2013

ELUCIDATION OF MECHANISMS GENERATING 5-HYDROXYMETHYLCYTOSINE (5hmC) IN MAMMALIAN MITOCHONDRIA

Prashant Thakkar
Virginia Commonwealth University

ELUCIDATION OF MECHANISMS GENERATING 5-HYDROXYMETHYLCYTOSINE

(5hmC) IN MAMMALIAN MITOCHONDRIA

A dissertation submitted in partial fulfillments of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

by

PRASHANT V. THAKKAR
Master of Science, University of Pune, India. 2007.

Advisor:- Shirley M. Taylor, PhD
Associate Professor,
Department of Microbiology and Immunology,

Virginia Commonwealth University,
Richmond, VA.
April 2013.

i

Acknowledgements

This dissertation would not have been possible without the invaluable guidance and help of many individuals, who have contributed to the preparation and completion of this study. First, I owe my deepest gratitude to my dissertation advisor, Dr. Shirley M. Taylor for her excellent mentorship, caring and patience during my graduate training. Her unwavering support and confidence in me has motivated me to work hard towards my dissertation. Her insights lead to the original proposal to look for presence of 5hmC and the role of Tet enzymes in their generation in mammalian mitochondria. She has helped me through difficult times during this project with data analysis and the writing of this dissertation. Her approach has helped me garner critical thinking and technical expertise, components that will continue to benefit me throughout my career. She has inspired not only as a scientist but even as a person. It has been an honor to work with her.

I would also like to thank members of my graduate committee, Dr. Daniel Conrad, Dr. Richard G. Moran, Dr. Joyce Lloyd and Dr. Gail Christie for their constant guidance and thoughtful criticisms throughout the course of my graduate career. Each member has brought unique perspective on my project and their advice and suggestions have helped shape the dissertation to what it is today.

I would like to thank Drs. Gail Christie and Dr. Dennis Ohmann, for letting me train in the graduate program of Molecular Biology and Genetics in the department of Microbiology and Immunology at VCU. I would also like to thank Dr. Gordon Ginder and Massey Cancer Center for their support and accommodation.

I would like to thank past and present members of Taylor laboratory, Erica Peterson, Dr. Lisa Shock, Joyce Balinang, Dr. Timothy Lochmann, Laura Burton, John Strang and Shannon Hedrick. I would also like to thank past and present members of the Moran laboratory, Dr. Shane Kasten, Dr. Alex Racanelli, Dr. Scott Rothbart, Dr. Scott Lawrence, Dr. Cortney Heyer, Lin Xie, Guoyan Gao, Dr. Chen Yang, Stuti Agarwal, Catherine Bell, William Buchwald and Audrey Thacker. I am deeply indebted to them for their companionship and extending their friendship beyond the environment of the laboratory. Interactions with each of them have provided comic relief, which made the lab environment even more enjoyable. On a professional level, this group has provided me with technical expertise, scientific knowledge and valuable discussion which has been critical to my success as a graduate student.

I would also like to thank my family and friends for their constant support, optimisim and encouragement, without which this achievement would not have been possible. I want to thank them for the immense faith they have shown in my capabilities and providing me with confidence and strength when I most needed them.

Table of contents   **Page**

Chapter

1: **Introduction and Overview**

2. **Functional significance of cytosine methylation by mtDNMT1 in mammalian mitochondria.**

Introduction

Materials and Methods

3. **Characterization of the role of Tet enzymes in generation of 5hmC in mtDNA**

<u>Introduction</u>

<u>Materials and Methods</u>

<u>Results</u>

Discussion

Further experiments are needed to explore the role of Tet2 and Tet3 in generation of

5hmC in mtDNA……………………………………………………………117

4. **Perspectives**

List of Tables

List of Figures:

List of Abbreviations

neoS	Neomycin sensitive
ES cells	Embryonic Stem cells
LIF	Leukemia inhibitory factor
Wt MEFs	Wild type Mouse Embryonic Fibroblasts
^{o}C	Degree Celsius
µl	microliter
ml	milliliter
µg	Microgram
ng	Nanogram
pg	Picogram
α-KG	α-ketoglutarate
β-GT	β- Glucosyltransferase enzyme
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
ATPase6	ATPase subunit 6

CHO	**Chinese hamster ovary cells**
Cox1	**Cytochrome c oxidase subunit 1**
Cox2	**Cytochrome c oxidase subunit 2**
D-loop	**Displacement loop**
DNMT1	**DNA methyltransferase 1**
GFP	**Green Fluorescent protein**
HCT116	**Human Colon Carcinoma cell line (ATCC – CCL247)**
HDAC	**Histone Deacteylases**
H-strand	**Heavy strand**
HSP1	**Heavy Strand Promoter 1**
HSP2	**Heavy Strand Promoter 2**
HSP70	**Heat shock protein 70 kDa**
HSP90	**Heat shock protein 90kDa**
H3K4me3	**Histone 3 Lysine 4 trimethyl**
IgG	**Immunoglobulin G**
L-strand	**Light strand**
LSP	**Light Strand Promoter**

mtDNA	**mitochondrial DNA**
MBD	**Methyl-binding domain**
MeCP2	**methyl CpG binding protein 2**
MPP	**Mitochondrial Processing Peptidase**
MTS	**Mitochondrial Targeting Sequence**
ND1	**NADH dehydrogenase subunit 1**
ND6	**NADH dehydrogenase subunit 6**
NRF1	**Nuclear respiratory factor 1**
O_H	**Origin of replication on Heavy strand**
O_L	**Origin of replication on Light strand**
POLRMT	**Mitochondrial RNA Polymerase**
POLG	**Polymerase Gamma**
PGC1α	**peroxisome proliferator-activated receptor γ-coactivator 1α**
SAM	**S-adenosyl methionine**
Tet	**Ten –Eleven Translocation**
TFB1M	**Mitochondrial Transcription Factor B1**
TFB2M	**Mitochondrial Transcription Factor B1**

TFAM   **Mitochondrial Transcription Factor A**

TOM   **Transporter of the outer mitochondrial membrane**

TIM   **Transporter of the inner mitochondrial membrane**

UDPG   **Uridine diphosphoglucose**

VDAC   **Voltage Dependent Anion Channel**

Wt   **Wild type**

ABSTRACT

ELUCIDATION OF MECHANISMS GENERATING 5-HYDROXYMETHYLCYTOSINE

(5hmC) IN MAMMALIAN MITOCHONDRIA

Prashant V. Thakkar, Shirley M. Taylor, PhD

A dissertation submitted in partial fulfillments of the requirements for the degree of

Doctor of Philosophy at Virginia Commonwealth University

Department of Microbiology and Immunology, Virginia Commonwealth University, Massey

Cancer Center, Richmond, Virginia – 23298.

DNA methylation plays a pivotal role in governing cellular processes including genomic imprinting, gene expression, and development. Recently, the Tet family of methylcytosine dioxygenases(Tet1, Tet2 and Tet3) was found to catalyze the oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), an intermediate in the pathway of DNA demethylation. Tet enzymes catalyze this hydroxylation in a 2-oxoglutarate and Fe^{2+} dependent manner.

We have recently reported significant levels of 5mC and 5hmC modification in immunoprecipitates of mammalian mitochondrial DNA(mtDNA). We provide the first evidence that a DNA Methyltransferase-1 isoform (mtDNMT1) translocates to the mitochondria using an N-terminal mitochondrial targeting sequence. mtDNMT1 expression is upregulated by NRF1

and PGC1α, master regulators of mitochondrial biogenesis and function, as well as by loss of p53. Altered mtDNMT1 expression asymmetrically affects mtDNA transcription. We are now pursuing the role of Tet proteins in generating 5hmC in mtDNA. Using an in vitro enzyme assay, we have successfully detected Tet activity in crude and percoll purified mitochondrial fractions of HCT116 cells. Mitoprot analysis on Tet family predicts that Tet1 may be translocated to the mitochondria. Immunoblot analysis indicates that a band of expected size(235kDa) is present on immunoblots of mitochondrial fraction from mouse embryonic stem cells with an antibody directed against Tet1. This band, however, is not protected from trypsin treatment of mitochondria indicating that Tet1 may not be transported to the mitochondrial matrix. The putative Tet1 mitochondrial targeting sequence (MTS) fails to carry heterologous protein to the mitochondria. Knock out of Tet1 in mouse ES cells also does not alter 5hmC signal in hydroxyMeDIP assay. We now seek to determine if Tet2/Tet3 may be involved in 5hmC generation. In the nucleus, 5hmC serves as an intermediate in the process of DNA demethylation through the combined action of cytidine deaminases and the base excision repair pathway. We plan to investigate if 5hmC holds the same functional significance in the mitochondria as it does in the nucleus. Our overall goal is to understand epigenetic regulation of normal mitochondrial function and changes that occur in diseases involving mitochondrial dysfunction such as ischemic heart disease, neurodegenerative diseases like Parkinsons disease, and cancer.

Chapter 1: Introdution and Overview

Epigenetic regulation of transcription in the nucleus

Epigenetics involves processes that cause heritable changes in gene expression without a change in DNA sequence [1]. DNA methylation, histone modification and chromatin remodeling are the three key players involved in epigenetic regulation of cellular processes (Figure 1-1) [2, 3]. DNA methylation entails methylation of cytosine residues present in the context of CpG dinucleotides. Histone modification involves a range of modifications on different amino acids present on the N-terminal tails of histones H3 and H4, including methylation, sumoylation, ubiquitination, acetylation and ADP-ribosylation and phosphorylation [4]. Chromatin remodelers such as Swi/Snf and RSC complexes use energy from ATP hydrolysis to alter histone - DNA interactions and therefore bring about unwrapping of DNA from histones temporarily, facilitating binding of transcription factors and movement or eviction of nucleosomes from the transcriptional start sites [5, 6]. Interplay between these three epigenetic players affect a wide variety of processes within the cell such as development, transcription, imprinting, DNA damage repair and chromatin condensation during cell cycle [2, 3]. In each of these processes these players affect regulation by modulation of native chromatin structure.

DNA in eukaryotic cells is packaged into higher order chromatin structure through interaction with various proteins. Approximately 147 bp of the DNA sequence wraps around a histone core, comprising two dimers of histone H2A and H2B and a molecule of H3 and H4 each[7]. This basic unit of chromatin structure, nucleosome, interacts with neighboring histones to form a solenoid structure[7]. Genes with their regulatory sequences buried in the compact
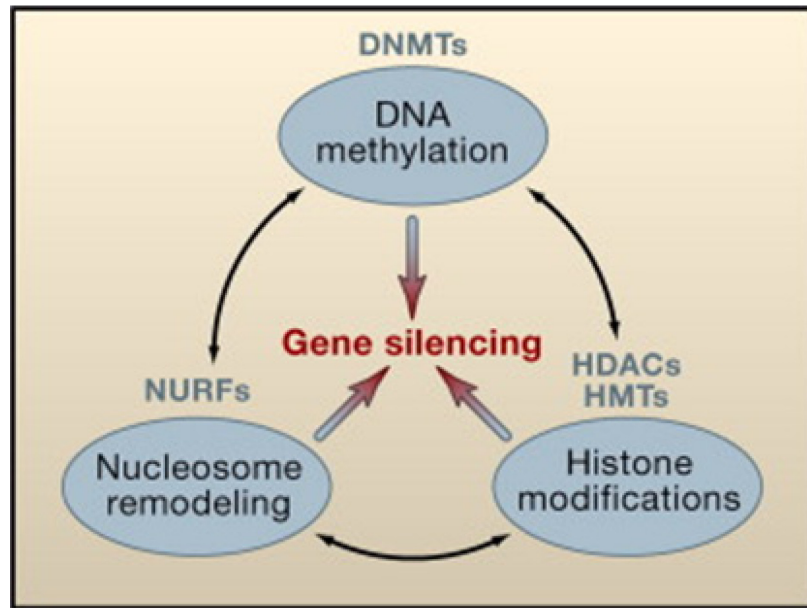
Figure 1-1: Interplay between DNA methylation, Histone modifications and Nucleosome remodeling factors.

The three key players of epigenetic regulation, DNA methylation, Histone modifications and Nucleosome alter histone-DNA interaction to modify local chromatin structure and regulate gene expression at the transcriptional level. Figure adapted from Jones and Baylin, Cell, 2007.

solenoid structure are repressed for expression and are said to be 'epigenetically silenced'. There has been a huge debate about the sequence of events involved in silencing of genes. One model proposes that, the process begins with methylation of the gene promoter by DNA methyltransferases (DNMT) [8]. Methyl DNA binding proteins such as MeCP2 and MBD2, then bind to methylated cytosines and recruit histone deacetylases (HDAC) and histone methyltransferases (HMT), which remove acetyl groups and add methyl groups respectively, on tails of histones H3 and H4 [9, 10]. These modifications, in combination, alter the net charge on histones, which brings about chromatin compaction and hence gene silencing [8]. The second theory dictates that histone modifications, at least in some cases, precedes DNA methylation [1, 11]. Evidence for this was obtained from the fact that cells lacking H3K9 methyltransferase Suv39h show reduced CpG methylation at centromeric regions [12]. The second model hypothesizes that histone modifications such as H3K9me are first laid on histone tails which in turn recruit DNMTs via heterochromatin protein 1 (HP1) [6, 11, 12]. DNMTs can then methylate cytosine in the DNA which then leads to recruitment of HDACs and thus brings about chromatin compaction [1, 11, 12]. Although the exact sequence of events needs to be elucidated, cues that underlie interplay between these three epigenetic players are more complex than initially perceived.

DNA methylation and Cancer

Generation of 5-methylcytosine (5mC) in the nuclear DNA occurs by covalent addition of a methyl group at the 5' position of cytosine present in the context of CpG dinucleotides in DNA. This reaction is catalyzed by a family of DNA methyltransferases (DNMTs) which uses S-adenosyl methionine (SAM) as a methyl donor [13].

Mammalian DNA is characterized by the presence of CpG sparse regions, rich in satellite DNA repeats, and regions dense in CpG dinucleotide content, referred to as CpG islands, most abundantly found in promoter regions of house keeping genes [14]. Somatic methylation patterns are established in the mammalian genome during early stages of development [15]. But before this process can occur, parental methylation patterns need to be removed and thus the genome undergoes a wave of demethylation [16]. Initial studies that assessed methylation patterns using methylation sensitive restriction enzymes indicated that, during development, mammalian DNA undergoes two stages of DNA demethylation, active demethylation of the paternal genome before formation of the zygote and passive loss of methylation through DNA replication post zygote formation until the morula stage [17, 18]. Patterns of methylation are then established in embryos just before implantation by the action of de novo methyltransferases DNMT3a and DNMT3b, where CpG sparse regions are hypermethylated and CpG islands remain devoid of methylation [15]. This distinct bimodal pattern of methylation is maintained in somatic cells by maintenance methylase, DNMT1. During cell division, DNMT1 maintains methylation patterns present on the parent strand by copying them on the newly synthesized daughter strand [19].

The bimodal pattern of methylation undergoes a complete reversal in cancer cells. Cancer cells are characterized by global hypomethylation and hypermethylated CpG islands [20]. This results in transcriptional repression of tumor suppressor genes and activation of oncogenes. For example, tumor suppressor genes such as p16, pRb often undergo silencing of one allele when the other is mutated [21]. Loss of methylation from repetitive elements also causes microsatellite instability, which makes the mammalian genome more susceptible to mutations [22]. Several genes involved in regulation of DNA repair and cell cyle often undergo promoter methylation e.g. RB, PTEN, MGMT, MLH1 etc. [21]. Thus, cancer can no longer be looked upon as simply a disease

caused by accumulation of genetic mutations; clearly disruption of epigenetic mechanisms is a major contributing factor.

Ten – Eleven translocation (Tet) family of enzymes and generation of 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxycytosine (5caC) in the nucleus:

Recently, three novel modifications of cytosine residues have been identified *in vivo*. These are 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxycytosine (5caC) and their formation is catalyzed by the Ten – Eleven translocation (Tet) family of proteins. Identification of these modifications has added another level of fine tuning in epigenetic regulation of cellular processes. The Tet family was initially identified as enzymes containing dioxygenase type enzyme activity during a search for mammalian homologues of the JBP family of enzymes [23, 24]. JBP enzymes are responsible for the synthesis of DNA base J (β-D-Glucosyl-Hydroxymethyluracil) in *Trypanosoma brucei* in a two step process, one of which is the Fe2+ and 2-oxoglutarate dependent hydroxylation of thymidine by thymidine hydroxylase domains of JBP1 and JBP2 [25-27]. Sequence comparisons revealed that the three paralogous Tet proteins contain the conserved hydroxylase domains of the JBP family [24]. The JBP/Tet family of enzymes is therefore classified as a family that can act on C-C bond present as a part of a side chain to an aromatic ring [23]. Structurally, this family of hydroxylases is characterized by the presence of a CXXC domain that acts as a Zn chelating domain essential for binding to chromatin, a cysteine rich domain that potentially acts as a redox center and an α-helix followed by a series of β-strand regions, defined as the double stranded β-helix (DSBH) domain [23, 24, 28].
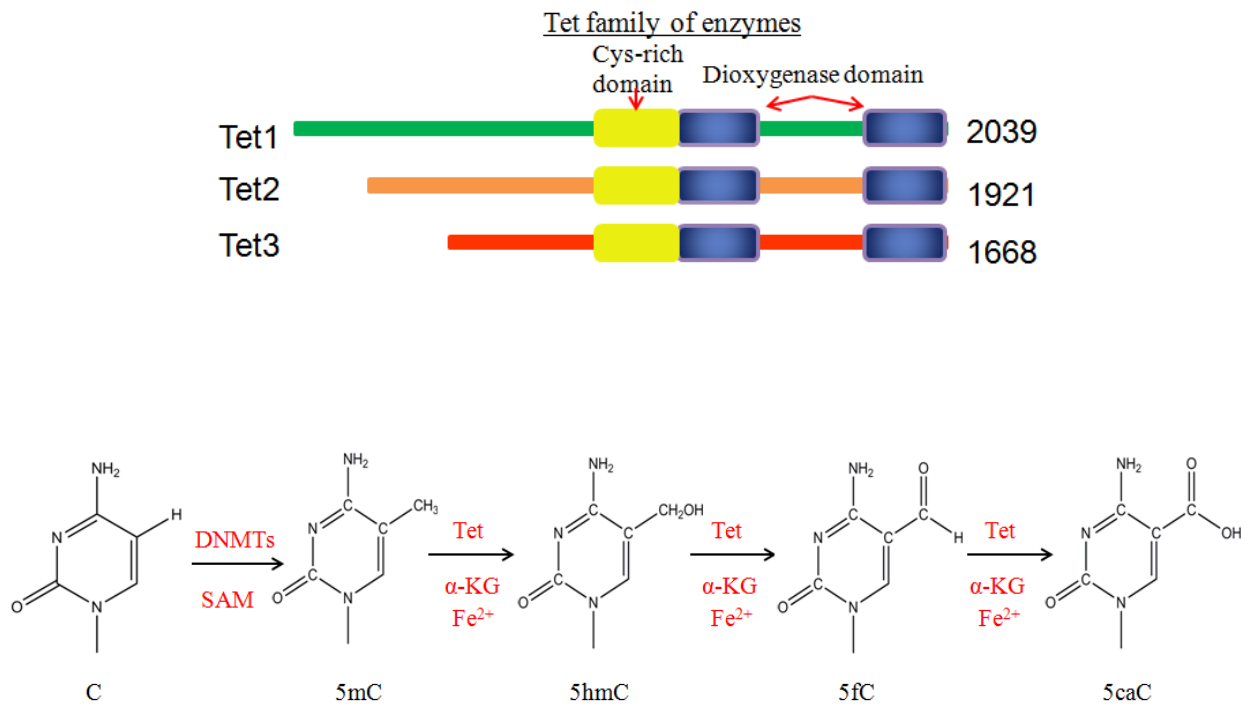
Figure 1-2:- Tet family of proteins catalyze conversion of 5mC to 5hmC, 5fC and 5caC.

The Tet family of proteins, in mammals, contains three paralogs namely Tet1, Tet2 and Tet3. The numbers on the right indicate the number of amino acids in the respective proteins. Each of the proteins contain a cysteine rich domain and a dioxyegenase domain, which are highly conserved. The three proteins show variability at the N-terminal domain. B) Schematic of the reaction catalyzed by the Tet proteins. DNA Methyltransferases (DNMTs) catalyze addition of methyl group using S-Adenosyl Methionine (SAM) as co-factor. 5mC is then oxidized by the Tet enzymes using α-ketoglutarate (α-KG) and Fe^{2+} as co-factors to generate 5hmC and subsequently 5fC and 5caC. Figure adopted from Ito et. al., Science, 2011.

Furthermore, sequence alignments showed that the DSBH domain in the Tet proteins had the following sequence characteristics conserved: a) HxD residues that help chelate Fe^{2+} b) the presence of a small residue like glycine downstream of the HxD motif that holds the active site arginine in place c) Hxs motif, where H helps chelate Fe^{2+} and s (s = small residue) interacts with α-ketoglutarate and d) A motif downsream of the Hxs motif, Rx5a, where R forms a salt bridge with α-ketoglutarate and a (a = aromatic residue) assists in proper positioning of the first iron chelating histidine [23, 24, 28]. These sequence and structural features made a strong case for Tet enzymes acting as Fe^{2+} and α-ketoglutarate dependent hydroxylases. Using *in vitro* binding assays and *in vitro* and *in vivo* enzyme assays combined with mutational studies, another independent study proved that these structural characteristics are indispensible for hydroxylase activity of Tet proteins [29].

There are three mammalian paralogs within the Tet family, namely Tet1, Tet2 and Tet3 (Figure 1-2). These three paralogs have a highly conserved C-terminal catalytic domain and show slight variability in the N-terminal DNA binding domain. All three of these enzymes, when expressed as either a full length enzyme or as only the C-terminal catalytic domain, can convert 5mC to 5hmC in a Fe^{2+} and α-ketoglutarate dependent manner (Figure 1-2) [24, 29, 30]. As opposed to methylation, hydroxylation by Tet proteins is primarily asymmetric in nature with only 21% of the 5hmCs showing symmetry, with 5mC on both strands modified to 5hmC [31]. Also, it shows strand bias in that a strand containing a high local content of guanosine is more likely to be hydroxymethylated [31]. Ito et. al. further proved that the Tet proteins, in addition to 5hmC, can also catalyze formation of 5fC and 5caC from 5mC (Figure 1-2). Purified Tet proteins were incubated with substrate containing 5mC or 5hmC and products analyzed using thin layer chromatography (TLC), were shown to contain 5fC and 5caC [32]. Further analysis using HPLC-

MS following hydrolysis of DNA revealed that an average of 3.0 x 10^4 5mC residues, 1.3 x 10^3 5hmC residues, 20 5fC residues and 3 5caC are present for every 10^6 cytosine residues in genomic DNA from mouse embryonic stem cells [32]. So far the exact mechanism of action of these enzymes is not known and needs to be elucidated.

Bisulphite sequencing does not distinguish between 5mC and 5hmC

Generation of 5hmC from 5mC makes it essential that methods used to study cytosine modifications can differentiate between methyl and hydroxymethyl modifications. The use of bisulphite sequencing to identify site specific methylation status of CpG dinucleotides has been well established for several years. Typically, the methylation status can be assessed by treating cytosine in the DNA with sodium bisulphite, which is then followed by PCR amplification and sequencing. Sodium bisulphite causes deamination of unmethylated cytosine thus converting it to uracil, which is then read as T (thymine) after PCR amplification [33]. Methylated cytosine, on the other hand, is resistant to deamination by sodium bisulphite, owing to the presence of the methyl group at the 5' position (Figure 1-3) [33]. As a result, whereas unmethylated cytosines are read as 'T's in a sequencing reaction, methylated ones are retained as 'C's. However, bisulphite sequencing cannot distinguish between 5mC and 5hmC (Figure 1-3) [34]. 5hmC upon reaction with sodium bisulphite forms a cytosine methylenesulphonate (CMS) adduct which is equally refractory to deamination, and therefore does not undergo C to T transition (Figure 1-3) [34, 35]. Thus, modified cytosines, both methylated and hydroxymethylated are indistinguishable in an ordinary bisluphite reaction.

Recently, a modified version of traditional bisulphite sequencing protocol has been employed to map 5hmC at a single nucleotide resolution. This method has been

8
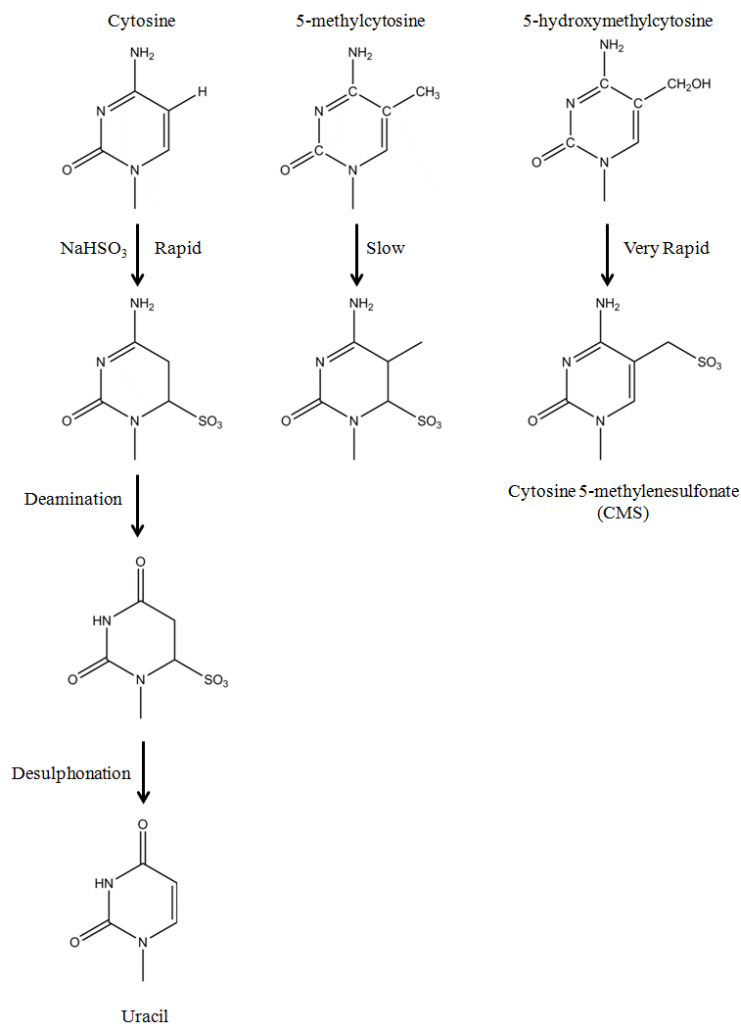
Figure 1-3: Bisulphite sequencing does not distinguish between 5mC and 5hmC.

Bisulphite treatment of cytosine adds HSO_3^- across double bond between carbons 5 and 6 of cytosine, which then undergoes deamination and desulfonation to form uracil. Conversion of 5mC to thymine upon bisulphite treatment can occur but at a rate two order of magnitude slower than that of cytosine. Bisulphite can quickly convert 5hmC to cytosine 5-methylenesulfonate adduct which is resistant to deamination. Figure adapted from Huang et. al., PLoS ONE, 2009.

termed oxidative bisulphite sequencing [36]. This method takes advantage of the fact that 5hmC, when treated with potassium perruthenate ($KRuO_4$), is converted to 5-formylcytosine (5fC), which upon bisulphite treatment allows conversion to uracil [36]. In an alternative method, 5hmC residues in the DNA are first protected by a glucosyltransferase reaction. This step is then followed by Tet mediated oxidation of 5mC to 5caC, which allows conversion of previously methylated residues to uracil upon bisulphite treatment, keeping glucosylated 5hmC residues protected from Tet mediated oxidation and subsequent bisulphite treatment [31]. Thus, bisulphite vs oxidative bisulphite treatment in comparison with the original DNA sequence can accurately help determine cytosines that are unmethylated, methylated or hydroxymethylated [36].

A number of other techniques have been assessed for their ability to identify 5hmC residues. Methyl DNA binding proteins, including all five proteins of the MBD family and $MeCP_2$, bind 5mC but are unable to bind 5hmC. Thus, although MBD/MeCP2 mediated precipitation of methylated DNA has been performed in the past, it cannot be used to precipitate 5hmC containing DNA [37, 38]. Methylation sensitive restriction enzymes, such as MspI, can cleave at the site CCGG, whether the internal cytosine is unmethylated, methylated or hydroxymethylated. This cleavage can be prevented by glucosylation of the 5hmC residue using β-glucosyltransferase in presence of UDPG [39]. On the other hand, HpaII can cleave at that site only when the cytosine is unmethylated and that cleavage is completely blocked by either methylated, hydroxymethylated or glucosylated cytosine. Thus, using a number of individual treatments, this method can be modified to study either methylation or hydroxymethylation status of DNA of interest. Additional methods include analysis by HPLC-MS post-hydrolysis of DNA to single nucleoside level [24, 32, 40]. This method is primarily used to study 5mC and 5hmC content in DNA.

Functional role of Tet proteins and 5hmC in the nucleus:

The initial discovery of 5hmC as an endogenous base in the cell led to numerous speculations about the possible functions of this modification in the nuclear genome. Since then, although quite a few have been validated to be correct, there are still others that lack experimental evidence. It has been well established that gene expression inversely correlates with methylation status of a promoter [41]. Considering 5hmC is generated from 5mC, it was obvious to test whether 5hmC counteracts the effects of methylation. It was shown that the Nanog promoter is a direct transcriptional target of Tet1 and knockdown of Tet1 in mouse embryonic stem cells reduced Nanog expression, leading to differentiation [30]. Furthermore, Tet1 acts to regulate gene transcription by preventing hypermethylation of the Nanog promoter, possibly by competing for the same site as DNMTs and converting 5mC to 5hmC [30]. Thus, by keeping the expression of nanog active in embryonic stem cells, Tet1 plays an important role in ES cell renewal and maintenance of pluripotency, mediating its effects through 5hmC [30]. Using a ChIP-Seq approach, it was found that Tet1 binds mainly to high to medium density CpG rich promoters and the majority of binding was detected at the TSS (transcription start site) of a gene and the gene body [42]. This, in conjunction with the presence of high 5hmC density around CpG rich TSS and gene bodies, strengthens the hypothesis that Tet1 and 5hmC are involved in activation of gene expression by conversion of 5mC to 5hmC, thus preventing hypermethylation of genes. Methyl binding domain containing proteins such as MBD2 and MeCP$_2$ bind to 5mC, recruit histone deacetylases (HDACs) and facilitate gene silencing through chromatin compaction. These methyl binding domain containing proteins bind 5mC but are unable to recognize 5hmC [37, 43]. This highlights a possible mechanism of action wherein conversion of

5mC to 5hmC prevents binding of methyl binding domain protein, HDACs are not recruited downstream and chromatin compaction and gene silencing is prevented.

Tet1 also plays a role in transcriptional silencing of genes. Tet1 forms a part of sin3A co-repressor complex and binds to a number of polycomb targeted genes important in regulation of development [42]. Tet1-Sin3A co-repressor complex facilitates this repression through recruitment of the PRC2 complex, and these transcriptionally repressed genes are also enriched in the repressive H3K27me3 mark [44]. In fact, more genes directly bound by Tet1 are up-regulated than are down-regulated upon Tet1 knockdown, emphasizing the role of Tet1 in transcriptional repression [42, 44].

The process of DNA demethylation is critical not only to the process of development but also epigenetic reprogramming in somatic cells [17, 45]. During development, once the zygote is formed, the entire paternal genome undergoes demethylation and methylation patterns are subsequently re-established [17]. So far the mechanism of DNA demethylation was unknown. It has now been proposed that 5hmC serves as an important intermediate in the process of DNA demethylation, which can either be passive or active. As a mechanism of active demethylation process, it was shown that 5hmC, generated by the action of Tet proteins, acts as a substrate for components of the base excision repair (BER) pathway [37, 46]. AID/APOBEC cytidine deaminases act on 5hmC to convert it to 5-hydroxyuracil (5-HU). 5-HU is then acted upon by DNA glycolases of the BER pathway, which excise it and replace it with unmethylated cytosine [37, 46]. This transcription-dependent but replication-independent process was shown to be most active in adult brain and played a role in region specific DNA demethylation, causing reactivation of silenced genes [46]. Alternatively it was shown that 5hmC may be further converted to 5-

carboxylcytosine (5-CaC) by the action of Tet enzymes, which can then be excised by thymine DNA glycosylases to replace it with unmethylated cytosine [32, 47]. As a possible mechanism for passive demethylation, it was proposed that failure of DNMT1 to recognize 5hmC would cause failure of maintenance methylation on the daughter strand. This would result in an eventual loss of methylation in the entire nuclear genome through repeated cell division. The hypothesis was supported by the fact that DNMT1 shows over 60 fold preference towards hemi-methylated DNA over hemi-hydroxymethylated DNA [37]. Furthermore, the oxidation of 5mC to 5hmC reduces DNMT1 selectivity for its substrate thereby preventing methylation of the target cytosine [43].

For a number of years it has been established that 85% of the CpGs in the mammalian genome are methylated and the other 15% of the genome that remains unmethylated comprises CpG islands. Several attempts were made to understand what keeps the CpG islands from being methylated by DNA methyltransferases. It has now come to be understood that Tet1 is enriched at CpG rich promoters or CpG islands keeping the DNMTs from methylating these regions [42]. Therefore, Tet1 contributes to DNA methylation fidelity and prevents aberrant methylation in the nuclear genome.

Thus, Tet proteins perform a wide variety of functions in the cell, some of which are mediated through 5hmC modifications while some others are mediated through protein – protein interactions of Tet proteins with their binding partners.

