Published in final edited form as:

Gene. 2016 January 10; 575(0): 339–352. doi:10.1016/j.gene.2015.09.003.

Methylation-dependent and independent regulatory regions in the Na,K-ATPase alpha4 gene (*Atp1a4*) may impact its testis-specific expression

Deepti L. Kumar¹, Priya L. Kumar¹, and Paul F. James^{*}

Department of Biology, Miami University, Oxford, OH, United States

Abstract

The $\alpha 4$ Na,K-ATPase is a sperm-specific protein essential for sperm motility and fertility yet little is known about the mechanisms that regulate its expression in germ cells. Here, the potential involvement of DNA methylation in regulating the expression of this sperm-specific protein is explored. A single, intragenic CpG island (M $\alpha 4$ -CGI) was identified in the gene encoding the mouse $\alpha 4$ Na,K-ATPase (*Atp1a4*), which displayed reduced methylation in mouse sperm (cells that contain $\alpha 4$) compared to mouse kidney (tissue that lacks $\alpha 4$ expression). Unlike the intragenic CGI, the putative promoter (the -700 to $+200$ region relative to the transcriptional start site) of *Atp1a4* did not show differential methylation between kidney and sperm nevertheless it did drive methylation-dependent reporter gene expression in the male germ cell line GC-1spg. Furthermore, treatment of GC-1spg cells with 5-Aza2-Deoxycytidine led to upregulation of the $\alpha 4$ transcript and decreased methylation of both the *Atp1a4* promoter and the M $\alpha 4$ -CGI. In addition, *Atp1a4* expression in mouse embryonic stem cells deficient in DNA methyltransferases suggests that both maintenance and *de novo* methylation are involved in regulating its expression. In an attempt to define the regulatory function of the M $\alpha 4$ -CGI, possible roles of the M $\alpha 4$ -CGI in regulating *Atp1a4* expression via methylation-dependent transcriptional elongation inhibition in somatic cells and via its ability to repress promoter activity in germ cells were uncovered. In all, our data suggests that both the promoter and the intragenic CGI could combine to provide multiple modes of regulation for optimizing the *Atp1a4* expression level in a cell type-specific manner.

Keywords

Epigenetics; Dnmts; transcriptional elongation; spermatozoa; motility

^{*}Correspondence to: Paul F. James, PhD, Department of Biology, Miami University, Oxford, OH 45056. jamespf@miamioh.edu, Telephone: 513-529-3129, Fax: 513-529-6900.

¹Current affiliation: Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

The Na,K-ATPase is an integral membrane protein present in the plasma membrane of all higher eukaryotic cells. This transporter moves three Na⁺ out of and two K⁺ into the cell utilizing the energy from the hydrolysis of each molecule of ATP, thus maintaining electrochemical gradients across the plasma membrane (Lingrel et al., 2003). The activity of the Na,K-ATPase is essential for many cellular processes including establishment of resting membrane potential, excitability of muscle and nerve, regulation of osmotic balance and secondary active transport of other ions like H⁺ and Ca⁺⁺ into and out of cells (Skou., 1990).

Structurally, the Na,K-ATPase is minimally composed of two subunits, α and β . The 112 kDa α is the catalytic subunit of the enzyme and contains binding sites for Na⁺, K⁺, ATP and inhibitory cardiac glycosides such as ouabain (Lingrel et al., 2003). The β subunit has a molecular mass of 35 kDa and is required for enzyme maturation and translocation to the plasma membrane (Lingrel and Kuntzweiler, 1994). Four isoforms of the α subunit (α 1, α 2, α 3, and α 4) and three isoforms of the β subunit (β 1, β 2 and β 3) have been identified in mammals. Each of these isoforms has a unique tissue distribution and developmental expression pattern (Sweadner K., 1989; Woo et al., 1999, McDonough et al., 1990; Geering, 1990).

