this process [20,109,111,151]. Thus, it is likely that GADD45A also acts as an adaptor for TET1, bridging TET1 and a yet unknown additional effector.

The first aim of my PhD thesis was to unravel the mechanism of how GADD45A enhances TET1 activity. For this purpose, I explored different paths: studying a) a GADD45A-TET1 fusion protein, b) the interplay between MAP3K4, GADD45A and TET1, c) the role of GADD45A in altering post-translational modifications of TET1, and d) the recruitment of additional interactors – members of the TRIM protein family – by GADD45A.

Unfortunately, the generated GADD45A-TET1CD fusion protein was not functional and could not be used to further elucidate the mechanism by which GADD45A enhances TET1 activity. Very recently, Taghbalout and colleagues could target a GADD45A-TET1CD fusion protein to a methylation-silenced gene using a newly developed methylation editing toolbox called Casilio-ME, which allows for recruitment of multiple effectors to a target site. Their fusion protein was active and could increase the expression of the methylation-silenced *MLH1* gene. Moreover, their GADD45A-TET1CD fusion protein was 3-fold more active than TET1CD alone [189]. I only tested if my GADD45A-TET1CD fusion protein was active regarding global 5mC oxidation and cannot rule out that it still might be active at a single locus. Another possibility is that Taghbalout's slightly longer GS-linker between GADD45A and TET1CD (16 amino acids compared to 10 amino acids) is necessary for the activity of the fusion protein. In any case, the GADD45A-TET1CD fusion protein by Taghbalout and colleagues will be of great interest to further unravel how GADD45A increases TET1 enzymatic activity.

All three TET family members are highly phosphorylated, mostly at their N-terminus and low complexity insert region, whereas only a few phosphorylation sites have been identified in the catalytic domain [96]. The impact of phosphorylation on TET activity has only been scarcely studied so far; also, the involved kinases remain largely unknown. Recently, AMPK was established as a kinase phosphorylating TET2 at S99. Mutation of this phosphorylation site destabilizes TET2 and leads to reduced global 5hmC levels [190]. Yet, this phosphorylation site is not conserved in TET1 and thus probably represents a TET2 specific mode of regulation.

In this study, the four identified MAPK-regulated phosphorylation sites were all located in the N-terminus of TET1. Two of these sites, S116 and S871, showed reduced phosphorylation upon inhibition of MAPK11/MAPK14. In addition, the SB202190 inhibitor treatment abolished the enhancement of TET1 activity by GADD45A (Figure 4.5). The orthologous site to human TET1

S871, mouse TET1 S854, was also identified as a phosphorylation site by Bauer and colleagues [96]. Nevertheless, since GADD45A enhances not only the activity of full-length TET1, but also the C-terminal catalytic domain alone, relevant phosphorylation sites for GADD45A-dependent TET1 regulation would be expected to lie within the catalytic domain of TET1. Two TET1 phosphorylation sites within the catalytic domain (S1605 and S1851) were indeed more heavily phosphorylated by the overexpression of GADD45A and MAP3K4. However, phosphorylation deficient serine to alanine mutations did not alter TET1 activity regarding global 5mC oxidation (Dr. Andrea Schäfer, personal communication), arguing against a role of these phosphorylation sites in regulating TET1 activity.

Of course, the cooperation between GADD45A and MAP3K4 could also affect TET1 not directly via phosphorylation, but via indirect means, e.g. by the recruitment of additional TET1 interactors. However, the co-expression of GADD45A and MAP3K4 rather reduced TET1 activity, instead of enhancing it like expression of GADD45A alone (Figure 4.2 and Figure 4.6). Thus, MAP3K4 is not the co-factor of GADD45A that is responsible for enhancing TET1 activity.

The role of phosphorylation in regulating TET1 activity remains elusive. We cannot exclude that GADD45A plays a role in the phosphorylation of TET1, but we could neither find any evidence that these phosphorylation changes affect TET1 activity on a global scale. The higher density of phosphorylation sites in the N-terminus of TET1 suggests that the recruitment to chromatin rather than the catalytic activity of TET1 could be affected by phosphorylation. In fact, the notion that TET1's N-terminus mediates chromatin binding is supported by the discovery of two different isoforms of mouse TET1: a full-length isoform that is exclusive to embryonic tissues (TET1FL), and a truncated isoform of TET1 (TET1s). TET1s lacks large parts of the N-terminus including the CXXC domain and was only found to be expressed in somatic tissues. Furthermore, TET1s has a reduced global catalytic activity and reduced global chromatin binding [70].

Perhaps, the N-terminus of human TET1 is also essential for the regulation by GADD45A, although this does not become evident with overexpressed proteins. Indeed, it is possible that the high non-physiological levels of TET1CD that were used in this study mask a possible deficiency of TET1CD in its global activity and chromatin binding. Thus, future studies should rather investigate TET1 phosphorylation in a physiological context, e.g. in embryonic tissues, to avoid unspecific effects caused by ectopic overexpression. In fact, whereas GADD45A increases 5hmC production via TET1FL as well as TET1CD, the enhanced removal of 5fC and 5caC is only observed for TET1FL, but not TET1CD (compare Figure 4.2 to Figure 4.10 and Figure 4.12) [104]. This argues that for GADD45A's full role in DNA demethylation, also the N-terminus of TET1 is required.

The third path that was explored to unravel how GADD45A enhances TET1 activity, was the role of GADD45A in altering ubiquitination or SUMOylation of TET1. All three TET proteins are monoubiquitylated by the E3 ubiquitin ligase complex CRL4 on a highly conserved lysine residue within the catalytic domain. This monoubiquitylation promotes the binding of TET to DNA and is essential for proper 5hmC production [93,94]. In this study, GADD45A reduced TET1

ubiquitination (Figure 4.7), which could indicate that GADD45A increases TET1 protein stability by preventing degradation via the proteasome, as it is well known that ubiquitin chains can target proteins to the 26S proteasome for degradation [191]. A CHX assay revealed that GADD45A indeed increases TET1 protein stability, but independently of the proteasome (Figure 4.8). Instead, a Calpain inhibitor stabilized TET1 protein levels to a similar extent as GADD45A, but only affected overexpressed human TET1 in HEK293T cells, not endogenous mouse TET1 in mESCs.

Yet, a different study, conducted by Wang and Zhang, discovered that mouse TET1 was targeted by CALPAIN1-mediated degradation also in mESCs. This study also showed that *Calpain1*-knockout cells had elevated 5hmC levels [46]. The discrepancy between the study by Wang and Zhang and ours could have different reasons. The main cause probably originates from the technical limitations of accurately quantifying proteins by Western blotting. The accurate quantification of proteins via Western blotting depends on many variables, such as protein isolation, membrane transfer, quality of the antibody, selection of a reference protein (serving as the “loading control”), and the method of quantification of band intensities [192]. Indeed, most of these variables differed in the two studies. For example, Wang and Zhang used TUBULIN as a reference protein for the relative quantification of TET1 protein levels. In this thesis, distinct bands of HA-TET1 could only be detected when nuclear protein lysates were used for Western blotting, requiring also a nuclear reference protein. HNRNPA1 was used for this purpose, as it gave more reliable signals than other tested loading controls (LAMINB1, Histone H3 and p68, data not shown). Still, Figure 4.8 shows that HNRNPA1 is not detected as a fully distinct band and this could affect protein quantification. As these technical limitations of Western blotting apply universally, any conclusions based solely on the Western blot method should be taken with caution.

Despite the limitations described above, this thesis provides first hints that GADD45A increases TET1 stability, possibly by preventing Calpain-mediated decay. To further explore this in the future, GADD45A-degron mESC lines could be used to monitor how an acute depletion of GADD45A affects TET1 protein stability in an endogenous TET1 context. In addition, it would be interesting to know if GADD45A only increases protein stability of full-length TET1, or also stabilizes TET1CD.

Still, the question of how GADD45A mediates stabilization of the TET1 protein remains unresolved. GADD45A has been shown to stabilize SMAD4 by preventing its ubiquitination and subsequent proteasomal degradation [193]. However, the proteasome inhibitor MG132 had no effect on TET1 stability in this study (data not shown). Another possibility would be that GADD45A does not directly affect protein stability, but rather increases mRNA stability. In fact, GADD45A was recently shown to stabilize memory-related mRNAs in mice, which requires the RNA-binding capacity of GADD45A [194]. Thus, GADD45A could also regulate TET1 activity by stabilizing its mRNA, for example via preventing binding and degradation by miRNAs. In this context, several miRNAs have been implicated in TET1 regulation, such as miRNA-22 [195], miRNA-26a [196], or

miRNA-29a [197]. Another RNA binding protein, Deleted in azoospermia-like (DAZL), has been shown to regulate TET1 by stabilizing its mRNA and enhancing its translation [198]. In the future, a possible stabilizing function of GADD45A on TET1 mRNA could be tested and compared to the effect of an RNA-binding deficient K45A mutant of GADD45A [199].

SUMOylation already has a proposed role in DNA demethylation: it enables TDG to be released from the DNA after excision; hence, increasing the turnover of TDG [168]. However, the role of SUMOylation has been a matter of debate and it may not be the only mechanism mediating product release [200,201]. Recently, SUMO was shown to not only coordinate TDG, but the whole TDG-base excision repair complex [202]. GADD45A was shown to interact with TDG and to promote removal of 5fC and 5caC by TDG, but the mechanism is yet unclear [111]. In fact, GADD45A itself can be SUMOylated, as recently shown by a mass spectrometry screen [169]. Therefore, GADD45A could help to assemble the DNA demethylation complex, for example by acting as a SUMO donor, similar to X-ray repair cross-complementing protein 1 (XRCC1), which acts as a SUMO donor for TDG during DNA demethylation [202].

In this light, I investigated if GADD45A promotes the removal of 5fC and 5caC by coordinating the SUMOylation of TDG and also TET1. Indeed, GADD45A strongly enhanced the SUMOylation of TDG (Figure 4.11), providing a first hint that this might be the mechanism by which GADD45A can enhance 5fC and 5caC removal by TDG. To my knowledge, TET1 had not been identified as a SUMO target at the beginning of this study, but TET3 had been found to be SUMOylated in a high resolution mass spectrometry screen [203]. Co-IPs of SUMO and TET1 revealed that also TET1 is SUMOylated. Surprisingly, GADD45A decreased the SUMOylation of full-length TET1, while it enhanced the SUMOylation of TET1CD (Figure 4.9).

I decided to focus on GADD45A's effect on the catalytic domain of TET1, as the activity of the isolated catalytic domain is also enhanced by GADD45A. Possible SUMOylation sites of TET1 were predicted by an identification of SUMO-consensus motifs and the predicted SUMO sites were mutated from lysine to alanine or arginine. However, none of the four tested TET1 mutations affected TET1 activity or the functional cooperativity with GADD45A (Figure 4.10). A recent study confirmed that TET1 can indeed be SUMOylated, and three SUMO sites were identified in human HeLa and U2OS cell lines: K540, K784 and K792 [169]. Of note, none of these identified sites lies in a SUMO-consensus motif, nor in the catalytic domain of TET1. However, since TET1 expression levels in somatic cell lines are generally low [41], less abundant SUMO sites might have been overlooked. Thus, it might be possible that 1) other SUMOylation sites exist in the catalytic domain, which lie outside of SUMO-consensus motifs and were thus not predicted and 2) SUMOylation, like phosphorylation, is mainly confined to the N-terminus of TET1. This thesis cannot provide evidence that SUMOylation plays a role in GADD45A-mediated TET1 regulation, but it can neither exclude it. *In vitro* SUMOylation of TET1 followed by a TET1 *in vitro* oxidation assay would give more insights into the role of SUMOylation in coordinating TET1 activity.

If not via post-translational modifications, how else could GADD45A regulate TET1? Interactors, regulating TET1 activity have been identified, which mostly affect TET1 activity by the recruitment

to specific loci (reviewed in [41]). In search for GADD45A-regulated TET1 interactors, the E3 Ubiquitin/ISG15 ligase TRIM25 and the E3 Ubiquitin/SUMO ligase TRIM28 were found to be required for GADD45A to enhance global TET1 activity (Figure 4.12). However, these two factors most likely affect TET1 enzymatic activity only indirectly, e.g. by enhanced recruitment to genomic loci, as the depletion of these factors did not affect TET1 activity in an *in vitro* oxidation assay (Figure 4.13). Interestingly, the connection between TRIM28 and TET proteins has been made before. TET proteins are required for efficient binding of TRIM28 to *MuERV-L* in order to repress *MuERV-L* transcription [50]. This, however, involves the repressive function of TET proteins, which is uncoupled from the catalytic activity [74]. In summary, this thesis could not identify any new GADD45A-mediated interactors of TET1. As a result, the main question remains unresolved: How does GADD45A enhance TET1 activity? Other post-translational modifications, apart from the ones discussed here, might play a role in this mechanism, including: acetylation, which can enhance TET2 activity [95], O-GlcNAcetylation, which was shown to be required for proper TET1 function [97], and PARylation, which also enhances TET1 activity [98].

Apart from recruiting co-factors to enhance the catalytic activity of TET1, GADD45A could also be involved in targeting TET1 to genomic loci. Several other proteins targeting TET1 to genomic loci have been identified, including NANOG [99], PRDM14 [100], PRC2 [103], LIN28A [102]. Indeed, GADD45A was recently shown to target TET1 to a subset of R-loops [166]. In addition, preliminary TET1 ChIP-seq results in wild-type vs. *Gadd45a/b/g* triple-knockout mESCs showed that a few hundred TET1 peaks were downregulated after GADD45A/B/G depletion (Dr. Khelifa Arab, personal communication). This suggests that GADD45 proteins are indeed involved in targeting TET1 to genomic loci.

Both TET1 and GADD45A belong to protein families composed of two additional members, TET2 and TET3; and GADD45B and GADD45G. In fact, those other family members are also involved in active DNA demethylation and can thus, at least partially, act redundantly [31,37,106,110]. The redundancy is also further stressed *in vivo*: single *Tet3* mouse mutants are viable [45], *Tet1/2* double knockout mutants are partially lethal [204], whereas *Tet1/2/3* triple-mutants are lethal due to gastrulation defects [52]. In addition, also GADD45 single knockouts of *Gadd45a, -b* or *-g* are viable [205–207]. One could envision that the functional interaction between GADD45A and TET1 is not exclusive. Other combinations of GADD45 and TET proteins could also cooperate in promoting DNA demethylation, which, in its whole, regulates the epigenome. As a first proof of this concept, co-expression of GADD45B and TET1, as well as GADD45G and TET1 also boosted 5hmC levels, similar to GADD45A co-expression (data not shown).

A general problem when studying GADD45A's role in TET1-mediated DNA demethylation is GADD45A's dual role of enhancing 5hmC production by TET1 and enhancing 5fC and 5caC removal by TDG (see e.g. Figure 4.2). To uncouple these two functions of GADD45A, one can make use of TDG mutants. This could reveal whether GADD45A indeed enhances the removal of 5fC and 5caC by TDG, or if GADD45A rather affects the substrate preference and processivity of TET1.

5.2 The interplay between chromatin structure and DNA demethylation

A remaining conundrum in the field of DNA demethylation is why TET binding sites are mostly found at promoters, whereas the oxidative derivatives of 5mC, which are generated by the TET enzymes, are mostly enriched at distal elements, such as enhancers. A simple model to explain this discrepancy is promoter-enhancer looping, as proposed by Wu and Zhang [41]. Interestingly, in a recent study ING1 (occupying promoters) was shown to promote local DNA demethylation via GADD45A (mostly bound at superenhancers), by long-range chromatin looping [151].

To test if promoter-enhancer looping plays a role in targeting TET activity to distal regulatory elements, I generated a RAD21-degron mESC-line, which was used to acutely and efficiently deplete RAD21 (Figure 4.15). Due to the complexity of identifying chromatin loops, I did not perform high-resolution genome-wide chromosome conformation capture (Hi-C) analysis to verify if the depletion of RAD21 resulted in a depletion of promoter-enhancer looping in the RAD21-degron mESC-line. However, a thorough analysis of acute RAD21-loss in the context of chromatin structure has been performed by Rao and colleagues in human HCT116 cells, also using an auxin-inducible RAD21-degron cell line. Hi-C analysis showed that already 40 min of auxin treatment depleted loops, whereas auxin withdrawal could restore the loops accordingly [132]. In addition, a recent study in mESCs confirmed the loss of loops 6 h after auxin-mediated RAD21 depletion [150], and another mESC study specifically detected a reduced contact frequency between the Sik1 promoter and its distal regulatory elements upon 24 h of RAD21 depletion [141]. For genome-wide analysis of 5fC deposition, I depleted RAD21 by auxin addition for a total of 13 h, with auxin renewal after the first 10 h. Thus, although not proven, it can be assumed that promoter-enhancer loops in my system were also depleted.

To test if the TET oxidation product 5fC is affected by loop depletion, specifically at enhancers, I performed a 5fC DIP-seq after RAD21 depletion. Surprisingly, differential 5fC peak analysis did not reveal any peaks that changed after RAD21 depletion (Figure 4.19 A and B). Yet, we are confident about having correctly mapped 5fC peaks, as 80 % of the peaks of a published 5fC dataset were also found in our dataset [81]. Also, the feature overlap analysis showed enrichment for 5fC in mESC enhancers and repetitive regions, which have been shown before to be enriched for 5fC (Figure 4.19 E) [81]. In addition, we were generally able to detect differential 5fC peaks in our experimental set-up, as revealed by 3,732 (-auxin) and 5,985 (+auxin) Vitamin C-induced 5fC peaks (Figure 4.19 C and D). The increase in the number of Vitamin C-induced peaks after RAD21 depletion (+auxin) is surprising. But since the additional peaks are random and are not enriched at any investigated genomic feature (data not shown), we think that this increase might just be a technical artefact. We thus conclude that RAD21 depletion, and therefore loop depletion, has no effect on 5fC distribution.

There are several caveats of the above-described experimental setup, which could explain our results: First, the experiment was performed in cells that express TDG. Consequently, the majority of newly produced 5fC might have been already removed by TDG. In fact, a 5fC DIP-seq study has revealed a large increase in overall 5fC peaks after TDG depletion, especially in distal

regulatory elements [81]. Also alternative methods to detect 5fC, termed 5-formylcytosine-selective chemical labeling (fC-Seal) or 5fC chemical modification-assisted bisulfite sequencing (fCAB-seq), revealed an increase in 5fC levels after TDG depletion, especially in poised enhancers, whereas the authors did not find any TDG dependent regulation for 5hmC [86]. Thus, it could be possible that 5fC distribution in the cell, in contrast to our model, is not mainly regulated by TET, but rather by TDG. In the future, one could study the distribution of 5fC after promoter-enhancer loop depletion in a TDG-deletion background. This would make it possible to study only the effects of TET on 5fC distribution, without disturbances caused by TDG removal. It also has to be considered that promoter-enhancer loop depletion could have a greater effect on 5hmC distribution at enhancers, since 5hmC is not removed by TDG. Although I did not observe any effects after 6 h of auxin depletion on the 5hmC levels at selected enhancers, a longer auxin treatment might still reveal differences at other enhancer regions (Figure 4.16).

Second, maybe the expected changes of 5fC peaks after loop depletion are too small to be detected by 5fC DIP-seq, although Vitamin C was added to enhance *de novo* 5fC formation. An average cell cycle of mESCs lasts 11 h [180]. Assuming there is no active removal of 5fC by TDG and no new production by TET enzymes, 50 % of “old” 5fC would still remain in the cell after 11 h due to DNA replication. In our experimental set-up, the cells underwent only one round of replication after loop depletion (Figure 4.18). Song and colleagues determined that 7 % of the total number of 5fC-containing peaks overlap with enhancers, which are predicted to link to promoters [86]. If 50 % of 5fC at these regions remained after DNA replication, only 3.5 % of the total 5fC, in the best scenario, would be predicted to change after RAD21 depletion, given our model is correct. In fact, “old” 5fC at the parental strand might serve as a mark to guide TET to the daughter strand during DNA replication, a concept similar to DNA methylation maintenance by DNMT1 [6]. This possible maintenance might make proper loop formation dispensable for new 5fC production. For example, TET might be recruited to enhancers not via promoter-enhancer looping, but via DNMT3A/B recruitment. Although this might seem surprising as DNMT3A/B are mainly known as *de novo* DNA methyltransferases, they have also been implicated in maintenance DNA methylation [209]. This concept would fit to findings of Gu and colleagues, who showed that TET1 occupancy at enhancers in *Dnmt3a* knock-out mESCs was decreased [210]. At the same time, TET1 occupancy at promoters slightly increased, suggesting that diverse modes of TET1 regulation exist, depending on the genomic context. In addition, Zhang and colleagues found that all human TET proteins interact with DNMT1, and TET2 also associates with DNMT3B. Furthermore, they propose that DNMT1 could be essential for recruiting TET2 to chromatin during oxidative stress [95].