Mitochondria: Structure and Function

Mitochondria are essential organelles that perform a wide variety of functions including apoptosis, thermogenesis, Ca^{2+} storage, carbohydrate, lipid and protein metabolism [48]. Individual mitochondria are spherical or cylindrical in shape, depending on the tissue of origin, and are approximately $0.2 - 5$ µm in diameter and 7-20 µm in length [49, 50]. Mitochondria are often present in a reticular network surrounding the nucleus. Each mitochondrion comprises soluble matrix surrounded by the inner and outer mitochondrial membranes, both of which are composed of phospholipid bilayers interspersed with protein complexes (Figure 1-4). The outer mitochondrial membrane contains lipid and proteins in about equal proportion [51]. It is permeable to ions and contains protein complexes such as the voltage dependent anion channel (VDAC) and the transporter of outer mitochondrial membrane (TOM) complexes. The inner mitochondrial membrane is composed of a protein to lipid ratio of 80:20 and protrudes into the mitochondrial matrix forming tubular structures called cristae [50, 51]. The cristae are enriched in protein complexes required for ATP generation through the process of oxidative phosphorylation [50]. The inner mitochondrial membrane is impermeable to ions and hence contains a range of transporters which allow exchange of ions and other metabolites. For example, the adenine nucleotide translocase transports ATP in exchange for cytosolic ADP, tricarboxylate carrier transports citrate/isocitrate in exchange for malate, Na^+/K^+ transporter transports the cations in exchange for a proton, among others [50, 51]. The outer and the inner membranes are separated by an inter-membrane space, which is aqueous in nature (Figure 1-4). The mitochondrial matrix is enriched in enzymes and substrates important for metabolic processes such as the citric acid cycle, lipid and amino acid oxidation, heme biosynthesis etc [51]. It also contains mitochondrial DNA, components important for its replication and
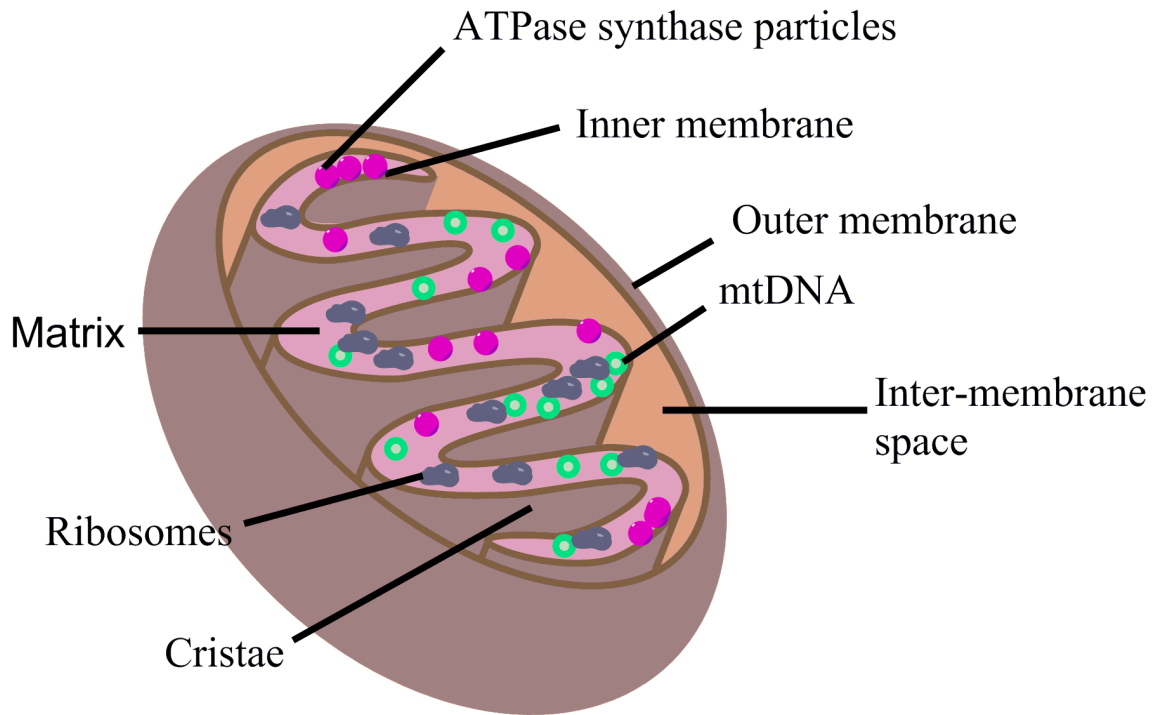
Figure 1-4: Internal structure and organization of mitochondria.

Mitochondria contain different compartments, outer membrane, inter-membrane space, inner membrane and mitochondrial matrix. Components in the mitochondrial matrix such as ribosomes, mtDNA and the respiratory chain complexes are shown. This figure is adapted from Perkins et. al, Micron, 2000.
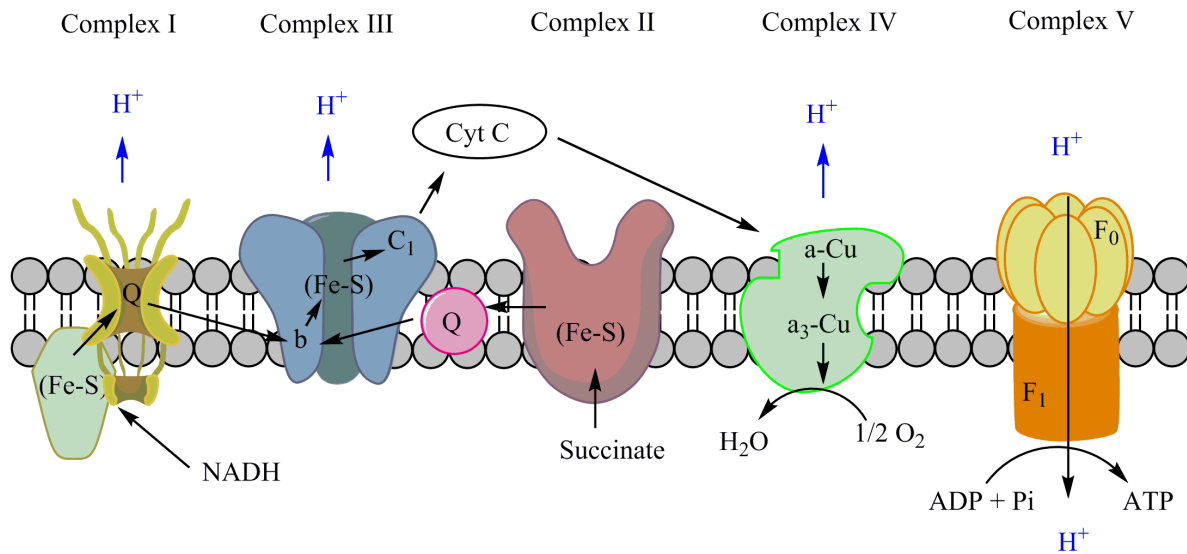
Figure 1-5: Schematic of complexes involved in oxidative phosphorylation.

Energy in mitochondria is generated by five protein complexes, Complex I, II, III, IV and V. Complex I, III and IV pump protons across the inner mitochondrial membrane during the process of electron transfer, which establishes an electrochemical gradient. Energy from the electrochemical gradient is then used by ATP synthase complex (Complex V) to synthesize ATP from ADP and Pi. This figure is adapted from Eng. et. al. Nature, 2003.

maintenance, and the transcription and translation machinery essential for expression of proteins encoded by the mitochondrial genome (Figure 1-4) [51].

Mitochondria are referred as 'power houses' of the cell because of their major function in generating ATP in the cell, using five different complexes of the respiratory chain embedded in the inner mitochondrial membrane, complex I, II, III and IV and V (Figure 1-5) [52]. These complexes bring about transfer of electrons from NADH and FADH to O_2 in three steps [52]. In the first step, electrons are transferred from NADH and FADH, generated during oxidation of carbohydrates and lipids, to co-enzyme Q (CoQ) via Complex I [51]. In the second step, Complex III facilitates transfers of electrons from reduced CoQ to cytochrome c (cyt c). In the third and the final step, cyt c transfers electrons to Complex IV which in turn transfers it to O^- radical, which is then reduced to water [51] (Figure 1-5). Each of these three steps of electron transfer are coupled to a proton pump, which transfer H^+ ions from mitochondrial matrix to the inter-membrane space. This process leads to formation of an electrochemical gradient across the inner mitochondrial membrane, owing to its ion impermeable nature [52]. The free energy of electron transport conserved in this electrochemical gradient is then used by the ATP synthase complex V to generate ATP from ADP and inorganic phosphate (Pi) (Figure 1-5). Four out of five of the above mentioned complexes contain at least one protein encoded on the mitochondrial genome, while also containing other proteins that are encoded in the nucleus [53]. This suggests that co-ordinated transcription and translation of the mitochondrial proteins, that are nuclear and mitochondrially encoded, is an important phenomenon for the process of energy generation.

Structural organization and replication of mtDNA

The mitochondrion is the only organelle in a mammalian cell, apart from nucleus, that carries its own genetic material. Mitochondrial DNA (mtDNA) is a circular, double stranded piece of DNA which is ~ 16,300 bp in mouse and ~16,600 bp in humans (Figure 1-6) [54]. It lacks introns and codes for 13 proteins, 22tRNAs and 2 rRNAs on both, the heavy and the light strand (Figure 1-6), called so due to difference in base composition and therefore their buoyant densities in CsCl gradient. Each mammalian cell contains approximately 1000 – 10,000 molecules of mtDNA [55]. One of the key differences between nuclear and mtDNA is that mtDNA is not wrapped around a histone core. However, mtDNA is packaged into higher order structures called nucleoids. Approximately 2-8 molecules of mtDNA are present in a sphere 70nm in diameter [56]. mtDNA associates with DNA binding proteins, such as TFAM, to form a highly compact rosette like structure. TFAM is known to coat the entire mtDNA, binding every 15-20 bp [57]. Other proteins that bind to mtDNA and are involved in the formation of the nucleoid structure are mitochondrial single stranded binding proteins (mtSSB), mitochondrial RNA polymerase (POLRMT), mitochondrial DNA polymerase (POLG), chaperone protein HSP60 and prohibitin [58, 59].

Replication of mtDNA begins at the origin of replication present on the heavy strand (O_H), using primers generated as a result of transcription from the light strand promoter (LSP) (Figure 1-7). There are three conserved sequence blocks, CSB I, CSB II and CSB III, present downstream of LSP (Figure 1-7). CSBII particularly acts to stabilize the RNA – DNA hybrid formed during initiation of mtDNA replication [55]. DNA synthesis from the heavy strand proceeds in only one direction and often stops prematurely 700 bp downstream of O_H forming 7S DNA [60].

At this point, a decision is made whether the replication of DNA should halt or proceed using as yet unknown mechanisms. Once replication on the heavy strand proceeds two-thirds of the way, another origin of replication on the light strand (O_L) is revealed and replication then begins on the light strand in the other direction [60]. Once replication has proceeded bidirectionally throughout the length of the mtDNA molecule, new mtDNA molecules are formed.

Mitochondrial inheritance and diseases:

Inheritance of mtDNA in mammals occurs uniparentally through the maternal germline. Oocyte contains 10^3- 10^4 times more mtDNA as compared to a mature sperm [61-63]. In the mature oocyte, just prior to fertilization, there is rapid amplification of mtDNA which results in an identical population of mtDNA molecules. Thus, mtDNA transmitted to the daughter cells is homoplasmic in nature [64]. To maintatin this homoplasmy mature sperm mtDNA is eliminated following fertilization. This is brought about by ubiquitination of sperm mitochondrial outer membrane, followed by recognition and digestion by the proteolytic enzymes in the oocyte [65]. Maintenance of homoplasmy is important since heteroplasmy acquired through paternal leakage and recombination of mtDNA causes mitochondrial myopathy [66]. Heteroplasmy can also result due to a 100 fold higher error rate of mitochondrial polymerase G and presence of limited DNA repair mechanisms in the mitochondria [67]. A mutated copy of mitochondrial DNA is tolerated in low copy number; however, upon segregation during meiosis homoplasmy can occur, where in only the mutated copy of mtDNA is present [68]. This can in turn result in mitochondrial diseases in the offspring that shows either homoplasmy for the mutated copy of mtDNA or carries the mutated copy in a high proportion relative to the wild type copy of mtDNA[68].

A number of mitochondrial diseases are known to occur in humans that result from acquired mutations in mitochondrial DNA. Leber's hereditary optic neuropathy (LHON) is known to occur due point mutations in the coding region of NADH dehydrogenase subunit 4 gene [69]. Mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes (MELAS) is another example where the disease phenotype is associated with a point mutation in tRNA (Leu) gene of mtDNA [70]. Mutation in tRNA (Leu) gene is also known to cause deafness and diabetes mellitus II [71]. In each of these cases, the onset of phenotype is determined by the proportion of mutant versus wild type molecules of mtDNA present in the affected tissue.

Transcription in Mitochondria:

Transcription in mitochondria occurs on both the Heavy (H) and the Light strand (L) encoding a total of 13 proteins, 2 ribosomal RNAs and 22 transfer RNAs [72]. The promoter for each strand, the Heavy strand promoter (HSP) and Light strand promoter (LSP), is present in the D-loop control region and serves as the site of transcription initiation [72]. Transcription of the heavy strand is initiated at two sites, HSP1 and HSP2, but only at one site (LSP) on the light strand promoter [73]. Transcription from HSP2 and LSP promoters proceeds in two different directions, each generating polycistronic transcripts corresponding to the entire heavy and light strands [74]. These polycistronic messages lack introns, and undergo processing by a mechanism known as the 'tRNA punctuation model' [75]. The polycistronic messages contain mRNA and rRNA sequences where each is immediately flanked by a tRNA sequence. These tRNA sequences adapt a clover leaf-like structure which acts as a substrate for RNAseP and other RNAses [75]. Processing by these RNAses then releases mRNAs and rRNA along with tRNAs. Transcription from HSP1 starts at a site located 16 bp upstream of the tRNA^Phe and terminates at
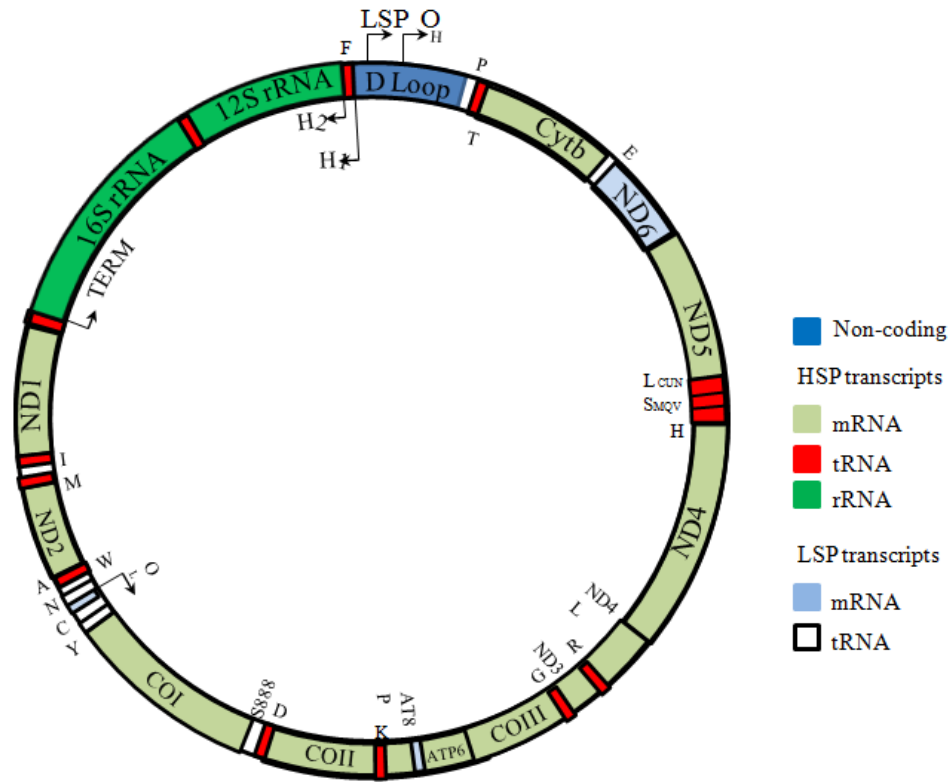
Figure 1-6: Schematic representing organization of the mitochondrial genome.

The mitochondrial genome comprises of the D-loop control region and contains 13 protein coding genes, 2 ribosomal rRNA genes and 22 tRNA genes. The transcription termination site is present downstream of the 16S rRNA genes. The figure is adapted from Falkenberg et. al. Ann Rev. Biochem 2007.

a site downstream of the 16S rRNA gene. This generates a transcript containing only 16S rRNA and 12S rRNA genes separated by tRNA[V] [74]. The mitochondrial transcription machinery involves mitochondrial DNA dependent RNA polymerase (POLRMT), Transcription factor of mitochondria B1, B2 and A (TFB1M, TFB2M and TFAM) and mitochondrial termination factor (MTERF) and these factors are sufficient for *in vitro* transcription [76, 77]. POLRMT is a 1230 amino acid protein coded in the nucleus. The C-terminal portion of the protein is highly conserved with the T-odd bacteriophage polymerase [78]. Similar to T7 polymerase, POLRMT contains a specificity loop that binds to the +10 to -4 region on the light strand. It forms a complex with TFB2M to form a two component core transcription machinery which is sufficient to initiate transcription from both strands [79].

TFB1M and TFAM serve as accessory proteins for POLRMT to perform its function. TFB1M, like TFB2M, forms a heterodimeric complex with POLRMT but functions less efficiently as a transcription factor when compared to TFB2M [80]. TFAM contains two high mobility group (HMG) domains separated by a 27 residue linker region. It also contains a 25 amino acid long C-terminal tail essential for specific DNA recognition and transcriptional activation [81, 82]. TFAM can bind, unwind and bend DNA, a characteristic of HMG domain containing proteins. Combined with the sequence specificity of the C-terminal tail, this property allows TFAM to bind specifically to the region between HSP and LSP promoters where it can bend and unwind the D-loop regulatory region, thus facilitating transcription initiation [83]. It was recently shown that transcription initiation requires only POLRMT and TFB2M and that in the presence of these two factors, transcription from HSP1 is more efficient than that from LSP [76]. Upon addition of TFAM there is a switch, with transcription from LSP becoming far more efficient than from HSP1 [76].

22
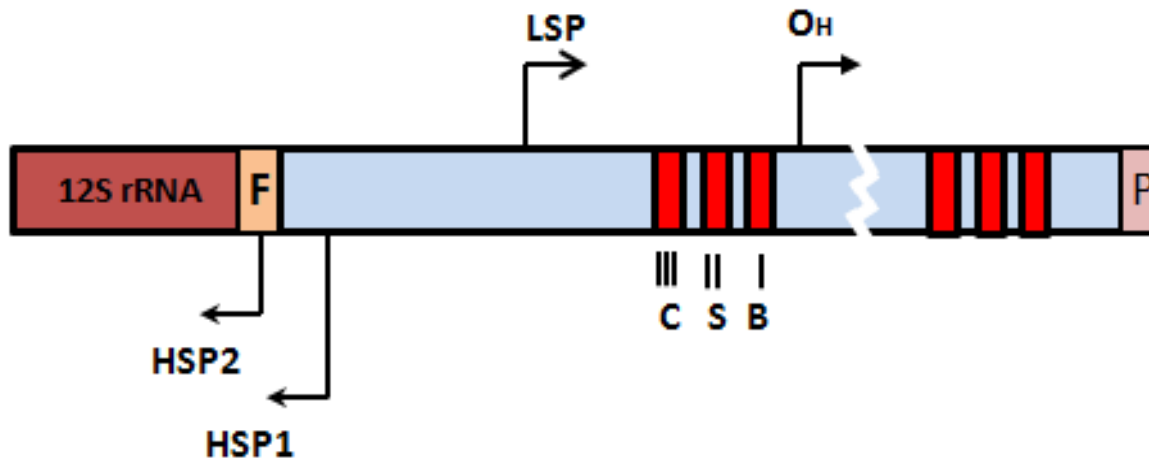
Figure 1-7: Architecture of the D-loop control region of the mitochondrial DNA.

Promoter regions of the Heavy and the light strand of the mitochondrial DNA are depicted. Transcription on the heavy strand can begin at two sites, HSP1 and HSP2, whereas transcription on light strand begins only at LSP. CSB I, II and III indicate conserved sequence blocks. Sequences transcribed within CSB regions serve as RNA primers for mtDNA replication. TAS serves as the termination site for mtDNA replication. Figure adapted from Falkenberg et. al., Ann rev Biochem, 2007.

Transcription begins at the three known sites, HSP1, HSP2 and LSP, yet there is only one known transcription termination site. This site is located downstream of the 16S rRNA gene and acts bi-directionally to terminate transcription from both the strands. Transcription termination factor M (MTERF), a 39-kDa protein, binds to this 28 bp region in a sequence specific manner to terminate transcription starting at HSP1 and LSP promoters [84]. Termination of transcription from HSP2 occurs at a site upstream of tRNA^{Phe}, yielding the polycistronic mRNA. A second binding site for MTERF was found near HSP1 [74]. It has been proposed that MTERF regulates initiation of transcription from HSP1 promoter by binding at that site [74]. It also regulates the rate of transcription from HSP1 promoter by binding at the two known binding sites, thus forming a transcription bubble [74]. This allows efficient recycling of the transcription machinery resulting in increased production of 12S rRNA and16S rRNA transcripts [74].

Import of proteins to the mitochondria:

The mitochondrial proteome is comprised of approximately 1500 proteins but only 13 of them are encoded by the mitochondrial genome [85]. The majority of mitochondrial proteins are encoded in the nucleus, synthesized by the free ribosomes present near the mitochondrial membrane and then transported to the mitochondria [86, 87]. The protein import machinery consists of multiple components including the mitochondrial targeting sequence (MTS), cytosolic and mitochondrial chaperone proteins, translocases of the outer and inner membrane, matrix metalloproteases as well as other less well characterized components.

Proteins that are destined for transport to the mitochondria carry sequences that serve as a code for their localization to the mitochondria. Proteins transported to the matrix usually contain MTSs that are 10 to 60 amino acids in length and form an amphipathic helix with one face

containing hydrophobic residues and the other containing positively charged residues [88]. These sequences are usually devoid of negatively charged residues [89]. The MTSs are usually N-terminal in position but proteins containing C-terminal MTSs, e.g. DNA helicase Hmi1 in yeast, or internal MTSs, e.g. BCS1, have also been described [90, 91].

Translocation of proteins across the mitochondrial membrane is facilitated by the translocases of the outer and inner membranes, i.e. TOM and TIM complexes respectively, in a two step process. In the first step, the precursor protein is transported across the outer mitochondrial membrane into the inter membrane space, which is then further transported in to the mitochondrial matrix by the TIM complex [92]. The TOM complex is composed of seven different subunits; Tom70, Tom40, Tom22, Tom20, Tom7, Tom6 and Tom5. Tom70 and Tom 20 serve as the receptors of the TOM complex that are involved in recognition of MTSs of the nuclear encoded mitochondrial proteins [92]. Proteins containing N-terminal MTSs are first recognized by AIP protein which then interacts with Tom20 receptor protein [93, 94]. The precursor protein is then transported through the pore of the complex, which is formed by Tom40, Tom7, Tom6 and Tom5 and has a diameter of 20 $^{\circ}$A [95].

Precursor protein from the TOM complex is then directed to the TIM23 complex, which represents the major translocase of the inner membrane. Its components can be divided into two major groups: the protein-conducting channel and the import motor. The protein conducting channel is comprised of three subunits namely, Tim50, Tim 23, Tim 21 and Tim17. Tim50 interacts with the incoming precursor protein, thus acting as a receptor for the TIM complex [96]. The precursor protein is then passed on to Tim23, which forms the channel of the complex. Tim23 contains a coiled-coiled domain that is essential for homodimerization and also for

interaction with the precursor protein [97]. The protein conducting channel uses energy from the membrane potential to help transfer only the MTS of the precursor protein, after which point the process is take over by the import motor component of the Tim23 complex [98]. The import motor component of the Tim23 complex consists of Tim44, Tim14 (Pam18), Tim 16 (Pam16), mtHsp70 and MgeI subunits [96]. Tim44 first binds to the precursor protein and transfers it to mtHSP70, recruited with the help of two DnaJ-like proteins, Tim12 and Tim16 [99]. mtHSP70 contains an ATP bound and an ADP bound form. In an ATP bound form, mtHSP70, which is in a complex with the precursor protein and Tim44, uses energy from hydrolysis of ATP to dissociate from Tim44 [100]. At this point, the protein moves further into the matrix. A series of such reactions ensure complete transport of precursor protein into the mitochondrial matrix [100]. Once the protein is entirely in the mitochondria, the MTS is recognized and cleaved by the matrix metalloproteases. The rest of the protein undergoes folding into its native conformation and is now functionally active.

Thus, using a series of complex and highly co-ordinated interactions, proteins are transported to the different compartments of the mitochondria.

Nuclear mitochondrial cross-talk:

Cross-talk between nucleus and mitochondria depends on the PGC1 family of co-activators which interact with transcriptional factors such as NRF1, ERR, PPARα, Sp1 etc. to regulate expression of genes involved in mitochondrial transcription, translation, replication and the respiratory chain complexes (Figure 1-8). Both NRF1 and PGC1α are involved in mitochondrial biogenesis in response to various environmental stimuli such as oxidative stress, reduced ATP levels, increased intracellular Ca^{2+} levels, exercise and growth factors [101-103].

PGC1α binds to NRF1 and activates nuclear genes that encode subunits that participate in the five complexes of the respiratory chain. They can also regulate the expression of TFAM, TFB1M and TFB2M mitochondrial transcription factors, thereby linking nuclear and mitochondrial regulatory networks (Figure 1-8).

Nitric oxide is a normal component of mammalian cell, generated by the action of nitric oxide synthase enzymes (NOS), and serves as an important signaling molecule in facilitating oxygen sensing by cells [104]. It was shown that in response to nitric oxide, PGC1α is upregulated along with NRF1, due to activation of guanylate cyclase and therefore increased cGMP dependent signaling, which in turn leads to increased mitochondrial biogenesis in various mammalian cells [101, 104] (Figure 1-8). Different tissues from NOS null mice displayed significantly reduced mtDNA content, low oxygen consumption and reduced ATP content. This suggested that NO induces mitochondrial biogenesis, primarily via PGC1α and NRF1 dependent activation of nuclear encoded mitochondrial protein expression [101, 104]. Furthermore, PGC1α is regulated at a transcriptional level by the cAMP dependent pathway [105, 106]. Elevated cAMP levels activate CREB via protein kinase A-mediated phosphorylation of CREB, which then binds to cAMP response element present in the PGC1α promoter [106]. CREB is also activated by other kinases such as Ca2$^+$/calmodulin-dependent protein kinase (CaMK) and MAPK, both of which cause activation of PGC1α expression in response to cytosolic calcium levels and exercise, respectively (Figure 1-8) [102, 105, 106]. In response to increased cAMP signaling, MAPK also phosphorylates and activates activating transcription factor-2 (ATF-2) and PGC1α, which in turn induces mitochondrial biogenesis. cAMP dependent activation of mitochondrial biogenesis and regulation of thermogenesis is important for norepinephrine dependent response to cold. PGC1α

is also known to activate genes involved in fatty acid oxidation in the mitochondria via PPARα [107].

Mitochondrial regulatory mechanisms feed back into the nuclear regulatory pathways through a retrograde pathway, as demonstrated in yeast [108]. Although such a retrograde pathway is yet to be identified in mammalian mitochondria, Ca^{2+} dependent signaling has been thought to play a role in the flow of information from mitochondria to the nucleus (Figure 1-8) [109]. Ca^{2+} is primarily stored in mitochondria and mitochondrial dysfunction or high nitric oxide content leads to increased cytosolic Ca^{2+} levels [109]. This, in-turn, causes activation of CaMK and cAMP dependent pathways to increase PGC1α expression, which ultimately results in increased mitochondrial biogenesis [109]. Thus, as described above, there are nuclear regulatory pathways, that regulate mitochondrial biogenesis and function, and mitochondrial regulatory pathways that feed back into the nucleus to communicate the functional state of mitochondria.

Methylation of DNA in the mitochondria: History

Several attempts have been made to prove the existence of methylation in mtDNA, with earliest reports dated as early as four decades ago. Nass et. al.,1973, reported the presence of 5mC in mtDNA and reported DNA methylase activity in mitochondrial fractions from the different cells studied. They used four different approaches to detect 5mC in the mitochondrial DNA, all of which were based on growing cells in a medium containing different radioactive compounds and looking for incorporation of radioactivity in 5mC fraction of mtDNA [110]. However, these studies were marred by a number of technical issues. Use of ^{3}H- L- methionine lead to nonspecific radioactivity in thymine and adenine fractions apart from 5mC [110]. Use of ^{3}H-
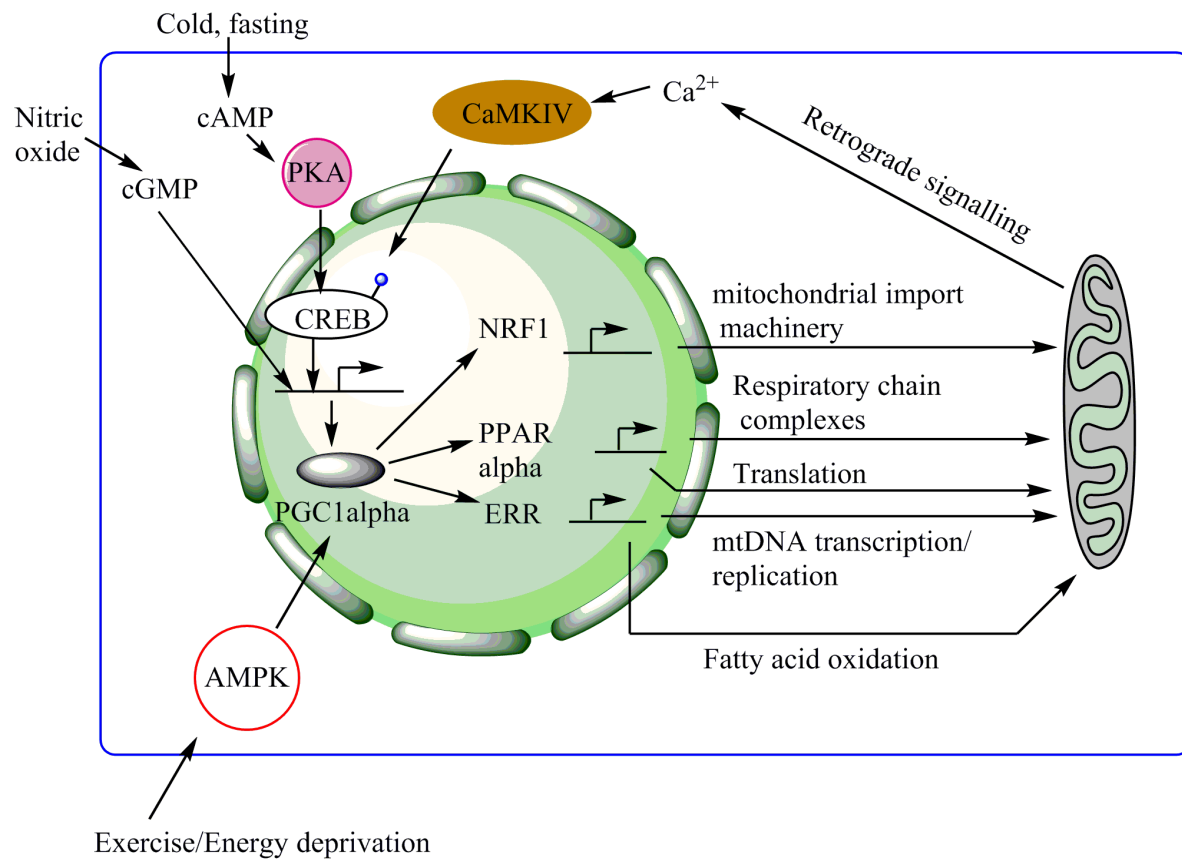
Figure 1-8: Different mechanisms underlying nuclear-mitochondrial cross-talk.

Flow of information between nucleus and mitochondria is bidirectional in nature as depicted above. PGC1 family of co-activators, along with various transcription factors co-ordinate expression of genes involved in multiple mitochondrial processes. Expression of PGC1 co-activators is regulated by different signaling pathways that respond to environmental stimuli. Signalling from mitochondria to the nucleus is mediated through Ca^{2+} dependent retrograde signaling**.** Figure adapted from Scarpulla et. al. Phys rev 2008.

deoxycytidine was complicated in that it mostly allowed detection of radioactivity in polymeric forms as opposed to nucleotide forms [110]. Even *in vitro* enzyme assays performed on the nuclear and mitochondrial fractions gave different activities based on whether heterologous or homologous DNA was used. The purity of the mitochondrial fractions used was determined simply on the basis of the differential reaction of nuclear and mitochondrial enzyme fractions to β-mercaptoethanol [110]. Overall, this report of 5mC in mtDNA failed to be convincing.

Another study by Pollack et. al., 1984, used nearest neighbor analysis to report the presence of 5mC in mouse mitochondrial DNA at levels similar to that reported by Nass et. al. They reported that 5mC was present in the context of CpG dinucleotide at a frequency of 3-5% of mitochondrial CpGs [111]. To gain an understanding of site specific methylation patterns in mitochondrial DNA, methylation sensitive restriction enzymes MspI and HpaII were employed [111]. This method looked at methylation status at only the MspI/HpaII restriction sites in mtDNA [111]. Also, no attempt was made to detect the presence of methylase activity in the mitochondria that might explain the mechanism generating 5mC nucleotide in mtDNA [111].

In 1994, Cardon et al. analyzed mitochondrial DNA from different species in order to study the representation of different dinucleotides. Similar to nuclear DNA, it was found that CpG dinucleotides were under-represented in all mitochondrial DNAs studied [112]. This suggested CpG suppression in mitochondrial DNA. CpG suppression is a phenomenon where the ratio of observed to the expected frequency of CpG dinucleotide in the genome is less than 1 [112]. The traditional explanation behind the occurrence of CpG suppression is the conversion of cytosine, present in the context of CpG dinucleotide, to 5-methylcytosine by the action of DNA methyltransferases [113]. This 5mC can then undergo deamination, either spontaneously or by

enzymatic action of deaminases, to thymidine [113, 114]. This changes the dinucleotide from CpG to TpG causing a mismatch with GpC present on the complementary strand. Through successive rounds of replication, TpG**.**GpC is changed to TpG.CpA causing a transition mutation in the genome [114]. Considering CpG suppression occurs using the above mentioned mechanism, it has been considered as a sign for presence methylation either currently or that it existed in the genome at some point. CpG suppression in the mitochondrial DNA was seen to the same extent as in nuclear DNA (Observed/Expected frequency for CpG dinucleotide was found to be 0.53 for mitochondrial DNA and 0.42 for nuclear DNA) [112]. This suggested that methylation was probably present in the mitochondrial DNA. CpG suppression is marked not only by under-representation of the CpG dinucleotide but also by over representation of the TpG dinucleotide. However, in mitochondrial DNA, even though CpGs were under represented, over representation of TpGs did not follow. Instead, over abundance of CpC.GpG dinucleotides was seen. This probably was due to the asymmetric composition of the two strands of the mitochondrial DNA [112].

Despite these early studies, the presence of 5mC in mtDNA has remained controversial.

Summary and Objectives

DNA methyltransferases play a central role in regulating multiple processes in the cell including development, differentiation and transcription, mediated by conversion of C to 5mC. The fidelity of methylation patterns is maintained by the Tet family of enzymes that prevent hypermethylation of CpG islands through catalytic conversion of 5mC to 5hmC. Loss of DNMT or Tet activities causes aberrant methylation in the nuclear genome, characteristic of diseases such as cancer. Thus, the balance between the two family of enzymes is critical for efficient and

complex epigenetic regulation of cellular processes. Although the importance of DNA methylation and hydroxymethylation has been studied extensively in the context of nuclear genome, their importance in the mitochondria remains to be explored. Previous studies aimed at detecting methylation in mitochondria were viewed with skepticism. Studies described in this dissertation were aimed at addressing the long standing debate on the presence of methylation in mtDNA. We have rigorously pursued the identification of 5-methylcytosine in mtDNA and its role in mitochondrial function. We demonstrated the presence of 5-hydroxymethylcytosine in mtDNA and elucidate mechanisms involved in its generation in mammalian mitochondria. Our goal was to explore mitochondrial epigenetic mechanisms that govern normal function of mitochondria and which may be disrupted in diseases characterized by mitochondrial dysfunction.

Chapter 2: Functional significance of cytosine methylation by mtDNMT1 in mammalian mitochondria.