Among the α isoforms, α 1 is expressed ubiquitously, α 2 is expressed mainly in the brain, heart and skeletal muscle, α 3 is expressed in the brain and ovaries and α 4 is expressed in the mature sperm flagellum (Woo et al., 2000; Hlivko et al., 2006). In rat, the α 4 isoform is localized mainly to the mid piece (Woo et al., 2000) and in humans and mice it is localized mainly to the principle piece of the sperm flagellum (Hlivko et al., 2006; our unpublished data). Additionally, immunocytochemical studies suggest that the α 4 isoform is expressed in the head of the bovine sperm (Newton et al., 2009). Further, homologs of the ATP1A4 gene have been identified in rhesus monkey and the dog (<http://www.ncbi.nlm.nih.gov/>) however no studies have been reported regarding the localization of ATP1A4 in these species.

Targeted disruption of the mouse α 4 Na,K-ATPase gene demonstrates that this Na,KATPase isoform is essential for male fertility: loss of the α 4 Na,K-ATPase isoform results in loss of basal mouse sperm motility as well as hyperactive motility during capacitation (Jimenez et al., 2011). Male α 4 Na,K-ATPase-null mice are sterile and their sperm are unable to fertilize oocytes *in vitro* (Jimenez et al., 2011). Although much is known about the function of the α 4 Na,KATPase, the mechanisms that regulate and limit the expression of this gene to male germ cells have not been completely defined.

While several transcription factors that drive germ cell-specific gene expression have been identified, some spermatogenic cell-specific genes such as *Ldhc*, *MAGE-A1*, *Pgk-2*, *tH2B*, *tnp1*, *Prn1*, *Prn2*, *Alf*, *Gsg2/Haspin* and *Pdha2* have also been shown to be regulated by DNA methylation (Ariel et al., 1991; Bonny et al., 1995; Choi et al., 1991; De Smet et al., 1999; Iannello et al., 2000; Muller et al., 2000; Pinheiro et al., 2012; Sato et al., 2011 Trasler et al., 1990; and Xie et al., 2002). DNA methylation is the modification of DNA via addition of a methyl group to the carbon atom at the fifth position of cytosine is one of the most extensively studied epigenetic modifications and is highly conserved in plants, animals, and

fungi (Feng et al., 2010). In mammals, the methylation of cytosine occurs primarily at CpG dinucleotides. Clusters of CpG dinucleotides occur in regions called CpG islands, defined as stretches of DNA rich in CpGs, which are frequently located in the promoter regions of genes. In fact, the majority of gene promoters include CpG islands and methylation of CpG islands typically leads to silencing of the promoter and represses gene expression. Examples of genes whose expression depends on DNA methylation status include developmentally regulated genes, tissue-specific genes, genes located on the inactive X chromosomes, germ cell-specific genes in somatic cells and imprinted genes (Hong et al., 2009; Deaton and Bird, 2011; Peter Jones, 2012).

Until recently, studies on gene regulation via DNA methylation have focused heavily on CpG islands located in the promoter regions. However, affinity purification of unmethylated CpGs from genomic DNA using the CXXC protein domain revealed that only half of CpG islands are located at annotated promoter regions/transcriptional start sites (TSSs) in the mouse and human genomes (Illingworth et al., 2008; Lister et al., 2009). The remaining CpG islands are found either within gene bodies (intragenic CpG islands) or between genes (intergenic CpG islands). These non-promoter CpG islands have been termed “orphan CpG islands”. While these so-called orphan CpG islands exhibit a high degree of tissue-specific methylation (Illingworth et al., 2010; Deaton and Bird, 2011), their functional significance with respect to gene regulation is only beginning to be deciphered. Studies have suggested that intragenic/intergenic CpG islands can regulate gene expression in several ways; they may serve as alternative promoters that initiate expression of regulatory transcripts such as non-coding RNA's, they may act as enhancer/silencer elements, or they may alter RNA polymerase II transcription elongation efficiency (Illingworth et al., 2010; Maunakea et al., 2010; Sleutels et al., 2002; Herzing et al., 1997; Rinn et al., 2007; Yu et al., 2013 and Lorincz et al., 2004).