To discriminate between newly produced 5fC and old 5fC, one could perform nascent sequencing of 5fC, via pulse labeling with the nucleotide analog ethynyl-deoxyuridine (EdU), as described for 5mC by Xu and Corces [154]. Using this method in combination with a base-resolution method to detect 5fC would also overcome the drawbacks of the 5fC DIP-seq method, of having a low spatial resolution, a strong density bias (discussed for 5mC- and 5hmC-antibodies in [82]) and the lack of absolute quantification [41]. An additional recently discussed problem of DIP-methods is off-target binding of IgG to short unmodified DNA repeats, which increases the noise and complicates the identification of subtle alterations of the modifications [211]. In fact, this could have been a problem in the 5fC DIP-seq in our and other studies, as simple repeats were enriched in our called

5fC peaks (Figure 4.19). Alternative base-resolution methods for 5fC detection include e.g. M.SssI methylase-assisted bisulfite sequencing (MAB-seq), which measures 5fC and 5caC [87], reduced bisulfite sequencing (redBS-seq) [212], or fCAB-seq [86]. Using one of these methods would increase the sensitivity and thus, the possibility to detect also subtle changes of 5fC levels.

A large number of TADs are cell specific [123,124] and in contrast to the population average, Cohesin depletion in single cells did not abolish domain separation [123]. Thus, perhaps also the changes of 5fC distribution are masked by cell-to-cell differences and single cell 5fC sequencing would be required to detect changes of 5fC distribution in individual cells.

Supposing that experimental caveats indeed masked changes of 5fC levels after RAD21 depletion, what evidence speaks for a relationship between the 3D chromatin structure and DNA demethylation? Xu and Corces discovered that a loss of DNMT3B reduces RAD21-mediated chromatin interactions in human ESCs. They propose that this is due to a loss of hemimethylation, perhaps because of an altered physical interaction with MeCP2 [154]. MeCP2 can interact with Cohesin and regulate chromatin looping [157,158]. Furthermore, a recent study by Wiehle and colleagues established CTCF as an insulator for DNA methylation spreading. They propose a methylation/nucleosome/CTCF switch, where TET1/2 deficiency leads to a gain of DNA methylation and increased nucleosome occupancy. This affects CTCF binding, which in turn changes DNA methylation patterns [159]. A study focusing on the role of CTCF in shaping the 3D chromatin also established CTCF as a boundary element, which is required for TAD insulation [147].

This gives CTCF a distinct role from RAD21 and the Cohesin complex, which is required for loop domain formation [132]. Still, both CTCF and Cohesin are involved in shaping the 3D chromatin structure. If TET recruitment to enhancers was indeed dependent on promoter-enhancer looping, maybe only the disruption of both, CTCF and the Cohesin complex would affect TET activity at enhancers. Although promoter-enhancer contacts are depleted after Cohesin loss, TADs still persist and compartmentalization is even enhanced [132]. Perhaps, the compartmentalization alone is sufficient to recruit TET to enhancers.

Another important question is also whether the 3D chromatin structure shapes the (de-)methylome as proposed by Xu and Corces [154], or whether the (de-)methylome shapes the 3D chromatin structure, as proposed by Wiehle and colleagues [159]. In fact, Fang and colleagues revealed additional evidence suggesting that the (de-)methylome might shape the 3D chromatin structure by showing that a combined *Tet2* and *Tet3* depletion disrupts YY1 binding and long-range chromatin interactions during mouse heart development [208].

Many recent studies propose that TADs and chromatin loops are dynamic structures [213]. Hence, it is not surprising that factors other than CTCF and Cohesin also play a role in shaping the 3D chromatin structure. In fact, Cohesin has many CTCF-independent binding sites, and CTCF is specifically not often found at promoter-enhancer elements, whereas Cohesin is [121,214]. Promoter-enhancer interactions might be mediated by Cohesin alone, by the interaction of Cohesin and the transcriptional co-activator complex Mediator [137], ESRRB [138], as well as the transcription factors YY1 [139], ZNF143 [140] and OCT4 [141], long noncoding RNAs [142] and the ChAHP complex [143]. Lately, also the histone chaperone Facilitates

chromatin transcription (FACT) was shown to contribute to the formation of Cohesin-dependent TADs in *Saccharomyces cerevisiae,* independently of transcription [215]*.* In addition, a recent preprint demonstrated that the acetyltransferase ESCO1 can contribute to boundary formation in chromatin looping, like CTCF [216]. Therefore, to completely abolish promoter-enhancer contacts, one would have to perhaps deplete far more proteins than only RAD21.

Despite the possibility that other factors might be involved in forming promoter-enhancer loops, TET enzymes might also be recruited to enhancers via other means, for example by PAR. Both CTCF and TET1 are targets of PARylation, and thus, PARylation might bring these two factors together [98,188]. In addition, PAR chains have been shown to decondense higher-order chromatin structures [217] and have been proposed to cause phase separation and thereby reorganize the 3D chromatin structure [218]. Therefore, I tested a possible involvement of PARylation in regulating TET activity at enhancers by inhibiting PARylation with Olaparib and PJ34. However, the inhibitors did not affect the enrichment of 5hmC at candidate enhancers (Figure 4.21 F), suggesting that PARylation does not play a major role in the recruitment of TET to these enhancers. Surprisingly, the inhibition of PARylation by PJ34 slightly increased global 5hmC levels (Figure 4.21 C). This is in contrast to what has been observed by Ciccarone and colleagues, who observed a decrease of global 5hmC after PJ34 treatment in HEK293T cells [98]. This difference might be due to the different cell types and different expression levels of TET proteins in mESC (high) and HEK293T (low), or different methods of 5hmC-detection (quantitative mass spectrometry used in this study vs. dot blot used by Ciccarone and colleagues). Although the PARylated Peroxisome proliferator-activated receptor γ (PPARγ) complex can recruit TET1 to PPAR response elements [219], this mechanism is likely not transferable to TET1 recruitment to enhancers in general.

As discussed in chapter 5.1, several other TET interactors could also be involved in the recruitment of TET to enhancers, among them GADD45A and the long noncoding RNA TARID [109,166]. This is consisent with the finding that GADD45A is bound at enhancers in mouse embryonic fibroblasts (MEFs) [151]. Transcription factor occupancy might also play an important role in TET recruitment, as transcription factors themselves bind to enhancers [220]. Examples include NANOG [99], PRDM14 [100] and EGR1 [101].

Another possibility is that TET proteins are not readily detected at enhancers due to technical limitations of the ChIP-seq method. In ChIP-seq, transient binding of TET to the DNA (e.g. to enhancers) might be undetectable and more stable binding sites of TET (e.g. promoters) might mask the transient binding sites. This issue has been raised for repressive transcription factors, which only transiently bind DNA and then vacate their binding regions [221]. Other techniques, such as DNA adenine methyltransferase identification (DamID)-seq might circumvent this problem, as DamID-seq does not only capture one snapshot of a protein binding to DNA, but monitors binding events over a specific timeframe. In addition, the method does not rely on protein-DNA crosslinking or antibody binding, which might introduce additional artefacts [222]. Recently, two new methods – Cleavage Under Targets and Release Using Nuclease (CUT&RUN) and Cleavage Under Targets and Tagmentation (CUT&Tag) – have been shown to have a better

resolution and a better signal-to-noise ratio than ChIP-seq; thus, also allowing detection of transient protein-DNA interactions [223,224]. As these methods have not been employed to investigate TET binding thus far, using these methods will give a more comprehensive understanding about TET binding sites in the future.

6 Material and Methods

6.1 Material

6.1.1 Equipment

-150°C freezer (Sanyo); -80°C freezer (Sanyo); agarose gel chambers (Bio-Rad); Agilent 6490 triple quadrupole mass spectrometer (Agilent Technologies); bacterial incubators (Thermo Scientific); bacterial shaker (Infors); balances (Sartorius); Bioanalyzer (Agilent); Bioruptor (Diagenode); cell counter (Bio-Rad); cell culture dishes and flasks (TPP); cell culture incubators (Thermo Scientific); centrifuges (Heraeus); ChemiDoc XR+ System (Bio-Rad), ChemiDoc XRS+ System (Bio-Rad), Confocal microscope TCS SP5 (Leica); coverslips (Thermo Fisher Scientific); Criterion TGX Precast Midi Protein Gels (Bio-Rad); Cryo-Safe Cooler (Belart); cryo tubes (Greiner Bio-One); DNA LoBind 1.5 ml Tubes (Eppendorf); DynaMag-2 Magnet rack (Invitrogen); E-Gel electrophoresis system (Invitrogen); extra thick blot filter paper (Bio-Rad); fridges (Liebherr); heating blocks (Eppendorf); laminar flow hoods (Dometic); LightCycler 480 (Roche); Low Binding 1.5 ml tubes (Biozym); LSRFortessa SORP cell analyzer (BD Biosciences); magnetic stirrer (Heidolph); microcentrifuges (Heraeus); microscope (Leica); micro tube (Covaris); microwave oven (Sharp); multidispenser pipette (Eppendorf); Nanodrop 2000 spectrophotometer (Thermo Scientific); NextSeq500 (Illumina); orbital shaker (Neolab); PAGE midigel chambers (Bio-Rad); PCR thermocyclers (Biometra); pH meter (Mettler Toledo); pipet boy (Integra); pipettes (Eppendorf); power supplies (Bio-Rad); PVDF transfer membrane (Neolab); Qubit (Thermo Fisher); rotator (Neolab); SpeedVac concentrator (Eppendorf); SuperFrost Plus Adhesion slides (Thermo Fisher Scientific); test tubes (Eppendorf, Falcon, Sarstedt); test tube rotator (Kisker Biotech), TPX microtubes (Diagenode); Trans-Blot SD semi-dry transfer cell (Bio-Rad); Trans-Blot Turbo (Bio-Rad); Trans-Blot Turbo Midi PVDF (Bio-Rad); tubes with cell-strainer cap (Falcon); ultrapure water purification system (Millipore); ultrasonicator S2/E210 (Covaris); vortexer (Scientific industries); water baths (Neolab)

6.1.2 Chemicals and pre-made Buffers

0.05 % Trypsin-EDTA (1x) (Gibco); 0.25 % Trypsin-EDTA (1x) (Gibco); 0.1 % gelatin in ultrapure water (Millipore); 15N dCTP (Silantes); agarose (Biozym); α-ketoglutaric acid disodium salt hydrate (Sigma-Aldrich); ammonium acetate (Sigma-Aldrich); ammonium iron(II) sulfate hexahydrate (Sigma-Aldrich); ampicillin (Sigma-Aldrich); ATP (NEB); boric acid (Sigma-Aldrich); bovine serum albumin (BSA) (Sigma-Aldrich); Buffer 2.1 (NEB); Calpeptin (Sigma); chloroacetamide (CAA) (Sigma-Aldrich); chloroform-isoamylalcohol (Roth); Cutsmart buffer (NEB); cycloheximide solution (Sigma-Aldrich); DAPI (Sigma-Aldrich); dialyzed FBS (Sigma-Aldrich); dimethyl sulfoxide (DMSO) (Sigma-Aldrich); dithiothreitol (DTT) (Sigma-Aldrich); DMEM for SILAC (Thermo Fisher Scientific); DMEM, HEPES (Sigma-Aldrich); high glucose, pyruvate, no glutamine (Gibco); DMEM, high glucose, no glutamine, no methionine, no cysteine (Gibco); DPBS, no calcium, no magnesium (Gibco); EDTA (Sigma-Aldrich); EGTA (Sigma-Adlrich); ES-grade FBS (PAN Biotech); ethanol (Sigma-Aldrich); ethidium bromide (Roth); fetal bovine serum (FBS) (Lonza); FLAG peptide (Sigma Aldrich); Fluorescence Mounting Medium (Dako); FreeStyle

293 Expression Medium (Gibco); gelatin (Sigma); GeneAmp PCR Buffer 1, 15 mM MgCl (Thermo Fisher Scientific); geneticin (Gibco); glycerol (Sigma-Aldrich); glycine (Sigma-Aldrich); hygromycin B (Sigma-Aldrich); hydrochloric acid (Sigma-Aldrich); indole-3-acetic acid sodium salt (Sigma-Aldrich); isopropanol (Sigma-Aldrich); L-ascorbic acid (Sigma-Aldrich); L-Cysteine (Sigma-Aldrich); LIF (Homemade cell supernatant); lithium chloride (Sigma-Aldrich); L-Glutamine 100x 200 mM (Gibco); L-Methionine-(methyl-$^{13}C,d_3$) (Sigma-Aldrich); magnesium chloride (Sigma-Aldrich); MEM Non-essential amino acids 100x (Gibco); methanol (Sigma-Aldrich); MG132 (Tocris); N-Ethylmaleimide (NEM) (Sigma-Aldrich); Nonidet P 40 Substitute (Sigma-Aldrich); nuclease-free water (Qiagen); NuPAGE LDS Sample Buffer (4X) (Invitrogen); Olaparib (AZD2281, Ku-0059436) (Selleckchem); Optimem (Gibco); paraformaldehyde (Merck Millipore); PEI MAX 4000 (Polysciences); penicillin/streptomycin 10,000 U/ml (PAN); phenol-chloroform-isoamylalcohol (Roth); Phosstop (Roche); PJ-34 hydrochloride hydrate (Sigma-Aldrich); Plasmocin (Invitrogen); potassium hydroxide (Sigma-Aldrich); propidium iodide (Sigma-Aldrich); Protease inhibitor cocktail tablets EDTA-free (Roche); puromycin (Sigma-Aldrich); Quick Start Bradford 1x Dye Reagent (Bio-Rad); random primers (Invitrogen); SB202190 (Sigma-Aldrich); skim milk powder (Sigma-Aldrich); SAM (NEB); sodium acetate (Sigma-Aldrich); sodium chloride (Sigma-Aldrich); sodium citrate tribasic dihydrate (Sigma-Aldrich); sodium deoxycholate (Sigma-Aldrich); sodium dodecyl sulfate (Sigma-Aldrich); sodium hydroxide (Sigma-Aldrich); sodium phosphate dibasic dihydrate (Sigma-Aldrich); sodium phosphate monobasic monohydrate (Sigma-Aldrich); sodium pyruvate 100 mM (Gibco); ß-mercaptoethanol (Sigma-Aldrich); thymidine (Sigma); Trizma base (Sigma-Aldrich); Trizma hydrochloride (Sigma-Aldrich); Triton X-100 (Sigma-Aldrich); Tween-20 (Sigma-Aldrich); U0126 (Sigma-Aldrich); urea (Sigma-Aldrich); xylene cyanol (Sigma-Aldrich); yeast tRNA (Roche)

6.1.3 Kits

ANTI-FLAG M2 Affinity Gel (Sigma-Aldrich); DNAse I (Roche); DNeasy 96 Blood & Tissue kit (Qiagen); DNeasy Blood & Tissue kit (Qiagen); Dynabeads Protein A and G (Thermo Fisher Scientific); GFP-Trap Agarose (Chromotek); Gibson Assembly Master Mix (NEB); High Sensitivity DNA Assay (Agilent); LightCycler 480 Probes Master (Roche); LightCycler 480 SYBR Green I Master (Roche); Lipofectamine 2000 (Invitrogen); PCR Mycoplasma Test Kit I/C (Promokine); Plasmocin (ant-mpt-1) (InvivoGen); Protein A Agarose (Roche); PvuII (Promega); Qiaprep Maxiprep kit (Qiagen); Qiaprep Midiprep kit (Qiagen); Qiaprep Miniprep kit (Qiagen); Qiaquick Gel extraction kit (Qiagen); QIAquick PCR purification kit (Qiagen); QIAshredder (Qiagen); Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific); QuikChange II XL Site-Directed Mutagenesis Kit (Agilent); Reverse Transcriptase (CF IMB); RNeasy 96 kit (Qiagen); RNeasy Mini Kit (Qiagen); Superscript II Reverse Transcriptase (Invitrogen); SuperSignal West Femto/Pico (Thermo Fisher Scientific); X-tremeGENE 9 (Roche)

6.1.4 Enzymes

AccuPrime Taq (Thermo Fisher Scientific); alkaline phosphatase (Fermentas); BamHI HF (NEB); CpG Methyltransferase (M.SssI) (NEB); MluI (NEB); NotI (NEB); nuclease P1 (Roche); Phusion High-Fidelity DNA Polymerase (NEB); Proteinase K (Roche); RNAse A (Qiagen); RNAse free

DNAse set (Qiagen); SacI (NEB); SalI (NEB); SmaI (Thermo Fisher Scientific); snake venom phosphodiesterase (Worthington), Taq (CF IMB); XbaI (NEB)

6.1.5 Buffers and Solutions

Solution	Components
Blastocyst Lysis Buffer	100 mM Tris HCl pH 8, 100 mM KCl, 0.02% gelatin, 0.45% Tween-20, 60 µg/ml yeast tRNA, 125 µg/ml Proteinase K
Blotting Buffer (Bjerrum Schafer Nielsen)	48 mM Tris, 39 mM glycine, pH 9.2, 20 % methanol
CL Lysis Buffer	150 mM NaCl, 50 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.5 % Triton X-100, 1x Protease inhibitor cocktail, 1x Phosstop, 1 mM Na_3VO_4, 20 mM NaF
Denaturing Lysis Buffer	1x PBS, 1 % (wt/vol) SDS, 5 mM EDTA, 5 mM EGTA, 10 mM NEM, 1x Protease inhibitor cocktail, 1x Phosstop, 1 mM Na_3VO_4, 20 mM NaF
DNA loading buffer (5x)	for 50 ml total volume in water: 25 ml glycerol, 0.5 ml EDTA (0.5 M), 0.5 ml Tris-HCl pH 8 (0.5 M), pinch of xylene cyanol,
dNTP mix	10 mM dATP, 10 mM dCTP, 10 mM dGTP, 10 mM dTTP
FLAG elution buffer	50 mM Tris-HCl pH 7.5,150 mM NaCl, 0.5 % Triton X-100, 100 µg/ml FLAG peptide; 1x Protease inhibitor cocktail, 1x Phosstop, 1 mM Na_3VO_4, 20 mM NaF
GFP-Trap Dilution Buffer	10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 10 mM NEM, 1x Protease inhibitor cocktail, 1x Phosstop, 1 mM Na_3VO_4, 20 mM NaF
IP Buffer (1x)	10 mM Na-phosphate buffer pH 7, 140 mM NaCl, 0.05 % Triton X-100
IVA3 Buffer	150 mM NaCl, 20 mM Tris-HCl pH 7.5, 0.5 % Nonidet P 40 Substitute, 10 % glycerol, 1x Protease inhibitor cocktail, 1 mM DTT
IVA5 Buffer	500 mM NaCl, 50 mM Tris-HCl pH 7.5, 1 % Nonidet P 40 Substitute, 1x Protease inhibitor cocktail, 1 mM DTT
Luria broth (LB)	for 1 l total volume in water: 10 g bactotryptone, 5 g yeast extract, 10 g NaCl pH 7.0, autoclaved
Lysis Buffer A	10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 1 mM DTT, 1x Protease inhibitor cocktail
Lysis Buffer B	20 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 420 mM NaCl, 0.5 mM EDTA, 25 % glycerol, 1 mM DTT, 1x Protease inhibitor cocktail
Na-phosphate buffer (1 M)	39 ml 2 M NaH_2PO_4 (276 g/L), 61 ml 2 M $NaHPO_4$ (284 g/L), 100 ml H_2O, pH 7
4x NuPAGE+DTT	4x NuPAGE LDS sample buffer including 400 mM DTT
PBS	140 mM NaCl, 2.7 mM KCl, 1.5 mM KH_2PO_4, 8.1 mM Na_2HPO_4, pH 7.4, autoclaved
PBS-0.05 % Tween	1x TBS, 0.05 % Tween-20
PBS-BSA 0.1 %	1x PBS, 0.1 % BSA
PEI MAX (1 µg/µl)	1 µg/µl PEI MAX in H_2O, pH 7.5, sterile filtered
PEX-AS2	150 mM NaCl, 20 mM Tris pH 7.5, 2 mM EDTA, 10 % Glycerol, 1 % Triton X-100, 1x Protease inhibitor cocktail Optional: 10 mM NEM, 1x Phosstop, 1 mM Na_3VO_4, 20 mM NaF