Introduction

The presence of cytosine methylation in mtDNA has been a matter of debate over the last three decades. Using nearest neighbor analysis and methyl-sensitive restriction enzymes, low levels of methylation were reported in mtDNA from different mammalian cells studied [110]. The lack of knowledge of purity of mitochondrial fractions made these results unconvincing [110]. In 1984, another study reported that 3-5% of CpGs in mtDNA are methylated, although no attempt was made to detect methylase activity in these mitochondria [111]. It was then reported that, similar to nuclear DNA, mtDNA shows CpG suppression, suggesting that 5mC can serve as a hot spot for mutations in mtDNA [112]. Although the presence of 5mC has been described in mtDNA, mechanisms generating 5mC in mtDNA were not known. The lack of a mitochondrially encoded DNA methyltransferase led us to explore the possibility of a nuclear encoded DNA methyltransferase that might be translocated to the mitochondria, and thus generating 5mC in mtDNA.

Mitochondrial DNA (mtDNA) is a circular, double stranded genome of ~ 16.5kb in mammals [54]. It comprises two strands, the heavy and the light strand encoding a total of 13 proteins, 2 ribosomal RNAs and 22 transfer RNAs [72]. Transcription of the heavy strand is initiated at HSP1 and HSP2 and at LSP on the light strand [73]. Transcription from HSP2 and LSP promoters proceeds in two different directions, each generating polycistronic transcripts corresponding to the entire heavy strand and most of the light strand [74]. Transcription from the HSP1 promoter yields 16S rRNA and 12S rRNA transcripts and is terminated at the MTERF

binding site, present at 3' end of 12SrRNA gene. All components essential for transcription of mtDNA, including POLRMT, TFAM, TFB1M and TFB2M are encoded in the nucleus and are transported to the mitochondria.

Regulation of DNMT1 by p53 in the nucleus:

p53 is a known tumor suppressor gene that is frequently mutated in colon, lung, liver and prostate cancers. It performs a wide variety of functions in the cell including cell cycle regulation, gene expression, induction of apoptosis, development and differentiation. In most cases, p53 mediates its effect by controlling expression of the key molecular species involved in each of these mechanisms. It can act as a transcriptional activator or as a repressor. Some of the examples of genes activated by p53 are Bax, Bak, Puma etc. whereas those repressed are Bcl-2, Bcl-X$_L$, MAP4 etc.

Our lab showed that p53 acts as a transcriptional repressor and negatively regulates expression of DNMT1 [115]. This was deduced from the fact that deletion of p53 caused ~6-fold upregulation of DNMT1 in an immunoblot and a northern blot assay. The p53 binding site is characterized by the presence of two half sites with RRRCWWGYY N(0-13) RRRCWWGYYY as the consensus sequence. Sequence analysis of the DNMT1 promoter revealed three putative p53 binding sites, at -1196 to -1174, -553 to -533 and +30 to +56 at the DNMT1 promoter. EMSA assays demonstrated high affinity binding of p53 at DNMT1 promoter at the site labeled +30 to +56 [115]. ChIP experiments further revealed p53 binds at this site *in vivo*, and activation and stabilization of p53, following ionizing radiation or VP16 treatment, reduced p53 binding at that site *in vivo*, which leads to ~5 fold upregulation of DNMT1 [115]. Thus, p53 inhibits DNMT1 at the transcriptional level, facilitated by direct binding of p53 at the DNMT1 promoter [115]. This

explained, at a molecular level, the phenomenon of CpG hypermethylation that occurs in most cancers that are also associated with loss of p53. Another mechanism that also contributed to the above mentioned phenotype was regulation of DNMT1 by pRb/E2F pathway [116].

Objectives:

Several mechanisms are involved in the bidirectional flow of information between the nucleus and mitochondria. The PGC1 family of co-activators, along with other transcription factors, activates expression of nuclear encoded mitochondrial genes that are involved in a large number of processes in the mitochondria. Using Ca^{2+} dependent signaling and controlled flow of molecules such as S-adenosyl methionine, signals from mitochondria feed back into the nuclear regulatory mechanisms co-ordinating the two genomes. Identification of 5mC in mtDNA and a mitochondrial isoform of DNMT1, led us to investigate the functional significance of cytosine modification by mtDNMT1 in mammalian mitochondria. Our goal was to determine whether mitochondrial epigenetic mechanisms help to co-ordinate expression of nuclear and mitochondrially encoded genes and therefore contribute to the cross-talk between the nuclear and mitochondrial genomes.

Materials and Methods:

Cell Culture

HCT116 cells (ATCC – CCL 247) were grown in RPMI 1640 medium (Gibco/Invitrogen) supplemented with 10% FBS (Gemini) at 37 $^{\circ}$C in 5% CO_2. Wild type and p53-/- mouse embryonic fibroblast cells (MEFs) were grown at 37 $^{\circ}$C in 10% CO_2 in DMEM (Gibco/Invitrogen) supplemented with 10% FBS (Gemini). Owing to the cell cycle regulated

expression of DNMT1, cells were fed 24 hrs prior to replating them for experimental set up. Cells were harvested 24 hours after replating thus ensuring the majority of cells were in S-phase and therefore maximally expressing of DNMT-1. Maintaining this procedure for all the experiments helped with consistency and reproducibility of results.

Preparation of lysates and Bradford assay to determine protein concentrations:

Whole cell, mitochondrial or cytoplasmic pellets were resuspended in SDS lysis buffer (2.25 ml EDTA free protease inhibitor tablets (Roche) resuspended in sample buffer (62.5mM Tris-HCl, pH 7.5, 5% glycerol, 2% SDS), 2.025 ml water, 225 µl β-Mercaptoethanol). The lysates were then passed through 21 gauge needle approximately 15 times to sheer chromatin or, alternatively, sonicated in a bath sonicator (Diagenode, Bioruptor) for 10 mins on High for 30 sec ON and 30 sec OFF. Lysates were stored at -80 °C until further use.

Protein concentration of lysates was determined using the Bradford assay. A standard curve was generated using 0, 2, 4, 6, 8, 10 µg/ml concentrations of BSA (Fisher, Cat no. 23210). 1 µl of SDS lysis buffer, from an aliquot saved after making lysates, was added to all the tubes containing standard. The final volume was adjusted to 800 µl with water. For sample lysates, 1 µl of lysate was added to 799 µl water. 200 µl of Bradford reagent (Biorad, Cat no. 5000006) was added to all the standards and the samples. Absorbance was measured at 595 nm using a Spectrophotometer.

Subcellular fractionation to isolate mitochondria from Cultured Cells

Cells were plated in 150 mm dishes and harvested 24 hrs later for experimental consistency. Depending on the cell type and the downstream application, a varying number of dishes were set

up. Dishes (150 mm) that were approximately 90% confluent were rinsed twice with cold 1 X PBS, pH 7.4 and scraped into a pre-weighed 15 ml conical bottom centrifuge tube (BD Falcon, Cat no.:- 352096), making sure the dishes were on ice at all times. Cells were collected by centrifuging at 3000 rpm for 10 minutes at 4 ^{O}C in a Beckman J6-MI centrifuge. The tubes were weighed to get an estimate of the whole cell pellet weight. For Immuno blot, Cells were resuspended in 5 ml cold 1 X PBS, pH 7.4 and 500 μl (~10%) was retained in a separate pre-weighed 1.5 ml centrifuge tube for preparation of the whole cell lysate. Whole cell lysate was prepared by pelleting cells at 900 x g for 5 minutes at 4 ^{O}C in a table top centrifuge, tubes were weighed to obtain the pellet weight and cells were resuspended in SDS lysis buffer (2.25 ml Protease Inhibitor solution, 2.025 ml water, 225 β-mercaptothanol) in a volume 10 times the amount of the pellet weight. The remaining 90% of the cells were centrifuged and resuspended in 3.5 ml of mitochondrial homogenization buffer (0.25 M sucrose, 10mM Tris-HCl, pH 7.0, 1mM EDTA pH 6.8) containing EDTA –free Protease inhibitor tablets (PI) (Roche, Cat no:- 11873580001, 1 tablet per 25 ml of homogenization buffer). The cells were transferred to a pre-chilled 7 ml Dounce homogenizer and incubated on ice for 5 minutes. The cells were dounce homogenized for 15 strokes, centrifuged at 2200 rpm (1300 x g) for 5 minutes at 4 ^{O}C in J6-MI centrifuge and the supernatant (Post-Nuclear supernatant, PNS) was collected in a separate 15 ml conical centrifuge tube. This Post-Nuclear supernatant contains the mitochondrial and the cytosolic fractions, while the nuclear and the unbroken cells form the pellet. The pellet was again resuspended in 3.5 ml of mitochondrial HB buffer and the above procedure was repeated a total of three times, collecting PNS each time. The post-nuclear supernatant, collected at the end of three rounds of douncing, was centrifuged at (2500 x g) 3000 rpm for 5 minutes at 4 ^{O}C in J6-MI centrifuge to pellet the remaining unbroken cells. The supernatant was collected, leaving behind

2 ml to avoid contamination from unbroken cells, and transferred to a 14 ml round bottom centrifuge tube (BD Falcon, Cat no:- 352059) and centrifuged at 9000 rpm (10,000 x g) for 15 minutes at 4 ^{O}C in a Beckman J2-HC centrifuge using a JA-17 rotor. The pellet is the mitochondrial fraction, which is then collected and washed twice with mitochondrial homogenization buffer with protease inhibitor tablets.

Immunoblotting:

To load equal cell equivalents, 75 µg of whole cell lysate, 25 µg of cytosplamic lysate and 18 µg of mitochondrial lysate was loaded. When loaded in these ratios, protein from lysates of these fractions gave an equal intensity of VDAC signal, used as a marker for the mitochondrial compartment. The protein concentration was determined by Bradford assay and volumes corresponding to the appropriate amount of protein were placed in a 1.5 ml eppendorf tube. An equal volume of sample buffer (475 µL laemulli sample buffer + 25 µl β-mercaptoethanol) was added to each of the samples. The final volume was adjusted using SDS lysis buffer to keep volume constant in all tubes. Samples were boiled for 5 mins and then loaded on to 4-15% gradient gel together with an aliquot of Precision plus dual colour protein ladder (Biorad- Cat no. 1610374). The gels were run at 150V for 60 mins in 1X running buffer (25mM Tris, 250mM glycine, 0.1% SDS).

For semi dry transfer, a sandwich was set up such that a gel was laid on top of 9 layers of whatman filter paper (Whatman 3mm Cat no. 3030917). Each sheet of whatmann paper was soaked in 1X towbin buffer (25mM Tris-HCl, pH 8.3, 192mM glycine, 10% methanol and 0.04% SDS) and rolled to squeeze out the bubbles after laying it on top of the previous layer. A PVDF membrane (Millipore, Cat No. IPVH00010), soaked in 100 % methanol, was then laid on

top of the gel followed by another six layers of whatman paper also soaked in 1X transfer buffer. The transfer sandwich was assembled on the bottom portion of the semi dry transfer apparatus (Hoefer Cat. No. TE77) and the lid was laid on top. Transfer was set up at a constant current of 80mA for 4 hours and was left overnight in the transfer apparatus once the transfer was complete. The transfer sandwich was removed the next day and the membrane was reactivated in 100% methanol before setting it up for blocking. Blocking of the PVDF membrane was performed in Starting block buffer (Thermo scientific, Cat no. 357543) for 1 hr at room temperature with shaking. The membrane was then washed thrice with 1X TBS (0.5M Tris-HCl, pH 7.5, 0.14M NaCl, 2.7mM KCl and 0.1% Tween 20) containing 0.05% Tween20, for 5 mins each, with shaking at high speed. The membrane was then incubated with the appropriate primary and HRP-conjugated secondary antibodies for 1 hr each at room temperature with shaking. For Immuno blots containing cell fractionation markers, blots were cut horizontally at the 75kDa, 50kDa, and 37kDa mark and the four parts, from top to bottom, were probed for DNMT1, tubulin, VDAC and H_3K_4me3 antibodies respectively. The membrane was washed again thrice with 1X TBS between primary and secondary antibody incubations and also after secondary antibody incubation. The blot was developed using West Pico chemiluminescent substrate (Pierce, Cat no. 34080). In certain cases, West Dura (Pierce, Cat. No. 34075), a chemiluminescent substrate of higher sensitivity had to be used. Blots were then developed by exposing the membranes to X-ray films (ISC Bioexpress, Cat no. F9024-8X10). A list of all the antibodies used and their optimized conditions are listed in table 2-1.

MeDIP/HydroxyMeDIP:-

Sonicated genomic/mitochondrial DNA (4 µg) was diluted with TE buffer to 450µl total volume, denatured for 10 minutes in boiling water and immediately cooled on ice for 10 minutes. To this mixture was added 51 µl of 10x IP buffer, followed by 2 µl (2 µg) of 5mC Antibody (Active Motif, Cat No. 39649)/ 5hmC Ab (Active Motif, Cat. No. 39791)/ Non-specific Mouse IgG (Millipore, Cat. No. 12-371)/ Non-specific Rabbit IgG (Santacruz, Cat. No. sc-2027) and samples were incubated overnight at 4 °C with end over end rotating.

In a separate tube, Protein G beads (20 µl per sample) were centrifuged at 4500 rpm for 5 min to get rid of 70% ethanol solution. The supernatant was carefully discarded by pipetting. To the pellet of protein G beads, 400 µl of 0.1% BSA in 1X PBS, pH 7.4 (10 mg of BSA was dissolved in 1 ml water which was then added to 9 ml 1X PBS, pH 7.4) was added. It is important that beads are always resupended by flicking the tube and never by vortexing or pipetting up and down. The beads were rotated for 5 mins at 4 °C and were later centrifuged at 4500 rpm for 5 mins at 4 °C. The supernatant was carefully discarded by pipetting. The wash with 0.1% BSA in PBS, pH 7.4 was performed for a total of three times. At the end of last wash, beads were resuspended in an equal volume of 1X IP buffer to make a 50% slurry and incubated with sonicated lambda DNA and BSA such that 5 µg of both were present per 30 µl of 50 % slurry. Beads were incubated for 3 hrs with rotation at 4 °C. At the end of 3 hrs, beads were collected by centrifugation and washed again thrice with 0.1% BSA in 1X PBS, pH 7.4 with rotation for 5 mins at 4 °C and centrifugation at 4500 rpm for 5 mins at 4 °C. At the end of the third wash, beads were resuspended in an equal volume of 1X IP buffer to form a 50% slurry. To

each sample, 20 µl of the 50% slurry was added and the mixture was incubated further for 2 hrs with rotation.

At the end of the 2 hr incubation, beads were collected by centrifugation and the supernatant was discarded. The beads were then washed by adding 500 µl 1X IP buffer and rotating the samples for 10 mins at 4 $^{\circ}$C. The beads were centrifuged at 4500 rpm for 2 mins at 4 $^{\circ}$C, to get rid of the unbound excess DNA which was present in the supernatant. The supernatant was discarded carefully by pipetting. The beads were such washed thrice. At the end of the third wash, the beads were resuspended in 250 ul of Proteinase K digestion buffer and 3.5 µl of Proteinase K (20 mg/ml) and incubated at 37 $^{\circ}$C for 3 hrs. The samples were then extracted with 250 ul of Phenol – Chloroform – Isoamyl alcohol (25:24:1) and centrifuging the samples at 13,000 rpm for 3 mins at 4 $^{\circ}$C in a table top centrifuge. This separates the sample into aqueous phase and organic phase. The aqueous phase, which contains the DNA, was collected in a separate tube. Chloroform (250 µl) was added to the collected aqueous phase, sample was vortexed and centrifuged at 13,000 rpm for 3 mins at 4 $^{\circ}$C in a table top centrifuge. This step ensures the removal of excess phenol from the first step. The aqueous layer from the second step was again collected in a separate tube and subjected to ethanol precipitation. Ethanol precipitation was performed by adding 25 µl of 3M sodium acetate pH 5.2, 20 units glycogen and 750 µl of 100% ethanol (3 volumes), mixed thoroughly and stored at -20 $^{\circ}$C overnight. The DNA was collected the next day by centrifugation at 10,000 rpm for 15 mins at 4 $^{\circ}$C. The supernatant was discarded and the DNA pellet was washed using 500 µl of 70% ethanol and spinning the samples at 10,000 rpm for 15 mins at 4 $^{\circ}$C. The DNA was resuspended in 75 ul TE and analyzed by qPCR. Standard deviations (SDs) were calculated using the formula:

SD = Avg(sample)/Avg(input) x ($\sqrt{((SD(sample)/Avg(sample))2+(SD(input)/Avg(input))2)}$).

Antigen	Primary Ab obtained from	Dilution used	Blocking buffer	Secondary Antibody used	Secondary Antibody dilution used	Expected protein size
N-terminal DNMT1	Abcam (Cat. No. ab16632)	1:1000	Starting Block	Goat anti-rabbit	1:15,000	185kDa
Tubulin	A gift from Dr. Larner	1:10,000	5% milk in 1X TBST	Goat anti-mouse	1:10,000	60kDa
VDAC	Calbiochem (Cat No. AP1059)	1:4000	Starting block	Goat anti-rabbit	1:10,000	32kDa
H_3K_4me3	Upstate (Cat. No. 07-463)	1:1000	Starting block	Goat anti-rabbit	1:10,000	17kDa

Table 2-1: List of antibodies used.

Antibodies used for immunoblot experiemtns are listed along with their optimized conditions.

Species	Primer Set	Forward	Reverse
Mitochondrial gene expression			
Mouse	ND1	5'- CAGGATGAGCCTCAAACTCCA-3'	5'-CGGCTCGTAAAGCTCCGA-3'
	ATP6	5'- ATTCCCATCCTCAAAACGCC – 3'	5'- TGTTGGAAAGAATGGAGACGGT-3'
	ND6	5'-GGTTATGTTAGAGGGAGGGATTGG-3'	5'- TACCCGCAAACAAAGATCACCCAG-3'
	Cox1	5'- TCGCAATTCCTACCGGTGTC-3'	5'- CGTGTAGGGTTGCAAGTCAGC-3'
	18S rRNA	5'- GTCTGTGATGCCCTTAGATG -3'	5'-AGCTTATGACCCGCACTTAC-3'
	12S rRNA	5'- ACACATGCAAACCTCCATAGACCG-3'	5'- TGGCTGGCACGAAATTTACCAACC -3'
	16S rRNA	5'- TCTGTTAACCCAACACCGGAATGC -3'	5'- TCGTTTAGCCGTTCATGCTAGTCC -3'
Methyl DNA Immunoprecipitation (Me-DIP) and HydroxyMeDIP			
Human	ATP6	5'-ATTCAACCAATAGCCCTGGCCG-3'	5'-ACGTAGGCTTGGATTAAGGCGAC-3'
	COX2	5'-ACAGATGCAATTCCCGGACGTC-3'	5'-TGGGCATGAAACTGTGGTTTGCTC-3'
	12S rRNA	5'-AGTTCACCCTCTAAATCACCACG-3'	5'-TGACTTGGGTTAATCGTGTGACC-3'
	16S rRNA	5'- ACCTTACTACCAGACAACCTTAGCC-3'	5'-TAGCTGTTCTTAGGTAGCTCGTCTGG-3'
	27	5'-GCATTTGGTATTTTCGTCTGGGGGG-3'	5'-GACCAACCCTGGGGGTTAGTATAGC-3'

Table 2-2: List of primers used.

Primers designed against mitochondrial DNA used for the analysis of mitochondrial transcription and MeDIP analyses are shown.

Design of primers for analysis by MeDIP:

Primers were designed across regions that defined the coding sequence for each of the genes. Multiple genes were selected for analysis, the two rRNA genes, three protein coding genes on the heavy strand and the only protein coding gene on the light strand, and the D-loop regulatory region. Mitochondrial genes lack introns and hence it was possible that amplification could result from psuedogenes present in genomic DNA. To overcome this problem, each of the primer sets were tested using an *in silico* PCR analysis tool on the UCSC database (http://genome.ucsc.edu/cgi-bin/hgPcr) for amplification of fragments from genomic DNA. Primer sets that amplified the same size fragments from both mitochondrial and genomic DNA were discarded. Only primer sets that amplified the expected size fragment from mitochondrial DNA exclusively were used for further analysis. Primers used are listed in table 2-2.

Validation of specificity of 5mC and 5hmC antibodies towards their respective modifications:

Unmethylated, methylated and hydroxymethylated substrate (Active Motif Cat no. 55008) (25 pg) was mixed with 500 ng of sonicated genomic DNA from HCT116 (ATCC CCL-247) cells. All three DNA templates obtained from Active Motif contain the Human APC promoter sequence and only differed in the modification of cytosine present. MeDIP and HydroxyMeDIP were then performed to validate the antibodies as described above.

Isolation of RNA:

A 100 mm dish was set up for RNA isolation such that cell were 70-80% after 24 hrs. For experimental consistency, cells were harvested at 24 hrs for RNA isolation. RNA was isolated from different cell lines using the Trizol reagent (Invitrogen, Cat. No.15596-026). Cells were washed twice with ice cold (1X) PBS, pH 7.4 and lysed in 4 ml Trizol and incubated on ice for 5mins to ensure complete lysis. The homogenate was then transferred to a 14 ml round bottomed falcon tube (BD, Cat. No.352059) and incubated at room temperature for 5 min to allow complete dissociation of the nucleoprotein complexes. Chloroform (0.2 ml) was added per ml of Trizol and tubes were shaken vigorously and incubated for another 2-3 mins at room temperature. The tubes were centrifuged at 12,000 x g for 15 mins in the Beckman J2 centrifuge (JA-17 rotor). This step causes the separation of aqueous phase and organic phase. The aqueous phase, which contains the RNA, was transferred to another 14 ml round bottomed tube. RNA was then further precipitated with 0.5 ml 100% isopropanol per ml of trizol, incubated for 10 mins at room temperature and centrifuged at 12,000 x g for 10 mins at 4 $^{\circ}$C. A RNA pellet could be seen at the end of this step. The pellet was further washed with 1 ml of 75% ethanol per ml trizol and centrifuged at 7,500 x g for 5 mins at 4 $^{\circ}$C. The pellet was allowed to air dry for 5 mins before resuspending in DEPC treated water. Concentration was measured using the nanodrop-1000 spectrophotometer. RNA was stored at -80 $^{\circ}$C in three volumes of ethanol until further use.

cDNA synthesis and analysis of mitochondrial transcription using analysis by qPCR:

Prior to reverse transcription, ~15 μg RNA was precipitated using $1/10^{th}$ volume of 3M sodium acetate and 20 units of glycogen (Ambion, Cat no. AM9516) and centrifuging the tubes

at 10,000 x g for 10 mins at 4 ^{O}C. A wash with 500 µl of 75% ethanol was performed by centrifugation at 10,000 x g for 10 mins before resuspending the pellet in DEPC treated water.

For studies on mitochondrial transcription, 15 µg RNA was DNAseI treated in a total volume of 75 µl with 4 units of DNAseI (Ambion, Cat no. AM2222) and 1X DNAseI buffer (Ambion) at 37 ^{O}C for 30 mins. At the end of 30 min incubation, EDTA was added to a final concentration of 5mM and samples were further incubated at 75 ^{O}C for 10 mins to inactivate DNaseI. RNA was ethanol precipitated to remove DNAseI and was resuspended in 20 ul DEPC treated water for use in cDNA synthesis. Concentration was determined by nanodrop and integrity of RNA was checked on an agarose gel.

cDNA was synthesized using First Strand Superscript III cDNA synthesis kit (Invitrogen, Cat no. 18080051) using 5 µg of total RNA. For synthesis using random hexamers, RNA was first incubated with 50 ng/µl random hexamers and 10mM dNTPs at 25 ^{O}C for 10 mins to allow annealing of primers to the RNA strands. cDNA synthesis mix (5X RT buffer, 10mM MgCl$_2$, 0.02mM DTT, 40 Units RNAse OUT, 200 Units Superscript III reverse transcriptase enzyme) was added and samples were incubated for 50 mins at 50 ^{O}C. The reaction was terminated at 85 ^{O}C for 5 min. cDNA was incubated with 2 units RNAseH (Ambion, Cat no. AM2292) at 37 ^{O}C for 30 mins to degrade RNA. For each sample, a reverse transcriptase negative control (-RT) was set up to control for amplification from residual contaminating DNA. For cDNA synthesis using gene specific primers, 2 µM ATP6 sense or antisense primers was used to synthesize cDNA exclusively from Light and Heavy strands of mitochondrial DNA respectively using the protocol described above. cDNA was then further analyzed using Quantitect SYBR green (Qiagen, Cat no. 204145).

For analysis of mitochondrial transcription, RNA was isolated from wild type (wt) and p53 -/- MEFs and cDNA was synthesized using the procedure mentioned above using both random hexamers and gene specific primers in two different experiments. Statistical analyses involved random effects analysis of variance (ANOVA) test using JMP version 7.0. Least square means for each gene analyzed from wt and p53-/- cells was obtained with standard errors and data was normalized to that of 18S rRNA. Three sets of biological replicates were combined, where each set contained three technical repeats. The ANOVA model included technical repeats as nested effects and biological repeats as random effects. The corresponding 95% confidence intervals were obtained by using the statistical method for transformations. The primers used for analysis of mitochondrial gene transcription are listed in Table 2-2. The cycling conditions for the qPCR are as follows:- Step 1: Activation at 95 ^{O}C for 15 mins, Step 2: Denaturation at 94 ^{O}C for 30 secs, Step 3: Annealing at 55-60 ^{O}C for 30 secs, Step 4: Extension at 72 ^{O}C for 30 secs, Step 5: Plate read, Step 6: Go to Step 2, 40 times, Step 7: Final extenstion at 72 ^{O}C for 5 mins, Step 8: Melting curve, 40 ^{O}C to 90 ^{O}C, read every 1 ^{O}C, hold 1 sec.

Results:

An isoform of DNMT1 is present in the mitochondria:

DNA methyltransferases are enzymes responsible for methylation of DNA in the nucleus. Sequence analysis of DNMT1, a member of DNA methyltransferase family, brought to light a sequence upstream of the known transcriptional start site that extended up to 303 bp in humans and 90 bp in mouse. This sequence contained two ATG codons that could potentially serve as the translational start codon and protein starting at any of these two ATGs would be in frame with the published translational start site for nuclear DNMT1, ATG_3. We asked if this upstream sequence, present as a part of DNMT1 protein, would facilitate its translocation to an organelle other than the nucleus. Mitoprot analysis of the DNMT1 protein sequence containing this upstream portion indicated a high probability of transport to the mitochondria and suggested that this upstream sequence would serve as the mitochondrial targeting sequence (MTS) [117]. Mitoprot analysis was also performed against DNMT1 sequences from various species including chimpanzee, rat and cow, and in each of the species one or more in-frame ATGs were found that would code for a peptide with a high probability of mitochondrial transport. Although these putative DNMT1 MTSs found in various species did not show sequence conservation, their functional conservation was demonstrated by their ability to operate across species [117].

To provide experimental evidence for these analyses, we asked whether an isoform of DNMT1 was present in the mitochondria. Mitochondria were isolated from HCT116 neoS cells and lysates were prepared using SDS lysis buffer. Immunoblots containing whole cell, cytoplasmic and mitochondrial lysates were probed with an antibody directed against DNMT1. This antibody identifies amino acid 1-10 of nuclear DNMT1 sequence and should also

successfully detect an isoform of DNMT1 containing the upstream sequence. Immunoblot analysis revealed the presence of full length DNMT1 protein and a smaller band (~60 kDa) in the mitochondrial fraction from these cells (Figure 2-1). The mitochondrial isoform of DNMT1 migrates at ~185kDa, similar to nuclear DNMT1, implying that the upstream sequence that serves as a putative MTS is cleaved upon migration of DNMT1 to the mitochondria. The presence of the 60 kDa band in the mitochondrial fractions indicates that full length mtDNMT1 may be proteolytically processed in the mitochondria. VDAC (Voltage-dependent ion channel), an outer mitochondrial membrane protein, Tubulin and H_3K_4me3 (histone H_3 carrying a tri-methyl mark on lysine 4) serve as the mitochondrial, cytoplasmic and nuclear compartment specific markers respectively and demonstrate the purity of each of the fractions. Comparison of VDAC signal in whole cell and mitochondrial fractions indicates that equal amounts of protein have been loaded as per cell equivalents (Figure 2-1). The absence of H_3K_4me3 signal in the mitochondrial fraction indicates that our mitochondrial preparations are free from nuclear contamination, a primary site of localization for the DNA methyltransferases (Figure 2-1). These studies were performed by Erica Peterson and are mentioned here for the sake of completeness.

p53 negatively regulates mitochondrial isoform of DNMT1:

As mentioned earlier, p53 can act as a transcriptional repressor and negatively regulates nuclear DNMT1 through direct binding to a p53 consensus sequence in the DNMT1 promoter [115]. Loss of p53 or activation of p53 by irradiation or VP16 treatment relieves this repression and increases expression of DNMT1 [115]. Lisa Shock, a former student in the lab, asked whether p53 could regulate expression of mtDNMT1. Lysates from mitochondrial fractions of wild type and p53-/- MEFs along with the respective whole cell and cytoplasmic
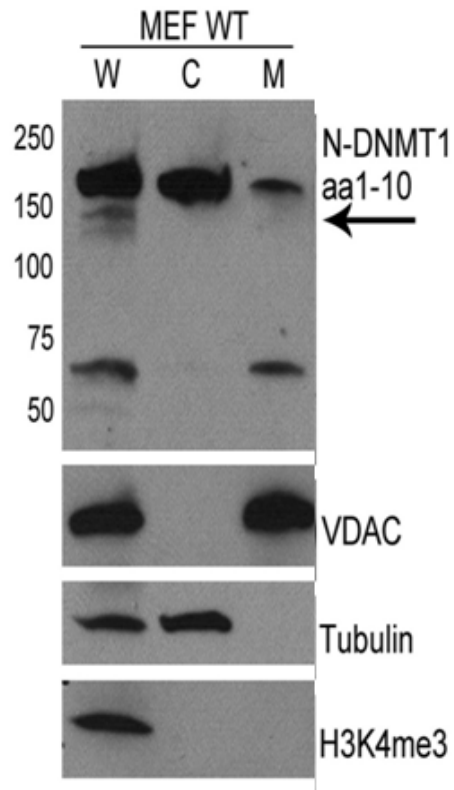
Figure 2-1: Detection of DNMT1 in mitochondrial fractions of HCT116 neoS cells.

Lysates from whole cell, cytoplasmic and mitochondrial fractions from HCT116 neoS cells were resolved on a 4-15% gradient and detected using an antibody against DNMT1. Band corresponding to ~185kDa, the expected size of DNMT1, could be detected in the mitochondrial fraction. This experiment was performed by Erica Peterson.
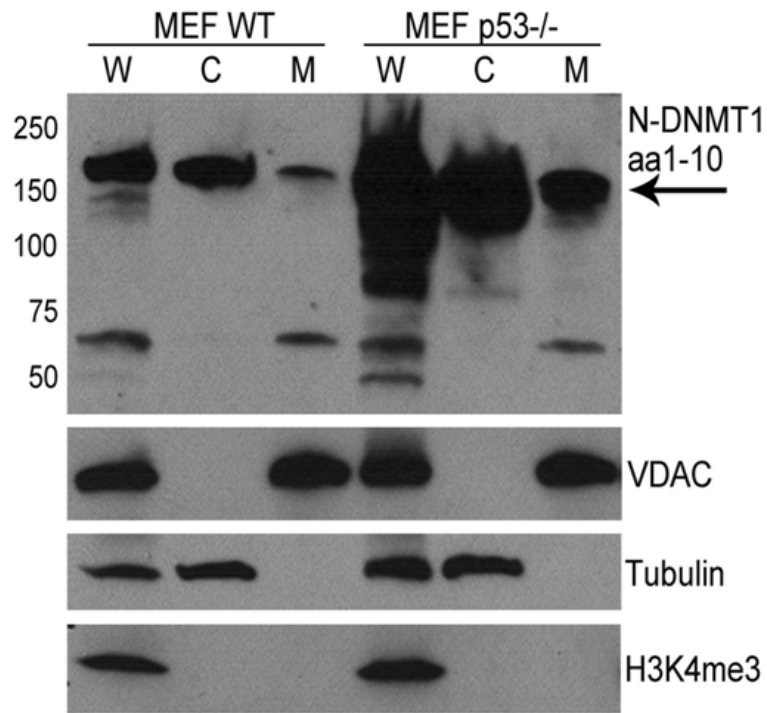
Figure 2-2: mtDNMT1 expression is regulated by p53:

In order to analyze if mtDNMT1 is regulated by p53, lysates from whole cell (W), cytoplasmic (C) and mitochondrial (M) fractions from wild type and p53-/- mouse embryonic fibroblasts (MEFs) were analyzed by Immuno blot. An approximately 5 fold increase in signal for mtDNMT1 could be seen in mitochondrial fractions of p53-/- cells as compared to wild type. This experiment was performed by Lisa Shock.

fractions were resolved on a 4-15% gradient gel. The blot was probed using an antibody against DNMT1. Analogous to its effect on nuclear DNMT1, the loss of p53 led to an approximately five fold increase in mtDNMT1 expression as seen by signal in mitochondrial fraction from p53-/- cells as compared to that in wild type (Figure 2-2). VDAC, Tubulin and H_3K_4me3 serve as the mitochondrial, cytoplasmic and nuclear compartment specific markers and signify the purity of each of the fractions (Figure 2-2). There is a predicted p53 binding site in the DNMT1 sequence upstream of ATG_3 [117]. Binding of p53 at this site has been shown by our lab [115]. Further ChIP analysis might be required in wild type and p53-/- MEFs to prove relief of p53 binding with simultaneous increase in mtDNMT1 expression.