In the study presented here, a differentially methylated intragenic CpG island in *Atp1a4* is identified and functionally analyzed. This CpG island displays reduced methylation in sperm compared to kidney hinting, as this may be a methylation pattern retained from earlier spermatogenic progenitors, that DNA methylation may be important for regulating the expression of this sperm-specific protein. Supporting the idea that DNA methylation regulate its expression, upregulation of *Atp1a4* expression following treatment of the male germ cell line GC-1spg (known to express low levels of $\alpha 4$ mRNA) with the DNA methylation inhibitor 5-Aza2-Deoxycytidine (5-azaC) was demonstrated. Furthermore, mouse embryonic stem cells lacking the DNA methyl transferase (Dnmt) enzymes Dnmt1, Dnmt3a, or Dnmt3b displayed increased *Atp1a4* expression compared to wild-type ES cells indicating the involvement of both maintenance methylation and *de novo* methylation in the regulation of $\alpha 4$ expression in these cells. Finally, in our attempt to identify the functional significance of DNA methylation of the *Atp1a4* CGI and promoter methylation-dependent/independent gene regulatory functions for both the promoter and the intragenic CpG island were uncovered.

2. Materials and Methods

2.1. DNA methylation analysis using bisulfite sequencing

Epididymal sperm and kidneys from adult male C57BL/6 mice were collected under protocols approved by the Miami University Institutional Animal Care and Use Committee (IACUC), rinsed in phosphate buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄ and 1.8 mM KH₂PO₄, pH 7.4) and either pelleted (sperm) by centrifugation or homogenized (kidneys) using a PowerGen700 homogenizer (Fisher Scientific). Genomic DNA from the sperm pellets and kidney lysates was extracted using the 5αPrime Archive pure DNA cell/tissue extraction kit according to the manufacturer's instructions. 200 ng of genomic DNA from kidney and sperm was bisulfite treated using the EZ DNA methylation kit (Zymo Research). Methyl Primer Express® Software v1.0 (Applied Biosystems Life technologies) was used to design primers for amplification of bisulfite treated DNA. The entire Mα4-CGI was amplified, in four overlapping fragments, using the primers listed in the Table 1.1. The Mα4-Promoter of *Atp1a4* was amplified, in three overlapping fragments using the primers listed in Table 1.1. Platinum Taq DNA polymerase (Life technologies) with high fidelity was used to amplify DNA fragments containing Mα4-CGI and the Mα4-Promoter region. The PCR conditions for amplification were as follows: 94 °C for 2 min, 94°C for 30 s, 40 cycles of 55 °C for 30 s, 68 °C for 1 min, and a final extension of 68 °C for 5 min. Following amplification, the PCR products were gel purified and cloned into the pCR 2.1 TOPO vector using the TOPO TA cloning kit (Life technologies). Twenty positive clones were picked for each of the primer pairs from two independent biological replicates. Plasmid DNA was extracted, sequenced, and the methylation pattern of individual CpGs in the DNA fragments were analyzed using the BiQ analyzer software (Max Planck institute for computer science, Germany).