Proteinase K digestion buffer	50 mM Tris HCl pH 8, 10 mM EDTA, 0.5 % SDS
Reaction buffer	7 ng/µl ^{15}N-labeled methylated DNA substrate, 50 mM HEPES pH 8.0, 75 µM ammonium iron(II) sulfate hexahydrate, 1 mM α-ketoglutaric acid disodium salt hydrate, 2 mM L-ascorbic acid, 50 mM NaCl, 1 mM DTT, 1 mM ATP
Reaction wash buffer	50 mM HEPES pH 8.0, 2 mM L-ascorbic acid, 100 mM NaCl, 1x Protease inhibitor cocktail, 1 mM DTT
RIPA buffer	20 mM Na-phosphate buffer, 150 mM NaCl, 1 % (vol/vol) Triton-X 100, 0.5 % (wt/vol) sodium deoxycholate, 0.1 % (wt/vol) SDS, 5 mM EDTA, 5 mM EGTA, 10 mM NEM, 1x Protease inhibitor cocktail, 1x Phosstop, 1 mM Na_3VO_4, 20 mM NaF
RIPA dilution buffer	20 mM Na-phosphate buffer, 150 mM NaCl, 1 % (vol/vol) Triton-X 100, 0.5 % (wt/vol) sodium deoxycholate, 5 mM EDTA, 5 mM EGTA, 10 mM NEM, 1x Protease inhibitor cocktail, 1x Phosstop, 1 mM Na_3VO_4, 20 mM NaF
SDS-Running buffer 10x	for 1 l total volume: 10 g SDS, 30.3 g Tris, 144.1 g glycine
TBE (10x)	1 M Trizma base, 1 M boric acid, 20 mM EDTA, autoclaved
TBS (20x)	3 M NaCl, 53.7 mM KCl, 839 M Trizma hydrochloride, 160 mM Trizma base pH 7.4, autoclaved
TBS-T	1x TBS, 0.1 % Tween-20
TE buffer	10 mM Tris-HCl pH 8.0, 1 mM EDTA, autoclaved
Wash Buffer 1	8 M urea, 1 % (wt/vol) SDS in 1x PBS
Wash Buffer 2	1 % SDS in 1x PBS

Table 6.1 Buffers and Solutions

6.1.6 Antibodies

Antibody	Host species	Company	Catalogue #	Dilution WB
5-Formylcytosine	rabbit	Active Motif	61227	-
5-Hydroxymethylcytosine	rabbit	Active Motif	39791	-
Alpha-Tubulin	mouse	Sigma	T5168	1:5000
FLAG	mouse	Sigma-Aldrich	F3165	1:1000
FLAG	rabbit	Sigma-Aldrich	F7425	1:2000
GADD5A (H-165)	rabbit	Santa Cruz	sc-797	1:2000
GFP	rabbit	Thermo Fisher Scientific	A11122	1:2000
HA, ChIP grade	rabbit	Abcam	ab9110	1:4000
HNRNPA1, clone 4B10	mouse	Santa Cruz	sc-32301	1:5000
IgG ChIP grade control	rabbit	Abcam	ab46540	-
Mouse IgG, Alexa Flour 488	goat	Thermo Fisher Scientific	A11029	1:1000 (IF)
Mouse HRP	goat	Dianova	115-035-146	1:10000
Myc, clone 9E10	mouse	Sigma-Aldrich	M4439	1:2000
PAR	rabbit	Trevigen	4336-BPC	1:2000
PARP-1 N-terminal	rabbit	Active Motif	39560	1:2000
Rabbit HRP	goat	Dianova	111-035-144	1:10000
RAD21 clone 53A303	mouse	Merck Millipore	05-908	1:2000 (WB)/ 1:1000 (IF)
SUMO 2,3,4 (C3)	mouse	Santa Cruz	Sc-393144	1:500
TATA binding protein TBP	mouse	Abcam	ab818	1:2000
Ubiquitin	mouse	Santa Cruz	sc-8017	1:1000
V5	mouse	Invitrogen	46-1157	1:2000

Table 6.2 Antibodies
WB, dilution used for Western blot; IF, dilution used for immunofluorescence staining

6.1.7 Oligonucleotide sequences

6.1.7.1 RT qPCR expression primers

Target gene	Forward primer (5'-3')	Reverse Primer (5'-3')	UPL Probe
GADD45A (human)	GCCAAGCTGCTCAACGTC	CTCTGTCGTCGTCCTCGTC	40
Gadd45a (mouse)	GCTGCCAAGCTGCTCAAC	TCGTCGTCTTCGTCAGCA	98
GAPDH	GCATCCTGGGCTACACTGAG	AGGTGGAGGAGTGGGTGTC	82
MAPK11	CATGCTCAACTGGATGCATTA	CGCTTCAGCTGGTCAATGTA	37
MAPK12	CTCCTTTGACGACGTTGACC	TGAAGCTGAGCACCTCTTTGT	47
MAPK13	CAGGACGTCAACAAGACAGC	TTCTTGATGGCCACCTTCTC	24
MAPK14	TGGTCTGTTGGACGTTTTTACA	CCATGAGATGGGTCACCAG	65
MAP2K3	GAGGACATCCTTGGGGAGAT	GTGGATCACCGACAGCTTG	1
MAP2K4	GGCCAAAGTATAAAGAGCTTCTGA	CAGCGATATCAATCGACATACAT	33
MAP2K6	CAAGGCTTGCATTTCTATTGG	CCAGTTCCATTATAGGCTCCA	8
MAP2K7	CGGAGGATCGACCTCAAC	GGGAGCTCTCTGAGGATGG	24
MAP3K5	CACGTGATGACTTAAAATGCTTG	AGTCAATGATAGCCTTCCACAGT	89
MAP3K6	GGCTCTGGATGCCTTCTACA	ACAGGGAGGGCTGTACCA	67
MAP3K7	AGCTTTATGGAGCCTGCTTG	AGAGCCCCCTTCAGCATATT	47
MTK1	GCCTTGAGAGTGACCCAAAG	GCTTCATTCTTCATCTGTGCAA	17
Tbp (mouse)	GGGGAGCTGTGATGTGAAGT	CCAGGAAATAATTCTGGCTCA	97
TET1CD	AATGGCTGTAAGTTTGGTAGAAGC	CGTGTAGCCAAACTCTGTAAGTTATC	45
TRIM24	ACGGTCCAGTCACCAAATTC	GGGGGTGTACACTAGTCATAGTAACAG	80
TRIM25 (human)	TCGAGAGCACTGGATGATGT	TTGTTGTAGCTGCTCCACCTT	18
Trim25 (mouse)	CCCTACGACCCTAAGTCAAGC	TGTGGCTGTGCATGATAGTG	85
TRIM28 (human)	TGGTCAATGATGCCCAGA	CTTGGTCATGGTCCAGTGC	1
Trim28 (mouse)	CAACCAGCGGAAATGTGAG	GGTACCACCAGGCTGCTC	2
TRIM33	TGCCCTGTCTTCACTCCTTC	GGGCACCGTATTACACCAAC	27

Table 6.3 RT qPCR expression primers

6.1.7.2 Genotyping primers

Genotyping PCR	Forward primer (5'-3')	Reverse primer (5'-3')
mmRAD21_mAID_neo	TATGACTGGGCACAACAGAC	CCTAGGCAGGTTTTCCTTTC
mmRAD21_mAID_neo_nested	AGGATCTCCTGTCATCTCAC	AGTCTTCTGGGTTCTTGGGC
mmRAD21_mAID_hygro	TGGCTGTGTAGAAGTACTCG	CCTAGGCAGGTTTTCCTTTC
mmRAD21_mAID_hygro_nested	GTCCGAGGGCAAAGGAATAG	AGTCTTCTGGGTTCTTGGGC
OsTir_Tigre_Cterm	TCACCTCGACCCATGGTAATA	TGTACCACAGGGCGTTTCA
OsTir_Tigre_Nterm	TCTGGCTAGACAGGGATCTTACA	ACTCCCACTGTCCTTTCCTAATAA
Tigre integration 1	GCTTCCCCAACGGTCACTTA	CCCAGGAGGGACGGTTACTA
Tigre integration 2	GGGGTGGGAATCAAGCCATA	GGCGTGAGTCGGAGAACTAA

Table 6.4 Genotyping primers

6.1.7.3 DIP-qPCR primers

Target site	Forward primer (5'-3')	Reverse Primer (5'-3')	UPL Probe
Negative control	GAAATCAAGGAGGCCTGAAA	TCTAAATAGCTCTGAGGGACCTG	68
Positive control	CTGAGTGGCAGGCGTATTC	AATCTCAAAAGTTGCCAAGCA	33
Adam15_enhancer	CCCCGGATTTTATCCTCGTA	CAGATGCTGTCGACAATCCA	53
Adam15_promoter	TTTGGACACCACAAGACTCG	GTGAGCCGGAAGAACTGG	16
Atxn7l3_enhancer	CCCTCAGGCATTCTTGCTAA	GGTGAGGGTGCTGAAAAGTC	95
Atxn7l3_promoter	TCCCGACCCCACTAGGTC	CCCCGGGACAATAGAGACA	55
Chd4_enhancer	GCCGGCTCATCACACAGT	GACCATCGGGGCCTATTC	69
Chd4_promoter	GGTGTTTCGGGCCCTTTA	GTCTCGCCGGTGTCACTC	17
Hsp90ab1_enhancer	CGAGGTAGCCATTGGAGAAG	CTTTAAGCATGACGCCTGGT	46
Hsp90ab1_promoter	AGCCCACTTCTTGGGAGACT	GTGGGAAGGGATTAAGACAGG	88
Myh9_enhancer	TGCAGTCTCTGCATTTCCTG	AGGAGCCTCCCTTAGTCACAG	99
Myh9_promoter	CTTCCGAGTGGACTTTCTCG	GCACGGAAGGCTCAAGAAC	62
Ppp1r9b_enhancer	TACGTGGCAGGCAGTTGA	CTCTTCCAACCCGCTGAC	33
Ppp1r9b_promoter	TCCCTTTAGGACCCAGTGGT	TCTAGCTCTGCCCGAACC	70

Table 6.5 DIP-qPCR primers

6.1.7.4 Cloning primers

Plasmid name	Fragment name	Primer name	Primer sequence (5'-3')
pEF FLAG-HA TET1 CD	FLAG-HA	FLAG_HA_fwd	TGGCTAGGTAAGCTTGGTACATGGACTACAAGGACGACGATG
		FLAG_HA_rev 2	AGGTGGGCAGTTCTGAGTCGAGTCCTCCGGCGTA
	TET1CD	TET1CD_fwd 2	AGGACTCGACTCAGAACTGCCCACCTGC
		TET1CD_rev	GCTTACCTTCGAAGGGCCCTTCAGACCCAATGGTTATAGGG
pEF FLAG-HA GADD45A-TET1 CD	pEF	pEF_fwd	TCTAGAGGGCCCTTCGAAG
		pEF_rev	GTCGAGTCCTCCGGCGTA
	GADD45A	GADD45A_fwd	GCCCGACTACGCCGGAGGACTCGACACTTTGGAGGAATTCTCGGCTGGAGAG
		GADD45A_rev	GCAGTTCTGAATTCGACCCACCACCGCCCGA
	TET1CD_3	TET1CD_fwd	CGGTGGTGGGTCGAATTCAGAACTGCCCACC
		TET1CD_rev	GCTTACCTTCGAAGGGCCCTCTAGATCAGACCCAATGGTTATAGG
pEF GFP-TEV TET1	pEF_2	pEF_fwd	TCTAGAGGGCCCTTCGAAG
		pEF_4_rev	GGTGGCGGTACCAAGCTT
	eGFP-TEV	eGFP_TEV_2_fwd	GCTAGGTAAGCTTGGTACCGCCACCATGGTGAGCAAGGGCGAG
		eGFP_TEV_2_rev	GGCGGGATCGAGAGCGATCGCCTCGAGATCT
	TET1	TET1_2_fwd	TCGAGGCGATCGCTCTCGATCCCGCCATGCA
		TET1_2_rev	GCTTACCTTCGAAGGGCCCTCTAGATCAGACCCAATGGTTATAGGGCC
pEF GFP-TEV-TET1CD	TET1CD in pEF	pEF_5_fwd	TCTGAACTGCCCACCTGC
		pEF_rev_4	GGTGGCGGTACCAAGCTTA
	eGFP_TEV	eGFP_TEV3_fwd	CTTGGTACCGCCACCATGGTGAGCAAGGGCGAG
		eGFP_TEV3_rev	GGTGGGCAGTTCAGAGCGATCGCCTCGAGATCT
pCS2 FLAG-TET1CD K1813A		TET1_K1813A_fwd	AGACCGTGCAACCTGAAGTAGCAAGTGAAACCGAACCCC
		TET1_K1813A_rev	GGGGTTCGGTTTCACTTGCTACTTCAGGTTGCACGGTCT
pCS2 FLAG-TET1CD K1842R		TET1_K1842R_fwd	TCCGCTCCTCACCCAGTGAGAGAGGCATCT
		TET1_K1842R_rev	AGATGCCTCTCTCACTGGGTGAGGAGCGGA
pCS2 FLAG-TET1CD K1863A		TET1_K1863A_fwd	CACACCAGCTCCACTGGCGAATGACGCAACAGCC
		TET1_K1863A_rev	GGCTGTTGCGTCATTCGCCAGTGGAGCTGGTGTG

pCS2 FLAG-TET1CD K2067R	TET1_K2067R_fwd	CCCAACATGGTTTTGAACTAAACAAGATTAGGTTTGAGGCTAAAGAAG
	TET1_K2067R_rev	CTTCTTTAGCCTCAAACCTAATCTTGTTTAGTTCAAAACCATGTTGGG

Table 6.6 Cloning Primers

6.1.7.5 Oligonucleotides used for CRISPR/Cas9 editing

guide RNA	Forward Oligo (5'-3')	Reverse Oligo (5'-3')
CRISPR_mmRad21_mAID	CACCGTTTTGGTGCCCTGCACACG	AAACCGTGTGCAGGGCACCAAAAC

Table 6.7 guide RNA sequence

Gene Synthesis fragment	Sequence (5'-3')
CRISPR_mmRad21_mAID	ATGGTCTTCAGCGAGCTCTTGCTAAAACTGGAGCAGAGTCTATCAGTTTGCTTGAGCTGTGTCGAAACACAAACCGAAAGCAGGCAGCAGCAAAGTTCTACAGCTTTTTGGTTCTTAAGAAGCAGCAAGCCATCGAGCTCACACAGGAAGAGCCGTACAGTGACATCATTGCAACCCCTGGACCACGGTTCCATATTATCGGATCCGGAGCTAGATGTGTTCGAGCTAGTGATAACTCACTAGTACATACAAATTGCCCCCGAGTCCAGTGCGCCTAAGCCCTTTAAGAAAGTTTTTAGATTTCTGTTTGTACAAAAATCTTTGCCTTTTCTTTCTTCTTTTTCCCCCCAGTGTTTCTAATTTTGTCAACCATATTTTTAAGGGAAACTGCTTATTTGGGTTGGGT

Table 6.8 Homology arm synthesis fragment

6.1.7.6 Miscellaneous Oligonucleotides

PCR reaction	Forward Primer (5'-3')	Reverse Primer (5'-3')
IVO biotin	[BtnTg]GATAAGCTTGGGATCTCTATAATCTCG	[BtnTg]CAGCTTGTGCCCCAGGAT

Table 6.9 Miscellaneous PCR Primers

6.1.8 siRNAs

Target Gene	siRNA ID	Supplier
GADD45A (human)	VHS40261, stealth siRNA	Thermo Fisher Scientific
Gadd45a (mouse)	Set of 3: MSS236318, MSS236319, MSS236320	Thermo Fisher Scientific
MAPK11	HSS183382, stealth siRNA	Thermo Fisher Scientific
MAPK12	HSS109467, stealth siRNA	Thermo Fisher Scientific
MAPK13	VHS40525, stealth siRNA	Thermo Fisher Scientific
MAPK14	VHS40416, stealth siRNA	Thermo Fisher Scientific
MAP2K3	VHS40371, stealth siRNA	Thermo Fisher Scientific
MAP2K4	HSS184677, stealth siRNA	Thermo Fisher Scientific
MAP2K6	HSS183392, stealth siRNA	Thermo Fisher Scientific
MAP2K7	HSS140876, stealth siRNA	Thermo Fisher Scientific
MAP3K5	VHS40812, stealth siRNA	Thermo Fisher Scientific
MAP3K6	VHS40987, stealth siRNA	Thermo Fisher Scientific
MAP3K7	VHS40542, stealth siRNA	Thermo Fisher Scientific
siRNA Negative Control, Med GC	12935113, stealth siRNA	Thermo Fisher Scientific
Non-Targeting siRNA Pool #1	D-001206-13-50, siGENOME	Dharmacon
TRIM24	HSS112966, stealth siRNA	Thermo Fisher Scientific
TRIM25 (human)	HSS111608, stealth siRNA	Thermo Fisher Scientific
Trim25 (mouse)	M-047483-01, siGENOME SMARTpool	Dharmacon
TRIM28 (human)	HSS115468, stealth siRNA	Thermo Fisher Scientific
Trim28 (mouse)	M-040800-01, siGENOME SMARTpool	Dharmacon
TRIM33	HSS122396, stealth siRNA	Thermo Fisher Scientific

Table 6.10 siRNAs

6.1.9 Plasmids

Name	Source	Description
pCAG-GFP-TEV-mTET1	Frauer et al [225]	N-terminally GFP tagged mouse full-length TET1, with TEV Protease cleavage site between GFP and mTET1
pcDNA3 HA-Ubiquitin	Kamitani et al [226]	HA-tagged Ubiquitin
pcDNA3 HA-SUMO2	Frauke Melchior	HA-tagged SUMO2
pCS2+_GFP	Andrea Schäfer	GFP overexpression plasmid
pCS2FLAG/hTET1FL	Lars Schomacher	N-terminally 2x FLAG tagged full-length human TET1
pCS2FLAG/hTET1CD	Lars Schomacher	N-terminally 2x FLAG-tagged catalytic domain of human TET1 (amino acids 1416-2136)
pCS2 FLAG-TET1CD K1813A	this study	N-terminally 2x FLAG-tagged catalytic domain of human TET1 (amino acids 1416-2136) with a mutation at amino acid 1813 from lysine to alanine
pCS2 FLAG-TET1CD K1842R	this study	N-terminally 2x FLAG-tagged catalytic domain of human TET1 (amino acids 1416-2136) with a mutation at amino acid 1842 from lysine to arginine
pCS2 FLAG-TET1CD K1863A	this study	N-terminally 2x FLAG-tagged catalytic domain of human TET1 (amino acids 1416-2136) with a mutation at amino acid 1863 from lysine to alanine
pCS2 FLAG-TET1CD K2067R	this study	N-terminally 2x FLAG-tagged catalytic domain of human TET1 (amino acids 1416-2136) with a mutation at amino acid 2067 from lysine to arginine
pCS2FLAG/hTDG	Lars Schomacher	N-terminally 2x FLAG-tagged human TDG
pCS2+myc-HsGadd45a	Mathias Gierl	N-terminally myc-tagged human GADD45A
pCS2-myc-hMTK1	Andrea Schäfer	N-terminally myc-tagged human MTK (MAP3K4)
Flag-HA-hTET1-pEF	Tahiliani et al [22]	Full-length human TET1 fused to a N-terminal Flag-HA tag
pEF FLAG-HA TET1CD	this study	Expression plasmid for the catalytic domain of human TET1, amino acids 1416-2136, with an N-terminal FLAG-HA tag
pEF FLAG-HA GADD45A-TET1CD	this study	Expression plasmid for a fusion construct of human GADD45A and the catalytic domain of human TET1 (amino acids 1416-2136) with an N-terminal FLAG-HA tag
pEF GFP-TEV TET1	this study	N-terminally GFP-tagged human full-length TET1
pEF GFP-TEV-TET1CD	this study	N-terminally GFP-tagged human TET1CD
pEN396_Tigre OsTir1	Nora et al [147]	Targeting vector to introduce an *OsTir1* expression cassette at the mouse *Tigre* acceptor locus using puromycin selection

pMA-T_mmRad21_homology arms	Thermo Fisher Scientific	Homology arms of *mmRad21* for C-terminal tagging with mAID-3xHA using CRISPR/Cas9
pMK_mmRad21_mAID_Neo	Pia Schäfer	Targeting vector for C-terminal tagging of *mmRad21* with mAID-3xHA using CRISPR/Cas9, contains neomycin resistance gene
pMK_mmRad21_mAID_Hygro	Pia Schäfer/this study	Targeting vector for C-terminal tagging of *mmRad21* with mAID-3xHA using CRISPR/Cas9, contains hygromycin resistance gene
pX330_Cas9_gRNA Rad21	Pia Schäfer	Expression plasmid for Cas9 and *Rad21* gRNA (5`-GTTTTGGTGCCCTGCACACG-3')
pX330_EN1201_Cas9_gRNA Tigre	Nora et al [147]	Expression plasmid for Cas9 and *Tigre* gRNA (5'-ACTGCCATAACACCTAACTT-3')

Table 6.11 Plasmids

6.2 Methods

Unless stated otherwise, incubations were performed at room temperature (~20 °C).