Antibodies against 5mC and 5hmC are specific for respective modifications

Since immunoblots performed in our lab indicated the presence of DNMT1 within the mitochondrial matrix, we sought to determine whether modified cytosine residues could be detected in mtDNA. While several assays are available to study the methylation status of DNA, very few have been useful to study 5hmC modification in DNA because of a limited ability to differentiate between 5mC and 5hmC modifications. However methylated DNA immunoprecipitation (MeDIP) and HydroxyMeDIP assay could in principle distinguish these modifications, provided the antibodies are specific for their respective modifications. It was therefore essential that we prove the specificity of antibodies against 5mC and 5hmC in order to use MeDIP and HydroxyMeDIP assay as a primary tool to assess methylation and hydroxymethylation status of mtDNA. We mixed 25pg of unmethylated, methylated or hydoxymethylated purified PCR templates (Active motif) with 500 ng of HCT116 neoS genomic DNA, sonicated to a size range of 200-700 bp, for immunoprecipitation with 5mC, 5hmC and
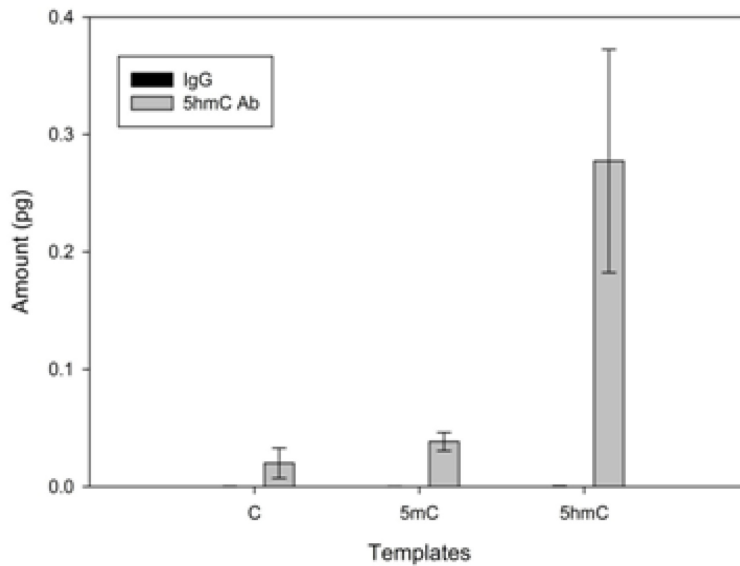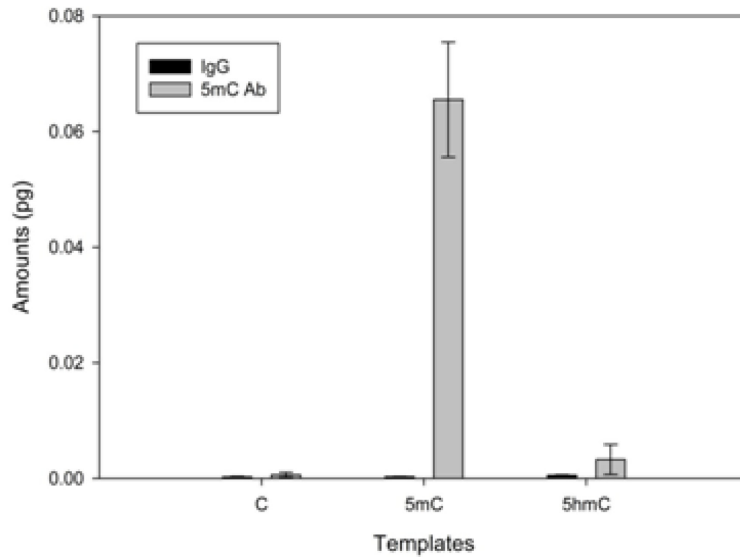
Figure 2-3: Validation of antibodies for specificity against the respective modifications.

5mC (A) and 5hmC (B) antibodies specifically pulldown methylated and hydroxymethylated substrate respectively. Data represents mean ± SD of three independent biological repeats.

IgG antibodies. Immunoprecipitated material was quanitated using qPCR and revealed that cross reactivity was less than 10% of the specific antigen- antibody pull-down (Figure 2-3). This degree of specificity was deemed sufficient to study the relative levels and types of cytosine modification in mtDNA.

Both 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) are present in mitochondrial DNA:

Since earlier studies claiming to demonstrate the presence of 5mC in mtDNA had been met with skepticism due to technical limitations, we sought to assess the presence of modified cytosine in mtDNA using antibodies specific for 5mC and 5hmC. We isolated mtDNA from HCT116 neoS, sonicated it to a size range of 200-700 bp and subjected it to immunoprecipitation with 5mC antibody. The immunoprecipitated DNA was analyzed by qPCR using primer sets against different regions of mitochondrial DNA. Significant pull down of mitochondrial DNA could be seen with 5mC antibody as compared to that of non-specific IgG (Figure 2-5, upper panel). Five different regions of mitochondrial DNA were analyzed and in each of them at least 10-20 fold higher signal was obtained with 5mC antibody than IgG (Figure 2-5, upper panel). These studies demonstrated conclusively that 5mC is present in purified mtDNA.

Having established presence of 5mC in the mitochondrial DNA, we then asked whether 5-hydroxymethylcytosine is present in mitochondrial DNA. HydroxyMeDIP was performed in a manner similar to that described above with 5hmC antibody along with its respective IgG control. Five different regions on mitochondrial DNA were analyzed for presence of 5hmC. They were mtIP27, representing the D − loop control region of mtDNA, 12S and 16S rDNA, encoding mitochondrial rRNAs, as well as Cox2 and ATP6, the two protein coding genes on the

54
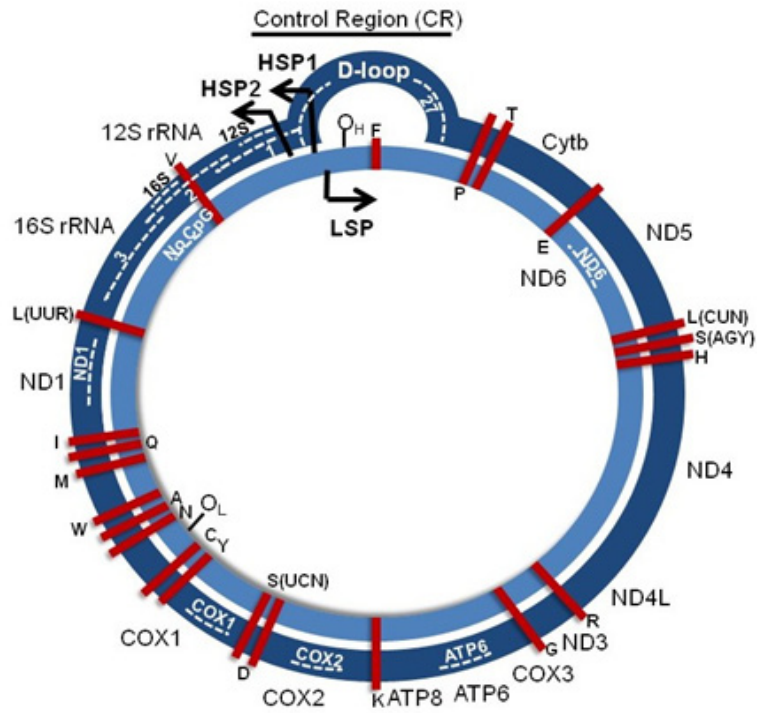
Figure 2-4: Representation of genes analyzed by Methyl DNA Immunoprecipitation (MeDIP) assay:

Five different regions on the mitochondrial DNA were analyzed for end point analysis of MeDIP assay. The amplicons are indicated by the white dotted lines within the blue strands of mitochondrial DNA. The analyzed regions were mtIP27, 12S rRNA, 16S rRNA, Cox2 and ATP6.
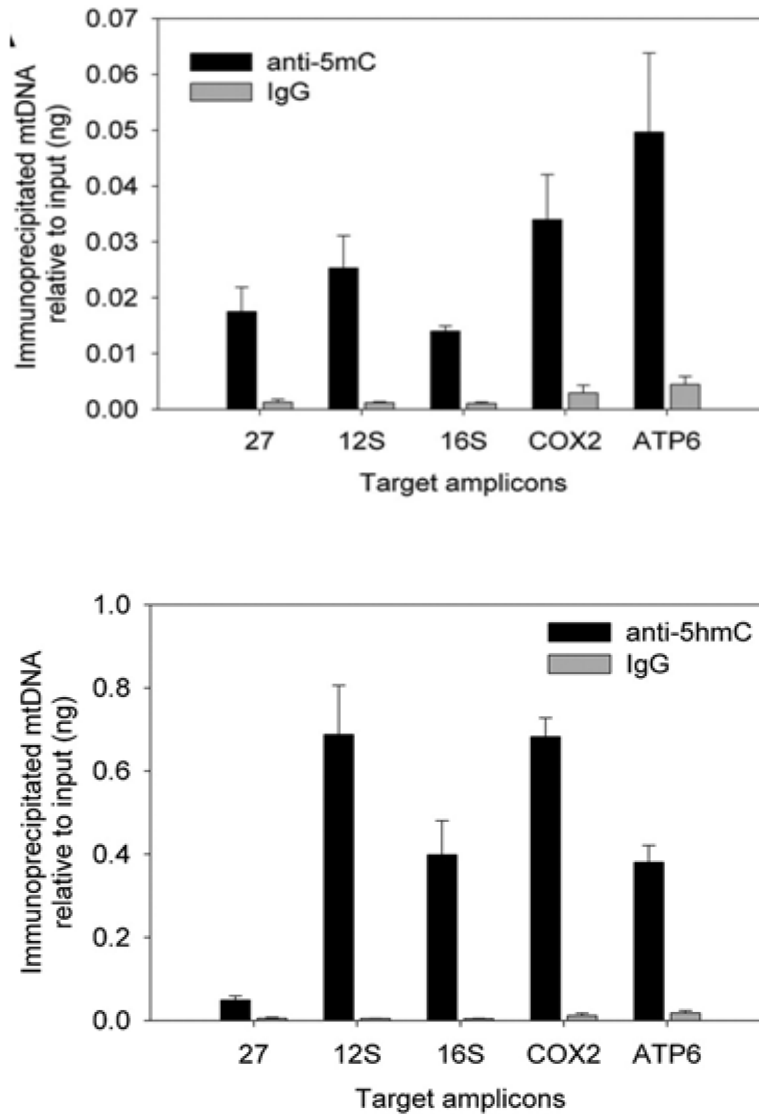
Figure 2-5: Detection of 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) in imunoprecipitates of mitochondrial DNA.

Mitochondrial DNA from HCT116 neoS cells were subject to MeDIP/HydroxyMeDIP assay using 5mC and 5hmC antibody. Significant pulldown could be seen in five different regions of mitochondrial DNA analyzed, with 5mC antibody (upper panel) and 5hmC antibody (lower panel) as compared to that of IgG control. Data represents mean ± SD of two independent biological repeats.

heavy strand of mtDNA (Figure 2-4). Figure 2-5 (lower panel) indicates that all of these regions showed significant enrichment of the mitochondrial DNA with 5hmC Ab as compared to IgG, ranging from 38-fold enrichment in the D-loop region to 580 fold in the 12S rDNA region. Thus, 5hmC modification is present in the mitochondrial genome and its detection is not the result of contaminating nuclear DNA. This indicated that both 5mC and 5hmC modifications were present in the mtDNA. This assay, however, does not allow quantitation of the abundance of 5mC and 5hmC in mitochondrial DNA (See discussion).

DNMT1 regulates transcription of mitochondrial genome in a gene specific fashion:

DNA methylation plays an important role in the regulation of transcription in the nucleus. Methylation of a promoter combined with action of HDACs and nucleosome remodeling factors causes silencing of the gene and results in the reduction in expression of its mRNA transcript [118]. Analogous to its role in the nucleus, we wondered if modulating levels of mtDNMT1 altered transcription of the mitochondrial genome. We decided to take advantage of the system available to us that already expressed higher levels of DNMT1. Increased mtDNMT1, as a result of loss of p53, is seen in p53-/- MEFs cells as compared to wild type MEFs. RNA was isolated from both of these cell types and cDNA was synthesized from each using random hexamers. The cDNA was then subject to qPCR using primer sets designed to analyze expression of different mitochondrial genes. Expression of three different protein coding genes on the heavy strand, i.e. ND1, Cox1 and ATP6, and the only protein coding gene on light strand, ND6, were chosen for analysis. No change in expression was seen for the Cox1 and ATP6, while ND1 showed a significant increase in expression in p53-/- cells as compared to wild type cells (Figure 2-6). In
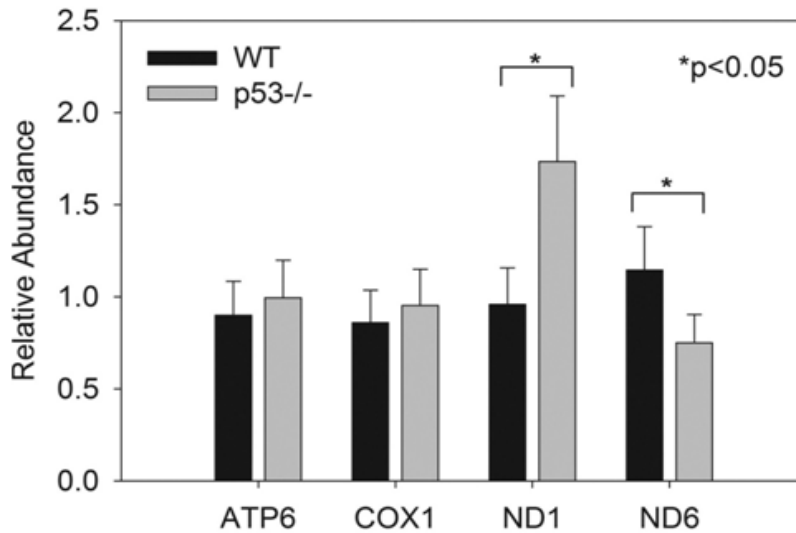
Figure 2-6: Effect of increased expression of mtDNMT1 on transcription of mitochondrial genes exhibits gene specific changes.

RNA isolated from wild type and p53-/- MEFs was converted to cDNA using random hexamers. The cDNA was analyzed for an effect on transcription by using primer sets against four different mitochondrial genes. An increase in ND1 gene expression in p53 -/- cells was siginifically higher (mean = 1.73, SE = 0.18, 95% CI = 1.37, 2.09) in wt MEFs (mean = 0.95, SE = 0.09, 95% CI = 0.76, 1.15). ND6 expression in p53 -/- cells was siginifically lower (mean = 0.75, SE = 0.07, 95% CI = 0.59, 0.90) than in wt MEFs (mean = 1.14, SE = 0.11, 95% CI = 0.91, 1.38). The values represent mean ± SE of three independent biological repeats.
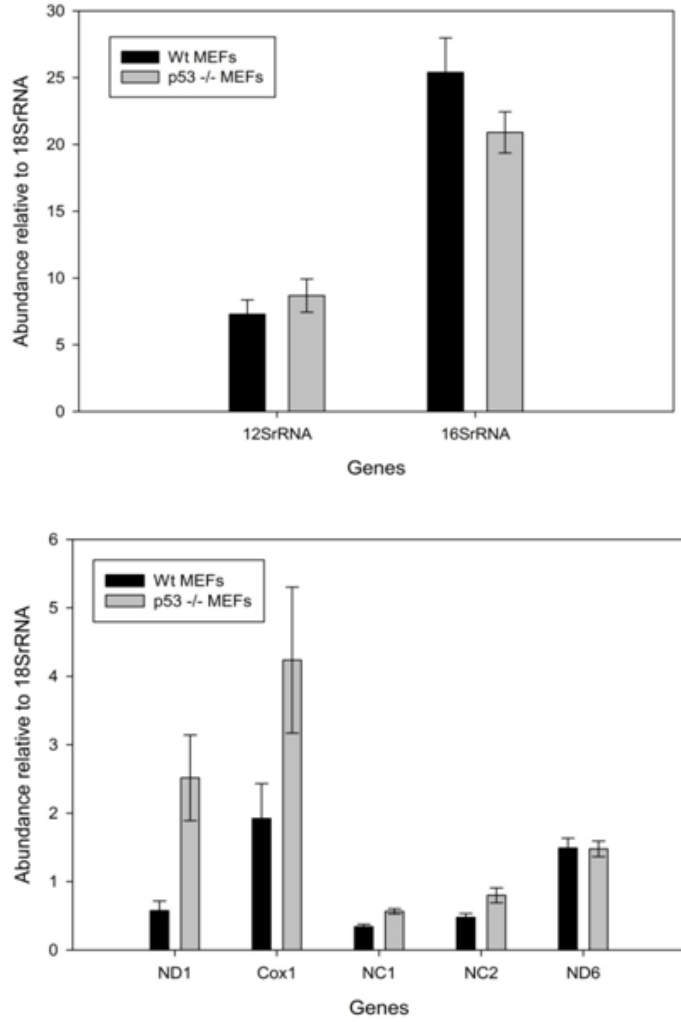
Figure 2-7: Careful analysis of effect of increased expression of mtDNMT1 on mitochondrial transcription reveals a strand specific effect.

RNA was isolated from wild type and p53-/- MEFs and cDNA was synthesized using gene specific primers instead of random hexamers. On the heavy strand, the two protein coding genes ND1 and Cox1 showed an increase and effect was more dramatic in comparison to using random hexamers (Figure 2-8). A previous decrease seen on ND6 was lost in this analysis. These studies were performed in collaboration with Lisa Shock. The data represents mean ± SD and is representative of two independent biological repeats.

contrast, ND6 showed a modest decrease in p53-/- cells in comparison to wild type (Figure 2-6). Student's t-test was applied to establish significance and the difference was found to be significant for ND1 and ND6 ($p < 0.05$). This result, an increase in ND1 with a simultaneous decrease in ND6 and no change in ATP6 and Cox1 expression levels, implied that the effect of increased mtDNMT1 on mitochondrial transcription occurs in a gene specific fashion.

DNMT1 regulates mitochondrial transcription in a strand specific fashion:

Transcription of the mitochondrial genome occurs individually on the heavy and the light strand through their respective promoters. Once transcription begins, a polycystronic transcript is generated that mimics the composition of its coding strand along its entire length on the mitochondrial genome. The nature of mitochondrial transcription revealed a technical flaw in our previous study on effect of mtDNMT1 overexpression on mitochondrial transcription. In our previous study, we synthesized cDNA from RNA using random hexamers. Thus, both strands were equally represented in the cDNA. Subsequent qPCR thus yielded an average of transcription from the two strands. Hence, an opposite effect on the two strands or an effect on one strand with none on the other would either neutralize or minimize any differences in our study and small differences would be entirely eliminated. To overcome this issue, we modified our strategy to use gene specific primers to synthesize cDNA from RNA of the two cell lines. Sense and antisense primers against the ATP6 gene were used to synthesize cDNA from the light and heavy strand of the mitochondrial genomes respectively. cDNA was analyzed by qPCR using the same primer sets as before. As indicated in Figure 2-7, ND1 shows an increase in expression in p53 -/- cells as compared to wild type, that is much more dramatic than seen in the previous analyses (Figure 2-6). Cox1, that failed to show any difference in earlier studies, also showed an increase in p53-/- cells as compared to wild type (Figure 2-7). It should be noted that

both ND1 and Cox1 are present downstream of the mitochondrial transcription terminator site, which is responsible for the production of high levels of rRNA synthesis required to support mitochondrial translation. No change was seen in ND6, a gene that showed reduction in expression in p53-/- cells in the previous studies (Figure 2-7). No change could be seen in either the 12S rRNA or the 16S rRNA gene. These results suggest a strand-specific effect of over expression of mtDNMT1 on mitochondrial transcription, perhaps due to altered promoter firing. Increased mtDNMT1 expression causes an increase in expression of protein coding genes on the heavy strand and does not appear affect the expression of genes on the light strand.

Discussion:

Previous studies might have under-estimated the methylation levels of 5mC in the mitochondrial DNA:

5-hydroxymethylcytosine (5hmC) is a new modification that has been identified in the nuclear genome. This modification is generated from 5mC by the action of Tet enzymes in the presence of α-ketoglutarate and Fe^{2+}. The fact that this nucleotide represents a new form of modified base and is derived from 5mC, prompted us to ask whether this modifications is present in the mitochondrial genome. As shown in Figure 2-5 and [117], we have successfully detected 5hmC in mitochondrial DNA. The presence of 5hmC as well as 5mC in mtDNA suggests that previous studies likely under-estimated the level of modified bases present in mitochondrial DNA. Previous studies using restriction enzyme digestion would not have distinguished 5mC from 5hmC, due of inability of restriction enzymes to differentially cleave sites that are either methylated or hydroxymethylated [111]. Besides, using restriction enzyme digestion only analyses cytosine modification at the enzyme recognition sites, and thus would only capture a small

percentage of modifications present in mtDNA. Nearest neighbor analysis, which involves hydrolyzing DNA to a single nucleotide level and specifically analyzing spot corresponding to 5mC on a TLC plate would not have identified presence of 5hmC in the mitochondrial DNA fraction because no standard was available for comparison at that time [110, 111].

Although both 5mC and 5hmC antibodies are highly specific, they likely display different affinities towards their antigens. As a result, the extent of pull down of mtDNA with each of the antibodies cannot be used at as a measure of relative abundances of the two modifications in the mitochondrial genome. In order to determine the abundance of each of the modifications in the mitochondrial DNA assays we plan to use LC-MS to quantitate the abundance of modified cytosine in mtDNA following hydrolysis to the single nucleoside level.

At the time of this experiment, only 5hmC had been reported as a newly identified base in the genome. Later, two additional cytosine modifications, 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) were identified as being present in the nuclear genome. These modifications are formed by sequential conversion of 5hmC to 5fC and then to 5caC in a reaction catalyzed by Tet enzymes [32]. It is yet to be determined if these are present in the mitochondrial DNA as well, but is likely since 5hmC was found in the mitochondrial fraction. Antibodies against 5fC and 5caC modifications have recently become commercially available (Active Motif). We plan to validate the specificity of these antibodies and then use them in a MeDIP type of assay to detect these modifications in the mitochondrial DNA. We hypothesize that 5mC is generated in the mitochondrial DNA by mtDNMT1, which then is further converted into 5hmC by as yet unknown mechanism. If Tet enzymes catalyze this conversion, then it is likely that 5fC and 5caC will also be present in mtDNA.

Role of mtDNMT1 in regulating mitochondrial transcription:

p53 has been shown to have an isoform that translocates to the mitochondria. Hence, it can be argued that the effect seen on mitochondrial transcription in HCT116 p53 -/- cells could be a result of loss of p53 rather than an increase in mtDNMT1 expression. p53 is important in regulation of mitochondrial biogenesis and loss of p53 has been shown to cause a decrease in mitochondrial copy number [119, 120]. Our lab has not been able to validate these results using wild type and p53-/- MEFs and we found a similar mitochondrial DNA copy number in both the cell lines (Lisa Shock, personal communication). Changes in mitochondrial transcription without changes in mitochondrial copy number in these two cell lines led us to believe that the changes are due to over-expression of mtDNMT1 rather than loss of p53. Furthermore, the effect of p53 on transcription would have resulted in an increase or decrease alone in expression of all the protein coding genes studied, subject to its role as a transcriptional repressor or activator in mitochondria. There is a p53 binding site predicted in the D-loop regulatory region near the heavy and light strand promoters. Whether p53 binds at this site and regulates transcription of mitochondrial genes by direct binding remains to be determined. In order to further tease apart the role of p53 vs DNMT1 in regulation of mitochondrial transcription, DNMT1 will be over-expressed in cells that maintain normal p53 levels.

It has been speculated that there is cross-talk between the nuclear and the mitochondrial genomes, although it is not known whether epigenetics plays a role in this process. To account for such a possibility, it was important that we modulate levels of mtDNMT1 without change in nuclear DNMT1 expression. We have cloned full length DNMT1 cDNA carrying the upstream mitochondrial leader peptide. The ATG start codon for nuclear DNMT1 was then mutated to

ATC. Such a clone is expected express a mitochondrial form of DNMT1 without expressing the nuclear DNMT1. Our attempt at stably expressing this clone in different cells has thus far not been successful. We then further mutated the sequences around ATG1 and ATG2 to put them in the context of an ideal Kozak consensus sequence [121]. This is expected to further enhance translation of mtDNMT1 from ATG1 and ATG2 giving us increased expression of the mitochondrial isoform without changing levels of nuclear DNMT1. We plan to use this clone to generate cell lines in which mtDNMT1 is stably over-expressed. Transcription patterns of different mitochondrial genes will be studied in these cells to elucidate the role of DNMT1 in mitochondrial transcription, in the absence of a change in p53 levels.

Model for regulation of mitochondrial transcription by mtDNMT1:

Our analysis on mitochondrial transcription was performed on cDNA synthesized using gene-specific primers (Figure 2-7). These studies represent steady-state levels of only the polycistronic transcripts, and fail to capture the contribution of mature processed mRNA molecules. To tease apart gene specific versus strand specific effects on mitochondrial transcription, we now plan to perform gene-specific and strand-specific transcription analyses. These studies will use gene-specific primers designed in every gene along both heavy and the light strand, to synthesize cDNA. The cDNA can then be analyzed for expression of the same gene as the one used for cDNA synthesis, using nested primers by qPCR. Such an analysis will account for steady state levels of both the mature mRNA molecules as well as the polycistronic transcripts.

Our results from studies on mitochondrial transcription suggest several possible mechanisms for regulation of mitochondrial transcription by mtDNMT1, which are not mutually

exclusive. Our results indicate that genes downstream of the MTERF binding site, ND1 and Cox1, were upregulated upon overexpression of mtDNMT1 due to loss of p53. Such a result could indicate that transcription termination by MTERF was impaired by increased mtDNMT1 expression, causing an increased read-through of the downstream genes by the transcription machinery. We hypothesize that mtDNMT1, either by methylating the MTERF binding sequence or through protein - protein interaction with MTERF, disrupts binding of MTERF at its recognition sequence (Figure 2-9). This would be expected to cause inefficient termination of HSP1 transcription but should also interfere with the formation of a transcription bubble (Figure 2-9). MTERF binds at two sites of the mitochondrial genome, one downstream of 16S rRNA gene and the other adjacent to the HSP1 promoter, forming a transcription bubble (Figure 2-8). This allows efficient recycling of transcription machinery for expression of 16S rRNA and 12S rRNAs genes expected to be most abundantly required for translation [74]. This efficient recycling machinery would be prevented and this might be expected to reduce 12S rRNA and 16S rRNA gene expression, while increasing read-through to downstream protein coding genes.

Mitochondrial transcription is unique in its mechanisms for transcription inititation, regulation and termination, the details of which have been discussed earlier. This is the first demonstration that epigenetics may be involved in regulation of mitochondrial transcription. In the nucleus, gene silencing through DNA methylation entails binding of MBD domain-containing proteins to methylated DNA, followed by recruitment of HDACs and chromatin remodeling factors to bring about compaction of chromatin. Such a mechanism takes advantage of the fact that DNA is wrapped around histones. Mitochondria, on the other hand, lack histones, although the genomes are complexed with a number of proteins including at least one HMG containing protein. Hence, DNA methylation might be expected to operate through mechanisms

different from those operating in the nucleus. One of the possible ways (discussed above) where binding of MTERF could be disrupted, bringing about lack of termination of transcription.

Another possible mechanism would be to modify the way in which TFAM binds mitochondrial DNA. TFAM is known to coat mitochondrial DNA across its entire length with a stoichiometry of one TFAM molecule per 15-20 bp of mitochondrial DNA [57]. This allows mitochondrial DNA to form the nucleoid structure. mtDNMT1 could potentially modify TFAM association with mitochondrial DNA, giving mtDNA a more relaxed conformation for more efficient transcription. Furthermore, TFAM has been shown to repress transcription from HSP2 promoter by binding at the TFAM response element (TRE) present near the HSP2 promoter and therefore limiting the recruitment of POLRMT-TFB2M transcription initiation complex at the HSP2 promoter [76, 122]. Methylation dependent disruption of TFAM binding at this site would thus be expected to cause increased recruitment of TFB2M-POLRMT complex at HSP2 promoter, resulting in higher levels of heavy strand protein coding genes, as seen in our studies (Figure 2-7).

The increased expression of heavy strand transcripts as seen in Figure 2-7, could also be the result of increased recruitment of the core transcription machinery, POLRMT and TFB2M, at the HSP2 promoter. The TFB2M homologue in yeast is involved in promoter recognition and transcription initiation along with RpoF1, the yeast homologue of POLRMT [123]. Although TFB2M in mammals is not involved in promoter recognition, it helps stabilize POLRMT and activates it to promote transcription [80]. It remains to be determined if the TFB2M-POLRMT heterodimer has higher affinity for methylated DNA as compared to unmethylated DNA. If so, increased expression of mtDNMT1 upon loss of p53 could result in promoter specific

methylation of HSP2 which in turn can cause increased recruitment of the initiation complex at the HSP2 promoter.

DNMT1 is the only known catalytically active DNA methyltransferase that has an isoform present in the mitochondria. Traditionally DNMT3a and DNMT3b act as *de novo* methyltransferases which establish methylation patterns that are maintained by DNMT1. Since DNMT3a and DNMT3b were not found in the mitochondrial fraction, it can be speculated that DNMT1 may be responsible for both generating and maintaining the methylation patterns in the mitochondria. Thus, by methylating specific sites in the D-loop control region mtDNMT1 can exert control of mitochondrial transcription.

Catalytic activity of mtDNMT1:

In studies performed by Lisa Shock, a cell line was created in which a Tandem Affinity Purification (TAP) tag was incorporated in one endogenous allele of DNMT1 at the C-terminal end. The site of incorporation was just before the stop codon allowing expression of full length DNMT1 containing a fusion TAP tag at its C-terminal. Use of the endogenous promoter allowed expression of protein at levels normally found in the cell. This cell line was created to allow efficient purification of DNMT1 protein under mild conditions which would essentially retain its catalytic activity. Using this particular cell line, Lisa performed a three step purification of DNMT1 exclusively from the mitochondrial compartment and found that mtDNMT1 associates with mtDNA. When subjected to an *in vitro* enzyme assay, she showed that mtDNMT1 is catalytically active even when purified from crude mitochondrial preparation treated with trypsin prior to lysis or pure mitochondrial fraction isolated using percoll gradient (Lisa Shock

Dissertation). These studies will further attest the role of mtDNMT1 in methylation of mitochondrial DNA and possible regulation of downstream processes.
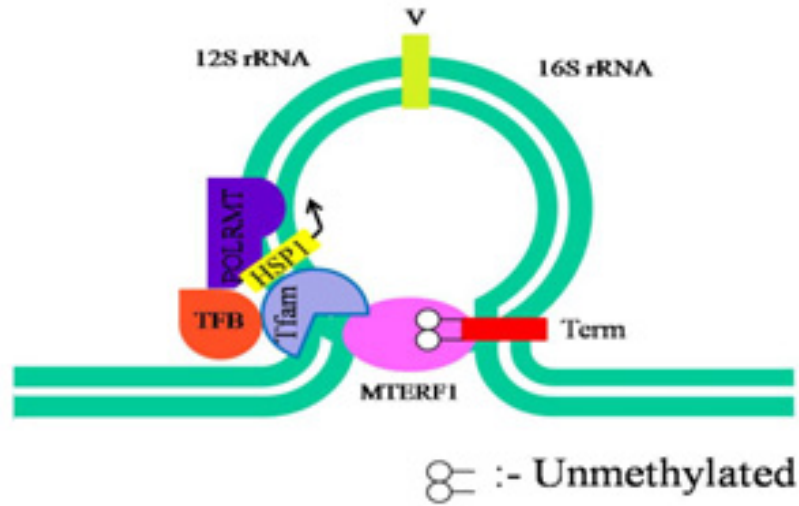
Figure 2-8: Proposed model for MTERF-mediated transcription termination:

Transcription from HSP1 is facilitated by formation of a transcription bubble as depicted in the diagram. This bubble is formed by MTERF binding at its two binding sites. This allows efficient recycling of the machinery to produce high levels of 12S rRNA and 16S rRNA transcripts [74].
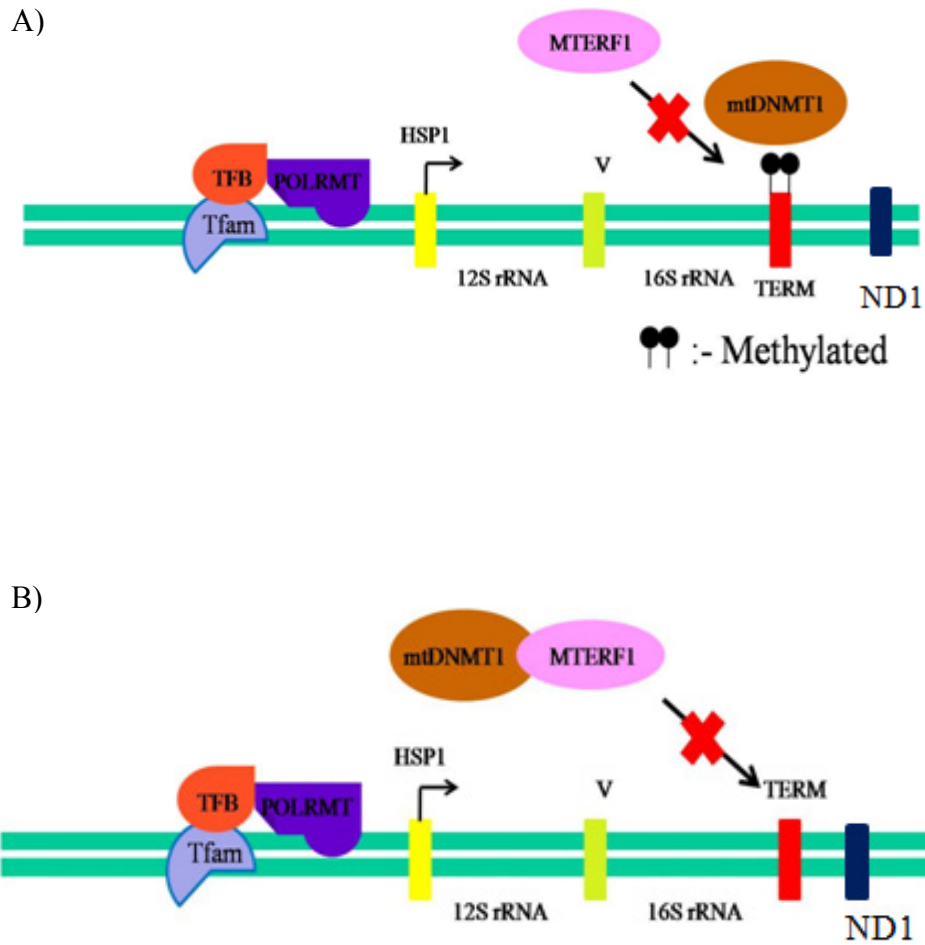
Figure 2-9: Model for regulation of mitochondrial transcription termination by mtDNMT1:

We speculate that mtDNMT1 may disrupt the formation of a transcription bubble (Figure2-8). It may either methylate MTERF binding site (A) or via protein – protein interaction with MTERF (B) and therefore prevent MTERF binding. This may allow increased read through of the downstream protein coding genes, ND1 and Cox1.

Chapter 3: Characterization of role of Tet enzymes in generation of 5hmC in mtDNA

Introduction

Recently a novel modification of cytosine, 5-hydroxymethylcytosine (5hmC), was identified *in vivo*, formation of which is catalyzed by Ten – Eleven translocation (Tet) family of proteins [24]. There are three mammalian paralogs within the Tet family, namely Tet1, Tet2 and Tet3. All three of these enzymes, expressed as either a full length enzyme or merely the C-terminal catalytic domain, can convert 5mC to 5hmC in a Fe^{2+} and α-ketoglutarate dependent manner *in vivo* as well as *in vitro* [24, 30, 32].

The mitochondrial proteome contains approximately 1500 proteins, only 13 of which are encoded by mtDNA. All other proteins are encoded by the nuclear genome and translocated to the mitochondrial using transport systems involving mitochondrial targeting sequences (MTS) [88] . Proteins that are destined for transport to the mitochondria carry sequences that serve as a code for their localization to the mitochondria. Proteins transported to the matrix usually contain MTSs that are 10 to 60 amino acids in length and form an amphipathic helix with one face containing hydrophobic residues and the other containing positively charged residues [88]. The MTSs are usually N-terminal in position but proteins containing C-terminal, e.g. DNA helicase Hmi1 in yeast, or internal MTSs, e.g. BCS1, have also been described [90, 91]. These MTSs are recognized by proteins that are part of TOM and TIM translocases, which then facilitate transport of proteins to the mitochondrial matrix.

Objectives

5hmC is among the newly identified cytosine modifications present in nuclear DNA. Similar to nuclear DNA, mtDNA contains 5mC and a nuclear encoded DNA methylase is responsible for its establishment and maintenance in the mitochondrial genome. Having validated the presence of 5mC, we explored the possibility that 5hmC was present in mtDNA. Here we describe our efforts to define the mechanisms involved in generation of 5hmC in mtDNA. We sought to explore the possibility that there is Tet mediated conversion of 5mC to 5hmC in mtDNA. Our overall goal is to identify mitochondrial epigenetic mechanisms, other than methylation, that either amplify or nullify effects of 5mC in mtDNA. These newly identified mechanisms will add another level of fine tuning to epigenetic regulation of mitochondrial processes and therefore mitochondrial function.

Materials and Methods

Cell Culture

HCT116 neoS cells were grown in RPMI 1640 medium (Gibco/Invitrogen) at 37 $^{\circ}$C in 5% CO_2. Wild type, DNMT1 +/- and DNMT1 -/- mouse ES cells were a kind gift of Masaki Okano and colleagues (Riken Center for Developmental Biology, Japan). Tet1+/- and Tet1 -/- ES cells were obtained from Rudolph Jaenisch and colleagues (Whitehead Institute of Technology, MA) [124]. Mouse ES cells were maintained at 37 $^{\circ}$C in 10% CO_2 on Feeder layers in a DMEM medium (Gibco/Invitrogen) containing 10% FBS (Gemini Biologicals), 2mM (1%) glutamine (Gibco/Invitrogen), 500 units/ml penicillin (1%) (Gibco/Invitrogen), 500 µg/ml streptomycin (1%) (Gibco/Invitrogen), 1mM (1%) sodium pyruvate (Gibco/Invitrogen), 1% non-essential

amino acids (Gibco/Invitrogen) and 0.2% β – mercaptoethanol. Leukemia inhibitory factor (LIF) was added to Mouse embryonic stem cells at a final concentration of 1 μg/ml, to maintain pluripotency. Feeder layers were made fresh every two weeks, by culturing wt MEFs in DMEM medium (Gibco/Invitrogen) containing 10% FBS at 37 $^{\circ}$C in 10% CO_2, irradiated to 40 Gray and plated on collagen coated dishes at least 24 hours prior to use as feeders. To avoid interference from MEF feeder layers, mouse ES cells were cultured on feeder free collagen coated dishes for at least two passages before they were used for experiments. Owing to cell cycle regulated expression of DNMT1 and possibly TETs, cells were fed 24 hrs prior to replating them for experimental set up. Cells were harvested 24 hours after replating thus ensuring the majority of cells were in S-phase and therefore maximally expressing DNMT-1. Maintaining this procedure for all the experiments ensured consistency and reproducibility of results.

Subcellular fractionation to isolate mitochondria from cells:

Mitochondria were isolated as described in Chapter 2. Depending on the downstream application, the mitochondrial pellet was treated differently as mentioned below:-

a) Immunoblots:- After two washes with mitochondrial homogenization buffer containing PI tablets, the mitochondrial pellet was collected in a pre-weighed 1.5ml centrifuge tube. The tube was weighed to determine the mitochondrial pellet weight. The pellet was suspended in SDS lysis buffer with a volume five times the weight of the pellet. The lysate was passed through a 21 gauge needle, 15 times, to sheer DNA and make the lysate less viscous. For small volumes, lysate was sonicated using the Bioruptor bath sonicator on half water- half ice mixture for 7 mins on High, with a 30 sec ON and 30 sec OFF cycle.

b) Trypsin treatment of mitochondrial pellet prior to lysis, for enzyme assay and immunoblot:- After two washes with mitochondrial homogenization buffer containing PI tablets, the mitochondrial pellet was washed once with mitochondrial homogenization buffer without PI and divided into two equal fractions. Mitochondrial pellet weights were obtained and the pellets were resuspended in Trypsin Digestion buffer (10 mM Hepes–KOH (pH 7.4), 250 mM sucrose, 0.5 mM EGTA, 2 mM EDTA, and 1 mM DTT), in a volume 20 times the weight of the pellet. Trypsin was added only to one fraction at a final concentration of 10 µg/ml, while the other fraction was treated as a control. Samples were incubated at room temperature for 20 minutes with occasional inversion to mix. At the end of the incubation, bovine trypsin inhibitor was added to both the control and the trypsin treated fraction to a final concentration of 10 µg/ml. Mitochondrial pellets were washed twice with trypsin digestion buffer containing protease inhibitors. For immunoblots, lysates were prepared as mentioned above by resuspending the mitochondrial pellet in SDS lysis buffer. For enzyme assays, mitochondrial pellet weights were recorded and pellets were resuspended in 25 µl Tet1 lysis buffer (20mM Tris, pH 8.0, 300mM NaCl, 0.1% Tween20, 5% Glycerol and Water. 1 tablet of EDTA free protease inhibitor (PI) tablets (Roche) was added for every 25 ml of lysis buffer. 2mM DTT was added freshly before use and stored at -20 °C as 50% glycerol stocks by adding 25 ul glycerol.

c) Percoll purification of mitochondria for enzyme assay:- The mitochondrial pellet obtained after trypsin treatment was washed twice with trypsin digestion buffer containing PI. For percoll purification, pellets were resuspended in 1 ml mitochondrial homogenization buffer containing PI. The resuspnded pellets were laid on top of a 10 ml percoll gradient, constructed by mixing 6.55 ml Percoll, 2.2 ml of 2.5 M sucrose, 12.25 ml of TE, pH 7.4 and using 10 ml of the mixture per gradient. After overlaying the mitochondrial pellet on the gradient, samples were

centrifuged at 60,000 x g for 1 hr at 4 °C in a Beckman Ultracentrifuge using a fixed angle rotor. The mitochondria migrate to the center of the gradient and appear as a white band. Mitochondria were collected using a 21 gauge needle to pierce the tubes and using a syringe to collect the white band. The collected mitochondria were then transferred to 14ml round bottom falcon tubes and washed thrice using 3 ml mitochondrial homogenization buffer containing protease inhibitor tables by centrguation at 10,000 x g for 15 mins at 4 °C in J2 centrifuge. After the final wash, the mitochondrial pellets were lysed in 25 µl of Tet1 lysis buffer (20mM Tris, pH 8.0, 300mM NaCl, 0.1% Tween20, 5% Glycerol, 2mM DTT and protease inhibitor tables) and stored as 50% glycerol stocks.

d) Isolation of mitochondrial DNA for MeDIP/HydroxyMeDIP:- The mitochondrial fraction was collected and washed twice with 1X PBS, pH7.4 with centrifugation at 10,000 x g in a table top centrifuge. At the end of the second wash, the mitochondrial pellet weights were noted and mitochondria were lysed in 200 µl mtIP lysis buffer (0.5% Triton X-100, 300mM NaCl, 50mM Tris-HCl, pH 7.4). Mitochondrial lysates were subjected to phenol - chloroform extraction to isolate mtDNA, which was then treated with RNAseA (1 µl of 10 mg/ml stock) at 37 °C for one hour followed by treatment with proteinase K (1 µl of 20 mg/ml stock and 0.25% SDS (final concentration)) at 37 °C overnight to eliminate contaminating RNA and protein respectively. The DNA was further purified by another two rounds of phenol – chloroform extraction and one round of ethanol precipitation. Finally, the DNA was resuspended in TE buffer, pH 8.0, and sonicated using Bioruptor bath sonicator on half water- half ice mixture in cycles of 15 mins on High, with a 30 sec ON and 30 sec OFF cycle until size range of 200-700 bp was obtained. The DNA was then used for MeDIP/HydroxyMeDIP assay.

e) Isolation of mitochondrial DNA for analysis by HPLC:- Post-Nuclear supernatant, collected at the end of three rounds of dounce homogenization, was centrifuged at (2500 x g) 3000 rpm for 10 minutes at 4 $^{\circ}$C in J6-MI centrifuge to pellet the remaining unbroken cells. The supernatant was collected, leaving behind 2 ml to avoid contamination from unbroken cells, and transferred to a 14 ml round bottom centrifuge tube (BD Falcon, Cat no:- 352059) and centrifuged at (10,000 x g) 9000 rpm for 15 minutes at 4 $^{\circ}$C in Beckman J2-HC centrifuge. The mitochondrial fraction was then collected and washed twice with 1X PBS, pH7.4 with centrifugation at 10,000 x g in a table top centrifuge. At the end of the second wash, mitochondrial pellet weights were noted and pellets were resuspended in 400 μl 1X PBS, pH7.4, 44 ul 10X DNAseI buffer (1X final conc.) and 8 ul DNAseI (16 units) to remove contaminating nuclear DNA. The reaction was incubated at room temperature for 30 mins and terminated by adding EDTA to a final concentration of 5mM. Thereafter, pellets were washed twice with IX PBS, pH 7.4 containing 25mM EDTA with centrifugation at 10,000 x g for 15 mins. At the end of the last wash, mitochondrial pellet weights were noted and mitochondria were lysed in 200 μl mtIP lysis buffer. Mitochondrial lysates were subjected to phenol - chloroform extraction to isolate mtDNA, which was then RNAseA (1 μl of 10 mg/ml stock) and proteinase K (1 μl of 20 mg/ml stock and 0.25% SDS (final concentration) treated to eliminate contaminating RNA and protein respectively. Samples were then further treated with 50 units of plasmid safe ATP dependent exonuclease (Epicenter Biotechnologies) in 250 μl total volume for 4 hrs to degrade contaminating genomic DNA. mtDNA was assessed for extent of contamination by nuclear DNA, by using qPCR to amplify intron 3 region of actin gene and normalizing it to 16S rDNA content. Only DNA containing less than 1% contaminating nuclear DNA was used further for hydrolysis and analysis by LC-MS.

β – Glucosyltransferase assay

β – Glucosyltransferase (β- GT) assay was performed using Quest 5-hmC detection kit (Zymo research Cat. No. D5410) as per the manufacturer's instructions. Briefly, 150 ng of human mitochondrial DNA or 80 ng of mouse mitochondrial or genomic DNA was added to a reaction mixture containing 1 X β-GT buffer, 1 X UDPG, 32mM S-adenosyl methionine and 4 units each of CpG and GpC methyltransferases (NEB Cat nos. M0226S and M0227S). Two units of β – glucosyltransferase enzyme was added to each of the samples and the reaction was mixed by pipetting up and down. A corresponding negative β- GT control was set up for each of the samples. The entire reaction was assembled on ice. The reaction was then incubated at 30 ^{O}C for 2 hrs in a thermocycler. At the end of the incubation, 1 unit of GlaI enzyme, diluted in GlaI-enzyme buffer (10 mM Tris-HCl (pH 8,5 at 25°C); 5 mM MgCl$_2$; 10 mM NaCl; 1 mM 2-mercaptoethanol), was added to each of the samples including the negative controls and the reactions were further incubated at 30 ^{O}C for 16 hrs. At the end of the 16 hr incubation, the extent of protection of mitochondrial DNA fragments were studied using end point PCR with primer sets designed agsinst different region of the mitochondrial DNA. DNA standards (Active Motif) were run alongside DNA samples from cells in each assay. The primers used are listed in Table 3-1. The cycling conditions for the qPCR are as follows:- Step 1: Activation at 95 ^{O}C for 15 mins, Step 2: Denaturation at 94 ^{O}C for 30 secs, Step 3: Annealing at 55-60 ^{O}C for 30 secs, Step 4: Extension at 72 ^{O}C for 30 secs, Step 5: Plate read, Step 6: Go to Step 2, 40 times, Step 7: Final extenstion at 72 ^{O}C for 5 mins, Step 8: Melting curve, 40 ^{O}C to 90 ^{O}C, read every 1 ^{O}C, hold 1 sec.

Tet Enzyme Assay

Trypsin treated crude or percoll purified mitochondrial lysate (5 µl), containing 50% glycerol, was incubated with 100 ng of methylated DNA containing the human APC gene promoter sequence (Active Motif, Cat. No. 55008) in the presence of TET enzyme assay buffer (100mM HEPES, pH 6.8, 150 uM Fe(NH4)$_2$(so4), 4mM Ascorbate, 2mM alpha-ketoglutarate and water) at 37 ^{O}C for 1 hr in a PCR thermo cycler. The final volume of the reaction was adjusted to 25 µl using HPLC water. No more than 5 ul of the lysate was used for the assay to limit glycerol content in the reaction to less than 10%. At the end of the 1 hr incubation, the reaction was terminated by quenching on ice for 5 mins. Boiling of samples to terminate the reaction or to kill the enzyme was not possible since non-specific heat-dependent oxidation of 5mC to 5hmC was seen. The reactions were then stored at -80 ^{O}C until further use for end point analysis using immunodetection on slot blots. Purified recombinant C-terminal domain of Tet1 protein (1 µg), shown to have enzyme activity (Active Motif, Cat no. 31363) was used as a positive control. For the negative control, 5 ul of the respective mitochondrial lysates were boiled for 10 mins before use. All reactions were performed alongside a corresponding no substrate control to account for signal intrinsically obtained from mitochondrial lysates due to presence of mtDNA.

Immunodetection of 5hmC using a slot blot

Slot blotting was used as the method for end point analysis of enzyme assay. DNA from enzyme assay reaction sample (20ng/5ul) was mixed with 30 ul of 20 X SSC (175.3g of NaCl, 88.2g of sodium citrate in 1L water, adjusted pH 7.0 with HCl). The final volume was adjusted to 100 µl with water, giving a final concentration of 6X SCC. The samples were boiled for

	16S-2	5'-AAAAGAGGGACAGCTCTTCTGGAACG-3'	5'-TCGTTTAGCCGTTCATGCTAGTCC-3'
Mouse	16S-3	5'-CGGTTTCTATCTATTTACGATTTCTCCC-3'	5'-GCCACCCTAATAACCTTCTCTAGG-3'
	ATP6/ COX3	5'-CCCACCAACAGCTACCATTAC-3'	5'-CTAGACCTGATGTTAGAAGGAGGG-3'
	ATP6	5'-ATTCAACCAATAGCCCTGGCCG-3'	5'-ACGTAGGCTTGGATTAAGGCGAC-3'
Human	12S rRNA	5'-AGTTCACCCTCTAAATCACCACG-3'	5'-TGACTTGGGTTAATCGTGTGACC-3'
	2	5'-GAGACAAGTCGTAACATGGTAAG-3'	5'-GGGTAAGGTTTGCCGAGTTCCT-3'

Table 3-1: Sequences of primers used for detection of 5hmC in mitochondrial DNA fragments from human and mouse.

These primers were used for end point analysis of β − Glucosyltransferase assay performed on human and mouse mitochondrial DNA samples.

10 mins and immediately quenched on ice for 10 mins. An equal volume (100 µl) of 20 X SSC solution was added to each of the samples. A nitrocellulose membrane, 3.5 cm x 12 cm in size, was first rinsed in water for 10 mins followed by incubation in 10 X SSC solution for 10 mins. A 3mm whatman paper was cut to the same dimensions as the membrane, soaked in 10X SSC and used as a support underneath the membrane to keep it from tearing under vacuum. The entire apparatus was assembled as per the manufacturer's instructions. All the bubbles were rolled out to avoid improper loading of samples. All the wells were rinsed with 500 µl of 20 X SSC and suction was applied slowly to drain all the SSC buffer. Samples were loaded on the membrane, with the vacuum closed, and again suction was applied slowly to immobilize the sample on the membrane. Precautions were taken to ensure that the membrane did not run dry. Once all the samples were spotted on the membrane, the apparatus was disassembled and membrane was rinsed in 2X SSC for 5 mins with slow shaking and then UV cross linked. The membrane was blocked using 5% milk for 1 hr. Primary Antibody (1:500 of anti-5hmC) (Active Motif, Cat. No. 39791), resuspended in 5% milk, and 1:15,000 of Goat anti-Rabbit secondary antibody was used. Chemiluminescent substrate (Thermo scientific) was used to develop the blot.

MeDIP/HydroxyMeDIP analysis of DNA from mouse embryonic stem cells:

MeDIP/HydroxyMeDIP assays were performed as described in Chapter 2. Total DNA was isolated from different mouse embryonic stem cells including wild type, DNMT1 +/-, DNMT1 -/-, Tet1 +/-, Tet1 -/- cells using Qiagen DNAeasy kit. DNA was sonicated to 200-700 bp size range and subjected to MeDIP/HydroxyMeDIP assay.

Cloning of Tet1 MTS- GFP fusion constructs:-

Total RNA was isolated from wild type mouse ES cells and cDNA was synthesized using random hexamers, which was then used for PCR amplification of different putative mouse Tet1 MTSs starting at ATG1, ATG3 and ATG4. The PCR products were separated on a 1% TAE/EtBr agarose gel, visualized using UV and the desired bands were excised from the gel using a clean razor blade. The bands were then gel-purified using a gel extraction kit (Qiagen). pLEGFP vector (Addgene) as well as gel-extracted PCR products were subjected to restriction enyme digestion using BglII and BamHI to generate sticky ends for ligation. After the digest, the desired bands were gel-purified as described above and ligation reactions were set up using purified linear fragments of vector and individual PCR products using T4 DNA ligase. The ligation reactions were transformed into TOP10 cells and selected on media containing appropriate antibiotics. DNA from different colonies was screened for presence and orientation of appropriate insert in pLFEGP vector using restriction enzyme digestion. Positive clones were further verified for correct insert using DNA sequencing. A total of three mouse Tet1-GFP fusion constructs were generated, including MTSs starting at ATG1 (F1R3 construct), ATG3 (F3R2 construct) and ATG4 (F4R3 construct) (Figure3-13), which were transfected into NIH3T3 cells for analysis using confocal microscopy.

Confocal Microscopy of GFP fusion proteins:-

Coverslips were prepared for confocal microscopy by incubating them in 1M Nitric acid at room temperature overnight with shaking. The next day, the coverslips were rinsed with water five to six times adding fresh water for each wash and letting them sit for approximately 45 mins to 1 hr with shaking. The coverslips were then sterilized by immersing them in 70% ethanol

before placing them in 35 mm dishes. The coverslips were then allowed to dry in the hood overnight and in presence of UV to keep them sterile. The coverslips could now be used for growth of cells.

NIH3T3 cells were plated in a 60 mm dish such that cells were ~70-80% confluent at the time of transfection. Cells were transfected with F1R3-GFP, F4R3-GFP and F3R2-GFP fusion constructs using Poly Jet transfection reagent (Signagen Cat no. SL100688) as per manufacturer's instructions. 24 hours post transfection cells were trypsinized and ~50,000 – 75,000 cells were plated on a 18 mm x 18 mm poly lysine coated coverslip. 48 hrs post-transfection cells were rinsed twice with 1X PBS, pH 7.4 and then fixed using 3.7% formaldehyde for 15 mins at 37 ^{O}C. Cells were again rinsed twice with 1X PBS, pH 7.4 and then stained with 25nM Mitotracker red dye by incubating the cells at 37 ^{O}C for 45 mins. Cells were rinsed thrice with 1X PBS, pH 7.4 before mounting them on slides using vectashield containing DAPI stain. The cover slips were sealed on the slides using clear nail polish. Samples were then analyzed using confocal microscopy on LSM-510 Meta laser scanning microscope.

Immunofluorescence:-

HeLa cells (50,000-75,000) were plated on a 18 mm x 18 mm poly lysine coated coverslip and grown in DMEM containing 10 % FBS overnight at 37 ^{O}C. The next day, cells were rinsed twice with 1X PBS, pH 7.4 and then fixed using 4% paraformaldehyde for 15 mins at 37 ^{O}C. Cells were rinsed twice with 1X PBS, pH 7.4 and the cell membrane was permeabilized by incubating cells with cold methanol – acetone mixture (1:1) for 15 mins at -20 ^{O}C. The cells were then rinsed again twice with 1X PBS, pH 7.4 and blocked with starting block buffer for 1 hr at room temperature. Following blocking, cells were incubated with Tet1 antibody

at 1:2000 and FITC conjugated Goat anti-Rabbit secondary antibody at 1:500 dilution for 1 hr each. Cells were washed thrice with 1X PBS, pH 7.4 then stained with 25nM Mitotracker red dye by incubating the cells at 37 ^{O}C for 45 mins. Cells were washed three times with 1X PBS, pH 7.4. and fixed again with 4% paraformaldehyde by incubating the coverslips for 15 mins at 37 ^{O}C. The cells were further washed with 1 X PBS, pH 7.4 twice and the coverslips were then mounted using vectashield containing DAPI stain. The coverslips were sealed on the slides using clear nail polish. The samples were analyzed using confocal microscopy on a LSM-510 Meta laser scanning microscope, using a 63x objective.

Results

I. 5-hydroxymethylcytosine is present in mtDNA:

The successful detection of 5hmC in immunoprecipitates of mitochondrial DNA using 5hmC antibody was demonstrated earlier in Chapter 2. In order to confirm the presence of 5hmC in mtDNA, an independent β-glucosyltransferase assay was performed by Dr. Shirley M. Taylor. This assay takes advantage of a viral enzyme, β-glucosyltransferase (β-GT), which adds glucose moiety specifically on 5-hydroxymethylcytosine residue in the DNA using Uridine diphosphoglucose (UDPG) as a substrate (Figure 3-1 A). Following the β-glucosyltransferase reaction, DNA is treated with CpG and GpC methyltransferases to methylate all the unmethylated CpG and GpC sites respectively in the DNA. DNA is then incubated with restriction enzyme GlaI, which cleaves at Pu(5mC)↑GPy (indicated by an arrow) only when cytosine present at the cleavage site is either fully methylated or hydroxymethylated. GlaI cannot cleave at its target site when cytosine is protected by the glucose moiety (Figure 3-1 B). DNA protected from GlaI digest is then analyzed by qPCR. This assay is best used to analyse site-specific hydroxymethylation of cytosine residues in the DNA.

As a proof of principle, DNA standards containing either unmethylated, methylated and hydroxymethylated cytosines (Active Motif) were subjected to the assay. Only the hydroxymethylated DNA reaction containing β-glucosyltransferase enzyme was protected as compared to the 'no β-GT enzyme' control, which showed no protection of hydroxymethylated DNA towards digestion by GlaI (Figure 3-1 C). Template DNA containing either methylated cytosines or that containing unmethylated cytosines, treated with methylases, template was also completely digested by GlaI.
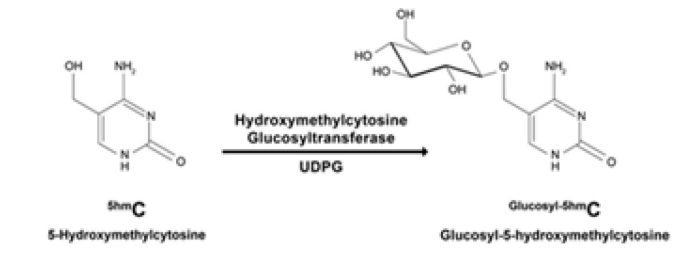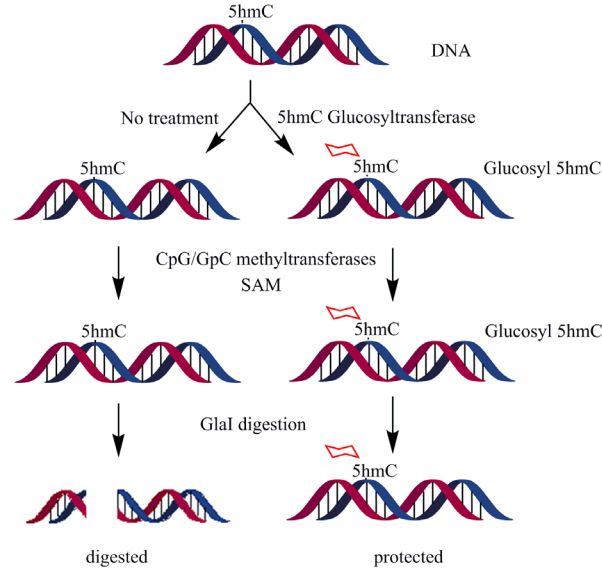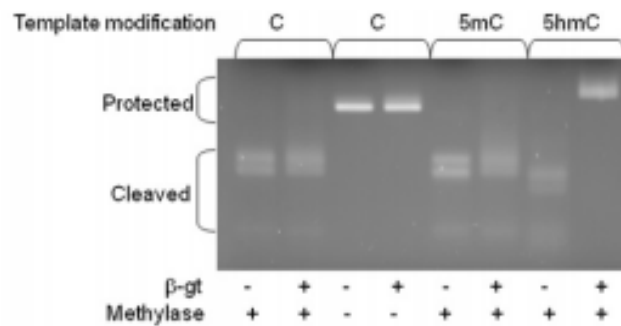
Figure 3-1: Overview of β-glucosyltransferase assay to detect presence of 5hmC.

A) β – glucosyltransferase enzume catalyzes transfer of glucose moeity from UDPG to exclusively to 5hmC. B) Schematic of β-glucosyltransferase assay. Transfer of glucose moiety to 5hmC protects 5hmC containing DNA from cleavage by GlaI as opposed to the sample not containing βGT enzyme. The protected DNA can then be analyzed by PCR. C) β-gluosyltranferase assay performed on unmethylated, methylated and hydroxymethylated substrates provide validate the use of this assay to detect 5hmC in DNA.
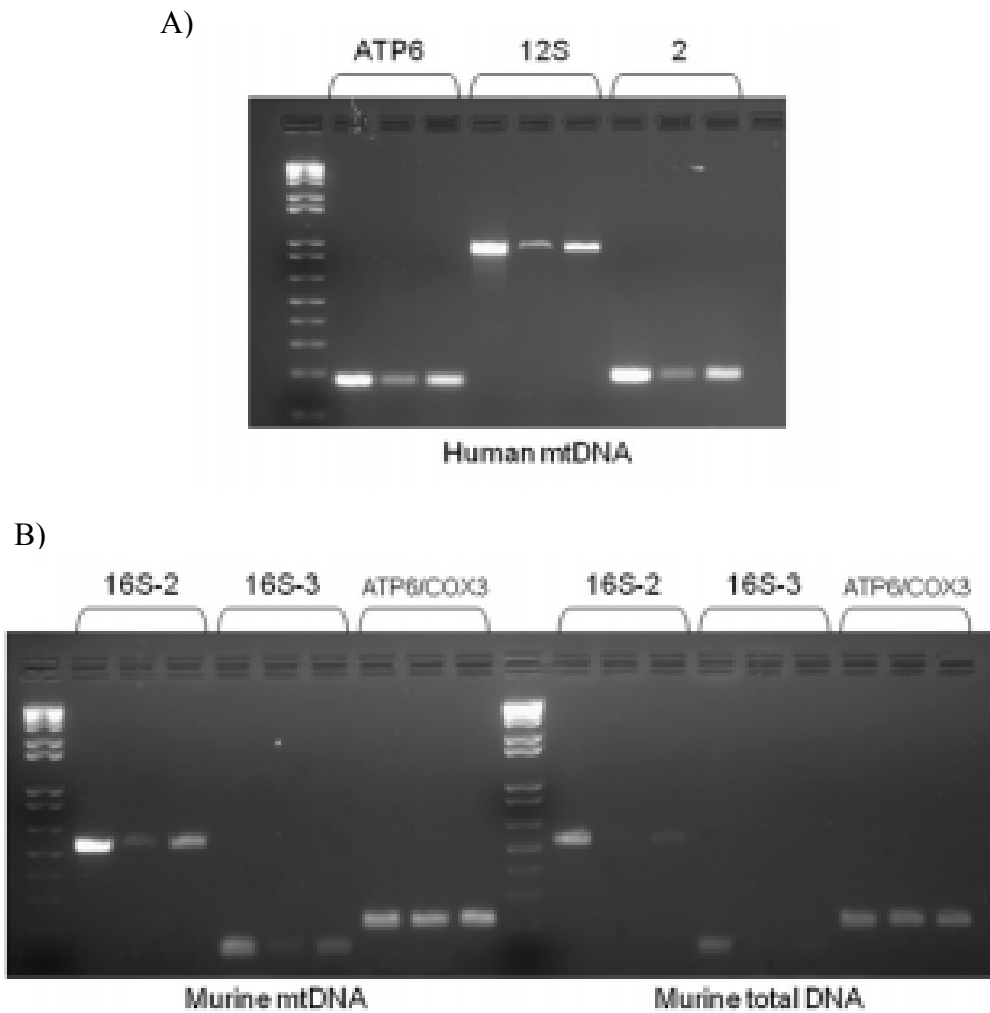
Figure 3-2: Agarose gel displaying results of β- glucosyltransferase assay as performed on Human and Mouse DNA.

Human mitochondrial DNA (A) or Mouse mitochondrial and genomic DNA (B), show protection of 5hmC-containing mitochondrial DNA fragments after protection by β-GT containing samples (Lane 3 for each amplicon) vsersus samples not treated with β-GT enzyme (Lane 2 for each amplicon). Lane 1 for each amplicon represents input DNA. ATP6/COX3 amplicon in mouse does not show any difference in the +β-GT and –βGT lanes. This region was shown to contain a mutation in GlaI cleavage site upon sequencing of the fragment. This experiment was performed by Dr. Shirley M. Taylor.

Protection of unmethylated susbtrate in the absence of CpG and GpC methyltransferases proves that GlaI can only act on substrates containing 5mC/5hmC in its cleavage site (Figure 3-1C).

Using this assay, we could successfully detect 5hmC in human and mouse mitochondrial DNA (Figure 3-2) using different amplicons containing the GlaI cleavage site. As shown in figure 3-2, three human mtDNA amplicons (amplicons ATP6, 12S and 2) and three mouse amplicons (amplicons 16S-2, 16S-3 and ATP6/Cox3) were studied. In each of the amplicons studied, a higher signal could be seen in +β-GT samples as compared to the − β-GT controls (compare lane 2 (-βGT) vs lane 3 (+ βGT) for each amplicon). Moreover, 5hmC could also be detected in mitochondrial DNA fragments using as little as 80 ng of total DNA from cells. This eliminates the need to purify mitochondrial DNA for use in this assay. APT6/COX3 amplicon in mouse failed to show any difference in the +β-GT and −βGT lanes. This fragment serves as an internal control since careful analysis of this fragment using DNA sequencing revealed a point mutation polymorphism in the GlaI cleavage site in wt MEF cells, which explains protection of DNA from GlaI cleavage in the absence of β-GT.

DNA Methyltransferase 1 (DNMT1) is not responsible for generating 5hmC in the mitochondria:

Cytosine methyltransferases, including DNMT1, can reversibly form 5-hydroxymethylcytosine from cytosine in presence of formaldehyde (Figure 3-3) [125]. This reaction has been shown to occur only under *in vitro* conditions and the physiological relevance of such a mechanism still needs to be investigated. However, since DNMT1, the only member of catalytically active methyltransferases, is transported to the mitochondria [117], it was important

Figure 3-3: Schematic of a proposed mechanism for formation of 5hmC from C by DNMTs.
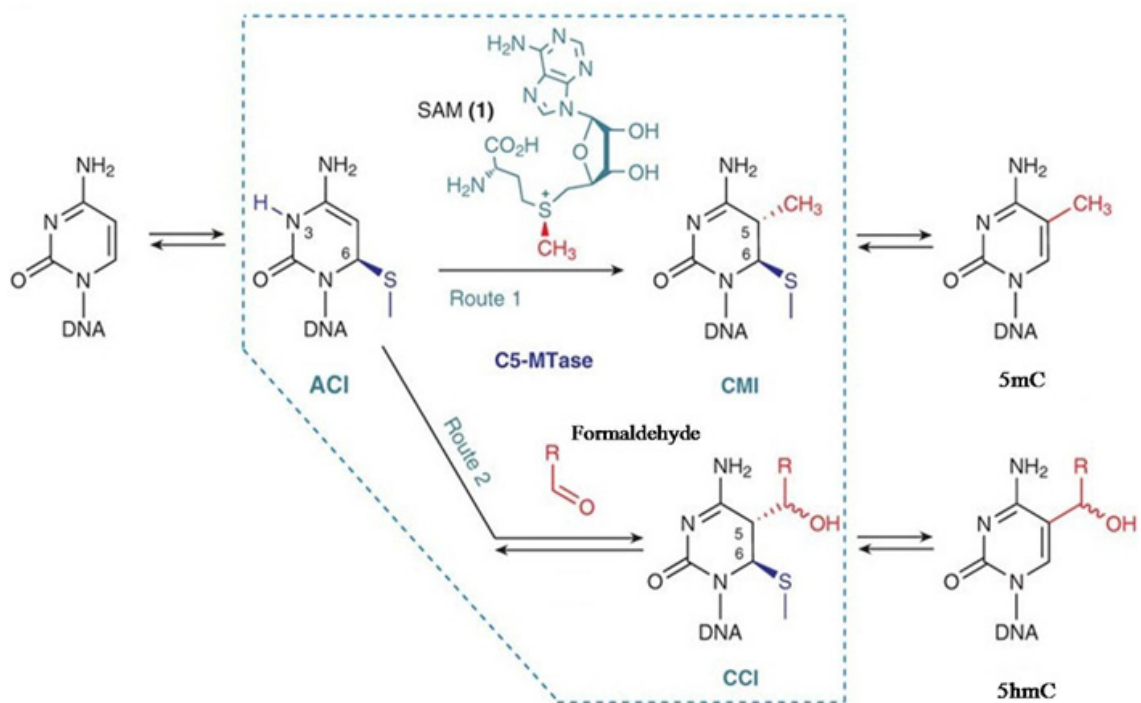
Liutkeviciute et al. proposed that DNMTs can catalyze *in vitro* formation of 5hmC from C in the presence of formaldehyde as a co-factor. Existence of this reaction under physiological conditions is yet to be proven. Figure adapted from Liutkeviciute et al., Nature Chemical Biology, 2009.

A)

B)
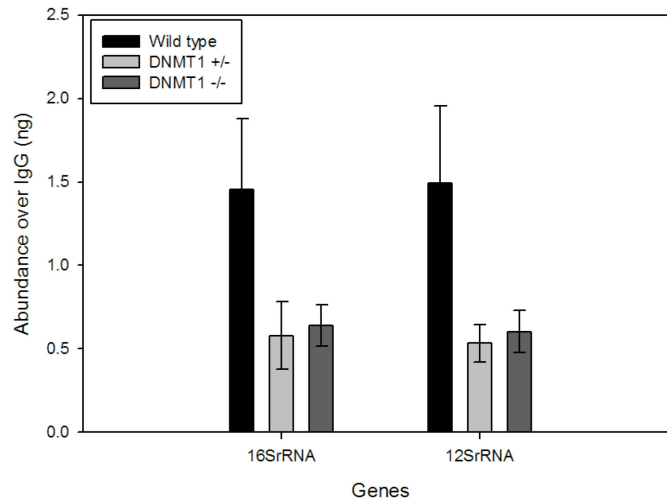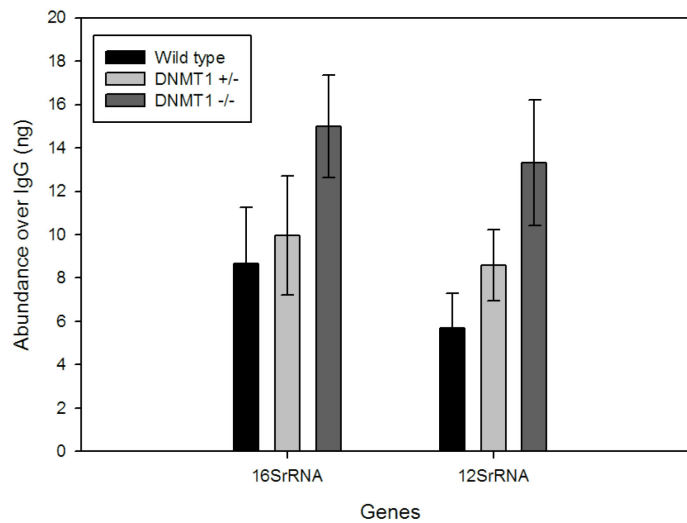
Figure 3-4: MeDIP/hydroxyMeDIP assay on wild type, DNMT1 +/- and DNMT1 -/- mouse embryonic stem cells.