6.2.1 Molecular Biology methods

Standard molecular biology methods including preparation of chemically competent XL1-blue *Escherichia coli* (*E. coli*) bacteria, plasmid amplification in *E. coli*, spectrophotometric quantification of DNA and RNA, restriction digests, PCR, agarose gel electrophoresis, and SDS-PAGE were carried out as previously described [227]. All oligonucleotides used in this study were synthesized by Sigma-Aldrich or Integrated DNA Technologies (IDT). Plasmid DNA was sequenced by GATC Biotech or StarSEQ.

6.2.1.1 Amplification and purification of plasmids from bacteria

For plasmid amplification, 0.5 µg plasmid DNA was transformed into freshly thawed 50 µl XL-1 blue *E. coli* chemically competent bacteria by a 30 s heat shock at 42 °C. This was followed by a 2 min incubation on ice, subsequent incubation in 950 µl LB medium, which contained an appropriate selection antibiotic, for 1 h at 37 °C. 100 µl of bacteria suspension was plated on an LB-Agar plate (containing an appropriate selection antibiotic) and incubated at 37 °C overnight. The next day, a single clone was picked and used to inoculate LB medium (volume depending on downstream applications), containing an appropriate selection antibiotic, and incubated overnight. Plasmid DNA was purified using Qiagen Miniprep, Midiprep or Maxiprep kits according to the manufacturer's recommendations. DNA amount and purity were measured on a Nanodrop 2000 spectrophotometer.

6.2.1.2 Cloning of plasmids

Plasmids generated for this study were cloned using the Gibson Assembly kit from NEB, according to the manufacturer's recommendations. In brief, PCR primers were designed using the NEBuilder Assembly Tool, PCRs were performed using Phusion polymerase, purified using Qiaquick Gel extraction kit, and ligated using the Gibson Assembly Master Mix. Plasmids were then amplified, purified, analyzed by restriction digest and sequenced to confirm correct insertion.

Site-directed mutagenesis was carried out using the QuikChange II XL Site-Directed Mutagenesis Kit, according to the manufacturer's recommendations.

6.2.1.3 Protein extraction and Western blot

Cultured cells were harvested either by Trypsin digestion, or by cell scraping in PBS-0.05 % Tween, transferred to a 15 ml or 50 ml falcon tube, and pelleted by centrifugation for 5 min at 300 g.

For whole cell lysates, cell pellets were resuspended in RIPA buffer or PEX-AS2 buffer. Cells were incubated on ice, and cell disruption was achieved by 5 times needle homogenization using a 20 g needle or by sonication for 20 cycles 15 s on/off, high setting, using a Bioruptor. Samples were cleared by centrifugation for 10 min at 21,000 g at 4 °C.

For nuclear and cytoplasmic fractionation, cells were resuspended in Lysis Buffer A and incubated for 10 min on ice, followed by centrifugation for 1 min at 21,000 g at 4 °C. The supernatant contained the cytoplasmic fraction and the pellet was resuspended in Lysis Buffer B, followed by a 20 min incubation on ice. Both fractions were sonicated for 20 cycles 15 s on/off, high energy, using a Bioruptor, followed by centrifugation for 10 min at 21,000 g at 4 °C to clear the lysates.

Protein concentration was determined by Bradford Assay, by incubating 1-5 µl of protein lysate and 250 µl Quick Start Bradford 1x Dye for 5 min and measuring the absorption on a Nanodrop 2000 spectrophotometer, which was compared to a standard curve of BSA samples with known protein concentrations, and was run in parallel to the samples of interest.

Defined protein amounts of cell lysates were mixed with 4x NuPAGE+DTT and incubated for 10 min at 70 °C. SDS-PAGE, transfer to PVDF membranes and Western blotting was performed according to standard protocols [227]. Signals were developed with SuperSignal West Pico or Femto Chemiluminescent Substrate and analyzed using a ChemiDoc with Image Lab software. Quantification of band intensity was performed using Image Lab Software.

6.2.1.4 Immunofluorescence staining and Microscopy

mESCs grown on coverslips in 12-well plates were rinsed with PBS and fixed for 15 min in 4 % paraformaldehyde on an orbital shaker at 45 rpm. Coverslips were washed 3 x 5 min in PBS + 1 mg/ml BSA and permeabilized with 0.5 % Triton X-100 in PBS + 1 mg/ml BSA for 1 h. After 3 x 5 min washes in PBS + 1 mg/ml BSA, the coverslips were blocked with 0.1 % Triton X-100 in PBS + 1 mg/ml BSA for 1 h. Subsequently, incubation with 100 µl primary antibody RAD21 was performed overnight at 4 °C in a humid chamber, at a 1:1000 dilution in 0.1 % Triton X-100 in PBS + 1 mg/ml BSA. Coverslips were washed 3 x 10 min in PBS + 1 mg/ml BSA and then incubated with 1:1000 diluted Alexa Fluor 488-coupled secondary antibody in 0.1 % Triton X-100 in PBS + 1 mg/ml BSA, followed by 3 x 10 min washes in PBS + 1 mg/ml BSA. To stain nuclei, coverslips were incubated with 1:10,000 DAPI in PBS + 1 mg/ml BSA for 1 h, followed by 2 x 5 min washes in PBS. Coverslips were mounted on glass slides with Dako mounting medium and sealed with nail polish after 10 min. Immunofluorescence stainings were analyzed with a Leica SP5 confocal microscope and Leica LAS image software.

6.2.1.5 RNA isolation, reverse transcription and qPCR

RNA was isolated using the QIAshredder and RNeasy Mini kits, or the RNeasy 96-well kit from Qiagen, according to the manufacturer's instructions. The optional DNAse I on-column digestion was carried out and RNA was eluted in nuclease-free water and measured on a Nanodrop 2000.

cDNA synthesis was carried out using Superscript II reverse transcriptase according to the manufacturer's instructions. Briefly, 1000 ng RNA, 1 µl 10 mM dNTPs and 360 ng random primers were mixed in a final volume of 12 µl. After denaturation for 5 min at 65 °C the mixture was cooled on ice for 2 min. 4 µl 5x FS buffer, 2 µl 0.1 M DTT, 1 µl Ribolock and 1 µl Superscript II Polymerase were added and samples were incubated for 10 min at 25 °C, 90 min at 42 °C for cDNA synthesis and 15 min at 70 °C for enzyme inactivation. Resulting cDNA was diluted 1:5 in TE buffer. All incubation steps were carried out in a PCR thermocycler.

For each 11 µl qPCR reaction, 5 µl cDNA, 5.5 µl 2x Probes Master, 0.11 µl 100 µM forward/reverse primer mixture, 0.11 µl UPL probe and 0.28 µl ddH_2O were mixed. PCR reactions were performed in technical duplicates in a 384-well format in the Roche Light Cycler 480 using the following PCR program:

Temperature (°C)	Time	Ramp rate	
95	10 min	4.8 °C/s	
95	10 s	4.8 °C/s	50 x
60	20 s	2.5 °C/s	
72	1 s	4.8 °C/s	
4	1 min	2.5 °C/s	

Table 6.12 qPCR program

Cp values were determined using the Roche LightCycler software and relative expression values were calculated according to the ddCp method [228], normalizing to *GAPDH* or *Tbp* as housekeeping genes and including normalization for primer efficiencies. UPL-compatible primers were designed using the Roche Assay Design center.

6.2.1.6 Genotyping PCRs

For CRISPR/Cas9 clone genotyping, cells from the equivalent of 1/2 well of a 96-well cell culture plate were transferred to a 96-well genotyping PCR plate and incubated in 50 µl blastocyst lysis buffer for 30 min at 37 °C, 10 min at 56 °C and 10 min at 95 °C. Afterwards, a PCR reaction was set up with the following components and PCR program:

Component	Volume (µl)
Cell lysate	5
10x PCR Buffer 1 including 15 mM $MgCl_2$	2
dNTPs (10 mM)	0.5
Primer forward (10 µM)	0.5
Primer reverse (10 µM)	0.5
Nuclease-free H_2O	11
Taq	0.5

Table 6.13 Genotyping PCR components

Temperature (°C)	Time	
94	30 s	
94	20 s	30 x
51-54	30 s	
72	1 min – 1 min 45 s	
72	10 min	
4	5 min	

Table 6.14 Genotyping PCR program

6.2.1.7 Genomic DNA isolation for Mass Spectrometry

DNA was isolated using the DNeasy Blood & Tissue kit or DNeasy 96 Blood & Tissue kit from Qiagen according to the manufacturer's instructions with the following modifications:

If intended for mass spectrometry, HEK293T cells were resuspended in 380 µl ATL and incubated with 20 µl Proteinase K for 2 h at 56 °C, shaking at 900 rpm. Next, 3 µl RNAse A + 27 µl ATL were added and incubated for 15 min. Subsequently, 820 µl premixed Buffer AL-Ethanol were added and further steps were performed as described in the kit.

mESCs were resuspended in 100 µl PBS including 5 µl of RNAse A, vortexed, followed by 10 min of 21,000 g centrifugation at 4 °C, further incubated for 15 min on ice and afterwards mixed with 100 µl of buffer ATL before proceeding with the kit's standard instructions.

After elution of DNA with AE buffer, DNA was precipitated with 3.3 volumes of 100 % ethanol, 0.1 volume of 7.5 M ammonium acetate, followed by centrifugation at 21,000 g, 4 °C for 1 h. The DNA pellet was washed once with 70 % ethanol, the supernatant was removed and the pellet was dried for at least 30 min at room temperature and resuspended in nuclease-free water. DNA was degraded to nucleosides with nuclease P1, snake venom phosphodiesterase and alkaline phosphatase. Quantification of 5mC and its oxidative derivatives was carried out by Dr. Michael Musheev as described in [201], using an Agilent 1290 Infinity Binary LC system (Agilent Technologies) with a ReproSil 100 C18 column (Jasco) coupled to an Agilent 6490 triple quadrupole mass spectrometer (Agilent Technologies).

6.2.1.8 Denaturing cell lysis followed by HA affinity purification

To detect SUMOylated or ubiquitinated proteins using overexpressed HA-tagged SUMO or Ubiquitin, denaturing cell lysis followed by HA affinity purification was performed. One confluent 15 cm dish of HEK293T cells was washed with cold PBS supplemented with 10 mM NEM and harvested in cold PBS supplemented with 10 mM NEM, using a cell scraper, and transferred to a 15 ml falcon. After centrifugation for 5 min at 300 g at 4 °C, the cell pellet was resuspended in

1.5 ml denaturing lysis buffer, transferred to a 1.5 ml tube and kept at room temperature until the lysates were sonicated for 28 cycles, 15 s on/off, high energy using a Bioruptor. Afterwards, 75 µl of a 1 M DTT solution were added to the lysates (50 mM final concentration), the lysates were boiled in a water bath at 97 °C for 10 min, and transferred to a 15 ml falcon. Subsequently, the lysates were diluted with 13.5 ml of cold RIPA dilution buffer and supplemented with 75 µl 2 M NEM. For the “input” sample, 40 µl of lysate were transferred to a 1.5 ml tube and supplemented with 15 µl 4x NuPAGE+DTT and incubated for 10 min at 70° C.

For each IP reaction, 70 µl Protein A Agarose beads were washed 3 times with 1 ml RIPA buffer for 10 min on a rotating wheel at 4 °C, followed by 3 min centrifugation at 300 g and 4 °C in a Heraeus centrifuge, and afterwards resuspended in 70 µl RIPA buffer. For pre-clearing, 20 µl beads were added to the lysates and incubated for 30 min at 4 °C on a rotating wheel. Lysates were cleared by centrifugation in a Heraeus centrifuge (using the same settings as above), and transferred to a new 15 ml falcon. The antibody incubation was performed overnight using 8 µl HA antibody on a rotating wheel at 4 °C, followed by incubation with 50 µl beads for 2 h on a rotating wheel at 4 °C. Subsequently, lysates were washed 3 times with 1 ml RIPA buffer for 10 min on a rotating wheel at 4 °C, followed by 3 min centrifugation at 300 g and 4 °C in a Heraeus centrifuge. Elution was performed by adding 35 µl of 1x NuPAGE, including 100 mM DTT, followed by a 10 min incubation at 70 °C and clearing for 10 min at 21,000 g. To detect immunoprecipitated proteins, 12 µl of the lysates were loaded on an SDS-PAGE and detected via Western blotting (see chapter 6.2.1.3).

6.2.1.9 GFP-Trap

To detect SUMOylated or ubiquitinated proteins without overexpression of SUMO or Ubiquitin, GFP-Trap with denaturing washes was performed. One confluent 15 cm dish of HEK293T cells was washed twice with cold PBS supplemented with 10 mM NEM, and harvested in cold PBS supplemented with 10 mM NEM, using a cell scraper, and then transferred to a 15 ml falcon. After centrifugation for 5 min at 300 g at 4 °C, the cell pellet was resuspended in 500 µl CL lysis buffer, transferred to a 1.5 ml tube, incubated on ice for 10 min, and lysates were cleared by centrifugation at 21,000 g for 10 min at 4 °C. The supernatant was transferred to a new tube and for the “input” sample, 50 µl were transferred to another 1.5 ml tube and supplemented with 18 µl 4x NuPAGE+DTT and incubated for 10 min at 70° C. Per IP reaction, 20 µl of GFP-Trap Agarose beads were washed 3 times with 1 ml CL lysis buffer for 10 min on a rotating wheel at 4 °C, followed by 1 min centrifugation at 1200 g and 4 °C, and afterwards resuspended in 20 µl CL lysis buffer. Afterwards, the beads were incubated with the lysates for 1 h on a rotating wheel at 4 °C. Washes were performed by inverting the tubes and spinning down at 1200 g and 4 °C, with 1 ml of the following wash buffers: 1x GFP-Trap dilution buffer, 3x Wash buffer 1, 1x Wash buffer 2. As a final step, 40 µl of 1x NuPAGE, including 100 mM DTT, were added to the beads, followed by a 10 min incubation at 70 °C and clearing for 10 min at 21,000 g. Samples were either used for SDS-PAGE and Western blotting or further processed for mass spectrometry.

For mass spectrometry, the last centrifugation step was omitted and instead, the samples were cooled down to room temperature and alkylated with 5.5 mM CAA for 30 min in the dark. Samples were then stored at -20 °C before being further processed by Matthias Ostermaier (Beli lab, IMB Mainz).

6.2.1.10 FLAG-TET1 immunoprecipitation for proteomic mass spectrometry analysis

To detect TET1 interacting proteins or phosphorylation sites of TET1, SILAC labeled HEK293T cells were transfected with plasmid encoding FLAG-TET1, as described in chapters 6.2.2.3 and 6.2.2.6. For one condition, two dense 10 cm dishes were harvested using a cell scraper and transferred to one 15 ml falcon. Cells were pelleted by centrifugation for 5 min at 300 g and washed with 10 ml PBS. Afterwards, the cell pellet was resuspended in 500 µl PEX-AS2 buffer, transferred to a 1.5 ml tube and cell disruption was achieved by 5 times needle homogenization using a 27 g needle, and 5 times needle homogenization using a 23 g needle, followed by sonication for 28 cycles 15 s on/off, high setting, using a Bioruptor. Lysates were cleared by centrifugation at 21,000 g, for 15 min at 4 °C. The supernatant was transferred to a low binding 1.5 ml tube and 30 µl pre-washed ANTI-FLAG M2 Affinity Gel (FLAG beads) were added to the supernatant. To wash the beads, 30 µl beads were washed 3 times with 1 ml PEX-AS2 buffer for 10 min on a rotating wheel at 4 °C, followed by 3 min centrifugation at 300 g and 4 °C in a Heraeus centrifuge, and afterwards resuspended in 30 µl PEX-AS2 buffer. Incubation of FLAG beads with the lysates was performed for 3 h on a rotating wheel at 8 rpm at 4 °C. Subsequently, lysates were washed 4 times with 1 ml PEX-AS2 buffer for 10 min on a rotating wheel at 4 °C, followed by 4 min centrifugation at 300 g and 4 °C in a Heraeus centrifuge. For phosphoproteomic analysis of TET1, the 4 washes were performed with 1 ml PEX-AS2 buffer containing 750 mM NaCl, followed by 1 wash with regular PEX-AS2 buffer. Then, “Light”, “Medium”, and “Heavy” conditions were combined by adding 100 µl PEX-AS2 to each condition and transferring the beads to one 1.5 ml tube. Samples were centrifuged again using the same settings as above and the last traces of buffer were removed using a gel loading tip. For elution, 100 µl FLAG elution buffer were incubated with the beads for 45 min, shaking at 800 rpm at 4 °C. Samples were cleared by centrifugation and the supernatant transferred to a new 1.5 ml tube. For downstream mass spectrometry analysis, 35 µl of 4x NuPAGE+DTT, were added to the eluates, followed by a 10 min incubation at 70 °C and alkylation with 5.5 mM CAA for 30 min in the dark. Samples were then stored at -20 °C before being further processed by Andrea Voigt (Beli lab, IMB Mainz).

6.2.1.11 5hmC and 5fC DNA Immunoprecipitation (DIP)

mESCs from two approximately 60 % confluent 15 cm dishes were harvested in 4 °C cold 8 ml PBS-0.05 % Tween, each, using a cell scraper, and were transferred to the same 50 ml falcon. After centrifugation for 5 min at 300 g at 4 °C, the cell pellet was resuspended in 2720 µl PBS+80 µl RNase A, incubated for 5 min and transferred to a 15 ml falcon. After the addition of 270 µl Proteinase K and 2700 µl Buffer AL, the lysates were incubated overnight, rotating at 56 °C. Next, 2700 µl ethanol were added, the contents were vortexed and distributed equally to 6 DNeasy spin columns, and the subsequent steps were performed as described in the DNeasy Blood & Tissue

kit. The DNA was eluted with 2 x 50 µl nuclease-free water and DNA concentration was measured on a Nanodrop 2000.

10 µg of DNA were sheared to ~200 bp using the Covaris S2/E210 focused ultrasonicator using 130 µl Covaris micro tubes and the following settings: Intensity 5, duty cycle 10 %, cycles per burst 200, treatment time 280 s, temperature 7 °C and water level 12. A total of 20 µg of sheared DNA were transferred to a DNA LoBind 1.5 ml tube and made up to a total volume of 900 µl using nuclease-free water. 2 % of the total volume were taken for the “input” (18 µl) and the remaining DNA was denatured for 10 min in a heat block at 95 °C, and cooled on ice for 10 min. For the IP reaction, 100 µl of 10x IP buffer and 20 µg of the appropriate antibody or IgG (as control) were added and incubated overnight at 4 °C on a rotator at 10 rpm. Before addition to the IP reaction, 80 µl of Protein G Dynabeads were pre-washed twice with 800 µl PBS-BSA 0.1 % for 5 min on a rotator, collected using a magnetic rack, resuspended in 80 µl of 1x IP buffer and added to the IP reactions. Samples were incubated with the beads for 2 h at 4 °C, rotating at 10 rpm and washed 3 times with 1 ml 1x IP buffer for 10 min. Finally, beads were collected using a magnetic rack and resuspended in 500 µl Proteinase K digestion buffer. After addition of 7 µl Proteinase K, the samples were incubated for 3 h at 56 °C in a heating block, shaking at 950 rpm, and afterwards the supernatant was transferred to a new tube. DNA was purified using the QIAquick PCR purification kit, using 2535 µl of Buffer PB and adjusting the pH by adding 6 µl sodium acetate (3 M) and following the manufacturer’s instructions. The DNA was eluted in 30 µl nuclease-free water and used either for qPCR (described in 6.2.1.5; eluted DNA was diluted 25 times and used instead of cDNA) or concentrated using a SpeedVac and used for next generation sequencing (6.2.1.12).