MeDIP (A) and hydroxyMeDIP (B) assay on total DNA from wild type, DNMT1+/- and DNMT1 -/- mouse embryonic stem cells reveal decrease in 5mC and a simultaneous increase in 5hmC in DNMT1-/- cells as compared to wild type cells. The 5mC and 5hmC signal obtained from each of the three cell lines were normalized to their respective IgG controls Data represents mean ± SD of two independent biological repeats.

we consider the possibility of mtDNMT1 generating 5hmC in the mitochondria. In support of this hypothesis, to date, no member of the Tet enzyme family has been found in the mitochondria. To test this hypothesis, genomic DNA was isolated from Wild type, DNMT1 +/- and DNMT1 -/- mouse embryonic stem cells (MES), sonicated to a size range of 200-700 bp and subjected to MeDIP/HydroxyMeDIP assay. The 5mC and 5hmC signal obtained from each of the three cell lines were normalized to their respective IgG controls. Figure 3-4A shows that DNMT1 +/- cells show a reduced signal with 5mC antibody, 50% to that seen in wild type cells. No further reduction in 5mC content could be seen in DNMT1 -/- cells compared to DNMT1+/- cells which could be attributed to the limit of detection of the assay. If DNMT1 was responsible for the generation of 5hmC in the mtDNA, a decrease in 5hmC content would be expected in cells partially or completely lacking DNMT1 compared to that of wild type cells. On the contrary, a modest increase in 5hmC content could be seen in DNMT1 -/- cells as compared to wild type cells (Figure 3-4 B).

Tet activity is present in trypsin treated crude and percoll purified fractions of mitochondria

Before setting out on a quest to determine which member of Tet enzyme family is present in the mitochondria, the detection of Tet enzyme activity in the mitochondrial fraction was essential. Mitochondria were isolated from HCT116 neoS cells, treated with trypsin and were either used as a crude preparation or were further purified using Percoll gradient. Trypsin treatment of mitochondrial fractions, prior to lysis, was done to prevent contamination resulting from nuclear or cytosolic Tet enzymes fraction adhering to the mitochondrial membrane. Lysates containing 50% glycerol were made from both crude and percoll purified mitochondria and used

in an *in vitro* Tet enzyme assay. The assay comprises DNA template containing methylated cytosines as the substrate, Fe^{2+} containing salt, a co-factor essential for activity of Tet enzymes, and crude or percoll purified mitochondrial lysate. This assay was set up with the premise that Tet enzyme, if present in the mitochondrial fraction, would act on the methylated cytosines in the substrate and convert it to a 5-hydroxymethylated form. The newly formed hydroxymethylated cytosines in the DNA could then be detected by spotting DNA on a slot blot and probing it with 5hmC antibody.

To explore the sensitivity of detection of hydroxymethylated DNA using 5hmC antibody on a slot blot, different amounts of hydroxymethylated substrate was spotted on a nitrocellulose membrane and probed with 5hmC Ab. (Figure 3-5) As little as 80 pg of hydroxymethylated substrate could be detected without any cross reactivity from 50 ng of either methylated or unmethylated substrate. This indicated that 5hmC Ab is highly specific, eliminating the possibility of false positive results. It also speaks to the sensitivity of the assay in that, by loading 20 ng of the substrate from the assay and obtaining a signal equivalent to 80 pg of hydroxymethylated substrate (Figure 3-5), only 1 in 200 cytosine residues need to be converted from methylated to the hydroxymethylated form. Thus, the assay will allow us to detect very small amounts of enzyme activity.

Optimal conditions for the assay were examined by incubating 200 ng methylated substrate with 1 μg of Tet1 recombinant protein (Active Motif) at 37 °C under two different pH conditions, pH 6.8 and pH 8, for reaction times of 5, 30, 60 and 180 mins. 20 ng of DNA from the reaction was spotted on a slot blot and probed with 5hmC antibody for end-point analysis. No generation of 5hmC in template DNA could be detected at pH 8.0 until the reaction was
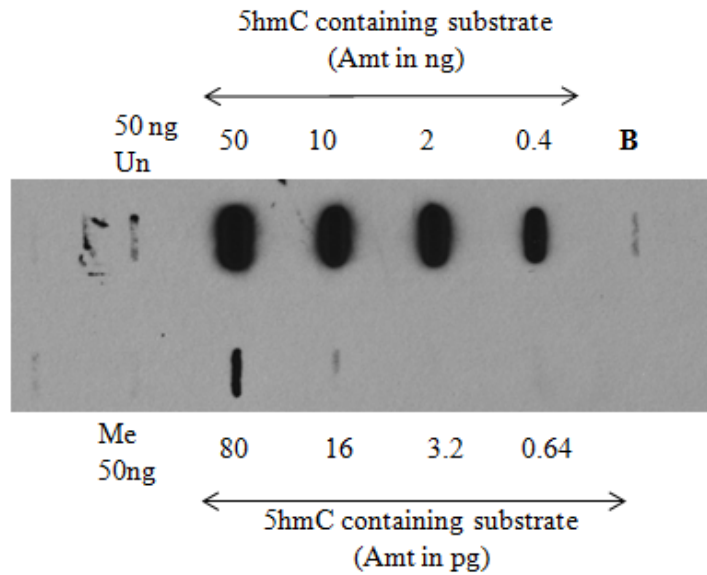
Figure 3-5: Optimization of 5hmC antibody for use in a slot blot assay.

Different amounts of 5hmC containing DNA were spotted on a Nitrocellulose membrane to test the sensitivity and the specificity of 5hmC antibody. As little as 80 pg of 5hmC containing DNA (Active Motif) could be detected without any cross reactivity against 50 ng of methylated (Active Motif) or unmethylated DNA (Active Motif).

Figure 3-6: Optimization of *in vitro* Tet enzyme assay

1 ug of purified recombinant C-terminal domain of Tet1 enzyme was incubated with 200 ng of methylated DNA (Active Motif) at pH6.8 or pH8.0 for the indicated incubation times. Incubation at pH6.8 for 1 hr were considered as optimal conditions for the assay. No signal could be detected using 5hmC antibody in '-ve' control either at pH 6.8 or at pH 8.0. 0.4ng of unmethylated, methylated and hydroxymethylated DNA were spotted as positive controls to asses reactivity of the 5hmC antibody.

Figure 3-7: In vitro enzyme assay performed on trypsin treated crude mitochondrial lysates from HCT116neoS cells.

A) Immunoblot indicating VDAC and H3K4me3 signals on Whole cells (W), mock treated (M) and trypsin treated (T) mitochondrial fractions. The mitochondtial fractions are free of nuclear contamination as indicated by lack of H3K4me3 signal B) 0.4 ng of unmethylated, methylated and hydroxymethylated DNA spotted on the nitrocellulose membrane acts as positive controls for the specificity of the 5hmC antibody. C) 1 µg of recombinant C-terminal domain of Tet1 (Active Motif) incubated with 100 ng of methylated DNA, serve as positive control for the *in vitro* Tet enzyme assay. D) Mock and trypsin treated mitochondrial fractions on incubation with 100 ng of methylated DNA indicate Tet activity when compared to corresponding 'no substrate' controls.

A)

HCT116 neoS

W   M

α – VDAC

α – H$_3$K$_4$me$_3$

B)

Un   Me   Hy

C)

1 hr   No enzyme

D)

w/ 100 ng substrate   No substrate

Mito Mito Boiled Mito Mito Boiled
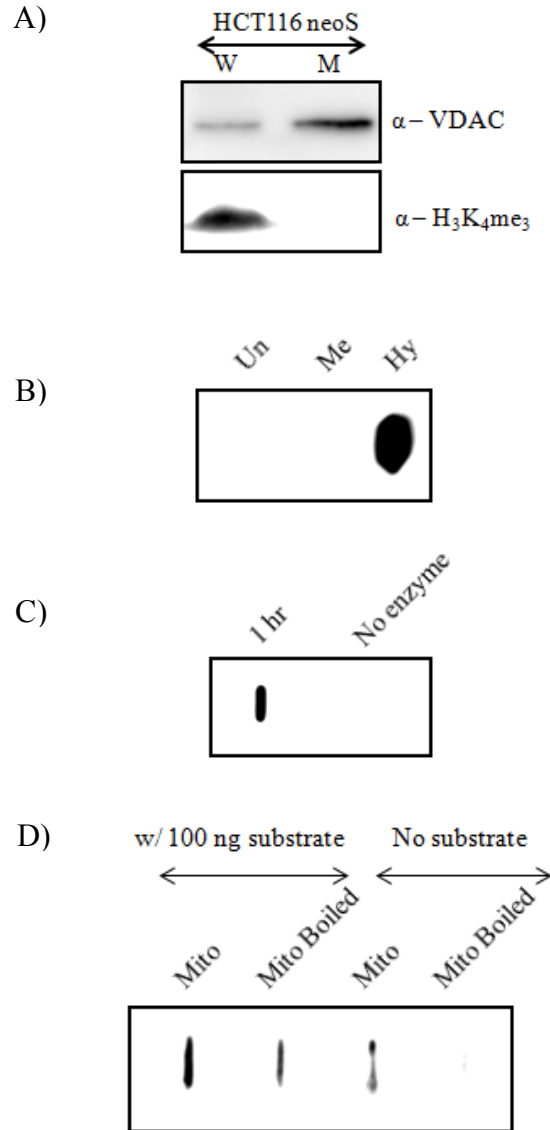
Figure 3-8: In vitro enzyme assay performed on trypsin treated, and percoll purified mitochondrial lysates from HCT116neoS cells.

A) Immuno blot indicating VDAC and H3K4me3 signals on Whole cells (W) and trypsin treated percoll purified (M) mitochondrial fractions. The mitochondrial fraction is free of nuclear contamination as indicated by H3K4me3 signal B) 0.4 ng of unmethylated, methylated and hydroxymethylated DNA spotted on the nitrocellulose membrane acts as positive controls for the specificity of the 5hmC antibody. C) 1 µg of recombinant C-terminal domain of Tet1 (Active Motif) incubated with 100 ng of methylated DNA, serve as positive control for the *in vitro* Tet enzyme assay. D) Percoll purified mitochondrial fractions on incubation with 100 ng of methylated DNA indicate Tet activity when compared to corresponding 'no substrate' controls.

incubated for 60 mins, and the signal was lost when the incubation time was extended to 180 mins (Figure 3-6). However, at pH 6.8, signal could be detected as early as 5 mins which increased at 30 mins and further increased at 60 mins (Figure 3-6). However, the signal decreased at 180 mins even at pH 6.8 which could be because of further conversion of 5hmC to the formylated and carboxylated forms [32]. No signal could be detected with 5hmC antibody in the no enzyme (-ve) control either at pH 6.8 or pH 8.0. It was evident that incubating the enzyme assay reaction at 37 $^{\circ}$C for 60 mins at pH 6.8 is optimum for the enzyme assay reaction (Figure 3-6) and these conditions were chosen for all further experiments.

Initially, crude mitochondrial preparations from HCT116neoS were tested for contamination from the nuclear fraction. Lysates from whole cell, mock and trypsin treated mitochondrial fractions were resolved on a 4-15% gradient gel and probed for antibodies against compartment-specific markers, VDAC (mitochondrial) and H_3K_4me3 (nuclear). Equal loading of fractions as per cell equivalents is indicated by the VDAC signal (Figure 3-7A). Absence of H_3K_4me3 signals in the mock and trypsin treated mitochondrial fractions indicates lack of nuclear contamination, the primary site of localization of Tet family of enzymes (Figure 3-7A). Thus, enzyme activity detected in the mock and trypsin treated mitochondrial fractions is expected to be the result of Tet enzyme present in the mitochondria rather than carryover of nuclear components.

5 ul of mock and trypsin treated crude mitochondrial fractions were incubated with 100 ng of methylated DNA substrate in the presence of Tet enzyme assay buffer under optimized conditions. 20 ng of DNA substrate from the enzyme assay was spotted on a nitrocellulose membrane and probed with 5hmC antibody. Figure 3-7 D shows a signal with 5hmC antibody in

lanes containing mock and trypsin treated mitochondrial lysates that was higher than that obtained in their corresponding 'no substrate' or 'boiled enzyme' controls. This suggests that Tet enzyme activity in present in both, mock and trypsin treated mitochondrial fractions and this activity is lost upon thermal denaturation. An assay performed with Tet1 recombinant protein for 60 mins acts as a positive control, which shows 5hmC signal higher than that of the 0 hr time point or the 'no enzyme' control (Figure 3-7C). Unmethylated, methylated and hydroxymethylated substrate (0.4 ng of each) serve as positive controls for the slot blot and for the specificity of the Ab (Figure 3-7B).

An experiment performed using percoll purified mitochondrial preparation also showed a similar result, confirming the presence of Tet activity in the mitochondria (Figure 3-8).

Mitoprot algorithm predicts with a high probability of Tet1 export to the mitochondria.

There are three members of the Ten- Eleven translocation (Tet) family of enzymes, all of which are capable of converting 5mC to 5hmC [24, 30]. Thus, all three of them could possibly generate 5hmC in mtDNA, provided they are translocated to the mitochondria. Initial prediction of mitochondrial transport was made using Mitoprot software using full length amino acid sequence for all three members of the Tet family. Analysis was also performed on amino acid sequences starting at each of the ATGs present downstream of the published translation start site. Performing the analysis in such a fashion would allow us to predict MTSs that are present internally or in the C-terminal region of the protein and help account for isoforms starting at ATGs downstream of the published translation start site. Mitoprot analysis on Tet1, Tet2 and Tet3 sequence revealed that Tet1 protein starting at ATG3 in human and ATG4 in mouse has a high probability of mitochondrial transport. The software also predicted a putative MTS

97

Species	ATG number	Probability of export to the mitochondria
Human	ATG 1	0.1793
	ATG 2	0.0076
	ATG 3	0.9541
Mouse	ATG 1	0.8294
	ATG 2	0.7732
	ATG 3	0.0079
	ATG 4	0.9300

Putative Human Mitochondrial leader sequence :- MALRSTSLSRRLSQPPLVV

Putative Mouse Mitochondrial leader sequence:- MALRRTSLSWRLSQRP

Figure 3-9: Mitoprot analysis of Tet1 enzyme predicts mitochondrial localization.

Mitoprot analysis was performed on Tet1/2 and 3 proteins to predict the probability of mitochondrial transport. The probabilities of mitochondrial transport for Tet1 protein starting at different ATGs are shown. Highlighted in green is the Tet1 protein starting at ATG3 in human and ATG4 in mouse, showing highest probabilities of mitochondrial transport.

expected to facilitate its transport, the sequence of which is highly conserved in both human and mouse (Figure 3-9). Such an analysis performed for Tet2 or Tet3 did not yield a high probability of mitochondrial transport for any of the sequences analyzed.

Tet1 may be present in the mitochondria:

The high probability of mitochondrial transport predicted for Tet1 made it a prime candidate for the enzyme responsible for generating 5hmC in mtDNA. To study the presence of Tet1 in the mitochondria, mitochondria were isolated from wild type, DNMT1 +/- and DNMT1 -/- mouse embryonic stem cells and lysates were prepared using SDS lysis buffer. Mouse embryonic stem cells were our system of choice because they express high levels of Tet1 mRNA when compared to that in different somatic tissues (Figure 3-10) [30]. Whole cell and mitochondrial lysates from all three cell lines were resolved on an immunoblot and probed with antibodies against Tet1, VDAC or H_3K_4me3.Figure 3-11 shows that immunoreactive material (~235 kDa), corresponding to the size of a full length Tet1protein, could be detected using Tet1 antibody in the mitochondrial fractions from all three cell lines. Levels of VDAC, which serves as a mitochondrial marker, indicate that cell equivalent amounts of lysate were loaded. Lack of H_3K_4me3 signal in any of the mitochondrial fractions indicates that mitochondrial fractions are free from contamination by nuclear material (Figure 3-11). This result suggested that Tet1 may be present in the mitochondria, consistent with published results by another group that showed that both Tet1 and Tet2 are present in the mitochondrial fractions of brain mitochondria [126].

mtDNA is organized in the mitochondrial matrix in the form of nucleoids [56]. This makes it essential for the enzyme acting on mtDNA to be transported to the mitochondrial matrix. Presence of the Tet1 enzyme in trypsin treated mitochondrial fractions was expected to firmly

establish its localization to the mitochondrial matrix. The rationale is that, if present in the mitochondrial matrix, Tet1 enzyme should be protected from cleavage by trypsin. On the contrary, any contaminating nuclear or cytosolic Tet1 enzyme adhering to the outer mitochondrial membrane, and therefore resulting in a false positive signal on the immunoblot, should be degraded by trypsin. Mitochondria isolated from wild type, DNMT1+/- and DNMT1 -/- mouse embryonic stem cells were treated with trypsin under hypotonic conditions and then lysed in SDS buffer. Whole cell, mock and trypsin treated mitochondrial lysates from the three cell lines were run on a polyacrylamide gel and probed with Tet1, VDAC or H_3K_4me3 antibodies. As previously mentioned, VDAC and H_3K_4me3 antibodies serve as appropriate loading and fractionation controls respectively (Figure 3-12). Full length Tet1 (~235kDa) could be detected using the Tet1 antibody whole cell lysates and mock treated mitochondrial fractions of wild type, DNMT1+/- and DNMT1 -/- cells (Figure 3-12). However, this immunoreactive material was absent in trypsin treated mitochondrial fractions from all three cell lines (Figure 3-12). This suggests that Tet1 protein might not be present in the mitochondrial matrix and the signal obtained with Tet1 antibody in the previous blot may be a result of carryover from the nuclear/cytosolic fraction (Figure 3-12) (see discussion). Absence of Tet1 immunoreactive material in trypsin treated mitochondrial fractions was not a result of mitochondrial lysis during trypsin treatment. The presence of mtDNMT1 in both mock and trypsin treated mitochondrial fractions (data not shown) serve as a control to prove that no lysis of mitochondria occurs during trypsin treatment, which would otherwise result in a false negative.
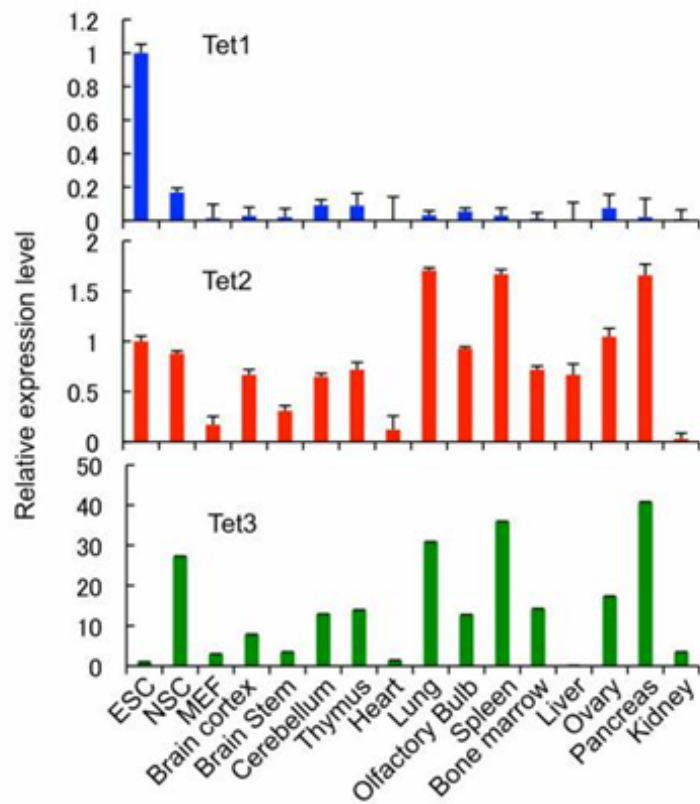
S Ito et al. Nature **000**, 1-5 (2010) doi:10.1038/nature09303

Figure 3-10: mRNA expression levels of Tet1/2 and 3 in different cells and tissue type from mouse.

Ito et. al. performed the mRNA expression analysis in different cells and tissue type from mouse. As indicated above, highest Tet1 expression levels are seen in mouse embryonic stem cells. This made mouse embryonic stem cells our system of choice for studies on Tet1. Figure obtained from Ito et. al. Nature 2010.
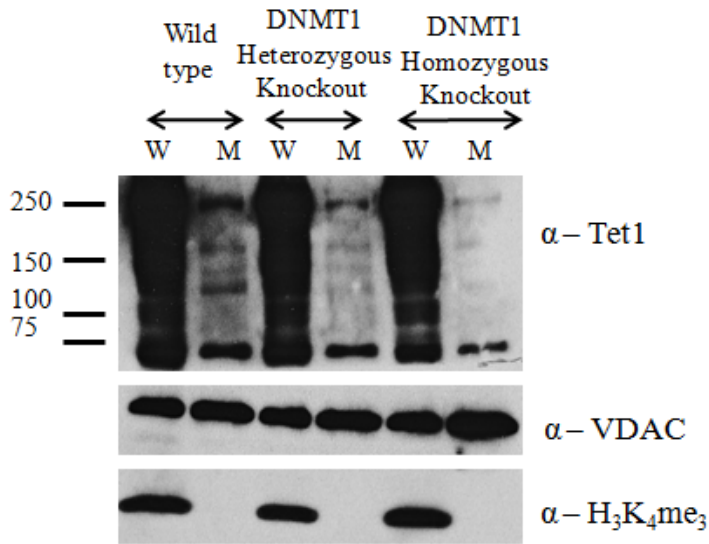
Figure 3-11: Immunoblot of whole cell and mitochondrial lysates from wild type, DNMT1 +/- and DNMT1 -/- cells.

Whole cell (W) and Mitochondrial (M) lysates from wild type, DNMT1 +/- and DNMT1 -/- mouse embryonic stem cells were run on a 4-15% gradient gel and probed for Tet1, VDAC and H3K4me3 antibodies. Absence of H3K4me3 signal in mitochondrial samples (lanes labelled M) indicates that mitochondrial fractions from all three cell lines are free from nuclear contamination. The presence of immunoreactive material in the lanes indicated 'M' suggests Tet1 may be present in the mitochondria from all three cell lines.
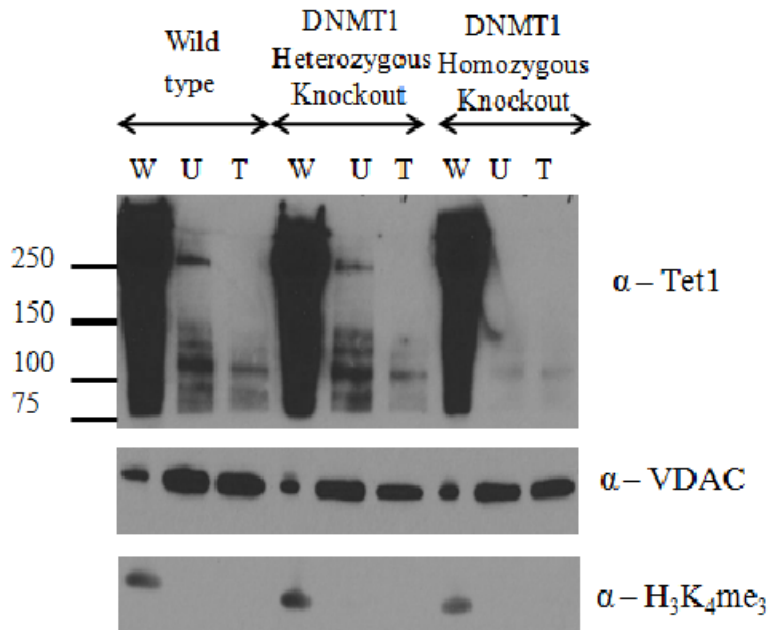
Figure 3-12: Trypsin treatment of mitochondrial lysates from wild type, DNMT1 +/- and DNMT1 -/- cells.

Whole cell (W) and Untreated (U) or trypsin treated mitochondrial (T) lysates from wild type, DNMT1 +/- and DNMT1 -/- mouse embryonic stem cells were probed for Tet1, VDAC and H3K4me3 antibodies. Absence of H3K4me3 signal in mitochondrial fractions (lanes labelled U and T) indicates that mitochondrial fractions from all three cell lines are free from nuclear contamination. Presence of a band corresponding to Tet1 in mitochondrial fractions untreated with trypsin (U) but its absence in trypsin treated mitochondrial fractions (T) indicated Tet1 is not transported to the mitochondrial matrix.

The putative mitochondrial leader sequence of Tet1 does not carry heterologous protein GFP to the mitochondria

If sequences predicted by the mitoprot algorithm serve as true MTSs, they should facilitate transport of any heterologous protein including GFP to the mitochondria. As a proof of principle, MTS from mtDNMT1 when fused to GFP and transfected into cells, localizes GFP to the mitochondria. [117]. We cloned putative murine Tet1 MTSs starting at ATG1, ATG3 and ATG4 fused in-frame with C-terminal GFP tag in pLEGFP vector (Figure 3-13) and constructs were transfected in NIH3T3 cells. The cells were stained with mitotracker red dye and nucleus specific DAPI stain and analyzed using confocal microscopy. Only sequences starting at either ATG1 or ATG4 were expected to translocate GFP to the mitochondria, as predicted by the mitoprot algorithm. No localization of GFP to the mitochondria was observed for any of the constructs, seen by the lack of convergence between GFP and mitotracker red dye (Figure 3-14). ATG1-GFP was exclusively localized in the nucleus as seen by overlap of green fluorescence with DAPI. ATG4-GFP construct was also localized to the nucleus but displayed green fluorescence in areas surrounding the nucleus (Figure 3-14). However, all of the green perinuclear stain failed to show overlap with mitotracker red dye, which otherwise results in a yellow merge. These results indicate that the sequences predicted for Tet1 protein by Mitoprot software do not serve as MTSs.

Immunofluorescence using Ab against Tet1 does not show co-localization of Tet1 with Mitotracker red.

In a different approach to establish presence of Tet1 in the mitochondria, immunostaining using Tet1 antibody was performed on Hela cells. Hela cells were grown on coverslips, fixed with

paraformaldehyde and permeabilized using cold methanol. These were then incubated with Tet1 antibody followed by FITC conjugated secondary antibody, stained with mitotracker red and visualized using confocal microscopy. Tet1 antibody bound primarily to the nucleus, as seen by overlap with nucleus specific DAPI stain (Figure 3-15). The reaction with Tet1 Ab surrounding the nucleus is punctate in nature but fails to show overlap with mitochondrial specific stain Mitotracker red (Figure 3-15). The presence of Tet1 in the mitochondria would be marked by Tet1 staining marked by green fluorescence combined with red staining of the mitotracker red dye giving a yellow merge. No such yellow colored merge could be seen. These results combined with absence of immunoreactive material in the trypsin treated mitochondrial fraction (Figure 3-12), made it difficult for us to convincingly demonstrate presence of Tet1 in the mitochondria.

Knockout of Tet1 does not alter hydroxymethylation of mtDNA

As an alternative approach to prove the presence of Tet1 in the mitochondria, we asked if altered levels of Tet1 change 5hmC levels in mtDNA. With this objective, we decided to look at 5hmC levels in cells that contain either heterozygous or homozygous knockout of Tet1. These cells were a gift of Dr. Rudolph Jaenisch at Whitehead Institute of Technology, MA. Tet1 was knocked-out in these cells using gene targeting by homologous recombination [124]. This strategy resulted in targeted deletion of exon 4 in the Tet1 gene and introduced a frame shift, such that a Tet1 transcript would express a truncated protein lacking the entire C-terminal catalytic domain [124]. Figure 3-16 shows validation of Tet1 knock-out in these cells. Tet1 +/- cells show 50% reduction in Tet1 protein expression compared to wild type cells, whereas Tet -/- cells completely lack immunoreactive material. Total DNA was isolated from V6.5 wild type,

A) Murine Tet1 sequence

F1
ATG TCT CGG TCC CGC CCC GCA AAG CCT TCC AAA TCA GTC AAA ACA AAG
CTA CAG AAA AAG AAA GAC ATC CAG ATG AAG ACG AAG ACA TCA AAG CAG
GCT GTC AGA CAT GGG GCA TCA GCG AAG GCA GTG AAC CCC GGA AAA CCG
AAG CAA TTA ATT AAA AGA AGA GAT GGT AAG AAG GAA ACG GAA GAC AAA
ACC CCG ACA CCG GCG CCG AGT TTT CTG ACG CGA GCC GGG GCC GCC CGA
F3
ATG AAT CGG GAT CGA AAC CAA GTT CTT TTT CAG AAC CCA GAC TCC TTA
                                F4
ACT TGC AAC GGG TTT ACA ATG GCT CTT CGA AGA ACC TCT CTT AGC TGG
                        R2
CGA CTC TCC CAG CGC CCA
                    R3

Probability of export to
the mitochondria

ATG1
F1R3-GFP   118aa   0.8294

ATG3
F3R2-GFP   22aa   0.0079

ATG4
F4R3-GFP   16aa   0.9300

Figure 3-13: Design and construction of MuTet1-GFP fusion constructs for analysis by confocal microscopy.

A) Nucleic acid sequence for the partial coding region of Tet1 from mouse is shown. Different mitochondrial leader sequences starting at different ATGs namely, ATG1 (highlighted in blue), ATG3 (highlighted in red) and ATG4 (highlighted in Green), predicted by the Mitoprot software are indicated. Also shown are the positions of the different primer sets used for amplification and cloning of each of the indicated MLSs. B) Different MLSs amplified in (A) were cloned in front of GFP to generate fusion constructs. Probabilities of mitochondrial transport for each of the constructs are indicated.
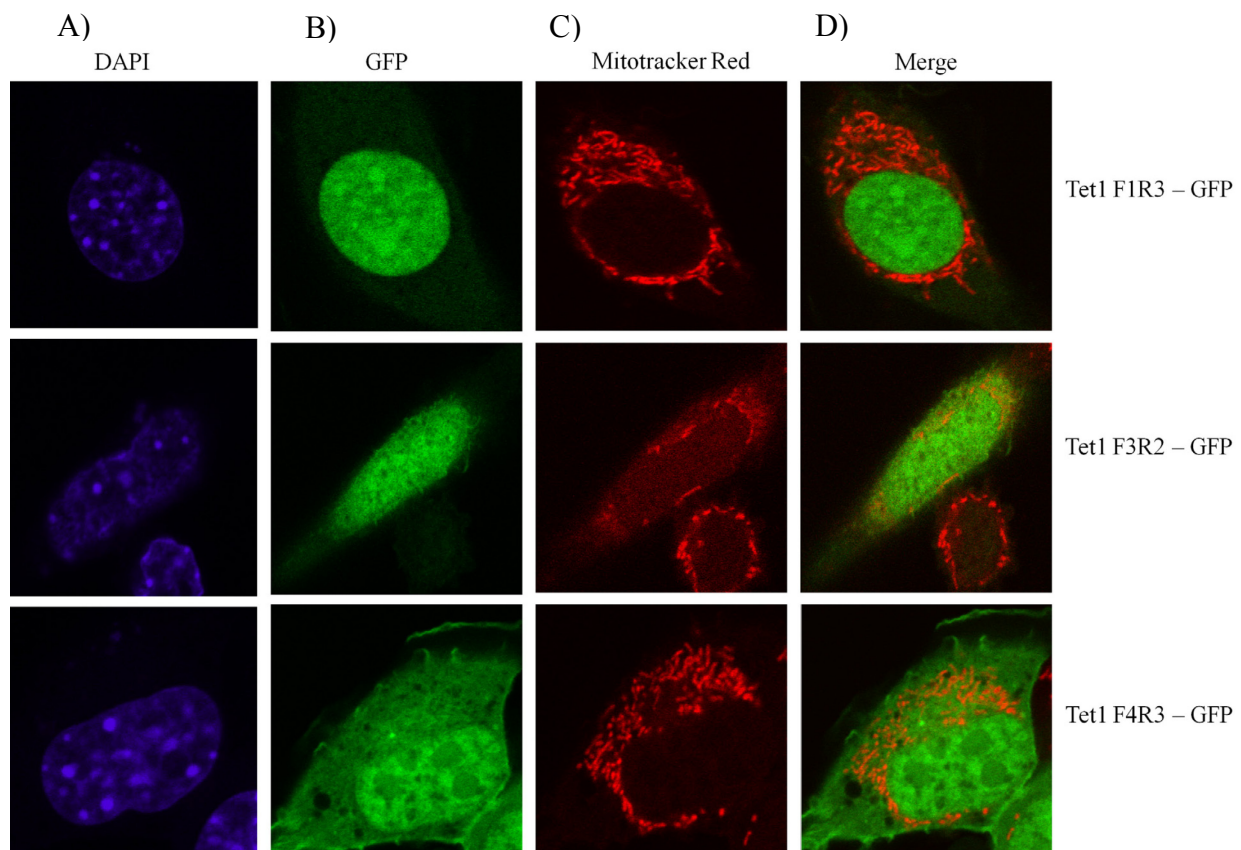
Figure 3-14: Confocal Microscopy of Tet1 MTS –GFP fusion constructs.

Confocal images of cell transfected with either Tet1F4R3 or Tet1F1R3- GFP fusion constructs are shown. Channels' showing specific stains: DAPI (A); GFP (B); Mitotracker Red (C); and merge (D); are shown. Neither Tet1F4R3 nor Tet1F1R3 carry heterologous protein GFP to the mitochondria.
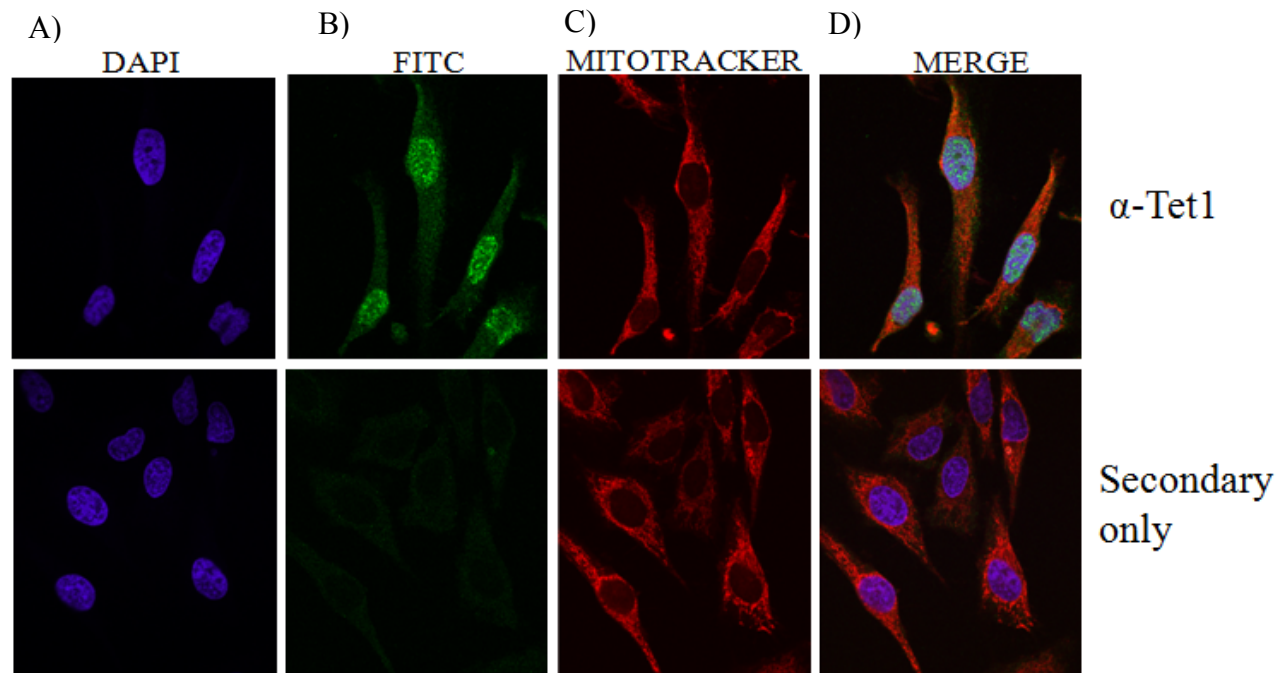
Figure 3-15: Immunofluorescence using Tet1 antibody as performed on HeLa cells.

Confocal images of immunofluorescence performed on HeLa cells are shown. Channels' showing specific stains: DAPI (A); GFP (B); Mitotracker Red (C); and merge (D) are indicated.Tet1 antibody failed to co-localize with mitochondrial specific stain mitotracker red, as seen in merge.

Tet1 mESCs

+/+ -/- +/-

Tet1
(~220kD)

*

Actin
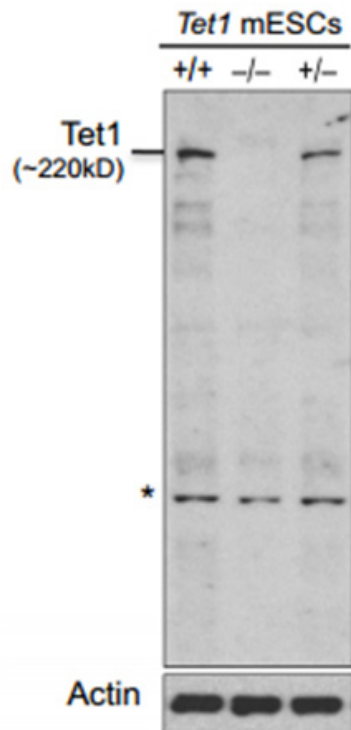
Dawlaty et. al. Cell Stem Cell, Vol9, Issue 2, 166-175, 2011

Figure 3-16: Immunoblots proving partial and complete knockouts of Tet1 in mouse embryonic stem cells.

Immunoblots of whole cell lysates from wild type, Tet1 +/- and Tet1 -/- cells using anti-Tet1 antibody (Millipore). Blots indicate 50% reduction in Tet1 protein in Tet1 heterozygous knockout cells and complete absence of Tet1 protein in Tet1 homozygous knockout cells, as compared to wild type.
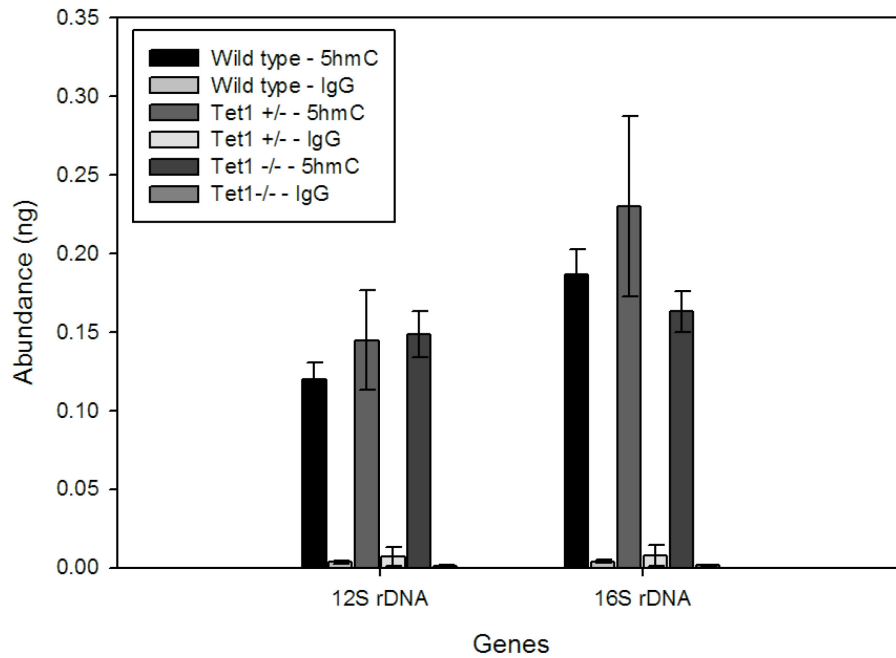
Figure 3-17: HydroxyMeDIP assay on wild type, Tet1 +/- and Tet1 -/- mouse embryonic stem cells.

HydroxyMeDIP assay on total DNA from wild type, Tet 1+/- and Tet1 -/- mouse embryonic stem cells reveal no change in 5hmC in Tet1-/- cells as compared to wild type cells. Data represents mean ± SD of two independent biological repeats.

Tet1 +/- and Tet1 -/- cells, sonicated to a size range of 200-700 bp and subjected to a HydroxyMeDIP assay. 16S rDNA and 12 rDNA regions from the mitochondrial DNA were analyzed by qPCR. Figure 3-17 shows that cells containing heterozygous knockout of Tet1 show no change in enrichment of mitochondrial by 5hmC Ab as compared to that of wild type. Even in Tet1 -/- cells, there was no change in 5hmC content as compared to Tet1 +/- or wild type cells (Figure 3-17). The signal obtained with 5hmC antibody was expected to reduce from wild type to Tet1 +/- cells, marked by further reduction in Tet1 -/- cells. We saw no such decrease in either Tet +/- or Tet1 -/- cells as compared to wild type. This suggested that Tet1 is not responsible for generation of 5hmC in mtDNA.

Discussion:

Having discovered the presence of 5mC in mtDNA and the underlying mechanism responsible for its generation [117], we first sought to determine if 5hmC, a product of oxidation of 5mC, was present in mtDNA. Just like 5mC, 5hmC has been described as a normal component of human genomic DNA [40]. At the time, the only assays that allowed successful and unambiguous detection of 5hmC in DNA were HydroxyMeDIP, β-glucosyltransferase assay and analysis by HPLC-MS assay after hydrolyzing DNA to single nucleoside level. Due to relative technical ease, we used the first two assays to establish beyond doubt that 5hmC is present in mitochondrial DNA.

Generation of 5hmC in mtDNA using mechanisms independent of 5mC:

The presence of 5hmC in mtDNA prompted us to investigate the mechanism responsible for its generation. Prior to detection of 5hmC in mtDNA, we had demonstrated that an isoform of

DNMT1 was translocated to the mitochondria. Other studies had shown that DNMT1 can generate 5hmC from C in DNA in the presence of formaldehyde *in vitro* [125]. We asked if mtDNMT1 might be responsible for direct generation of 5hmC in mtDNA. We used the HydroxyMeDIP assay to answer this question and we observed a modest increase in 5hmC containing DNA levels in cells lacking DNMT1-/- as compared to wild type. This suggested that DNMT1 was not responsible for direct generation of 5hmC from C.

These results raise the possibility of a DNA methyltransferase enzyme other than DNMT1, being present in the mitochondria. This enzyme would compensate for loss of mtDNMT1 by increasing 5mC content in the mtDNA which in turn could result in increased levels of 5hmC possibly by the action of Tet enzymes or an otherwise unknown mechanism. We have extensively looked for the presence of DNMT3a and DNMT3b in the mitochondria. To the best our knowledge, DNMT1 is the only member of catalytically active DNA methyltransferases that is present in the mitochondria [117]. An alternate explanation for our observation could be that 5hmC is incorporated in mtDNA as a pre-modified base and that its generation is independent of 5mC. Such a pathway has been described in the T-even bacteriophages [127]. As demonstrated in Figure 3-18, dCMP is acted upon by dCMP hydroxymethylase to form d5hmCMP, which is then directly incorporated into DNA. The reaction uses 5, 10 – methylene tetrahydrofolate as the 1-carbon group donor. This essentially makes the process of 5hmC generation independent of 5mC but dependent on mitochondrial folate metabolism. Both dCMP and 5,10-methylene tetrahydrofolate are present in mammalian mitochondria [128]. If an enzyme homologous to dCMP hydroxymethylase were to be found, the existence of such a pathway in mammalian mitochondria is not impossible. Mammalian mitochondria are known to contain enzymes that are homologous to those in bacteriophages. For example, mitochondrial
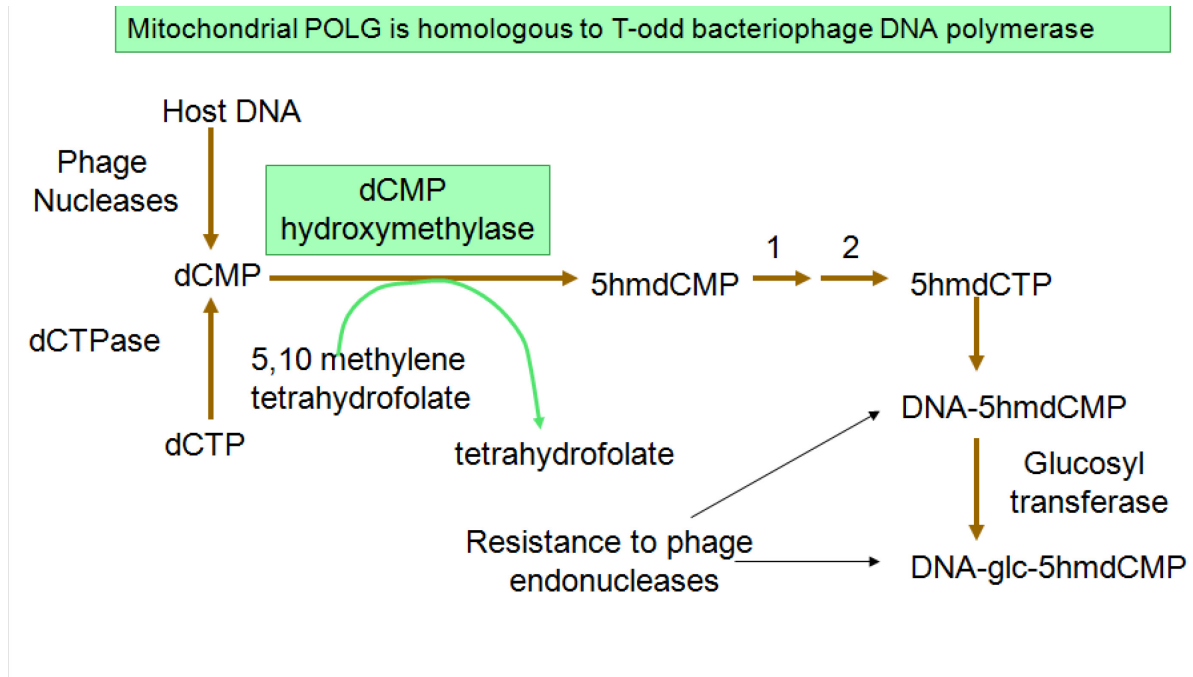
Figure 3-18: Schematic of the pathway generating 5hmC in T-even bacteriophages.

polymerase γ (POLG), the enzyme responsible for replication of mtDNA, is homologous to DNA polymerase in T-odd bacteriophages. It will be interesting to know if an enzyme homologous to viral dCMP hydroxymethylase is present mammalian mitochondria.

To test if 5hmC generation in the mitochondria was folate dependent and therefore was generated independent of 5mC, we analyzed 5hmC content in wild type chinese hamster ovary cells (CHO), the GlyB mutant CHO cells, which specifically contains a mutant and hence a non-functional form of mitochondrial folate transporter (MFT), and the GlyB-huMFT CHO cells, in which wild type form of human MFT was restored. We isolated total DNA from the three cell lines and performed a hydroxyMeDIP assay. Dependence of 5hmC generation in mammalian mitochondria would be indicated by a reduction in 5hmC signal in a hydroxyMeDIP assay in GlyB cells when compared to that in wild type CHO cells. Furthermore, restoring the wild type form of human MFT in GlyB cells would then be expected to bring 5hmC levels in these cells similar to that of wild type CHO cells. Although, the 5hmC signal in GlyB-huMFT cells was similar to that of the wild type, we did not see the expected reduction in 5hmC signal in GlyB cells compared to the wild type cells in the 12S rDNA region analyzed (Figure 3-19). We plan to extend our analysis to other regions of mtDNA from these cells; however, these results suggested that 5hmC generation in mtDNA is not folate dependent and that an enzyme catalyzing a reaction similar to dCMP hydroxymethylase may not be present in mammalian mitochondria.

Mitoprot algorithm failed to accurately predict translocation of Tet1 to the mitochondria:

Around the time of our discovery of 5hmC in mammalian mitochondrial DNA, there were three reports in the literature showing that Tet enzymes generate 5hmC in the nucleus [24, 29, 30]. As obvious candidates for generation of 5hmC in mtDNA, we asked
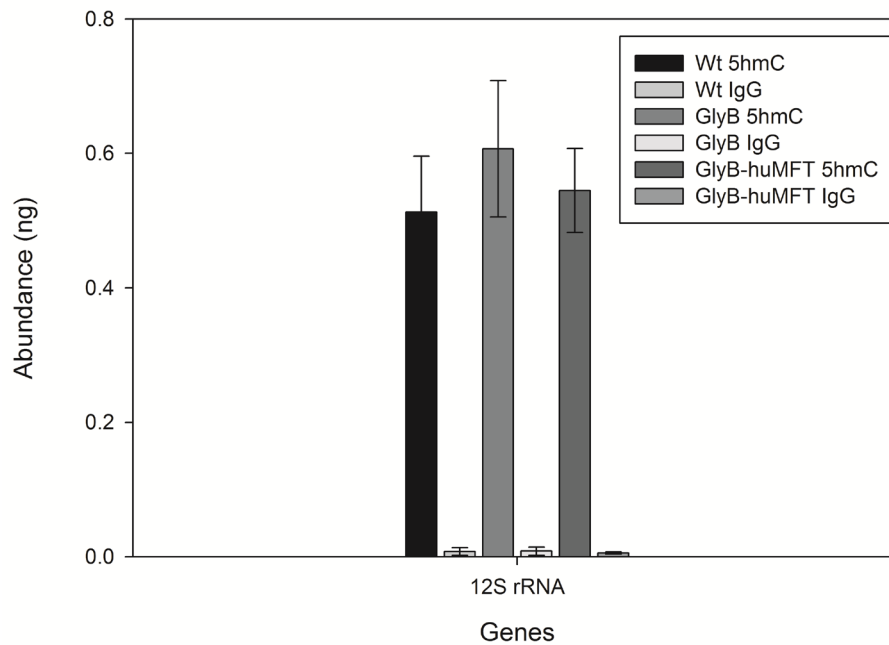
Figure 3-19: HydroxyMeDIP assay on wild type, GlyB and GlyB-huMFT chinese hamster ovary (CHO) cells.

HydroxyMeDIP assay on total DNA from wild type, GlyB and GlyB-huMFT Chinese hamster ovary (CHO) cells reveal no change in 5hmC in GlyB cells as well as GlyB-huMFT cells when compared to wild type cells. Data represents mean ± SD of two independent biological repeats.

if any member of the Tet family of dioxygenases is present in the mitochondria. Similar to our analysis for DNMT1 [117], we performed mitoprot analysis on all three Tet enzymes. The software performs a discriminant functional analysis using approximately 47 parameters [129]. Some of the parameters used by the software to predict a potential localization to the mitochondria are a) ZoneTo, which is the maximum number of amino acid residues located at the N-terminal that can be considered for putative mitochondrial leader sequence. Only the sequence N-terminal of the first acidic residue present in the sequence is considered. This is due to the fact that MTSs are comprised of hydrophobic and positively charged residues which form an amphipathic helix and are devoid of negatively charged residues [129] b) Coef20, which indicates whether abundance of particular residues in first 20 amino acids is in accordance with that seen for MTSs. c) Cleavsite, which defines the last putative cleavage site for mitochondrial peptidase in a region defined by ZoneTo. d) The number of positively charged residues in a region defined by ZoneTo [129]. The use of such parameters limits the prediction of mitochondrial transport to those proteins that contain an N-terminal MTS. The authors acknowledge that accuracy of prediction of mitochondrial transport reduces drastically for proteins lacking the N-terminal leader peptide [129]. Thus, any proteins carrying an internal or a C-terminal MTS would be missed. Furthermore, it does not take into account any isoforms that do not carry the traditional N-terminal end as the full length protein. To overcome this problem, we inititally performed analysis on the full length sequence of Tet proteins. We also performed analysis using amino acid sequence starting at each of the ATGs downstream of the published translation start site, while maintaining the C-terminal end. Such an analysis predicted a high probability of mitochondrial transport for Tet1, starting at ATG3 in human and ATG4 in mouse. However, neither Tet2 nor Tet3 showed any probability of mitochondrial transport. Similar to the scenario for mtDNMT1, we have examined the genomic

sequences upstream of the published transcriptional start site, to determine whether sequence in-frame with Tet protein coding sequence could encode a MTS. No such sequences were found for any of the Tet proteins. Unless there were an upstream exon that remains to be discovered for any of the Tet proteins, the Tet family of proteins appear to lack an N-terminal MTS. 5' RACE experiments will need to be performed in order to determine the true 5' start site for all the Tet enzymes.

Mitoprot software, unlike the analysis of mtDNMT1, failed to accurately predict mitochondrial transport of Tet1. Our analysis of the putative Tet1 MTS fused to GFP showed that the predicted MTS did not carry heterologous protein GFP to the mitochondria (Figure 3-14). MTSs are characterized by formation of an amphipathic helix that shows clustering of positively charged residues on one face and hydrophobic residues on the other. To determine whether putative Tet1 MTS formed such an amphipathic helix, we performed helical wheel analysis. In comparison with the MTSs of the known mitochondrial proteins, Tet1 MTS formed a very poor amphipathic helix implying that the prediction by mitoprot was possibly a false positive. There are many nuclear encoded mitochondrial proteins that lack MTS but are transported to the mitochondria with the help of chaperone proteins. Thus, if no MTS can be found for Tet proteins, its association with chaperones will need to be investigated by performing co-immunoprecipitation experiments.

Tet1 may not be transported to the mitochondrial matrix:

Based on mitoprot analysis, we elected to pursue the possibility of transport of Tet1 to the mitochondria. As a first measure, we decided to look for Tet1 in mitochondrial fractions from mouse embryonic stem cells using immunoblots. Immunoreactive material was found using Tet1

117

antibody in mitochondrial fractions from wild type, DNMT1+/- and DNMT1-/- cells (Figure 3-11).

Our initial finding was in agreement with another independent study, that showed bands corresponding to Tet1 and Tet2 on immunoblots containing mitochondrial fractions from the brain, using anti-Tet1 and anti-Tet2 antibodies (Santacruz) [126]. However, we found that our immunoreactive material was not protected from trypsin treatment of intact mitochondria (Figure 3-12). Dzitoyeva and colleagues did not perform this rigorous test for intramitochondrial localization. Mitochondrial DNA is present in the mitochondrial matrix arranged as nucleoids [56]. For an enzyme to be able to act on mtDNA, it must be transported into the mitochondrial matrix. Our results suggest that Tet1 enzyme may not be present in the mitochondrial matrix. Even in an immunofluorescence assay, antibody directed against Tet1 failed to show co-localization with mitotracker red dye. Thus, our conclusions do not agree with the published results [126].

Disagreement with the published results of Dzitoyeva et. al. could be attributed to certain technical differences in the experiments performed. During mitochondrial lysate preparation, proteins in the mitochondrial fraction are concentrated. To reflect the proportion of protein present in the organelle *in vivo* in comparison to the whole cell and therefore to make a true comparison between them, we load 75 μg of whole cell lysate while loading only 18 μg of mitochondrial lysate. As seen by equal VDAC signal in the mitochondrial and the whole cell fraction, our protein loading reflects equal amounts of protein loaded as per cell equivalence. On the contrary, Dzitoyeva et al. (2012) have loaded 30 μg of protein for the nuclear as well as the mitochondrial fractions. Thus, in their case, equal amounts of protein loaded do not reflect the equal amounts of proteins calculated from cell equivalents and their mitochondrial fraction is

likely to be overloaded, causing a false positive. Furthermore, the authors fail to determine whether Tet1 and Tet2 were protected from trypsin/proteinase K digestion of intact mitochondria. Their immunofluorescence data also shows very little, if any, co-localization with the mitochondria specific mitotracker red dye. Hence, it is uncertain whether Tet enzymes are translocated to the mitochondrial matrix. Based on our results and considering the technical shortcomings of the experiments performed in the referenced publication, the published results are unconvincing and need careful re-examination.

Further experiments are needed to explore the role of Tet2 and Tet3 in generation of 5hmC in mtDNA:

Our search for Tet2 in the mitochondria was hindered by the lack of success in optimization of the Tet2 antibody used by Dzitoyeva et al.. Our attempt at optimizing this antibody had preceeded the publication. The authors used mitochondria from mouse brain as opposed to wild type mouse embryonic fibroblasts used by us. In hindsight, the choice of tissue/cells may have made the difference. It is known that 5hmC is abundant in the brain [40] and therefore it might be expected that Tet expression levels are higher in the brain as opposed to other tissue/cell types. We plan to re-optimize the antibody and explore the possibility of Tet2 in the mitochondria.

Complete knockout of Tet1 did not show any change in 5hmC abundance in immunoprecipitates of mitochondrial DNA (Figure 3-17). If Tet1 were responsible for 5hmC generation in the mitochondria, it would be expected that knockout of Tet1 would result in a decrease in 5hmC content. Thus, the results we obtained indicate that Tet1 may not be involved in generating 5hmC in mtDNA. Dawalty et al. showed that complete knockout of Tet1 in mouse

embryonic stem cells resulted only in 35% decrease in the 5hmC content in the regions analyzed in nuclear DNA. All the mitochondrial DNA fragments analyzed for MeDIP/hydroxyMeDIP assay contain 6-8 CpGs that could be potentially methylated or hydroxymethylated. Even at 35% loss in hydroxylation, there could still be a significant proportion of CpGs that would retain 5hmC modification, which will be precipitated in a MeDIP/hydroxyMeDIP, thus, undermining small differences. Taking this into consideration, an alternative method of analysis should be used to confirm the results obtained in the MeDIP assay. We have optimized analysis of 5mC and 5hmC content in the mitochondrial DNA using HPLC-MS. This assay will be more informative since it provides a quantitative measure of each nucleosides, including cytosines modified to 5mC and 5hmC in mitochondrial DNA of various cells. We plan to use this assay for a comprehensive and quantitative analysis of the role of Tet proteins in 5hmC generation in the mitochondria.

Tet1 and Tet2 knock out mouse models and cells have been generated [124, 130]. These systems containing Tet1/2 knockouts have been requested from the authors and they have kindly agreed to provide it to us. Mitochondria from Tet1/2 knockout cells and also from tissues from Tet1/2 knockout mice can be isolated and 5hmC content can be studied using HPLC-MS. We have successfully cloned full length MuTet1, starting at both ATG1 and ATG4, MuTet2 and C-terminal domain of HuTet3 in a pInducer22 lentiviral expression vector. Studies on knockout mice/cells will be complemented by studies utilizing cells overexpressing each of these enzymes in HCT116 cells. mtDNA from these cells over expressing Tet enzymes will be analyzed for change in 5hmC content using HPLC-MS analysis.

Thus, overall we conclude that current data is insufficient to define the mechanism responsible for 5hmC generation in the mitochondria. With the above set of well-defined experiments we plan to continue this pursuit.

Chapter 4: Perspectives

Model for epigenetic regulation of transcription in mitochondria.

As presented in Chapter 2, altered levels of mtDNMT1 due to loss of p53 alters mitochondrial transcription. Our initial analysis using cDNA synthesized using random hexamers indicated up-regulation of ND1 on heavy strand and down-regulation of ND6 on the light strand in p53-/- cells compared to wild type (Figure 2-6). We then refined our studies by using cDNA synthesized using gene-specific primers which revealed a more pronounced upregulation of ND1, similar upregulation of Cox1 but no change in ND6 expression in p53-/- cells compared to wild type (Figure 2-7). These two sets of analyses have led us to question whether mtDNMT1 affects mitochondrial transcription in a gene-specific (Figure 2-6) or a strand-specific manner (Figure 2-7). Furthermore, analyses performed on cDNA synthesized using gene-specific primers (Figure 2-7) represents steady-state levels of only the polycistronic transcript, and fails to capture the contribution of mature processed mRNA molecules. To gain insight, we now plan to perform gene-specific and strand-specific transcription analyses. These studies will use gene-specific primers designed in every gene along both heavy and the light strand, to synthesize cDNA. The cDNA can then be analyzed for expression of the same gene as the one used for cDNA synthesis, using nested primers by qPCR. Such an analysis will account for steady state levels of both the mature mRNA molecules and the polycistronic transcript and hopefully help us tease apart gene specific versus strand specific effects on mitochondrial transcription.

The effect on mitochondrial transcription, due to up-regulation of mtDNMT1, potentially involves a number of different mechanisms and we plan to test each of these individually.

Firstly, an increase in expression of the heavy strand genes (Figure 2-7) could be due to disruption of MTERF binding at the terminator site. We would expect that disruption of MTERF binding at the terminator site would result in inefficient termination of HSP1 transcription, therefore allowing increased read-through to heavy strand genes located downstream of the terminator site by POLRMT. This would result in an increase in expression of heavy strand genes, similar to the result observed in our studies (Figure 2-7). It would also be expected to cause disruption of the transcription bubble and therefore might cause reduction in 12SrRNA and 16SrRNA levels. It is possible that MTERF binding at its terminator site is disrupted via protein-protein interaction with mtDNMT1, which could be tested by co-immunopreciptation studies. MTERF binding may also be disrupted due to increased methylation at the terminator site by upregulated mtDNMT1 in p53-/- cells. In such a case, the methylation status at the terminator site could be analyzed initially by either MeDIP or bisulphite sequencing. We would predict to see higher methylation at the terminator site in p53 -/- cells compared to wild type. Next, we could test if methylation of DNA alters affinity of MTERF binding at its site. For this, we plan to purify recombinant MTERF and test its affinity with unmethylated or methylated oligonucleotide substrates, containing the terminator site sequence, using Biacore or isothermal titration calorimetry (ITC). Laura Burton, a current Ph.D. student in the lab, is now actively pursuing these studies.

A second possibility is that TFAM, which binds to the region between the HSP2 and LSP promoters, is no longer able to bind because of increased methylation by mtDNMT1 in the region between the two promoters. TFAM has been shown to repress transcription from HSP2 promoter by binding at the TFAM response element (TRE) present near the HSP2 promoter [76, 122]. This repression of HSP2 transcription by TFAM can be overcome by increasing

concentrations of POLRMT and TFB2M, implying that binding of TFAM to the TRE element limits the recruitment of POLRMT-TFB2M transcription initiation complex at the HSP2 promoter in a competitive manner [122]. Methylation dependent interference of TFAM binding at this site would thus be expected to cause increased recruitment of TFB2M-POLRMT complex at HSP2 promoter. This would in turn cause increased expression of heavy strand protein coding genes, an effect observed in our studies on mitochondrial transcription (Figure 2-7). ChIP analysis using purified mitochondrial lysates should help us determine relative TFAM binding in the D-loop region in wild type versus p53 -/- cells which can then be correlated to the methylation status of the binding site in the two cells.

A third possibility might be the increased recruitment of TFB2M-POLRMT initiation complex in a methylation dependent manner at the HSP2 promoter, which would also explain increased expression of heavy strand protein coding genes seen in our analysis (Figure 2-7). We have yet to determine if POLRMT-TFB2M complex has a higher affinity towards methylated or DNA relative to unmethylated DNA. Alternatively, relative occupancy of core transcription machinery TFB2M and POLRMT, can be determined at the two promoters to answer the possibility of efficient promoter firing from HSP2 promoter as compared to LSP in cells overexpressing mtDNMT1.

It remains to be confirmed whether the effect seen on mitochondrial transcription is a result of loss of p53 or upregulation of mtDNMT1. To unambiguously answer this question, we plan to use genetic approaches to over-express and knockout the mitochondrial specific isoform of DNMT1. The 5' region of DNMT1, upstream of the translational start site for the nuclear isoform (ATG3), codes for the mitochondrial leader peptide. Two stop codons can be inserted

upstream of ATG3 by homologous recombination using targeting constructs containing antibiotic resistance cassettes. Using two different antibiotic resistance cassettes, e.g: neomycin and hygromycin, both alleles can be targeted to completely eliminate expression of mtDNMT1, while keeping nuclear DNMT1 expression unchanged. A similar strategy has been used by others to completely knock out DNMT1 in HCT116 cells, which results in a distinct phenotype where cells undergo G2 arrest and eventual cell death due to mitotic catastrophy [131]. It remains to be seen if complete loss of mtDNMT1 yields a similar response. If so, it would provide evidence that mtDNMT1, similar to nuclear DNMT1, plays an essential role in cell survival. A conditional knockout strategy would then be used to elucidate role of mtDNMT1 on mitochondrial transcription.

To over-express the mtDNMT1 specific isoform, we have cloned full-length DNMT1 cDNA carrying the upstream mitochondrial leader peptide in an expression vector. Furthermore, ATG3 has been mutated to ATC and sequence around ATG1 and ATG2 has been modified to put them in a context of an ideal Kozak consensus sequence [121], thus allowing over-expression of mtDNMT1 exclusively. This construct can be transfected or transduced to generate cell lines stably over-expressing mtDNMT1. Using both these cell lines and by performing gene-specific and strand-specific transcription studies we hope to definitively establish role of mtDNMT1 in regulating mitochondrial transcription.

IDH mutants and disruption of Tet function.

Mutations in the IDH family were first found in brain tumors such as astrocytomas and glioblastomas, with almost 70% of the patients showing lesions in IDH1 and IDH2 genes [132-134]. These mutations were also found in several myeloid malignancies such as acute myeloid

125

leukemias (AML), myelodysplastic syndrome and myeloproliferative neoplasms [135]. The most commonly found mutations, R132 of IDH1 and R140 and R172 of IDH2, are heterozygous in nature implying a gain of function in these mutant enzymes. The normal function of these enzymes is the $NADP^+$ dependent conversion of isocitrate to α-ketoglutarate (α-KG). The novel enzymatic property of IDH1 and IDH2, as a result of mutations, allows them to convert α-KG to 2-hydroxyglutarate (2-HG) in a NADPH dependent fashion [136, 137]. 2-HG is a structural analog of α-KG, differing in the presence of a hydroxyl group for a keto group, and thus can act as an inhibitor of Tet1 and Tet2 enzymes, both of which use α-KG as a substrate for conversion of 5mC to 5hmC [138]. Indeed, AML patients characterized with IDH1/IDH2 mutations, exhibit reduced Tet2 enzyme activity combined with low 5hmC levels in their nuclear genome and display a hypermethylator phenotype [139, 140]. Loss of function genetic mutations have also been identified in the Tet2 gene in AML and other myeloid malignancies [141]. Interestingly, these mutations are mutually exclusive to the IDH1/2 mutations [140]. These results indicate that IDH1/2 mutations and Tet2 mutations have redundant roles in AML disease phenotype.

IDH1/IDH2 mutants can be used to verify the role of Tet enzymes in 5hmC generation in the mitochondria. 2-HG is generated in both mitochondria and the cytosplasm, and has been shown to affect 5hmC generation in the nucleus [140]. Whether it can affect 5hmC in mtDNA in a similar fashion remains to be seen. Hannah Gardner, a former summer student in the lab, has successfully over-expressed wild type and mutant forms of IDH1 and IDH2 in HCT116 cells. mtDNA will be isolated from these cells and analyzed for 5hmC content using MeDIP and HPLC-MS. Reduction in 5hmC content in mtDNA, with simultaneous accumulation of 2-HG in mitochondria of cells expressing the mutant versus wild type forms, will indicate inhibition of Tet-catalyzed generation of 5hmC in the mitochondria. This inhibitory effect can then be rescued

by co-expressing L-2-Hydroxyglutarate dehydrogenase (L2HGDH) and D-2-hydroxyglutarate dehydrogenase (D2HGDH) genes, known to convert 2-HG back to α-KG. Both IDH2 and D2HGDH primarily localize to the mitochondria and thus the effect of mutant IDH2 of 5hmC in mtDNA and its rescue by D2HGDH will be most crucial.

The fact that mutations in IDH2, the mitochondrial homolog, and Tet2 are mutually exclusive in AML patient samples with no such mutations found in either Tet1 or Tet3 genes, hints at potential role of Tet2 in 5hmC generation in the mitochondria. However, the effect of 2-HG accumulation on the activity of Tet1 or Tet3 cannot be excluded. The above mentioned experiments will only prove that 5hmC generation in mtDNA is Tet mediated. Therefore, to define the specific Tet enzyme involved, analysis of 5hmC content in mtDNA from Tet1, Tet2 or Tet3 null cells will need to be performed. It is also possible that more than one Tet enzyme may be translocated to the mitochondria and therefore may be involved in conversion to 5hmC. These studies are now being actively pursued by John Strang, a current graduate student in the lab.

Potential roles of 5hmC in the mitochondria:

There has been no evidence as yet about the functional role of 5hmC in mtDNA. Hence, one can only speculate about the potential roles of 5hmC in mtDNA. As mentioned in Chapter 1, 5hmC can form an important intermediate in the process of demethylation. Upon conversion of 5mC to 5hmC/5fC/5caC by Tet enzymes, modified cytosines undergo demethylation following deamination and base excision repair. Mitochondria are characterized by high ROS production, which in turn induces damage in both free nucleotides and double stranded mtDNA. In addition, POLG is intrinsically error prone with an approximately 10 fold higher error rate than nuclear DNA polymerase. Mitochondria utilize a base excision repair pathway, similar to the nucleus, to

remove ROS-induced DNA damage as well as POLG-induced errors from mtDNA. If 5hmC is generated in mtDNA from 5mC, it could serve as an intermediate in the mitochondrial base excision repair mechanism and thus, would be expected to play an important role in mtDNA demethylation. This can, in effect, minimize effects of ROS induced DNA damage and will help maintain integrity of mtDNA sequence.

Nuclear chromatin is characterized by CpG islands and CpG sparse regions, both of which show contrasting patterns of methylation and hydroxymethylation. CpG islands are enriched for cytosine hydroxymethylation with low amounts of cytosine methylation with the opposite being true for CpG sparse regions [15, 42]. Thus, it seems like hydroxymethylation of cytosine residues at CpG islands prevents hypermethylation of the promoter regions and thereby maintains DNA methylation fidelity. In contrast to the nucleus, mtDNA lacks CpG islands and therefore methylation or hydroxymethylation at specific residues is likely to be more important in the functional role of these cytosine modifications. Site specific hydroxymethylation could play a role in maintaining fidelity of methylation in mtDNA, which in turn might be crucial for maintaining normal function of mitochondrial processes. Using oxidative bisulphite sequencing, methylation versus hydroxymethylation of specific sites in mtDNA could be determined which can then be tested for their functional significance. Proteins that bind to 5mC do not bind to 5hmC [43]. Identification of proteins that differentially bind methylated versus hydroxymethylated cytosine residues in mammalian mitochondria will give us insights into significance of site specific modification of cytosine residues.