6.2.1.12 5fC DNA Immunoprecipitation followed by sequencing

Library preparation and sequencing was performed by the Genomics Core Facility (IMB Mainz), and bioinformatic analysis was performed by Dr. Emil Karaulanov (Bioinformatic Core Facility, IMB Mainz) and Dr. Debasish Mukherjee (Niehrs group, IMB Mainz). 5fC DIP-seq was performed with *Rad21-AID-3xHA* cell line F5_F6, using triplicates of the following conditions: “-auxin” (control), “+auxin” (RAD21-depleted), “-auxin/+Vitamin C” (Vitamin C-treated), and “+auxin/+Vitamin C” (Vitamin C-treated and RAD21-depleted). Libraries of the 12 DIP-samples and four inputs, one pool from each condition, were prepared using a modified chromatin immunoprecipitation followed by sequencing (ChIP-seq) library prep protocol. Subsequently, the samples were sequenced on a NextSeq500, using the paired-end 42 mode, yielding an average of 2.4×10^7 reads per sample. Reads were mapped to the mm10 mouse genome using the Bowtie2 read alignment tool [229]. Since mapping of the paired-end reads resulted in many discordant and singleton reads in the DIP samples, single reads (read 1 from each pair) were mapped to the mm10 mouse genome using the Bowtie1 aligner [230]. Peaks were called using MACS2 peak finder algorithm [231], which yielded 203,404 common 5fC peaks (present in at least two DIP samples).

6.2.2 Cell culture

6.2.2.1 Thawing and Freezing of cells

Frozen cells were quickly thawed in a water bath and added to 10 ml of the respective cell culture medium in a 15 ml falcon. The cell suspension was centrifuged 3 min at 300 g, then cells were resuspended in fresh cell culture medium and plated on a cell culture dish.

For cryopreservation, cells were detached from their culture dish with Trypsin-EDTA, resuspended in the respective cell culture medium and pelleted for 3 min at 300 g. Cells were resuspended in a small volume of cell culture medium, and supplemented with the same volume of 2x freezing medium (20 % DMSO, 50 % FBS, 30 % cell culture medium). 1 ml aliquots of cells in freezing medium were transferred to cryo tubes and slowly frozen to -80 °C at a cooling rate of 1 °C/min in a Cryo-Safe Cooler filled with isopropanol. Cells were transferred to -150 °C for long-term storage.

6.2.2.2 General culture of HEK293T

HEK293T cells were cultured in supplemented DMEM medium (DMEM with 10% FBS, 2 mM L-Glutamine and 50 U/ml penicillin/streptomycin), at 37 °C, 5 % CO_2 and 21 % O_2, without changing medium between passaging. Cells were passaged every 2 days, once reaching approximately 80 % confluency, by washing cells with PBS once, detaching cells with 0.05 % Trypsin-EDTA for 2 min, quenching Trypsin with supplemented DMEM medium and plating 1/6 of the cells on a new cell culture flask.

6.2.2.3 DNA transfection of HEK293T

For plasmid DNA transfection, HEK293T cells were transfected with X-tremeGENE 9 24 h after plating cells in 600 µl supplemented DMEM medium without antibiotics. For a 12-well format, 480 ng of plasmid DNA were diluted in 20 µl TE buffer, and 1.6 µl X-tremeGENE 9 were mixed with 18.4 µl OptiMEM. Both solutions were vortexed and incubated for 2 min. Afterwards, both solutions were mixed and incubated for 30 min. Subsequently, the transfection mix was added dropwise to the cells and distributed by gentle swirling of the cell culture dish. Cells were usually harvested 48 h *post* DNA transfection.

6.2.2.4 siRNA transfection of HEK293T

For siRNA transfection, HEK293T cells were transfected with Lipofectamine 2000. For a 12-well format, 1.5 µl 20 µM siRNA were mixed with 13.5 µl OptiMEM, and 3 µl Lipofectamine 2000 were mixed with 12 µl OptiMEM. Both solutions were vortexed and incubated for 5 min. Afterwards, the two solutions were mixed and incubated for 20 min. Meanwhile, cells were plated on 12-well dishes in 600 µl supplemented DMEM medium without antibiotics and subsequently, the transfection mix was added dropwise to the cells and distributed by gentle swirling of the cell culture dish. The cell culture medium was exchanged the next morning. Cells were usually harvested 72 h *post* siRNA transfection.

6.2.2.5 Isotopic labeling of 5mC in HEK293T

To isotopically label 5mC in cells, HEK293T cells were cultured for at least 14 days (see 6.2.2.2) with the following medium: DMEM (high glucose, no glutamine, no methionine, no cysteine) supplemented with 0.2 mM L-Methionine-(methyl-$^{13}C,d_3$), 0.2 mM L-Cysteine, 1 mM sodium pyruvate, 10 % dialyzed FBS, 2 mM L-Glutamine.

6.2.2.6 Stable isotope labeling by amino acids in cell culture (SILAC) of HEK293T

In order to quantify proteins via mass spectrometry, HEK293T cells were cultured for at least 14 days (see 6.2.2.2) with stable isotopes of L-Arginine and L-Lysine, using “light”, “medium” or “heavy” DMEM:

Light DMEM: DMEM for SILAC, 10 % dialyzed FBS, 2 mM L-Glutamine, 50 U/ml penicillin/ streptomycin, 42 µg/ml L-Arginine-0, 73 µg/ml L-Lysine-0

Medium DMEM: DMEM for SILAC, 10 % dialyzed FBS, 2 mM L-Glutamine, 50 U/ml penicillin/ streptomycin, 42 µg/ml L-Arginine-6, 73 µg/ml L-Lysine-4

Heavy DMEM: DMEM for SILAC, 10 % dialyzed FBS, 2 mM L-Glutamine, 50 U/ml penicillin/ streptomycin, 42 µg/ml L-Arginine-10, 73 µg/ml L-Lysine-8

6.2.2.7 General culture of 293F

293F suspension cells were cultured in FreeStyle 293 Expression Medium in a volume of 30 ml in a 125 ml bottle, or 60-90 ml in a 250 ml bottle, shaking at 37 °C, 8 % CO_2 and 21 % O_2, without changing medium between passaging. Cells were passaged every 2 days, once reaching a density of 1-3 x 10^6 cells/ml, and diluted in fresh medium to a new density of 0.4 x 10^6 cells/ml.

6.2.2.8 DNA transfection of 293F

For plasmid DNA transfection, 293F suspension cells were transfected with PEI MAX. One day before transfection, cells were split to a final density of 1 x 10^6 cells/ml. At the day of transfection, 9 x 10^7 cells were collected by a 10 min centrifugation at 300 g, resuspended in 45 ml fresh FreeStyle 293 Expression Medium, transferred to a 250 ml bottle, and incubated in the cell culture incubator for 3-4 h. Then, 180 µg DNA were diluted to 0.5 µg/µl in fresh FreeStyle 293 Expression Medium and added directly to the cells. Cells were vortexed for 5 s and incubated in the cell culture incubator for another 5 min. Afterwards, 360 µl 1 µg/µl PEI MAX solution were mixed with 360 µl FreeStyle 293 Expression Medium by vortexing, and the mixture was also added to the cells. After 24 h, the cell culture incubator was switched to 34 °C and cells were harvested for TET1 *in vitro* oxidation assay 48 h *post* transfection.

6.2.2.9 General culture of mESCs

E14tg2a mESCs were cultured on gelatinized cell culture dishes. For this purpose, the surface area of a cell culture dish was completely covered in 0.1 % gelatin in ultrapure water (Millipore) and incubated for at least 15 min. Gelatin was aspirated immediately before plating cells. mESCs were cultured in mESC medium at 37 °C, 5 % CO_2 and 21 % O_2, with daily medium changes. Cells were passaged every 2 days once reaching approximately 60 % confluency, by washing

the cells with PBS once, detaching cells with 0.25 % Trypsin-EDTA for 2 min at 37 °C, quenching Trypsin-EDTA with mESC medium, spinning down the cells for 3 min at 300 g, resuspending cells in an adequate amount of mESC medium and plating 1/8 of the cells on a new cell culture dish.

mESC medium: DMEM supplemented with 15 % ES-grade FBS, 2 mM L-Glutamine, 50 U/ml penicillin/streptomycin, 1x MEM Non-essential amino acids, 1 mM sodium pyruvate, 100 µM β-mercaptoethanol, 8 % LIF

6.2.2.10 Mycoplasma Treatment of mESCs

As the original E14tg2a mESC-line was mycoplasma positive, cells were treated with antibiotics to remove mycoplasma contamination before performing experiments. Treatment was started one day after thawing cells, by seeding 1.8 x 10^5 cells on a gelatinized 6 cm dish, in 3 ml regular mESC medium containing 25 µg/ml Plasmocin. Cells were passaged every 2 days and each time, 1.8 x 10^5 cells were reseeded on a new 6 cm dish, in 3 ml regular mESC medium containing 25 µg/ml Plasmocin, without changing medium between passaging. After a total of 14 days, cells were frozen as described in 6.2.2.1 and success of the treatment was confirmed using the PCR Mycoplasma Test Kit I/C.

6.2.2.11 DNA transfection of mESCs

For plasmid DNA transfection, mESCs were transfected with Lipofectamine 2000. For a 10 cm dish, 1 x 10^6 mESCs were plated in 9.5 ml mESC medium without antibiotics and incubated in the cell culture incubator for 3-4 h. Generally, 6 µg of plasmid DNA were diluted in 250 µl OptiMEM and 12 µl of Lipofectamine 2000 were mixed with 238 µl OptiMEM medium. Both solutions were vortexed and incubated for 5 min. Afterwards, both solutions were mixed and incubated for 30 min. Subsequently, the transfection mix was added dropwise to the cells and distributed by gentle swirling of the cell culture dish. The cell culture medium was exchanged with 10 ml regular mESC medium the next morning. Cells were usually harvested 48 h *post* DNA transfection.

6.2.2.12 siRNA transfection of mESCs

For siRNA transfection, mESCs were transfected with Lipofectamine 2000. For a 6 cm dish, 3.5 x 10^5 mESCs were plated in 2 ml mESC medium without antibiotics and incubated in the cell culture incubator for 2-3 h. 2 µl 20 µM siRNA were mixed with 48 µl OptiMEM and 2 µl of Lipofectamine 2000 were mixed with 48 µl OptiMEM medium. Both solutions were vortexed and incubated for 5 min. Afterwards, both solutions were mixed and incubated for 20 min. Subsequently, the transfection mix was added dropwise to the cells and distributed by gentle swirling of the cell culture dish. The cell culture medium was exchanged with 3 ml regular mESC medium the next morning. Cells were usually harvested 48 h *post* siRNA transfection.

6.2.2.13 Double Thymidine Block of mESCs

To arrest mESCs at the G1/S boundary, a double thymidine block was performed [232]. For a 15 cm dish format, 5 x 10^6 cells were plated in 20 ml regular mESC medium and incubated in the cell culture incubator for 8 h. Then, the medium was exchanged with 20 ml mESC medium

containing 2.5 mM thymidine, and cells were incubated in the cell culture incubator for 14 h. Afterwards, cells were washed twice with 15 ml PBS and 20 ml fresh regular mESC medium were added to the cells, followed by a 9 h incubation in the cell culture incubator. For the second thymidine block, the medium was exchanged once more with 20 ml mESC medium containing 2.5 mM thymidine, and cells were incubated in the cell culture incubator for another 14 h. Finally, cells were washed twice with 15 ml PBS and 20 ml fresh regular mESC medium were added to the cells for release. Cells were harvested at suitable time points after release for downstream analyses such as 5fC and 5hmC DIP (6.2.1.11), mass spectrometry of genomic DNA (6.2.1.7), flow cytometry (6.2.2.15), or Western blot (6.2.1.3).

6.2.2.14 Auxin-Inducible Degron tagging of *Rad21* in mESCs via CRISPR/Cas9

Auxin-Inducible Degron tagging of *Rad21* in mESCs was performed in two steps, 1) both *Rad21* genes were tagged with the auxin-inducible degron and 2) *OsTir1* was inserted into the *Tigre* locus.

For the first step, 1 x 10^6 E14tg2a mycoplasma-free mESCs were seeded on 10 cm dishes in 9.5 ml mESC medium without antibiotics, and transfected after 3 h with 5 µg pMK_mmRad21_mAID_Neo, 5 µg pMK_mmRad21_mAID_Hygro, and 5 µg pX330_Cas9_gRNA Rad21 plasmids, using 45µl Lipofectamine 2000 in a volume of 250 µl OptiMEM, but otherwise as described in chapter 6.2.2.11. The cell culture medium was exchanged with 10 ml regular mESC medium the next morning and in the evening, mESCs were passaged from 1 x 10 cm to 2 x 15 cm dishes. 48 h *post* transfection, cells were selected by the addition of 500 µg/ml Geneticin and 500 µg/ml Hygromycin B to the cell culture medium, and kept under selection until the day of freezing.

Single mESC colonies were picked 12 days after transfection. Using the microscope, single colonies were detached using a 10 µl pipette, set to its maximum volume, by a combination of gentle pushing and pipetting-induced medium currents. Colonies were subsequently transferred to 96-well plates containing 50 µl 0.25 % Trypsin-EDTA, incubated for 5 min at room temperature, diluted with 2 x 100 µl mESC medium including selection antibiotics, and transferred to a new, gelatinized 96-well plate. Medium of cells was changed every day, until the cells reached approximately 60 % confluency after 3-4 days. At this time, cells were split and transferred to 24-well plates by washing with 100 µl PBS, incubating with 50 µl 0.25 % Trypsin-EDTA for 2 min at 37 °C, quenching and resuspending with 100 µl mESC medium including selection antibiotics. Then, 75 µl of the cell suspension were transferred to a gelatinized 24-well plate, containing 500 µl mESC medium including selection antibiotics. The other half of the cell suspension was used for genotyping PCRs, as described in chapter 6.2.1.6.

Double-positive cells, as determined by genotyping PCRs (see chapter 6.2.1.6), were expanded to 6-well plates after approximately 2-3 days and further to 10 cm dishes after another 3-5 days, when having reached approximately 60 % confluency, with daily medium changes. The cells of ½ of a 10 cm dish were frozen in two cryo tubes as described in chapter 6.2.2.1. The other ½ was pelleted by centrifugation and snap-frozen in liquid nitrogen and stored at -80 °C until further processed for Western blot analysis (see chapter 6.2.1.3).

For the second step, the two positive *Rad21-AID-3xHA* cell lines, A1 and F5, were transfected with 10 µg pEN396_Tigre OsTir1 and 5 µg pX330_EN1201_Cas9_gRNA Tigre, using Lipofectamine 2000 and the same protocol as for the first round of transfection described above, with the exception that 2 µg/ml Puromycin were used for antibiotic selection. To identify positive clones, genotyping PCRs were performed and positive clones were expanded and frozen as described above.

To verify an efficient and acute depletion of RAD21, approximately 60 % confluent cells on a 6 cm dish were treated with 500 µM of the auxin analog indole-3-acetic acid sodium salt (referred to as "auxin" in this thesis) and cells were harvested and processed for Western blot analysis (see chapter 6.2.1.3).

6.2.2.15 Cell cycle analysis via flow cytometry

mESCs were grown on a 6 cm cell culture dish to 80 % confluency. After detachment with 1 ml 0.25 % Trypsin-EDTA for 4 min at 37° C and quenching with mESC medium, cells were transferred to a 15 ml falcon tube and centrifuged at 300 g for 3 min. Cells were resuspended in 3 ml PBS+1 % FBS, counted using a cell counter, and centrifuged one more time at 300 g for 3 min. The supernatant was discarded and cells were fixed by dropwise addition of 1 ml ice-cold ethanol per 6.7 x 10^6 cells, under constant vortexing. Fixed cells were stored at -20 °C for at least 1 h. For propidium iodide staining, fixed cells were washed twice with ice-cold PBS containing 100 µg/ml RNAse A, by inverting the falcon tubes and centrifuging at 500 g for 5 min at 4 °C, before addition of 200 µl propidium iodide staining solution (1x PBS, 50 µg/ml propidium iodide, 1 mM EDTA) and incubation for 10 min in the dark. The cell suspension was pipetted through a cell-strainer and propidium iodide intensity was analyzed by flow cytometry using an LSRFortessa SORP system using FACSDiVa software. Signals from 50 000 cells were recorded and analyzed with FlowJo software v10.6.1.

6.2.3 *In vitro* TET1 oxidation assay

For the TET1 oxidation assay, FLAG-tagged TET1CD was overexpressed in 293F (see chapter 6.2.2.8) or HEK293T cells (see chapter 6.2.2.3). Volumes are referring to the conditions for 293F cells, and volumes for HEK293T cells are indicated in brackets if deviating. For cell extracts from 293F cells, 30 ml cell suspension were used per condition, while one dense 10 cm dish was used for HEK293T. Cells were pelleted by centrifugation for 5 min at 300 g and washed with 10 ml PBS. Afterwards, the cell pellet was resuspended in 10 ml (2 ml) IVA5 buffer and cell disruption was achieved by 5 times needle homogenization using a 20 g needle followed by sonication for 28 cycles 15 s on/off, high setting, using a Bioruptor. Lysates were transferred to a 50 ml (15 ml) falcon, followed by the addition of 20 ml (8 ml) IVA5 buffer. To immunoprecipitate FLAG-tagged TET1CD, 70 µl (10 µl) ANTI-FLAG M2 Affinity Gel (FLAG beads) was washed 3 times with 1 ml IVA5 buffer for 10 min on a rotating wheel at 4 °C, followed by 3 min centrifugation at 300 g and 4 °C in a Heraeus centrifuge, and afterwards resuspended in 70 µl (10 µl) IVA5 buffer. Subsequently, the FLAG beads were added to the lysates and incubation was performed overnight on a rotating wheel at 8 rpm and 4 °C. Subsequently, lysates were washed twice with

10 ml IVA3 buffer and once with 10 ml Reaction wash buffer for 10 min on a rotating wheel at 4 °C, followed by 5 min centrifugation at 870 g and 4 °C in a Heraeus centrifuge. Then, 1 ml Reaction wash buffer was added to the FLAG beads and the beads were transferred to a 1.5 ml tube. After another centrifugation for 5 min at 870 g and 4 °C, the last traces of buffer were removed using a gel loading tip and the *in vitro* TET1 oxidation assay was started by adding 100 µl freshly prepared Reaction buffer to the samples. Samples were incubated for 3, 12 or 24 min, shaking at 450 rpm at 37 °C, and the reaction was stopped by the addition of 16 µl 0.5 M EDTA. Beads were collected by centrifugation for 3 min at 870 g and 4 °C, and the supernatant was transferred to a new 1.5 ml tube. DNA was precipitated by the addition of 10 µl 7.5 M ammonium acetate and 330 µl ethanol, and further processed for mass spectrometry as described in chapter 6.2.1.7.

6.2.3.1 Generation of an isotopically labelled and *in vitro* methylated substrate for the TET1 oxidation assay

To generate the substrate for the TET1 oxidation assay, a PCR was set up with the following components and PCR program:

Component	Volume (µl)
10x Reaction Buffer HM, including 15 mM $MgCl_2$	5
$MgCl_2$ (25 mM)	2.5
dNTPs (with 15N-dCTP) (5 mM)	2.5
Primer IVO biotin forward+reverse (10µM)	1.5
Nuclease-free H20	32.25
AccuPrime Taq	1.25
IVO2 template (100 nM)	1.5

Table 6.15 PCR components

Temperature (°C)	Time	
94	7 min	
94	15 s	30 x
57	15 s	
68	20 s	
68	2 min	
4	2 min	

Table 6.16 PCR program

DNA was purified using the QIAquick PCR purification kit. 10 µg of DNA were then *in vitro* methylated using the following reagents:

Component	Volume (µl)
Purified DNA	x (10 µg)
10x Reaction Buffer 2	40
SAM (32 mM)	8
M.SssI (20,000 U/µl)	6
Nuclease-free H_2O	346-x µl

Table 6.17 In vitro methylation of purified DNA

Incubation was performed overnight at 37 °C, then 6 µl SAM and 4 µl M.SssI were added to the reaction mix and incubated for additional 4 h at 37 °C. The substrate was purified using phenol-chloroform extraction followed by ammonium-acetate/ethanol precipitation.

7 References

1. Ruppel WG (1898). Zeitschrift für physikalische Chemie (26): 218.
2. WYATT GR (1950) Occurrence of 5-methylcytosine in nucleic acids. Nature 166 (4214): 237–238.
3. Greenberg MVC, Bourc'his D (2019) The diverse roles of DNA methylation in mammalian development and disease. Nature reviews. Molecular cell biology.
4. Chen L, MacMillan AM, Chang W, Ezaz-Nikpay K, Lane WS, Verdine GL (1991) Direct identification of the active-site nucleophile in a DNA (cytosine-5)-methyltransferase. Biochemistry 30 (46): 11018–11025.
5. Okano M, Bell DW, Haber DA, Li E (1999) DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. Cell 99 (3): 247–257.
6. Hermann A, Goyal R, Jeltsch A (2004) The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. The Journal of biological chemistry 279 (46): 48350–48359.
7. Otani J, Nankumo T, Arita K, Inamoto S, Ariyoshi M, Shirakawa M (2009) Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX–DNMT3–DNMT3L domain. EMBO reports 10 (11): 1235–1241.
8. Guo X, Wang L, Li J, Ding Z, Xiao J, Yin X, He S, Shi P, Dong L, Li G, Tian C, Wang J, Cong Y, Xu Y (2015) Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. Nature 517 (7536): 640–644.
9. Dhayalan A, Rajavelu A, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, Jeltsch A (2010) The Dnmt3a PWWP Domain Reads Histone 3 Lysine 36 Trimethylation and Guides DNA Methylation. The Journal of biological chemistry 285 (34): 26114–26120.
10. Bostick M, Kim JK, Estève P-O, Clark A, Pradhan S, Jacobsen SE (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science (New York, N.Y.) 317 (5845): 1760–1764.
11. Rothbart SB, Dickson BM, Ong MS, Krajewski K, Houliston S, Kireev DB, Arrowsmith CH, Strahl BD (2013) Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation. Genes & development 27 (11): 1288–1298.
12. Qin W, Wolf P, Liu N, Link S, Smets M, La Mastra F, Forné I, Pichler G, Hörl D, Fellinger K, Spada F, Bonapace IM, Imhof A, Harz H, Leonhardt H (2015) DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination. Cell research 25 (8): 911–929.
13. Nishiyama A, Yamaguchi L, Sharif J, Johmura Y, Kawamura T, Nakanishi K, Shimamura S, Arita K, Kodama T, Ishikawa F, Koseki H, Nakanishi M (2013) Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. Nature 502 (7470): 249–253.
14. Ishiyama S, Nishiyama A, Saeki Y, Moritsugu K, Morimoto D, Yamaguchi L, Arai N, Matsumura R, Kawakami T, Mishima Y, Hojo H, Shimamura S, Ishikawa F, Tajima S,

Tanaka K, Ariyoshi M, Shirakawa M, Ikeguchi M, Kidera A, Suetake I, Arita K, Nakanishi M (2017) Structure of the Dnmt1 Reader Module Complexed with a Unique Two-Mono-Ubiquitin Mark on Histone H3 Reveals the Basis for DNA Methylation Maintenance. Molecular cell 68 (2): 350-360.e7.

15. Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature reviews. Genetics 13 (7): 484–492.
16. Bochtler M, Kolano A, Xu G-L (2017) DNA demethylation pathways: Additional players and regulators. BioEssays : news and reviews in molecular, cellular and developmental biology 39 (1): 1–13.
17. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T (2000) Demethylation of the zygotic paternal genome. Nature 403 (6769): 501–502.
18. Seisenberger S, Peat JR, Reik W (2013) Conceptual links between DNA methylation reprogramming in the early embryo and primordial germ cells. Current opinion in cell biology 25 (3): 281–288.
19. Howell CY, Bestor TH, Ding F, Latham KE, Mertineit C, Trasler JM, Chaillet JR (2001) Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. Cell 104 (6): 829–838.
20. Barreto G, Schäfer A, Marhold J, Stach D, Swaminathan SK, Handa V, Doderlein G, Maltry N, Wu W, Lyko F, Niehrs C (2007) Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. Nature 445 (7128): 671–675.
21. Zhu J-K (2009) Active DNA demethylation mediated by DNA glycosylases. Annual review of genetics 43: 143–166.
22. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science (New York, N.Y.) 324 (5929): 930–935.
23. He Y-F, Li B-Z, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song C-X, Zhang K, He C, Xu G-L (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science (New York, N.Y.) 333 (6047): 1303–1307.
24. Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science (New York, N.Y.) 324 (5929): 929–930.
25. Maiti A, Drohat AC (2011) Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. The Journal of biological chemistry 286 (41): 35334–35338.
26. Kohli RM, Zhang Y (2013) TET enzymes, TDG and the dynamics of DNA demethylation. Nature 502 (7472): 472–479.
27. Jin S-g, Guo C, Pfeifer GP (2008) GADD45A does not promote DNA demethylation. PLoS genetics 4 (3): e1000013.
28. Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK (2004) Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. The Journal of biological chemistry 279 (50): 52353–52360.

29. Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM (2010) Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature 463 (7284): 1042–1047.
30. Hogenbirk MA, Heideman MR, Velds A, van den Berk PCM, Kerkhoven RM, van Steensel B, Jacobs H (2013) Differential programming of B cells in AID deficient mice. PloS one 8 (7): e69815.
31. Rai K, Huggins IJ, James SR, Karpf AR, Jones DA, Cairns BR (2008) DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. Cell 135 (7): 1201–1212.
32. Fritz EL, Rosenberg BR, Lay K, Mihailović A, Tuschl T, Papavasiliou FN (2013) A comprehensive analysis of the effects of the deaminase AID on the transcriptome and methylome of activated B cells. Nature immunology 14 (7): 749–755.
33. Shimoda N, Hirose K, Kaneto R, Izawa T, Yokoi H, Hashimoto N, Kikuchi Y (2014) No evidence for AID/MBD4-coupled DNA demethylation in zebrafish embryos. PloS one 9 (12): e114816.
34. Franchini D-M, Chan C-F, Morgan H, Incorvaia E, Rangam G, Dean W, Santos F, Reik W, Petersen-Mahrt SK (2014) Processive DNA demethylation via DNA deaminase-induced lesion resolution. PloS one 9 (7): e97754.
35. Lorsbach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR (2003) TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). Leukemia 17 (3): 637–641.
36. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 466 (7310): 1129–1133.
37. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science (New York, N.Y.) 333 (6047): 1300–1303.
38. Hackett JA, Sengupta R, Zylicz JJ, Murakami K, Lee C, Down TA, Surani MA (2013) Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. Science (New York, N.Y.) 339 (6118): 448–452.
39. Inoue A, Shen L, Dai Q, He C, Zhang Y (2011) Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. Cell research 21 (12): 1670–1676.
40. Inoue A, Zhang Y (2011) Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. Science (New York, N.Y.) 334 (6053): 194.
41. Wu X, Zhang Y (2017) TET-mediated active DNA demethylation. Mechanism, function and beyond. Nature reviews. Genetics 18 (9): 517–534.
42. Weber AR, Krawczyk C, Robertson AB, Kuśnierczyk A, Vågbø CB, Schuermann D, Klungland A, Schär P (2016) Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. Nature communications 7: 10806.

43. Dianov GL, Hübscher U (2013) Mammalian base excision repair: the forgotten archangel. Nucleic acids research 41 (6): 3483–3490.
44. Schomacher L, Niehrs C (2017) DNA repair and erasure of 5-methylcytosine in vertebrates. BioEssays : news and reviews in molecular, cellular and developmental biology 39 (3).
45. Gu T-P, Guo F, Yang H, Wu H-P, Xu G-F, Liu W, Xie Z-G, Shi L, He X, Jin S-g, Iqbal K, Shi YG, Deng Z, Szabó PE, Pfeifer GP, Li J, Xu G-L (2011) The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature 477 (7366): 606–610.
46. Wang Y, Zhang Y (2014) Regulation of TET protein stability by calpains. Cell reports 6 (2): 278–284.
47. Amouroux R, Nashun B, Shirane K, Nakagawa S, Hill PWS, D'Souza Z, Nakayama M, Matsuda M, Turp A, Ndjetehe E, Encheva V, Kudo NR, Koseki H, Sasaki H, Hajkova P (2016) De novo DNA methylation drives 5hmC accumulation in mouse zygotes. Nat Cell Biol 18 (2): 225–233.
48. Vincent JJ, Huang Y, Chen P-Y, Feng S, Calvopiña JH, Nee K, Lee SA, Le T, Yoon AJ, Faull K, Fan G, Rao A, Jacobsen SE, Pellegrini M, Clark AT (2013) Stage-Specific Roles for Tet1 and Tet2 in DNA Demethylation in Primordial Germ Cells. Cell stem cell 12 (4): 470–478.
49. Yamaguchi S, Shen L, Liu Y, Sendler D, Zhang Y (2013) Role of Tet1 in erasure of genomic imprinting. Nature 504 (7480): 460–464.
50. Lu F, Liu Y, Jiang L, Yamaguchi S, Zhang Y (2014) Role of Tet proteins in enhancer activity and telomere elongation. Genes & development 28 (19): 2103–2119.
51. Dawlaty MM, Breiling A, Le T, Barrasa MI, Raddatz G, Gao Q, Powell BE, Cheng AW, Faull KF, Lyko F, Jaenisch R (2014) Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. Developmental cell 29 (1): 102–111.
52. Dai H-Q, Wang B-A, Yang L, Chen J-J, Zhu G-C, Sun M-L, Ge H, Wang R, Chapman DL, Tang F, Sun X, Xu G-L (2016) TET-mediated DNA demethylation controls gastrulation by regulating Lefty-Nodal signalling. Nature 538 (7626): 528–532.
53. Zhang R-R, Cui Q-Y, Murai K, Lim YC, Smith ZD, Jin S, Ye P, Rosa L, Lee YK, Wu H-P, Liu W, Xu Z-M, Yang L, Ding Y-Q, Tang F, Meissner A, Ding C, Shi Y, Xu G-L (2013) Tet1 regulates adult hippocampal neurogenesis and cognition. Cell stem cell 13 (2): 237–245.
54. Rudenko A, Dawlaty MM, Seo J, Cheng AW, Meng J, Le T, Faull KF, Jaenisch R, Tsai L-H (2013) Tet1 is critical for neuronal activity-regulated gene expression and memory extinction. Neuron 79 (6): 1109–1122.
55. Li X, Wei W, Zhao Q-Y, Widagdo J, Baker-Andresen D, Flavell CR, D'Alessio A, Zhang Y, Bredy TW (2014) Neocortical Tet3-mediated accumulation of 5-hydroxymethylcytosine promotes rapid behavioral adaptation. Proceedings of the National Academy of Sciences of the United States of America 111 (19): 7120–7125.
56. Morris-Blanco KC, Kim T, Lopez MS, Bertogliat MJ, Chelluboina B, Vemuganti R (2019) Induction of DNA Hydroxymethylation Protects the Brain After Stroke. Stroke: STROKEAHA119025665.

57. Kafer GR, Li X, Horii T, Suetake I, Tajima S, Hatada I, Carlton PM (2016) 5-Hydroxymethylcytosine Marks Sites of DNA Damage and Promotes Genome Stability. Cell reports 14 (6): 1283–1292.
58. Yang J, Guo R, Wang H, Ye X, Zhou Z, Dan J, Wang H, Gong P, Deng W, Yin Y, Mao S, Wang L, Ding J, Li J, Keefe DL, Dawlaty MM, Wang J, Xu G, Liu L (2016) Tet Enzymes Regulate Telomere Maintenance and Chromosomal Stability of Mouse ESCs. Cell reports 15 (8): 1809–1821.
59. Zhong J, Li X, Cai W, Wang Y, Dong S, Yang J, Zhang J'a, Wu N, Li Y, Mao F, Zeng C, Wu J, Xu X, Sun ZS (2017) TET1 modulates H4K16 acetylation by controlling auto-acetylation of hMOF to affect gene regulation and DNA repair function. Nucleic acids research 45 (2): 672–684.
60. Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, An J, Lamperti ED, Koh KP, Ganetzky R, Liu XS, Aravind L, Agarwal S, Maciejewski JP, Rao A (2010) Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature 468 (7325): 839–843.
61. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, Li Y, Bhagwat N, Vasanthakumar A, Fernandez HF, Tallman MS, Sun Z, Wolniak K, Peeters JK, Liu W, Choe SE, Fantin VR, Paietta E, Löwenberg B, Licht JD, Godley LA, Delwel R, Valk PJM, Thompson CB, Levine RL, Melnick A (2010) Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer cell 18 (6): 553–567.
62. Rampal R, Alkalin A, Madzo J, Vasanthakumar A, Pronier E, Patel J, Li Y, Ahn J, Abdel-Wahab O, Shih A, Lu C, Ward PS, Tsai JJ, Hricik T, Tosello V, Tallman JE, Zhao X, Daniels D, Dai Q, Ciminio L, Aifantis I, He C, Fuks F, Tallman MS, Ferrando A, Nimer S, Paietta E, Thompson CB, Licht JD, Mason CE, Godley LA, Melnick A, Figueroa ME, Levine RL (2014) DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia. Cell reports 9 (5): 1841–1855.
63. Wang Y, Xiao M, Chen X, Chen L, Xu Y, Lv L, Wang P, Yang H, Ma S, Lin H, Jiao B, Ren R, Ye D, Guan K-L, Xiong Y (2015) WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation. Molecular cell 57 (4): 662–673.
64. López-Moyado IF, Tsagaratou A, Yuita H, Seo H, Delatte B, Heinz S, Benner C, Rao A (2019) Paradoxical association of TET loss of function with genome-wide DNA hypomethylation. Proceedings of the National Academy of Sciences of the United States of America 116 (34): 16933–16942.
65. Hu L, Li Z, Cheng J, Rao Q, Gong W, Liu M, Shi YG, Zhu J, Wang P, Xu Y (2013) Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. Cell 155 (7): 1545–1555.
66. Hu L, Lu J, Cheng J, Rao Q, Li Z, Hou H, Lou Z, Zhang L, Li W, Gong W, Liu M, Sun C, Yin X, Li J, Tan X, Wang P, Wang Y, Fang D, Cui Q, Yang P, He C, Jiang H, Luo C, Xu Y (2015) Structural insight into substrate preference for TET-mediated oxidation. Nature 527 (7576): 118–122.

67. Xu Y, Xu C, Kato A, Tempel W, Abreu JG, Bian C, Hu Y, Di Hu, Zhao B, Cerovina T, Diao J, Wu F, He HH, Cui Q, Clark E, Ma C, Barbara A, Veenstra GJC, Xu G, Kaiser UB, Liu XS, Sugrue SP, He X, Min J, Kato Y, Shi YG (2012) Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. Cell 151 (6): 1200–1213.
68. Xu Y, Wu F, Tan L, Kong L, Xiong L, Deng J, Barbera AJ, Zheng L, Zhang H, Huang S, Min J, Nicholson T, Chen T, Xu G, Shi Y, Zhang K, Shi YG (2011) Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. Molecular cell 42 (4): 451–464.
69. Ko M, An J, Bandukwala HS, Chavez L, Äijö T, Pastor WA, Segal MF, Li H, Koh KP, Lähdesmäki H, Hogan PG, Aravind L, Rao A (2013) Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. Nature 497 (7447): 122–126.
70. Zhang W, Xia W, Wang Q, Towers AJ, Chen J, Gao R, Zhang Y, Yen C-A, Lee AY, Li Y, Zhou C, Liu K, Zhang J, Gu T-P, Chen X, Chang Z, Leung D, Gao S, Jiang Y-H, Xie W (2016) Isoform Switch of TET1 Regulates DNA Demethylation and Mouse Development. Molecular cell 64 (6): 1062–1073.
71. Jin S-g, Zhang Z-M, Dunwell TL, Harter MR, Wu X, Johnson J, Li Z, Liu J, Szabó PE, Lu Q, Xu G-L, Song J, Pfeifer GP (2016) Tet3 Reads 5-Carboxylcytosine through Its CXXC Domain and Is a Potential Guardian against Neurodegeneration. Cell reports 14 (3): 493–505.
72. Wu H, Zhang Y (2014) Reversing DNA methylation: mechanisms, genomics, and biological functions. Cell 156 (1-2): 45–68.
73. Chen Q, Chen Y, Bian C, Fujiki R, Yu X (2013) TET2 promotes histone O-GlcNAcylation during gene transcription. Nature 493 (7433): 561–564.
74. Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PAC, Rappsilber J, Helin K (2011) TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature 473 (7347): 343–348.
75. Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, Sun YE, Zhang Y (2011) Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. Genes & development 25 (7): 679–684.
76. Zhang Q, Zhao K, Shen Q, Han Y, Gu Y, Li X, Zhao D, Liu Y, Wang C, Zhang X, Su X, Liu J, Ge W, Levine RL, Li N, Cao X (2015) Tet2 is required to resolve inflammation by recruiting Hdac2 to specifically repress IL-6. Nature 525 (7569): 389–393.
77. Vella P, Scelfo A, Jammula S, Chiacchiera F, Williams K, Cuomo A, Roberto A, Christensen J, Bonaldi T, Helin K, Pasini D (2013) Tet Proteins Connect the O-Linked N-acetylglucosamine Transferase Ogt to Chromatin in Embryonic Stem Cells. Molecular cell 49 (4): 645–656.
78. Deplus R, Delatte B, Schwinn MK, Defrance M, Méndez J, Murphy N, Dawson MA, Volkmar M, Putmans P, Calonne E, Shih AH, Levine RL, Bernard O, Mercher T, Solary E, Urh M, Daniels DL, Fuks F (2013) TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. The EMBO journal 32 (5): 645–655.

79. Kaas GA, Zhong C, Eason DE, Ross DL, Vachhani RV, Ming G-l, King JR, Song H, Sweatt JD (2013) TET1 controls CNS 5-methylcytosine hydroxylation, active DNA demethylation, gene transcription, and memory formation. Neuron 79 (6): 1086–1093.
80. Ito K, Lee J, Chrysanthou S, Zhao Y, Josephs K, Sato H, Teruya-Feldstein J, Zheng D, Dawlaty MM, Ito K (2019) Non-catalytic Roles of Tet2 Are Essential to Regulate Hematopoietic Stem and Progenitor Cell Homeostasis. Cell reports 28 (10): 2480-2490.e4.
81. Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung H-L, Zhang K, Zhang Y (2013) Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. Cell 153 (3): 692–706.
82. Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, McLoughlin EM, Brudno Y, Mahapatra S, Kapranov P, Tahiliani M, Daley GQ, Liu XS, Ecker JR, Milos PM, Agarwal S, Rao A (2011) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature 473 (7347): 394–397.
83. Sun Z, Terragni J, Jolyon T, Borgaro JG, Liu Y, Yu L, Guan S, Wang H, Sun D, Cheng X, Zhu Z, Pradhan S, Zheng Y (2013) High-resolution enzymatic mapping of genomic 5-hydroxymethylcytosine in mouse embryonic stem cells. Cell reports 3 (2): 567–576.
84. Sun Z, Dai N, Borgaro JG, Quimby A, Sun D, Corrêa IR, Zheng Y, Zhu Z, Guan S (2015) A sensitive approach to map genome-wide 5-hydroxymethylcytosine and 5-formylcytosine at single-base resolution. Molecular cell 57 (4): 750–761.
85. Yu M, Hon GC, Szulwach KE, Song C-X, Zhang L, Kim A, Li X, Dai Q, Shen Y, Park B, Min J-H, Jin P, Ren B, He C (2012) Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell 149 (6): 1368–1380.
86. Song C-X, Szulwach KE, Dai Q, Fu Y, Mao S-Q, Lin L, Street C, Li Y, Poidevin M, Wu H, Gao J, Liu P, Li L, Xu G-L, Jin P, He C (2013) Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. Cell 153 (3): 678–691.
87. Wu H, Wu X, Shen L, Zhang Y (2014) Single-base resolution analysis of active DNA demethylation using methylase-assisted bisulfite sequencing. Nature biotechnology 32 (12): 1231–1240.
88. Carey BW, Finley LWS, Cross JR, Allis CD, Thompson CB (2015) Intracellular α-ketoglutarate maintains the pluripotency of embryonic stem cells. Nature 518 (7539): 413–416.
89. Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim S-H, Ito S, Yang C, Wang P, Xiao M-T, Liu L-x, Jiang W-q, Liu J, Zhang J-y, Wang B, Frye S, Zhang Y, Xu Y-h, Lei Q-y, Guan K-L, Zhao S-m, Xiong Y (2011) Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. Cancer cell 19 (1): 17–30.
90. Yin R, Mao S-Q, Zhao B, Chong Z, Yang Y, Zhao C, Zhang D, Huang H, Gao J, Li Z, Jiao Y, Li C, Liu S, Wu D, Gu W, Yang Y-G, Xu G-L, Wang H (2013) Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. Journal of the American Chemical Society 135 (28): 10396–10403.
91. DiTroia SP, Percharde M, Guerquin M-J, Wall E, Collignon E, Ebata KT, Mesh K, Mahesula S, Agathocleous M, Laird DJ, Livera G, Ramalho-Santos M (2019) Maternal vitamin C

regulates reprogramming of DNA methylation and germline development. Nature 23 (Suppl): 588S.