Site specific hydroxymethylation in the D-loop region could be important for regulation of transcription of mitochondrial genes. 5hmC may affect recruitment of core transcription

machinery, TFB2M and POLRMT, and binding of other transcription factors such as TFAM towards methylated versus hydroxymethylated D-loop region. It may even simply act to prevent the effects of 5mC on mitochondrial transcription, as has been shown for the nanog promoter in the nucleus [30]. Once we understand how methylation affects mitochondrial transcription, as proposed before, *in vitro* transcription studies using defined modified substrates can be performed to understand the role of 5hmC in regulation of mitochondrial transcription.

We have yet to prove whether 5hmC is generated from 5mC and if Tet enzymes play a role in this conversion. If any of the Tet enzyme family members are found to generate 5hmC in mtDNA, it will be interesting to determine if mtDNMT1 and the putative mitochondrial Tet enzyme interact with each other and therefore, act in concert to maintain a balance between 5mC and 5hmC in mtDNA. In contrast, they may also compete for the same recognition sites to bring about modification of cytosines, which will prevent hypermethylation of mtDNA.

5hmC levels as a potential biomarker:

Tet enzymes have been associated with different types of leukemia in the past. Tet1 was identified as a MLL fusion partner, resulting from translocation between chromosomes 10 and 11, in acute myeloid and lymphocytic leukemias (AML and ALL) [142, 143]. As described previously, Tet2 is often mutated in AML, myelodysplastic and myeloproliferative leukemia. IDH1/2 mutations in AML mediate their effect by inhibiting the activity of Tet enzymes [140, 144]. This implies that Tet enzymes act as tumor suppressors and that disruption of Tet function leads to oncogenesis. Since then numerous studies have been directed towards quantitation of 5hmC levels in various normal versus cancer tissues and cells. It was found that cancer cells/tissues show reduced Tet activity and therefore have low levels of 5hmC compared to normal

counterparts [145-147]. It was also shown that 5hmC levels are reduced in melanoma [148]. This implies that reduced 5hmC levels can correspond to type and severity of various tumors and can serve as a potential biomarker for cancer prognosis.

Various neurodegenerative diseases such as Parkinsons's and Alzheimer's disease as well as other ischemic heart disease, diabetes and cancer are all characterized by mitochondrial dysfunction combined with increased ROS production [149, 150]. It is not yet known if mitochondrial epigenetics play a causative role in these diseases. However, tissue samples from these patients can be compared to tissues from normal age-matched controls by quantitation of nuclear as well as mitochondrial 5hmC levels. These studies can help establish a correlation between 5hmC and various disease states, thus allowing the use of 5hmC levels as a potential biomarker for diagnosis and prognosis. Further studies need to be performed to elucidate mechanisms inhibiting Tet activity in cancer cells. Inhibitors/drugs can then be designed to restore Tet activity and therefore 5hmC levels in cancer cells which in turn might restore the normal phenotype.

Summary and Conclusions:

Mitochondria are essential organelles that perform a wide variety of functions including ATP production, apoptosis, thermogenesis, Ca^{2+} storage, carbohydrate, lipid and protein metabolism [48]. Mitochondria contain their own genetic material, but encode only a small fraction of proteins present in the mitochondria [85]. Most of the proteins essential for a variety of mitochondrial functions are coded in the nucleus and are transported to the mitochondria [86]. Studies performed decades ago detected the presence of cytosine modification in the mitochondria. Our studies were aimed at validating presence of cytosine modification, 5mC in mtDNA, identifying the

mechanism involved in its generation and establishing its functional role in mammalian mitochondria. We demonstrated that the newly identified epigenetic modification, 5hmC, was present in mtDNA. We further attempted to define nuclear enzymes that may be transported to the mitochondrial and hence may be responsible for generation of 5hmC in mtDNA.

Studies performed decades ago identified the presence of 5-methylcytosine but failed to detect any DNA methylase activity in mitochondrial fraction. Our laboratory pursued the possibility that a nuclear enzyme is responsible for mitochondrial DNA methylation. Cell fractionation combined with immunoblotting, identified a DNMT1 isoform that is translocated to the mitochondria. Using the MeDIP assays, we validated the presence of 5mC in mtDNA, obtaining 10-20 fold enrichment of mtDNA with 5mC antibody over IgG. Analogous to its role in the nucleus, we studied the effect of altered mtDNMT1 levels on mitochondrial transcription and found that upregulation of mtDNMT1, due to loss of p53, altered mitochondrial transcription. These studies have established the functional significance of mtDNMT1 in the mitochondria.

Using a hydroxy-MeDIP assay, we show for the first time that 5hmC is present in mtDNA, as evidenced by 38-580 fold higher enrichment of mtDNA with 5hmC antibody over that of IgG. This finding was confirmed using an alternative site specific assay for 5hmC. We next asked how the 5hmC modification was generated in mtDNA. Tet enzymes have been shown to catalyze the formation of 5hmC from 5mC in the nucleus. We considered the possibility that at least one of the three Tet enzymes was present in the mitochondria. We optimized an in vitro Tet enzyme assay to detect Tet-like enzyme activity in crude as well as percoll purified, trypsin treated mitochondrial fractions from HCT116 cells. Mitoprot analysis predicted that, Tet1 enzyme

starting at ATG3 in human and ATG4 in mouse would be translocated to the mitochondria. Upon further investigation, we found a band corresponding to full length Tet1 in immunoblots containing mitochondrial fractions from mouse embryonic stem cells. This band, however, was not protected from trypsin treatment of intact mitochondria indicating that Tet1 was not transported to the mitochondrial matrix. The MTS predicted for Tet1 by Mitoprot software, although conserved in human and mouse, did not carry heterologous protein GFP to the mitochondria. Immunofluorescence using Tet1 antibody also failed to show localization with mitotracker red dye. We attempted to study the role of Tet1in 5hmC generation at a functional level, by performing hydroxyl-MeDIP assay on cells containing either partial or complete knock-out of Tet1. Knock-out of Tet1 did not alter 5hmC signal in mtDNA. Overall, these results did not allow us to ascertain the role of Tet1 in 5hmC generation in mtDNA.

Overall, our data has identified some of the key players and epigenetic mechanisms that exist in mammalian mitochondria. We propose that these mitochondrial epigenetic mechanisms are one of the many processes that underlie nuclear-mitochondrial cross talk. Furthermore, these mechanisms are involved in regulation of several mitochondrial processes including mitochondrial transcription, disruption of which may cause mitochondrial dysfunction in a broad spectrum of disorders such as neurodegerative diseases, auto immune diseases, ischemic heart disease, diabetes and cancer. Estimation of nuclear and mitochondrial 5hmC levels, in each of the disease samples in comparison to normal, will allow use of 5hmC as a biomarker in diagnosis and prognosis.

References

1. Egger G, Liang G, Aparicio A, Jones P. **Epigenetics in human disease and prospects for epigenetic therapy**. Nature 2004; 429:457-63.

2. En L. **Chromatin modification and epigenetic reprogramming in mammalian development**. Nature Review Genetics 2002; 3:662-73.

3. Sharma S, Kelly T, Jones P. **Epigenetics in cancer**. Carcinogenesis 2010; 31:27-36.

4. Kouzarides T. **Chromatin modifications and their function**. Cell 2007; 128:693-705.

5. Saha A, Wittmeyer J, Cairns B. **Chromatin remodelling: The industrial revolution of DNA around histones**. Nature Reviews 2006; 7:437-47.

6. Hirschhorn J, Brown S, Clark C, Winston F.
Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin s tructure. Genes and Development 1992; 6:2288-98.

7. Kornberg R. **Chromatin structure: A repeating unit of histones and DNA.**. Science 1974; 184:868-71.

8. Robertson K, Wolffe A. **DNA methylation in health and disease**. Nature Review Genetics 2000; 1:11-9.

9. Nan X, Meehan R, Bird A. **Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2**. Nucleic Acids Research 1993; 21:4886-92.

10. Ng H, Zhang Y, Hendrich B, Johnson C, Turner B, Erdjument-Bromage H, et al. **MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex.**. Nature Genetics 1999; 23:58-61.

11. Fuks F. **DNA methylation and histone modifications: Teaming up to silence genes**
 . Current Opinion in Genetics and Development 2005; 15:1-6.

12. Lehnertz B, Ueda Y, Derijck A, Braunschweig U, Perez-Burgos L, Kubicek S, et al. **Suv39h-mediated Histone H3 Lysine 9 Methylation Directs DNA Methylation to Major Satellite Re peats atPericentric Heterochromatin**. Current Biology 2003; 13:1192-200.

13. Laird P. **The power and the promise of DNA methylation markers.**. Nature Reviews Cancer 2003; 3:253-66.

14. Takai D, Jones P. **Comprehensive analysis of CpG islands in human chromosomes 21 and 22**. Proceedings of the National Academy of Sciences 2002; 99:3740-5.

15. Okano M, Bell D, Haber D, Li E. **DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development.**. Cell 1999; 99:247-57.

16. Monk M, Boubelik M, Lehnert S. **Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ celllineages during mouse embryo development**. Development 1987; 99:371-82.

17. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. **Demethylation of the zygotic paternal genome.**. Nature 2000; 403:501-2.

18. Carlson L, Page A, Bestor T. **Properties and localization of DNA methyltransferase in preimplantation mouse embryos: Implications for genomic imprinting.**. Genes and Development 1992; 6:2536-41.

19. Li E, Bestor T, Jaenisch R. **Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.**. Cell 1992; 69:915-26.

20. Taylor S. **p53 and deregulation of DNA methylation in cancer**. Cell Science Reviews 2006; 2:1742.

21. Esteller M. **Cancer epigenomics: DNA methylomes and histone-modification maps.**. Nature Review Genetics 2007; 8:286-98.

22. Cunningham J, Christensen E, Tester D, Kim C, Roche P, Burgart L, et al. **Hypermethylation of the *hMLH1* Promoter in colon cancer with microsatellite instability**. Cancer Research 1998; 58:3455-60.

23. Iyer L, Tahiliani M, Rao A, Aravind L. **Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids**. Cell cycle 2009; 8:1698-710.

24. Tahiliani M, Koh K, Shen Y, Pastor W, Bandukwala H, Brudno Y, et al. **Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1.**. Science 2009; 324:930-5.

25. Leeuwen F, Kleft R, Cross M, Borst P. **Biosynthesis and function of the modified DNA base b-D-glucosyl-hydroxymethyluracil in trypanosoma brucei**. 1998 Molecular and Cellular Biology; 18:5643-51.

26. Yu Z, Genest A, Riet B, Sweeney K, DiPaolo C, Kieft R, et al. **The protein that binds to DNA base J in trypanosomatids has features of a thymidine hydroxylase**. Nucleic Acids Research 2007; 35:2107-15.

27. Borst P, Sabatini R. **Base J: Discovery, biosynthesis, and possible functions**. Annual Review of Microbiology 2008; 62:235-51.

28. Aravind L, Koonin E. **The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases**. Genome Biology 2001; 2:research 0007.1,research 0007.8.

29. Zhang H, Zhang X, Clark E, Mulcahey M, Huang S, Shi G. **TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine**. Cell Research 2010; 20:1390-3.

30. Ito S, D'Alessio A, Taranova O, Hong K, Sowers L, Zhang Y. **Role of tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification**. Nature 2010; 466:1129-33.

31. Yu M, Hon G, Szulwach K, Song C, Jin P, Ren B, et al. **Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine.**. Nature Protocols 2012; 7:2159-70.

32. Ito S, Shen L, Dai Q, Wu S, Collins L, Swenberg J, et al. **Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine.**. Science 2011; 333:1300-3.

33. Frommer M, McDonald L, Millar D, Collis C, Watt F, Grigg G, et al. **A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands**. Proceedings of the National Academy of Sciences 1992; 89:1827-31.

34. Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A. **The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing**. PLoS one 2010; 5:e8888.

35. Hayatsu H, Wataya Y, Kai K, Iida S. **Reaction of sodium bisulfite with uracil, cytosine, and their derivatives.**. Biochemistry 1970; 9:2858-65.

36. Booth M, Branco M, Ficz G, Oxley D, Krueger F, Reik W, et al. **Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution.**. Science 2012; 336:934-7.

37. Hashimoto H, Liu Y, Upadhyay A, Chang Y, Howerton S, Vertino P, et al. **Recognition and potential mechanisms for replication and erasure of cytosine hydroxymeth ylation.**. Nucleic Acids Research 2012; 40:4841-9.

38. Valinluck V, Tsai H, Rogstad D, Burdzy A, Bird A, Sowers L. **Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) ofmethyl-CpG binding protein 2 (MeCP2).**. Nucleic Acids Research 2004; 32:4100-8.

39. Szwagierczak A, Bultmann S, Schmidt C, Spada F, Leonhardt H. **Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA**. Nucleic Acids Research 2010; 38.

40. Kriaucionis S, Heintz N. **The nuclear DNA base 5-hydroxymethylcytosine is present in purkinje neurons and the brain.**. Science 2009; 324:929-30.

41. Baylin S, Esteller M, Rountree M, Bachman K, Schuebel K, Herman J. **Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer.**. Human Molecular Genetics 2001; 10:687-92.

42. Williams K, Christensen J, Pedersen M, Johansen J, Cloos P, Rappsilber J, et al. **TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity.**. Nature 2011; 473:343-8.

43. Valinluck V, Sowers L. **Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1.**. Cancer Research 2007; 67:946-50.

44. Wu H, D'Alessio A, Ito S, Xia K, Wang Z, Cui K, et al. **Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells.**. Nature 2011; 473:389-93.

45. Surani M, Hayashi K, Hajkova P. **Genetic and epigenetic regulators of pluripotency.**. Cell 2007; 128:747-62.

46. Guo J, Su Y, Zhong C, Ming G, Song H. **Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain.**. Cell 2011; 145:423-34.

47. He Y, Li B, Li Z, Liu P, Wang Y, Tang Q, et al. **Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA.**. Science 2011; 333:1303-7.

48. McBride H, Neuspiel M, Wasiak S. **Mitochondria: More than just a powerhouse.**. Current Biology 2006; 16:551-60.

49. Perkins G, Frey T. **Recent structural insight into mitochondria gained by microscopy**. Micron 2000; 31:97-111.

50. Frey T, Mannella C. **The internal structure of mitochondria**. Trends in biochemical sciences 2000; 25:319-24.

51. Krauss S. **Mitochondria: Structure and role in respiration**. Encyclopedia of Life Sciences 2001:1-6.

52. Papa S, Martino P, Capitanio G, Gaballo A, De RD, Signorile A, et al. **The oxidative phosphorylation system in mammalian mitochondria.**. Advances in experimental medicine and biology 2012; 942:3-37.

53. Mokranjac D, Neupert W. **Protein import into mitochondria.**. Biochemical society transactions 2005; 33:1019-23.

54. Smeitink J, van den Heuvel L, DiMauro S. **The genetics and pathology of oxidative phosphorylation**. Nature Review Genetics 2001; 2:342-52.

55. Falkenberg M, Larsson N, Gustafsson C. **DNA replication and transcription in mammalian mitochondria.**. Annual Review of Biochemistry 2007; 76:679-99.

56. Iborra F, Kimura H, Cook P. **The functional organization of mitochondrial genomes in human cells.**. BMC biology 2004; 2.

57. Alam T, Kanki T, Muta T, Ukaji K, Abe Y, Nakayama H, et al. **Human mitochondrial DNA is packaged with TFAM**. Nucleic Acids Research 2003; 31:1640-5.

58. Wang Y, Bogenhagen D. **Human mitochondrial DNA nucleoids are linked to protein folding machinery and metabolic enzymes at themitochondrial inner membrane.**. Journal of Biological Chemistry 2006; 281:25791-802.

59. Bogenhagen D, Rousseau D, Burke S. **The layered structure of human mitochondrial DNA nucleoids[*]**. Journal of Biological Chemistry 2008; 283:3665-75.

60. Berk A, Clayton D. **Mechanism of mitochondrial DNA replication in mouse L-cells: asynchronous replication of strands,segregation of circular daughter molecules, aspec ts of topology and turnover of an initiation sequence.**. Journal of Molecular Biology 1974; 86:801-24.

61. Kmiec B, Woloszynska M, Janska H. **Heteroplasmy as a common state of mitochondrial genetic information in plants and animals**. Curr Genet 2006; 50:149-59; PMID: 16763846; DOI: 10.1007/s00294-006-0082-1.

62. Hauswirth WW, Laipis PJ. **Mitochondrial DNA polymorphism in a maternal lineage of holstein cows**. Proc Natl Acad Sci U S A 1982; 79:4686-90; PMID: 6289312.

63. Koehler CM, Lindberg GL, Brown DR, Beitz DC, Freeman AE, Mayfield JE, et al. **Replacement of bovine mitochondrial DNA by a sequence variant within one generation**. Genetics 1991; 129:247-55; PMID: 1682213.

64. Giles RE, Blanc H, Cann HM, Wallace DC. **Maternal inheritance of human mitochondrial DNA**. Proc Natl Acad Sci U S A 1980; 77:6715-9; PMID: 6256757.

65. Thompson WE, Ramalho-Santos J, Sutovsky P. **Ubiquitination of prohibitin in mammalian sperm mitochondria: Possible roles in the regulation of mitochondrial inheritance and sperm quality control.** Biol Reprod 2003; 69:254-60; PMID: 12646488; DOI: 10.1095/biolreprod.102.010975.

66. Schwartz M, Vissing J. **Paternal inheritance of mitochondrial DNA**. N Engl J Med 2002; 347:576-80; PMID: 12192017; DOI: 10.1056/NEJMoa020350.

67. Bopenhagen D. **Repair of mtDNA in vertebrates**. American journal of Human genetics 1999; 64:1276-81.

68. Ashley MV, Laipis PJ, Hauswirth WW. **Rapid segregation of heteroplasmic bovine mitochondria**. Nucleic Acids Res 1989; 17:7325-31; PMID: 2798094.

69. Wallace DC, Singh G, Lott MT, Hodge JA, Schurr TG, Lezza AM, et al. **Mitochondrial DNA mutation associated with leber's hereditary optic neuropathy**. Science 1988; 242:1427-30; PMID: 3201231.

70. Goto Y, Nonaka I, Horai S. **A mutation in the tRNA(leu)(UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies**. Nature 1990; 348:651-3; PMID: 2102678; DOI: 10.1038/348651a0.

71. Van den Ouweland JM, Lemkes HH, Ruitenbeek W, Sandkuijl LA, de Vijlder MF, Struyvenberg PA, et al. **Mutation in mitochondrial tRNA(leu)(UUR) gene in a large pedigree with maternally transmitted type II diabetes mellitus and deafness**. Nat Genet 1992; 1:368-71; PMID: 1284550; DOI: 10.1038/ng0892-368.

72. Park C, Larsson N. **Mitochondrial DNA mutations in disease and aging.**. The Journal of Cell Biology 2011; 193:809-18.

73. Montoya J, Christianson T, Levens D, Rabinowitz M, Attardi G. **Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA.**. Proceedings of the National Academy of Sciences 1982; 79:7195-9.

74. Martin M, Cho J, Cesare A, Griffith J, Attardi G. **Termination factor-mediated DNA loop between termination and initiation sites drives mitochondrial rRNA synthesis.**. Cell 2005; 123:1227-40.

75. 'jala D, Montoya J, Attardi G.
tRNA punctuation model of RNA processing in human mitochondria.. Nature 1981; 290:470-4.

76. Shutt T, Lodeiro M, Cotney J, Cameron C, Shadel G.
Core human mitochondrial transcription apparatus is a regulated two-component system in vitro.. Proceedings of the National Academy of Sciences 2010; 107:12133-8.

77. Shi Y, Dierckx A, Wanrooij P, Wanrooij S, Larsson N, Wilhelmsson L, et al.
Mammalian transcription factor A is a core component of

the mitochondrial transcription machinery.. Proceedings of the National Academy of Sciences 2012; 109:16510-165151.

78. Masters B, Stohl L, Clayton D. **Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7.**. Cell 1987; 51:89-99.

79. Gaspari M, Falkenberg M, Larsson N, Gustafsson C. **The mitochondrial RNA polymerase contributes critically to promoter specificity in mammalian cells**. EMBO 2004; 23:4606-14.

80. Falkenberg M, Gaspari M, Rantanen A, Trifunovic A, Larsson N, Gustafsson C. **Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA.**. Nature Genetics 2002; 31:289-94.

81. Dairaghi D, Shadel G, Clayton D. **Addition of a 29 residue carboxyl-terminal tail converts a simple HMG box-containing protein into a transcriptional activator.**. Journal of Molecular Biology 1995; 249:11-28.

82. Ohgaki K, Kanki T, Fukuoh A, Kurisaki H, Aoki Y, Ikeuchi M, et al. **The C-terminal tail of mitochondrial transcription factor a markedly strengthens its general binding to DNA.**. Journal of Biochemistry 2007; 141:201-11.

83. Fisher R, Lisowsky T, Parisi M, Clayton D. **DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein.**. Journal of Biological Chemistry 1992; 267:3358-67.

84. Kruse B, Narasimhan N, Attardi G. **Termination of transcription in human mitochondria: Identification and purification of a DNA binding protein factor that promotes termination.**. Cell 1989; 58:391-7.

85. Lopez M, Kristal B, Chernokalskaya E, Lazarev A, Shestopalov A, Bogdanova A, et al. **High-throughput profiling of the mitochondrial proteome using affinity fractionation and automation.**. Electrophoresis 2000; 21:3427-40.

86. Attardi G, Schatz G. **Biogenesis of mitochondria.**. Annual Review of Cell Biology 1988; 4:289-333.

87. Hartl F, Pfanner N, Nicholson D, Neupert W. **Mitochondrial protein import.**. Biochimica et Biophysica acta 1989; 988:1-45.

88. Vögtle F, Wortelkamp S, Zahedi R, Becker D, Leidhold C, Gevaert K, et al. **Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability.**. Cell 2009; 139:428-39.

89. Schatz G. **Signals guiding proteins to their correct locations in mitochondria**. European Journal of Biochemistry 1987; 165:1-6.

90. Lee C, Sedman J, Neupert W, Stuart R. **The DNA helicase, Hmi1p, is transported into mitochondria by a C-terminal cleavable targeting signal.**. Journal of Biological Chemistry 1999; 274:20937-42.

91. Stan T, Brix J, Schneider-Mergener J, Pfanner N, Neupert W, Rapaport D. **Mitochondrial protein import: Recognition of internal import signals of BCS1 by the TOM complex.**. Molecular and Cellular Biology 2003; 23:2239-50.

92. Neupert W. **Protein import into mitochondria.**. Annual Review of Biochemistry 1997; 66:917.

93. Wu Y, Sha B. **Crystal structure of yeast mitochondrial outer membrane translocon member Tom70p.**. Nature structural and molecular biology 2006; 13:589-93.

94. Abe Y, Shodai T, Muto T, Mihara K, Torii H, Nishikawa S, et al. **Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20.**. Cell 2000; 100:551-60.

95. Ahting U, Thieffry M, Engelhardt H, Hegerl R, Neupert W, Nussberger S. **Tom40, the pore-forming component of the protein-conducting TOM channel in the outer membrane ofmitochondria.**. The Journal of Cell Biology 2001; 153:1151-60.

96. Mokranjac D, Neupert W. **The many faces of the mitochondrial TIM23 complex.**. Biochimica et Biophysica acta 2010; 1797:1045-54.

97. Yamamoto H, Esaki M, Kanamori T, Tamura Y, Nishikawa S, Endo T. **Tim50 is a subunit of the TIM23 complex that links protein translocation across the outer and inner mitochondrial membranes**. Cell 2002; 111:519-28.

98. Milisav I, Moro F, Neupert W, Brunner M. **Modular structure of the TIM23 preprotein translocase of mitochondria.**. Journal of Biological Chemistry 2001; 276:25856-61.

99. Schiller D, Cheng Y, Liu Q, Walter W, Craig E. **Residues of Tim44 involved in both association with the translocon of the inner mitochondrial membrane and regulation of mitochondrial Hsp70 tethering.**. Molecular and Cellular Biology 2008; 28:4424-33.

100. Ungermann C, Neupert W, Cyr D. **The role of Hsp70 in conferring unidirectionality on protein translocation into mitochondria.**. Science 1994; 266:1250-3.

101. Nisoli E, Clementi E, Paolucci C, Cozzi V, Tonello C, Sciorati C, et al. **Mitochondrial biogenesis in mammals: The role of endogenous nitic oxide**. Science 2003; 299:896-9.

102. Wu H, Kanatous S, Thurmond F, Gallardo T, Isotani E, Bassel-Duby R, et al. **Regulation of mitochondrial biogenesis in skeletal muscle by CaMK**. Science 2002; 296:349-52.

103. Zong H, Ren J, Young L, Pypaert M, Mu J, Birnbaum M, et al. **AMP kinase is required for mitochondrial biogenesis in skeletal muscle in response to chronic energydeprivation.**. Proceedings of the National Academy of Sciences 2002; 99:15983-7.

104. Nisoli E, Falcone S, Tonello C, Cozzi V, Palomba L, Fiorani M, et al. **Mitochondrial biogenesis by NO yields functionally active mitochondria in mammals.**. Proceedings of the National Academy of Sciences 2004; 101:16507-12.

105. Cao W, Daniel K, Robidoux J, Puigserver P, Medvedev A, Bai X, et al. **p38 mitogen-activated protein kinase is the central regulator of cyclic AMP-dependent transcription of the brown fat uncoupling protein 1 gene.**. Molecular and Cellular Biology 2004; 24:3057-67.

106. Thonberg H, Fredriksson J, Nedergaard J, Cannon B. **A novel pathway for adrenergic stimulation of cAMP-response-element-binding protein (CREB) phosphorylation: Mediation via alpha1-adrenoceptors and protein kinase C activation.**. The Biochemical Journal 2002; 364:73-9.

107. Gulick T, Cresci S, Caira T, Moore D, Kelly D. **The peroxisome proliferator-activated receptor regulates mitochondrial fatty acid oxidative enzyme gene expression.**. Proceedings of the National Academy of Sciences 1994; 91:11012-6.

108. Scarpulla R. **Transcriptional paradigms in mammalian mitochondrial biogenesis and function.**. Physiological Reviews 2007; 88:611-38.

109. Biswas G, Adebanjo O, Freedman B, Anandatheerthavarada H, Vijayasarathy C, Zaidi M, et al. **Retrograde Ca2+ signaling in C2C12 skeletal myocytes in response to mitochondrial genetic and metabolic stress: A novel mode of inter-organelle crosstalk.**. EMBO 1999; 18:522-33.

110. Nass M. **Differential methylation of mitochondrial and nuclear DNA in cultured mouse, hamster and virus-transformed hamster cells. in vivo and in vitro methylation.**. Journal of Molecular Biology 1973; 80:155-75.

111. Pollack Y, Kasir J, Shemer R, Metzger S, Szyf M. **Methylation pattern of mouse mitochondrial DNA.**. Nucleic Acids Research 1984; 12:4811-24.

112. Cardon L, Burge C, Clayton D, Karlin S. **Pervasive CpG suppression in animal mitochondrial genomes.**. Proceedings of the National Academy of Sciences 1994; 91:3799-803.

113. Cooper D, Krawczak M. **Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes.**. Human Genetics 1989; 83:181-8.

114. Elango N, Kim S, Vigoda E, Yi S. **Mutations of different molecular origins exhibit contrasting patterns of regional substitution rate variation**. PLoS computational biology 2008; 4.

115. Peterson E, Bögler O, Taylor S. **p53-mediated repression of DNA methyltransferase 1 expression by specific DNA binding**. Cancer Research 2003; 63:6579-82.

116. McCabe M, Davis J, Day M. **Regulation of DNA methyltransferase 1 by the pRb/E2F1 pathway.**. Cancer Research 2005; 65:3624-32.

117. Shock L, Thakkar P, Peterson E, Moran R, Taylor S. **DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria.**. Proceedings of the National Academy of Sciences 2011; 108:3630-5.

118. Jones P, Baylin S. **The epigenomics of cancer**. Cell 2007; 128:683-92.

119. Lebedeva M, Eaton J, Shadel G. **Loss of p53 causes mitochondrial DNA depletion and altered mitochondrial reactive oxygen species homeostasis.**. Biochimica et Biophysica acta 2009; 1787:328-34.

120. Kulawiec M, Ayyasamy V, Singh K. **p53 regulates mtDNA copy number and mitocheckpoint pathway**. Journal of Carcinogenesis 2009; 8.

121. Kozak M. **An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs**. Nucleic Acids Research 1987; 15:8125-48.

122. Lodeiro M, Uchida A, Bestwick M, Moustafa I, Arnold J, Shadel G, et al. **Transcription from the second heavy-strand promoter of human mtDNA is repressed by transcription factorA in vitro.**. Proceedings of the National Academy of Sciences 2012; 109:6513-8.

123. Matsunaga M, Jaehning J. **Intrinsic promoter recognition by a "core" RNA polymerase.**. Journal of Biological Chemistry 2004; 279:44239-42.

124. Dawlaty M, Ganz K, Powell B, Hu Y, Markoulaki S, Cheng A, et al. **Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development.**. Cell Stem cell 2011; 9:166-75.

125. Liutkeviciute Z, Lukinavicius G, Masevicius V, Daujotyte D, Klimasauskas S. **Cytosine-5-methyltransferases add aldehydes to DNA.**. Nat Chem Biol 2009; 5:400-2.

126. Dzitoyeva S, Chen H, Manev H. **Effect of aging on 5-hydroxymethylcytosine in brain mitochondria**. Neurobiol Aging 2012; 33:2881-91.

127. Wyatt G, Cohen S. **A new pyrimidine base from bacteriophage nucleic acids.**. Nature 1952; 170:1072-3.

128. Anderson D, Quintero C, Stover P. **Identification of a de novo thymidylate biosynthesis pathway in mammalian mitochondria.**. Proceedings of the National Academy of Sciences 2011; 108:15163-8.

129. Claros M, Vincens P. **Computational method to predict mitochondrially imported proteins and their targeting sequences.**. European Journal of Biochemistry 1996; 241:779-86.

130. Li Z, Cai X, Cai C, Wang J, Zhang W, Petersen B, et al. **Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies.**. Blood 2011; 118:4509-18.

131. Chen T, Hevi S, Gay F, Tsujimoto N, He T, Zhang B, et al. **Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells.**. Nature Genetics 2007; 39:391-6.

132. Parsons D, Jones S, Zhang X, Lin J, Leary R, Angenendt P, et al. **An integrated genomic analysis of human glioblastoma multiforme**. Science 2008; 321:1807-12.

133. Hartmann C, Meyer J, Balss J, Capper D, Mueller W, Christians A, et al. **Type and frequency of IDH1 and IDH2 mutations are related to astrocytic and oligodendroglial differentiation and age: A study of 1,010 diffuse gliomas**. Acta Neuropathologica 2009; 118:469-74.

134. Yan H, Parsons D, Jin G, McLendon R, Rasheed B, Yuan W, et al. **IDH1 and IDH2 mutations in gliomas.**. New Engalnd Journal of Medicine 2009; 360:765-73.

135. Ko M, Huang Y, Jankowska A, Pape U, Tahiliani M, Bandukwala H, et al. **Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2.**. Nature 2010; 468:839-43.

136. Dang L, White D, Gross S, Bennett B, Bittinger M, Driggers E, et al. **Cancer-associated IDH1 mutations produce 2-hydroxyglutarate.**. Nature 2009; 462:739-44.

137. Ward P, Patel J, Wise D, Abdel-Wahab O, Bennett B, Coller H, et al. **The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate.**. Cancer Cell 2010; 17:225-34.

138. Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim S, et al. **Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases.**. Cancer Cell 2011; 19:17-30.

139. Jin G, Reitman Z, Duncan C, Spasojevic I, Gooden D, Rasheed B, et al. **Disruption of wild-type IDH1 suppresses D-2-hydroxyglutarate production in IDH1-mutated gliomas.**. Cancer Research 2013; 73:496-501.

140. Figueroa M, Abdel-Wahab O, Lu C, Ward P, Patel J, Shih A, et al. **Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, andimpair hematopoietic differentiation.**. Cancer Cell 2010; 18:553-67.

141. Abdel-Wahab O, Mullally A, Hedvat C, Garcia-Manero G, Patel J, Wadleigh M, et al. **Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies**. Blood 2009; 114:144-7.

142. Lorsbach R, Moore J, Mathew S, Raimondi S, Mukatira S, Downing J. **TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23)**. Leukemia 2003; 17:637-41.

143. Lee S, Cho S, Kim M, Oh S, Cho E, Lee S, et al. **Genomic breakpoints and clinical features of MLL-TET1 rearrangement in acute leukemias**. Haemotologica 2012.

144. Lu C, Ward P, Kapoor G, Rohle D, Turcan S, Abdel-Wahab O, et al. **IDH mutation impairs histone demethylation and results in a block to cell differentiation**. Nature 2012; 483:474-8.

145. Jin S, Jiang Y, Qiu R, Rauch T, Wang Y, Schackert G, et al. **5-hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1mutations.**. Cancer Research 2011; 71:7360-5.

146. Li W, Liu M. **Distribution of 5-hydroxymethylcytosine in different human tissues.**. Journal of Nucleic acids 2011; 2011.

147. Yang H, Liu Y, Bai F, Zhang J, Ma S, Liu J, et al. **Tumor development is associated with decrease of TET gene expression and 5-methylcytosinehydroxylation.**. Oncogene 2013; 32:663-9.

148. Lian C, Xu Y, Ceol C, Wu F, Larson A, Dresser K, et al. **Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma.**. Cell 2012; 150:1135-46.

149. Wallace D. **A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: A dawn forevolutionary medicine.**. Annual Review of Genetics 2005; 39:359-407.

150. Pieczenik S, Neustadt J. **Mitochondrial dysfunction and molecular pathways of disease.**. Experimental and Molecular Pathology 2007; 83:84-92.

VITA

Prashant Vijay Thakkar was born on December 4th, 1984 in Ahmednagar, Maharashtra, India. He graduated with a Bachelor of Science (B.Sc.) degree in Chemistry - Vocational biotechnology from Ramnarain Ruia College (University of Mumbai) in 2005. He then went onto pursue his graduate studies at Ahmednagar College (University of Pune). In summer of 2007, he earned his Master of Science (M.Sc.) degree in Biochemistry. In Fall of 2007, he entered the Master's program in Molecular Biology and Genetics and was subsequently transferred into the PhD program in Molecular Biology and Genetics in summer of 2008. Later that same year, he joined the laboratory of Shirley M. Taylor, PhD, where he has been pursuing his graduate research career for the past six years. He has presented his work through posters at the Atlantic Coast Chromatin Conference in October 2010 and the NHLBI Mitochondrial Biology Symposium in April 2011, for which he was presented a travel award, jointly with Lisa Shock.

Manuscripts resulting from present graduate research:-

DNA Methyltransferase-1, cytosine methylation and cytosine hydroxymethylation in mammalian mitochondria.
Lisa S. Shock*, **Prashant V. Thakkar***, Erica Peterson*, Richard G. Moran and Shirley M. Taylor.
*These authors have contributed equally to this work.
Proceedings of the National Academy of Science, 2011 March;108(9):3630-3635