92. Zhao B, Yang Y, Wang X, Chong Z, Yin R, Song S-H, Zhao C, Li C, Huang H, Sun B-F, Wu D, Jin K-X, Song M, Zhu B-Z, Jiang G, Rendtlew Danielsen JM, Xu G-L, Yang Y-G, Wang H (2014) Redox-active quinones induces genome-wide DNA methylation changes by an iron-mediated and Tet-dependent mechanism. Nucleic acids research 42 (3): 1593–1605.
93. Yu C, Zhang Y-L, Pan W-W, Li X-M, Wang Z-W, Ge Z-J, Zhou J-J, Cang Y, Tong C, Sun Q-Y, Fan H-Y (2013) CRL4 complex regulates mammalian oocyte survival and reprogramming by activation of TET proteins. Science (New York, N.Y.) 342 (6165): 1518–1521.
94. Nakagawa T, Lv L, Nakagawa M, Yu Y, Yu C, D'Alessio AC, Nakayama K, Fan H-Y, Chen X, Xiong Y (2015) CRL4(VprBP) E3 ligase promotes monoubiquitylation and chromatin binding of TET dioxygenases. Molecular cell 57 (2): 247–260.
95. Zhang YW, Wang Z, Xie W, Cai Y, Xia L, Easwaran H, Luo J, Yen R-WC, Li Y, Baylin SB (2017) Acetylation Enhances TET2 Function in Protecting against Abnormal DNA Methylation during Oxidative Stress. Molecular cell 65 (2): 323–335.
96. Bauer C, Göbel K, Nagaraj N, Colantuoni C, Wang M, Müller U, Kremmer E, Rottach A, Leonhardt H (2015) Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT). The Journal of biological chemistry 290 (8): 4801–4812.
97. Shi F-T, Kim H, Lu W, He Q, Liu D, Goodell MA, Wan M, Songyang Z (2013) Ten-eleven translocation 1 (Tet1) is regulated by O-linked N-acetylglucosamine transferase (Ogt) for target gene repression in mouse embryonic stem cells. The Journal of biological chemistry 288 (29): 20776–20784.
98. Ciccarone F, Valentini E, Zampieri M, Caiafa P (2015) 5mC-hydroxylase activity is influenced by the PARylation of TET1 enzyme. Oncotarget 6 (27): 24333–24347.
99. Costa Y, Ding J, Theunissen TW, Faiola F, Hore TA, Shliaha PV, Fidalgo M, Saunders A, Lawrence M, Dietmann S, Das S, Levasseur DN, Li Z, Xu M, Reik W, Silva JCR, Wang J (2013) NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. Nature 495 (7441): 370–374.
100. Okashita N, Kumaki Y, Ebi K, Nishi M, Okamoto Y, Nakayama M, Hashimoto S, Nakamura T, Sugasawa K, Kojima N, Takada T, Okano M, Seki Y (2014) PRDM14 promotes active DNA demethylation through the ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells. Development (Cambridge, England) 141 (2): 269–280.
101. Sun Z, Xu X, He J, Murray A, Sun M-A, Wei X, Wang X, McCoig E, Xie E, Jiang X, Li L, Zhu J, Chen J, Morozov A, Pickrell AM, Theus MH, Xie H (2019) EGR1 recruits TET1 to shape the brain methylome during development and upon neuronal activity. Nature communications 10 (1): 3892.
102. Zeng Y, Yao B, Shin J, Lin L, Kim N, Song Q, Liu S, Su Y, Guo JU, Huang L, Wan J, Wu H, Qian J, Cheng X, Zhu H, Ming G-l, Jin P, Song H (2016) Lin28A Binds Active Promoters and Recruits Tet1 to Regulate Gene Expression. Molecular cell 61 (1): 153–160.

103. Neri F, Incarnato D, Krepelova A, Rapelli S, Pagnani A, Zecchina R, Parlato C, Oliviero S (2013) Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. Genome biology 14 (8): R91.
104. Kienhöfer S, Musheev MU, Stapf, Helm, Mark, Ulrike, Schomacher L, Niehrs C, Schäfer A (2015) GADD45a physically and functionally interacts with TET1.
105. Niehrs C, Schäfer A (2012) Active DNA demethylation by Gadd45 and DNA repair. Trends in cell biology 22 (4): 220–227.
106. Ma DK, Jang M-H, Guo JU, Kitabatake Y, Chang M-L, Pow-Anpongkul N, Flavell RA, Lu B, Ming G-l, Song H (2009) Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. Science (New York, N.Y.) 323 (5917): 1074–1077.
107. Schmitz K-M, Schmitt N, Hoffmann-Rohrer U, Schäfer A, Grummt I, Mayer C (2009) TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. Molecular cell 33 (3): 344–353.
108. Schäfer A, Karaulanov E, Stapf U, Döderlein G, Niehrs C (2013) Ing1 functions in DNA demethylation by directing Gadd45a to H3K4me3. Genes & development 27 (3): 261–273.
109. Arab K, Park YJ, Lindroth AM, Schäfer A, Oakes C, Weichenhan D, Lukanova A, Lundin E, Risch A, Meister M, Dienemann H, Dyckhoff G, Herold-Mende C, Grummt I, Niehrs C, Plass C (2014) Long noncoding RNA TARID directs demethylation and activation of the tumor suppressor TCF21 via GADD45A. Molecular cell 55 (4): 604–614.
110. Schüle KM, Leichsenring M, Andreani T, Vastolo V, Mallick M, Musheev MU, Karaulanov E, Niehrs C (2019) GADD45 promotes locus-specific DNA demethylation and 2C cycling in embryonic stem cells. Genes & development 33 (13-14): 782–798.
111. Li Z, Gu T-P, Weber AR, Shen J-Z, Li B-Z, Xie Z-G, Yin R, Guo F, Liu X, Tang F, Wang H, Schar P, Xu G-L (2015) Gadd45a promotes DNA demethylation through TDG. Nucleic acids research 43 (8): 3986–3997.
112. Schäfer A (2013) Gadd45 proteins: key players of repair-mediated DNA demethylation. Advances in experimental medicine and biology 793: 35–50.
113. Chang L, Karin M (2001) Mammalian MAP kinase signalling cascades. Nature 410 (6824): 37–40.
114. Kyriakis JM, Avruch J (2012) Mammalian MAPK signal transduction pathways activated by stress and inflammation: a 10-year update. Physiological reviews 92 (2): 689–737.
115. Miyake Z, Takekawa M, Ge Q, Saito H (2007) Activation of MTK1/MEKK4 by GADD45 through induced N-C dissociation and dimerization-mediated trans autophosphorylation of the MTK1 kinase domain. Molecular and cellular biology 27 (7): 2765–2776.
116. Bulavin DV, Kovalsky O, Hollander MC, Fornace, Albert J Jr (2003) Loss of oncogenic H-ras-induced cell cycle arrest and p38 mitogen-activated protein kinase activation by disruption of Gadd45a. Molecular and cellular biology 23 (11): 3859–3871.
117. Annunziato AT (2008) DNA Packaging: Nucleosomes and Chromatin. Nature Education 1 (1): 26.
118. Oudet P, Gross-Bellard M, Chambon P (1975) Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell 4 (4): 281–300.

119. Szabo Q, Bantignies F, Cavalli G (2019) Principles of genome folding into topologically associating domains. Science advances 5 (4): eaaw1668.
120. Bompadre O, Andrey G (2019) Chromatin topology in development and disease. Current opinion in genetics & development 55: 32–38.
121. Matthews BJ, Waxman DJ (2018) Computational prediction of CTCF/cohesin-based intra-TAD loops that insulate chromatin contacts and gene expression in mouse liver. eLife 7.
122. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science (New York, N.Y.) 326 (5950): 289–293.
123. Bintu B, Mateo LJ, Su J-H, Sinnott-Armstrong NA, Parker M, Kinrot S, Yamaya K, Boettiger AN, Zhuang X (2018) Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. Science (New York, N.Y.) 362 (6413).
124. Stevens TJ, Lando D, Basu S, Atkinson LP, Cao Y, Lee SF, Leeb M, Wohlfahrt KJ, Boucher W, O'Shaughnessy-Kirwan A, Cramard J, Faure AJ, Ralser M, Blanco E, Morey L, Sansó M, Palayret MGS, Lehner B, Di Croce L, Wutz A, Hendrich B, Klenerman D, Laue ED (2017) 3D structures of individual mammalian genomes studied by single-cell Hi-C. Nature 544 (7648): 59–64.
125. Moronta-Gines M, van Staveren TRH, Wendt KS (2019) One ring to bind them - Cohesin's interaction with chromatin fibers. Essays in biochemistry 63 (1): 167–176.
126. Michaelis C, Ciosk R, Nasmyth K (1997) Cohesins: Chromosomal Proteins that Prevent Premature Separation of Sister Chromatids. Cell 91 (1): 35–45.
127. Haering CH, Löwe J, Hochwagen A, Nasmyth K (2002) Molecular Architecture of SMC Proteins and the Yeast Cohesin Complex. Molecular cell 9 (4): 773–788.
128. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, Yahata K, Imamoto F, Aburatani H, Nakao M, Imamoto N, Maeshima K, Shirahige K, Peters J-M (2008) Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature 451 (7180): 796–801.
129. Rowley MJ, Corces VG (2018) Organizational principles of 3D genome architecture. Nature reviews. Genetics 19 (12): 789–800.
130. Vian L, Pękowska A, Rao SSP, Kieffer-Kwon K-R, Jung S, Baranello L, Huang S-C, El Khattabi L, Dose M, Pruett N, Sanborn AL, Canela A, Maman Y, Oksanen A, Resch W, Li X, Lee B, Kovalchuk AL, Tang Z, Nelson S, Di Pierro M, Cheng RR, Machol I, St Hilaire BG, Durand NC, Shamim MS, Stamenova EK, Onuchic JN, Ruan Y, Nussenzweig A, Levens D, Aiden EL, Casellas R (2018) The Energetics and Physiological Impact of Cohesin Extrusion. Cell 173 (5): 1165-1178.e20.
131. Brackley CA, Johnson J, Michieletto D, Morozov AN, Nicodemi M, Cook PR, Marenduzzo D (2017) Nonequilibrium Chromosome Looping via Molecular Slip Links. Physical review letters 119 (13): 138101.

132. Rao SSP, Huang S-C, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon K-R, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID, Huang X, Shamim MS, Shin J, Turner D, Ye Z, Omer AD, Robinson JT, Schlick T, Bernstein BE, Casellas R, Lander ES, Aiden EL (2017) Cohesin Loss Eliminates All Loop Domains. Cell 171 (2): 305-320.e24.
133. Gassler J, Brandão HB, Imakaev M, Flyamer IM, Ladstätter S, Bickmore WA, Peters J-M, Mirny LA, Tachibana K (2017) A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. The EMBO journal 36 (24): 3600–3618.
134. Schwarzer W, Abdennur N, Goloborodko A, Pekowska A, Fudenberg G, Loe-Mie Y, Fonseca NA, Huber W, Haering CH, Mirny L, Spitz F (2017) Two independent modes of chromatin organization revealed by cohesin removal. Nature 551 (7678): 51–56.
135. Haarhuis JHI, van der Weide RH, Blomen VA, Yáñez-Cuna JO, Amendola M, van Ruiten MS, Krijger PHL, Teunissen H, Medema RH, van Steensel B, Brummelkamp TR, Wit E de, Rowland BD (2017) The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. Cell 169 (4): 693-707.e14.
136. Wutz G, Várnai C, Nagasaka K, Cisneros DA, Stocsits RR, Tang W, Schoenfelder S, Jessberger G, Muhar M, Hossain MJ, Walther N, Koch B, Kueblbeck M, Ellenberg J, Zuber J, Fraser P, Peters J-M (2017) Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. The EMBO journal 36 (24): 3573–3599.
137. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, Taatjes DJ, Dekker J, Young RA (2010) Mediator and cohesin connect gene expression and chromatin architecture. Nature 467 (7314): 430–435.
138. Sun F, Chronis C, Kronenberg M, Chen X-F, Su T, Lay FD, Plath K, Kurdistani SK, Carey MF (2019) Promoter-Enhancer Communication Occurs Primarily within Insulated Neighborhoods. Molecular cell 73 (2): 250-263.e5.
139. Weintraub AS, Li CH, Zamudio AV, Sigova AA, Hannett NM, Day DS, Abraham BJ, Cohen MA, Nabet B, Buckley DL, Guo YE, Hnisz D, Jaenisch R, Bradner JE, Gray NS, Young RA (2017) YY1 Is a Structural Regulator of Enhancer-Promoter Loops. Cell 171 (7): 1573-1588.e28.
140. Bailey SD, Zhang X, Desai K, Aid M, Corradin O, Cowper-Sal Lari R, Akhtar-Zaidi B, Scacheri PC, Haibe-Kains B, Lupien M (2015) ZNF143 provides sequence specificity to secure chromatin interactions at gene promoters. Nature communications 2: 6186.
141. Liu NQ, Maresca M, van den Brand T, Braccioli L, Schijns MMGA, Teunissen H, Bruneau BG, Nora EP, Wit E de (2019) WAPL maintains dynamic cohesin to preserve lineage specific distal gene regulation. Biorxiv.
142. Tan JY, Smith AAT, Ferreira da Silva M, Matthey-Doret C, Rueedi R, Sönmez R, Ding D, Kutalik Z, Bergmann S, Marques AC (2017) cis-Acting Complex-Trait-Associated lincRNA Expression Correlates with Modulation of Chromosomal Architecture. Cell reports 18 (9): 2280–2288.

143. Kaaij LJT, Mohn F, van der Weide RH, Wit E de, Bühler M (2019) The ChAHP Complex Counteracts Chromatin Looping at CTCF Sites that Emerged from SINE Expansions in Mouse. Cell 178 (6): 1437-1451.e14.
144. Jung YH, Sauria MEG, Lyu X, Cheema MS, Ausio J, Taylor J, Corces VG (2017) Chromatin States in Mouse Sperm Correlate with Embryonic and Adult Regulatory Landscapes. Cell reports 18 (6): 1366–1382.
145. Rubin AJ, Barajas BC, Furlan-Magaril M, Lopez-Pajares V, Mumbach MR, Howard I, Kim DS, Boxer LD, Cairns J, Spivakov M, Wingett SW, Shi M, Zhao Z, Greenleaf WJ, Kundaje A, Snyder M, Chang HY, Fraser P, Khavari PA (2017) Lineage-specific dynamic and pre-established enhancer-promoter contacts cooperate in terminal differentiation. Nature genetics 49 (10): 1522–1528.
146. Visel A, Rubin EM, Pennacchio LA (2009) Genomic views of distant-acting enhancers. Nature 461 (7261): 199–205.
147. Nora EP, Goloborodko A, Valton A-L, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG (2017) Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell 169 (5): 930-944.e22.
148. Deng W, Rupon JW, Krivega I, Breda L, Motta I, Jahn KS, Reik A, Gregory PD, Rivella S, Dean A, Blobel GA (2014) Reactivation of developmentally silenced globin genes by forced chromatin looping. Cell 158 (4): 849–860.
149. Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, Huber W, Furlong EEM (2014) Enhancer loops appear stable during development and are associated with paused polymerase. Nature 512 (7512): 96–100.
150. Rhodes JDP, Feldmann A, Hernández-Rodríguez B, Díaz N, Brown JM, Fursova NA, Blackledge NP, Prathapan P, Dobrinic P, Huseyin M, Szczurek A, Kruse K, Nasmyth KA, Buckle VJ, Vaquerizas JM, Klose RJ (2019) Cohesin disrupts polycomb-dependent chromosome interactions. Biorxiv.
151. Schäfer A, Mekker B, Mallick M, Vastolo V, Karaulanov E, Sebastian D, Lippen C von der, Epe B, Downes DJ, Scholz C, Niehrs C (2018) Impaired DNA demethylation of C/EBP sites causes premature aging. Genes & development 32 (11-12): 742–762.
152. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D, Tiwari VK, Schübeler D (2011) DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature 480 (7378): 490–495.
153. Flavahan WA, Drier Y, Liau BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, Suvà ML, Bernstein BE (2016) Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature 529 (7584): 110–114.
154. Xu C, Corces VG (2018) Nascent DNA methylome mapping reveals inheritance of hemimethylation at CTCF/cohesin sites. Science (New York, N.Y.) 359 (6380): 1166–1170.
155. Mellén M, Ayata P, Dewell S, Kriaucionis S, Heintz N (2012) MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell 151 (7): 1417–1430.

156. Skene PJ, Illingworth RS, Webb S, Kerr ARW, James KD, Turner DJ, Andrews R, Bird AP (2010) Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. Molecular cell 37 (4): 457–468.
157. Kernohan KD, Vernimmen D, Gloor GB, Bérubé NG (2014) Analysis of neonatal brain lacking ATRX or MeCP2 reveals changes in nucleosome density, CTCF binding and chromatin looping. Nucleic acids research 42 (13): 8356–8368.
158. Kernohan KD, Jiang Y, Tremblay DC, Bonvissuto AC, Eubanks JH, Mann MRW, Bérubé NG (2010) ATRX partners with cohesin and MeCP2 and contributes to developmental silencing of imprinted genes in the brain. Developmental cell 18 (2): 191–202.
159. Wiehle L, Thorn GJ, Raddatz G, Clarkson CT, Rippe K, Lyko F, Breiling A, Teif VB (2019) DNA (de)methylation in embryonic stem cells controls CTCF-dependent chromatin boundaries. Genome research 29 (5): 750–761.
160. Takekawa M, Saito H (1998) A family of stress-inducible GADD45-like proteins mediate activation of the stress-responsive MTK1/MEKK4 MAPKKK. Cell 95 (4): 521–530.
161. Nemoto S, Xiang J, Huang S, Lin A (1998) Induction of apoptosis by SB202190 through inhibition of p38beta mitogen-activated protein kinase. The Journal of biological chemistry 273 (26): 16415–16420.
162. Favata MF, Horiuchi KY, Manos EJ, Daulerio AJ, Stradley DA, Feeser WS, van Dyk DE, Pitts WJ, Earl RA, Hobbs F, Copeland RA, Magolda RL, Scherle PA, Trzaskos JM (1998) Identification of a novel inhibitor of mitogen-activated protein kinase kinase. The Journal of biological chemistry 273 (29): 18623–18632.
163. Wu T-S, Yu F-Y, Su C-C, Kan J-C, Chung C-P, Liu B-H (2005) Activation of ERK mitogen-activated protein kinase in human cells by the mycotoxin patulin. Toxicology and applied pharmacology 207 (2): 103–111.
164. Zang M, Gong J, Luo L, Zhou J, Xiang X, Huang W, Huang Q, Luo X, Olbrot M, Peng Y, Chen C, Luo Z (2008) Characterization of Ser338 phosphorylation for Raf-1 activation. The Journal of biological chemistry 283 (46): 31429–31437.
165. Kao S-H, Wang W-L, Chen C-Y, Chang Y-L, Wu Y-Y, Wang Y-T, Wang S-P, Nesvizhskii AI, Chen Y-J, Hong T-M, Yang P-C (2015) Analysis of Protein Stability by the Cycloheximide Chase Assay. Bio-protocol 5 (1).
166. Arab K, Karaulanov E, Musheev M, Trnka P, Schäfer A, Grummt I, Niehrs C (2019) GADD45A binds R-loops and recruits TET1 to CpG island promoters. Nature genetics 51 (2): 217–223.
167. Zhao X (2018) SUMO-Mediated Regulation of Nuclear Functions and Signaling Processes. Molecular cell 71 (3): 409–418.
168. Hardeland U, Steinacher R, Jiricny J, Schar P (2002) Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. The EMBO journal 21 (6): 1456–1464.
169. Hendriks IA, Lyon D, Young C, Jensen LJ, Vertegaal ACO, Nielsen ML (2017) Site-specific mapping of the human SUMO proteome reveals co-modification with phosphorylation. Nature structural & molecular biology 24 (3): 325–336.

170. Ozato K, Shin D-M, Chang T-H, Morse HC3 (2008) TRIM family proteins and their emerging roles in innate immunity. Nature reviews. Immunology 8 (11): 849–860.
171. Nishimura K, Fukagawa T, Takisawa H, Kakimoto T, Kanemaki M (2009) An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nature methods 6 (12): 917–922.
172. Zeng H, Horie K, Madisen L, Pavlova MN, Gragerova G, Rohde AD, Schimpf BA, Liang Y, Ojala E, Kramer F, Roth P, Slobodskaya O, Dolka I, Southon EA, Tessarollo L, Bornfeldt KE, Gragerov A, Pavlakis GN, Gaitanaris GA (2008) An inducible and reversible mouse genetic rescue system. PLoS genetics 4 (5): e1000069.
173. Madisen L, Garner AR, Shimaoka D, Chuong AS, Klapoetke NC, Li L, van der Bourg A, Niino Y, Egolf L, Monetti C, Gu H, Mills M, Cheng A, Tasic B, Nguyen TN, Sunkin SM, Benucci A, Nagy A, Miyawaki A, Helmchen F, Empson RM, Knöpfel T, Boyden ES, Reid RC, Carandini M, Zeng H (2015) Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. Neuron 85 (5): 942–958.
174. Michaelis C, Ciosk R, Nasmyth K (1997) Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. Cell 91 (1): 35–45.
175. Hardeland U, Kunz C, Focke F, Szadkowski M, Schar P (2007) Cell cycle regulation as a mechanism for functional separation of the apparently redundant uracil DNA glycosylases TDG and UNG2. Nucleic acids research 35 (11): 3859–3867.
176. Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, Wagner U, Dixon J, Lee L, Lobanenkov VV, Ren B (2012) A map of the cis-regulatory sequences in the mouse genome. Nature 488 (7409): 116–120.
177. Raiber E-A, Beraldi D, Ficz G, Burgess HE, Branco MR, Murat P, Oxley D, Booth MJ, Reik W, Balasubramanian S (2012) Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. Genome biology 13 (8): R69.
178. Kong L, Tan L, Lv R, Shi Z, Xiong L, Wu F, Rabidou K, Smith M, He C, Zhang L, Qian Y, Ma D, Lan F, Shi Y, Shi YG (2016) A primary role of TET proteins in establishment and maintenance of De Novo bivalency at CpG islands. Nucleic acids research 44 (18): 8682–8692.
179. Sahlén P, Abdullayev I, Ramsköld D, Matskova L, Rilakovic N, Lötstedt B, Albert TJ, Lundeberg J, Sandberg R (2015) Genome-wide mapping of promoter-anchored interactions with close to single-enhancer resolution. Genome biology 16: 156.
180. White J, Dalton S (2005) Cell Cycle Control of Embryonic Stem Cells. SCR 1 (2): 131–138.
181. Jin C, Lu Y, Jelinek J, Liang S, Estecio MRH, Barton MC, Issa J-PJ (2014) TET1 is a maintenance DNA demethylase that prevents methylation spreading in differentiated cells. Nucleic acids research 42 (11): 6956–6971.
182. Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, Marques CJ, Andrews S, Reik W (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature 473 (7347): 398–402.

183. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15 (12): 550.
184. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics (Oxford, England) 26 (1): 139–140.
185. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, Ali S, Chin S-F, Palmieri C, Caldas C, Carroll JS (2012) Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature 481 (7381): 389–393.
186. Lienhard M, Grimm C, Morkel M, Herwig R, Chavez L (2014) MEDIPS: genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments. Bioinformatics (Oxford, England) 30 (2): 284–286.
187. Krishnakumar R, Kraus WL (2010) The PARP side of the nucleus: molecular actions, physiological outcomes, and clinical targets. Molecular cell 39 (1): 8–24.
188. Zhao H, Sifakis EG, Sumida N, Millán-Ariño L, Scholz BA, Svensson JP, Chen X, Ronnegren AL, Mallet de Lima CD, Varnoosfaderani FS, Shi C, Loseva O, Yammine S, Israelsson M, Rathje L-S, Németi B, Fredlund E, Helleday T, Imreh MP, Göndör A (2015) PARP1- and CTCF-Mediated Interactions between Active and Repressed Chromatin at the Lamina Promote Oscillating Transcription. Molecular cell 59 (6): 984–997.
189. Taghbalout A, Du M, Jillette N, Rosikiewicz W, Rath A, Heinen CD, Li S, Cheng AW (2019) Enhanced CRISPR-based DNA demethylation by Casilio-ME-mediated RNA-guided coupling of methylcytosine oxidation and DNA repair pathways. Nature communications 10 (1): 4296.
190. Di Wu, Di Hu, Chen H, Shi G, Fetahu IS, Wu F, Rabidou K, Fang R, Tan L, Xu S, Liu H, Argueta C, Zhang L, Mao F, Yan G, Chen J, Dong Z, Lv R, Xu Y, Wang M, Ye Y, Zhang S, Duquette D, Geng S, Yin C, Lian CG, Murphy GF, Adler GK, Garg R, Lynch L, Yang P, Li Y, Lan F, Fan J, Shi Y, Shi YG (2018) Glucose-regulated phosphorylation of TET2 by AMPK reveals a pathway linking diabetes to cancer. Nature 559 (7715): 637–641.
191. Komander D, Rape M (2012) The ubiquitin code. Annual review of biochemistry 81: 203–229.
192. Gorr TA, Vogel J (2015) Western blotting revisited: critical perusal of underappreciated technical issues. Proteomics. Clinical applications 9 (3-4): 396–405.
193. Li D, Kang N, Ji J, Zhan Q (2015) BRCA1 regulates transforming growth factor-β (TGF-β1) signaling through Gadd45a by enhancing the protein stability of Smad4. Molecular oncology 9 (8): 1655–1666.
194. Aparisi Rey A, Karaulanov E, Sharopov S, Arab K, Schäfer A, Gierl M, Guggenhuber S, Brandes C, Pennella L, Gruhn WH, Jelinek R, Maul C, Conrad A, Kilb W, Luhmann HJ, Niehrs C, Lutz B (2019) Gadd45α modulates aversive learning through post-transcriptional regulation of memory-related mRNAs. EMBO reports 20 (6).
195. Song SJ, Poliseno L, Song MS, Ala U, Webster K, Ng C, Beringer G, Brikbak NJ, Yuan X, Cantley LC, Richardson AL, Pandolfi PP (2013) MicroRNA-antagonism regulates breast

cancer stemness and metastasis via TET-family-dependent chromatin remodeling. Cell 154 (2): 311–324.

196. Fu X, Jin L, Wang X, Luo A, Hu J, Zheng X, Tsark WM, Riggs AD, Ku HT, Huang W (2013) MicroRNA-26a targets ten eleven translocation enzymes and is regulated during pancreatic cell differentiation. Proceedings of the National Academy of Sciences of the United States of America 110 (44): 17892–17897.
197. Zhang P, Huang B, Xu X, Sessa WC (2013) Ten-eleven translocation (Tet) and thymine DNA glycosylase (TDG), components of the demethylation pathway, are direct targets of miRNA-29a. Biochemical and biophysical research communications 437 (3): 368–373.
198. Welling M, Chen H-H, Muñoz J, Musheev MU, Kester L, Junker JP, Mischerikow N, Arbab M, Kuijk E, Silberstein L, Kharchenko PV, Geens M, Niehrs C, van de Velde H, van Oudenaarden A, Heck AJR, Geijsen N (2015) DAZL regulates Tet1 translation in murine embryonic stem cells. EMBO reports 16 (7): 791–802.
199. Sytnikova YA, Kubarenko AV, Schafer A, Weber, Alexander N R, Niehrs C (2011) Gadd45a is an RNA binding protein and is localized in nuclear speckles. PloS one 6 (1): e14500.
200. McLaughlin D, Coey CT, Yang W-C, Drohat AC, Matunis MJ (2016) Characterizing Requirements for Small Ubiquitin-like Modifier (SUMO) Modification and Binding on Base Excision Repair Activity of Thymine-DNA Glycosylase in Vivo. The Journal of biological chemistry 291 (17): 9014–9024.
201. Schomacher L, Han D, Musheev MU, Arab K, Kienhöfer S, Seggern A von, Niehrs C (2016) Neil DNA glycosylases promote substrate turnover by Tdg during DNA demethylation. Nature structural & molecular biology 23 (2): 116–124.
202. Steinacher R, Barekati Z, Botev P, Kuśnierczyk A, Slupphaug G, Schär P (2019) SUMOylation coordinates BERosome assembly in active DNA demethylation during cell differentiation. The EMBO journal 38 (1).
203. Hendriks IA, D'Souza RCJ, Yang B, Verlaan-de Vries M, Mann M, Vertegaal ACO (2014) Uncovering global SUMOylation signaling networks in a site-specific manner. Nature structural & molecular biology 21 (10): 927–936.
204. Dawlaty MM, Breiling A, Le T, Raddatz G, Barrasa MI, Cheng AW, Gao Q, Powell BE, Li Z, Xu M, Faull KF, Lyko F, Jaenisch R (2013) Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. Developmental cell 24 (3): 310–323.
205. Lu B, Yu H, Chow C-w, Li B, Zheng W-p, Davis RJ, Flavell RA (2001) GADD45γ Mediates the Activation of the p38 and JNK MAP Kinase Pathways and Cytokine Production in Effector TH1 Cells. Immunity 14 (5): 583–590.
206. Lu B, Ferrandino AF, Flavell RA (2004) Gadd45beta is important for perpetuating cognate and inflammatory signals in T cells. Nature immunology 5 (1): 38–44.
207. Hollander MC, Sheikh MS, Bulavin DV, Lundgren K, Augeri-Henmueller L, Shehee R, Molinaro TA, Kim KE, Tolosa E, Ashwell JD, Rosenberg MP, Zhan Q, Fernández-Salguero PM, Morgan WF, Deng CX, Fornace AJ (1999) Genomic instability in Gadd45a-deficient mice. Nature genetics 23 (2): 176–184.

208. Fang S, Li J, Xiao Y, Lee M, Guo L, Han W, Li T, Hill MC, Hong T, Mo W, Xu R, Zhang P, Wang F, Chang J, Zhou Y, Sun D, Martin JF, Huang Y (2019) Tet inactivation disrupts YY1 binding and long-range chromatin interactions during embryonic heart development. Nature communications 10 (1): 4297.
209. Rhee I, Jair KW, Yen RW, Lengauer C, Herman JG, Kinzler KW, Vogelstein B, Baylin SB, Schuebel KE (2000) CpG methylation is maintained in human cancer cells lacking DNMT1. Nature 404 (6781): 1003–1007.
210. Gu T, Lin X, Cullen SM, Luo M, Jeong M, Estecio M, Shen J, Hardikar S, Sun D, Su J, Rux D, Guzman A, Lee M, Qi LS, Chen J-J, Kyba M, Huang Y, Chen T, Li W, Goodell MA (2018) DNMT3A and TET1 cooperate to regulate promoter epigenetic landscapes in mouse embryonic stem cells. Genome biology 19 (1): 88.
211. Lentini A, Lagerwall C, Vikingsson S, Mjoseng HK, Douvlataniotis K, Vogt H, Green H, Meehan RR, Benson M, Nestor CE (2018) A reassessment of DNA-immunoprecipitation-based genomic profiling. Nature methods 15 (7): 499–504.
212. Booth MJ, Marsico G, Bachman M, Beraldi D, Balasubramanian S (2014) Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. Nature chemistry 6 (5): 435–440.
213. Hansen AS, Cattoglio C, Darzacq X, Tjian R (2018) Recent evidence that TADs and chromatin loops are dynamic structures. Nucleus (Austin, Tex.) 9 (1): 20–32.
214. Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, Carroll JS, Flicek P, Odom DT (2010) A CTCF-independent role for cohesin in tissue-specific transcription. Genome research 20 (5): 578–588.
215. Garcia-Luis J, Lazar-Stefanita L, Gutierrez-Escribano P, Thierry A, Cournac A, García A, González S, Sánchez M, Jarmuz A, Montoya A, Dore M, Kramer H, Karimi MM, Antequera F, Koszul R, Aragon L (2019) FACT mediates cohesin function on chromatin. Nature structural & molecular biology 26 (10): 970–979.
216. Wutz G, Ladurner R, St. Hilaire B, Stocsits R, Nagasaka K, Pignard B, Sanborn A, Tang W, Várnai C, Ivanov M, Schoenfelder S, van der Lelij P, Huang X, Dürnberger G, Roitinger E, Mechtler K, Davidson IF, Fraser P, Aiden EL, Peters J-M (2019) ESCO1 and CTCF enable formation of long chromatin loops by protecting cohesin STAG1 from WAPL. Biorxiv.
217. Benabdallah NS, Williamson I, Illingworth RS, Kane L, Boyle S, Sengupta D, Grimes GR, Therizols P, Bickmore WA (2019) Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation. Molecular cell.
218. Altmeyer M, Neelsen KJ, Teloni F, Pozdnyakova I, Pellegrino S, Grøfte M, Rask M-BD, Streicher W, Jungmichel S, Nielsen ML, Lukas J (2015) Liquid demixing of intrinsically disordered proteins is seeded by poly(ADP-ribose). Nature communications 6: 8088.
219. Fujiki K, Shinoda A, Kano F, Sato R, Shirahige K, Murata M (2013) PPARγ-induced PARylation promotes local DNA demethylation by production of 5-hydroxymethylcytosine. Nature communications 4: 2262.

220. Feldmann A, Ivanek R, Murr R, Gaidatzis D, Burger L, Schübeler D (2013) Transcription factor occupancy can mediate active turnover of DNA methylation at regulatory regions. PLoS genetics 9 (12): e1003994.
221. Melamed P, Yosefzon Y, David C, Tsukerman A, Pnueli L (2018) Tet Enzymes, Variants, and Differential Effects on Function. Frontiers in cell and developmental biology 6: 22.
222. Marshall OJ, Southall TD, Cheetham SW, Brand AH (2016) Cell-type-specific profiling of protein-DNA interactions without cell isolation using targeted DamID with next-generation sequencing. Nature protocols 11 (9): 1586–1598.
223. Skene PJ, Henikoff JG, Henikoff S (2018) Targeted in situ genome-wide profiling with high efficiency for low cell numbers. Nature protocols 13 (5): 1006–1019.
224. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S (2019) CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nature communications 10 (1): 1930.
225. Frauer C, Rottach A, Meilinger D, Bultmann S, Fellinger K, Hasenöder S, Wang M, Qin W, Söding J, Spada F, Leonhardt H (2011) Different binding properties and function of CXXC zinc finger domains in Dnmt1 and Tet1. PloS one 6 (2): e16627.
226. Kamitani T, Kito K, Nguyen HP, Yeh ET (1997) Characterization of NEDD8, a developmentally down-regulated ubiquitin-like protein. The Journal of biological chemistry 272 (45): 28557–28562.
227. Sambrook J, Fritsch EF, Maniatis TP, Russell DW, Green MR (NY : Cold Spring Harbor Laboratory Press) Molecular cloning. A laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
228. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods (San Diego, Calif.) 25 (4): 402–408.
229. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nature methods 9 (4): 357–359.
230. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology 10 (3): R25.
231. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). Genome biology 9 (9): R137.
232. Chen G, Deng X (2018) Cell Synchronization by Double Thymidine Block. Bio-protocol 8 (17).

8 List of Abbreviations

53BP1	p53-binding protein 1
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
a.a.	Amino acid
AID	Auxin-Inducible Degron
α-KG	α-ketoglutarate
AP	Abasic
ADD	ATRX-DNMT3-DNMT3L
AID	Activation-induced deaminase
AML	Acute Myeloid Leukaemia
APEX1	DNA-(apurinic or apyrimidinic site) lyase
BER	Base-Excision Repair
bp	Base pair
BSA	Bovine serum albumin
cDNA	Complementary DNA
C/EBPβ	CCAAT/enhancer-binding protein beta
CF IMB	Core-Facility IMB
ChAHP	CHD4-ADNP-HP1 complex
ChIP-Seq	Chromatin Immunoprecipitation followed by sequencing
CHX	Cycloheximide
Co	Control
Co-IP	Co-Immunoprecipitation
CpG	Cytosine-Guanine dinucleotide
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRL4	Cullin-ring finger ligase-4
CTCF	CCCTC-binding factor
CUT&RUN	Cleavage Under Targets and Release Using Nuclease
CUT&TAG	Cleavage Under Targets and Tagmentation
DamID	DNA adenine methyltransferase Identification
DAPI	4′,6-diamidino-2-phenylindole
DAZL	Deleted in azoospermia-like
DIP	DNA Immunoprecipitation
DIP-seq	DIP followed by sequencing
DDB1	Damaged DNA binding protein 1
ddCp	Delta delta Crossing point
DMEM	Dulbecco’s Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
DNMT	DNA methyltransferase
dNTP	Nucleoside triphosphate
DSBH	Double-stranded β-helix domain
dTAG	Degradation tag
DTT	Dithiothreitol
EGR1	Early growth response protein 1
ESRRB	Estrogen-related receptor b

EZH2	Enhancer of zeste homolog 2
FACT	Facilitates Chromatin Transcription
FBS	Fetal Bovine Serum
fCAB-seq	5fC chemical modification-Assisted Bisulfite sequencing
fC-Seal	5-formylcytosine-selective chemical labeling
FDR	False Discovery Rate
G45A	Growth arrest and DNA-damage-inducible protein 45 alpha
GADD45A	Growth arrest and DNA-damage-inducible protein 45 alpha
GADD45B	Growth arrest and DNA-damage-inducible protein 45 beta
GADD45G	Growth arrest and DNA-damage-inducible protein 45 gamma
GFP	Green fluorescent protein
GlcNAc	attachment of O-linked N-acetylglucosamine
γH2AX	Phosphorylated histone H2AX
gRNA	Guide RNA
GS-linker	Glycine-serine-linker
H3K4me1,2,3	Histone 3 mono-/di-/trimethylated at lysine 4
H3K9me1,2,3	Histone 3 mono-/di-/trimethylated at lysine 9
H3K27me3	Histone 3 trimethylated at lysine 27
H3K36me3	Histone 3 trimethylated at lysine 36
H4K16	Histone 4 lysine 16
HA	Human influenza hemagglutinin
HDAC1/2	Histone deacetylase 1/2
HEK293F/T	Human Embryonic Kidney 293F/T cells
IDH	Isocitrate dehydrogenase
IF	Immunofluorescence
IgG	Immunoglobulin G
IMB	Insitute of Molecular Biology
ING1	Inhibitor of growth protein 1
IP	Immunoprecipitation
KLF4	Krueppel-like factor 4
LB	Luria Broth
LIF	Leukemia inhibitory factor
LIG3	DNA ligase 3
LINE	Long Interspersed Nuclear Elements
MAB-seq	M.SssI Methylase-Assisted Bisulfite sequencing
MAP3K4	Mitogen-activated protein kinase kinase kinase 4
MAPK	Mitogen-activated protein kinase
MAPKK	Mitogen-activated protein kinase kinase
MAPKKK	Mitogen-activated protein kinase kinase kinase
MBD4	Methyl-CpG-binding domain protein 4
MeCP2	Methyl-CpG-binding protein 2
MEF	Mouse Embryonic Fibroblast
mESC	Mouse Embryonic stem Cell
min	Minute
miRNA	microRNA
MLL	Mixed Lineage Leukaemia
mm	Mus musculus
mRNA	Messenger RNA
MTase	Methyltransferase
MuERV-L	Murine endogenous retrovirus-like

MW	Molecular Weight
N/D	Non Detectable
NEM	N-Ethylmaleimide
NER	Nucleotide Excision Repair
NIPBL	Nipped-B-like protein
ns	Non significant
OGT	O-linked N-acetylglucosamine transferase
OsTIR1	Oryza sativa Transport inhibitor response 1-like protein
PAR	poly [ADP-ribose]
PARP1	poly [ADP-ribose] polymerase 1
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PGC	Primordial Germ Cell
Pol β	DNA polymerase β
PPARγ	Peroxisome proliferator-activated receptor γ
PRC1/2	Polycomb Repressive Complex 1/2
PRDM14	PR domain zinc finger protein 14
qPCR	Quantitative Polymerase Chain Reaction
resBS-seq	Reduced Bisulfite sequencing
RFTS	Replication Foci Targeting Sequence
RNA	Ribonucleic Acid
rpm	Revolutions per minute
RT-qPCR	Reverse Transcription followed by quantitative Polymerase Chain Reaction
SAM	S-adenosyl-L-methionine
SD	Standard Deviation
SDS-PAGE	Sodium dodecyl sulfate Polyacrylamide Gel Electrophoresis
SILAC	Stable Isotope Labeling with Amino acids in Cell culture
siRNA	small interfering Ribonucleic Acid
SRA	SET- and RING-Associated
SMC	Structural maintenance of chromosomes
SUMO	Small ubiquitin-related modifier
SUZ12	Suppressor of zeste 12 protein homolog
TAD	Topologically Associating Domain
TAF12	Transcription initiation factor TFIID subunit 12
TARID	TCF21 Antisense RNA Inducing Promoter Demethylation
TBP	TATA-box-binding protein
TBS	Tris-Buffered Saline
TDG	Thymine DNA glycosylase
TDG-S	SUMOylated TDG
TET	Ten-Eleven translocation
TET1CD	Catalytic domain of Ten-Eleven translocation 1
TET1/3FL	Full-length isoform of TET1/3
TET1/3s	Short isoform of TET1/3
Tigre	Tightly regulated
TRIM	Tripartite motif-containing protein
UBL	Ubiquitin-like
UPL	Universal probe library
VitC	Vitamin C
VprBP	DDB1- and CUL4-associated factor 1

WAPL	Wings apart-like protein homolog
WB	Western blot
WT	Wildtype
WT1	Wilms tumor protein 1
XRCC1	X-ray repair cross-complementing protein 1
YY1	Yin Yang 1
ZNF143/750	Zinc finger protein 143/750

9 Acknowledgements

10 Lebenslauf