2.2. Culturing of GC-1spg cells and mouse ES cells

GC-1spg cells (ATCC) were grown in Dulbecco's Modified Eagles media (DMEM) with high glucose (Hyclone) supplemented with 10% Stasis Gemcell Fetal Bovine Serum (FBS) (Gemini Biosciences) and 1% penicillin and streptomycin solution (GIBCO). The mouse J1 wild type (wt) ES cells from the 129S4/Sv Jae strain of mice, Dnmt1 knockout (KO) ES cells, Dnmt3a KO ES cells, Dnmt3b KO ES cells and Dnmt3a/3b double knockout (DKO) ES cells (kind gifts from Dr. Taiping Chen, MD Anderson cancer center, Texas with permission from Dr. En Li, R&D Head, Novartis China) were grown on a feeder layer of mouse embryonic fibroblasts in DMEM (Sigma Aldrich) supplemented with 10% Stasis stem cell FBS (Gemini Biosciences) and 1% penicillin and streptomycin solution (GIBCO) and 1000 units/ml of mouse leukemia inhibitory factor (LIF) (ESGRO® Millipore). For methylation analysis and qPCR analysis the ES cells were subsequently sub-cultured on gelatin coated plates for three to four passages in order to remove feeder cell contamination. Genomic DNA from the GC-1spg cells and ES cells was extracted using the 5αPrime Archive pure DNA cell/tissue extraction kit for bisulfite sequencing described above. Ten clones were picked for each of the primer pairs for the CGI and the promoter region. RNA from GC-1spg cells and ES cells was extracted for quantitative real time PCR using the Direct-zol™ RNA MiniPrep Kit (Zymo Research) according to the manufacturer's instructions.

2.3. 5-azaC and TSA treatment of cells and quantitative real time PCR

GC-1spg cells were plated at 1×10^6 cells/ml in 60 mm tissue culture dishes. The following day, the cells were treated with 5 μ M 5-azaC (Sigma Aldrich) alone, 50 nM TSA (Sigma Aldrich) alone or 5 μ M 5-azaC and 50 nM TSA and or their respective vehicles (distilled water for 5-azaC and ethanol for TSA). Each day, fresh inhibitor or vehicle was added to the cultures and after 3 days (72 h), the cells were washed with PBS (pH 7.4) and cell pellets were harvested (modified from Xie et al., 2002). DNA was extracted using 5 α Prime Archive pure DNA cell/tissue extraction kit. Genomic DNA from vehicle (0 μ M) and 5 μ M 5-azaC treated GC-1spg cells were subjected to bisulfite sequencing as above. RNA was extracted using the TRI REAGENT[®] (Molecular Research Center, Inc.) method according to the manufacturer's instructions. Following extraction, first strand cDNA synthesis was performed using the ImProm-II[™] Reverse Transcription System for Two-Step RT-PCR (Promega). The cDNA was then subjected to quantitative real time PCR using the RT² SYBR Green Fluor qPCR Mastermix (Qiagen) on the iCycler Thermal Cycler (Bio-Rad). GAPDH was used as an internal control. Table 1.2 lists the primers used for qPCR. Fold change in mRNA levels were calculated relative to the controls using the C_t method (Livak and Schmittgen, 2001). The qPCR was performed in triplicate from each of three independent biological replicates. A two-tailed student t-test was performed to calculate the p values.

2.4. Generation of M α 4-CGI and M α 4-Promoter reporter constructs

A DNA fragment containing the M α 4-CGI was PCR amplified from genomic DNA, isolated from C57BL/6 tails, as the template using the primers 5' TATTGGCAAGATGCCCTA 3' and Rev: 5' GCTAGCAGTCACTGAGGTGCCCTCT 3'. The reverse primer has an NheI site incorporated in the primer sequence (underlined). Following PCR amplification, the CGI fragment was sub-cloned into pCR 2.1 TOPO vector using the TOPO TA cloning kit (Life technologies) and sequenced. The CGI fragment was subsequently inserted into the NheI and HindIII sites of the pGL3-Basic firefly luciferase reporter vector (Promega). To test the promoter activity of the putative M α 4-Promoter, the DNA fragment was PCR amplified using the following primers; Fwd: 5' GGTGGTGCTAGCGTGCACACATATGGCCA 3' and Rev: 5' GGTGGTAAGCTTCAGCCAAGGAGTGGGAAGG 3' with NheI and HindIII sites incorporated in the primer sequences (underlined). Following PCR amplification the promoter fragment was gel purified and digested with NheI and HindIII restriction enzymes and inserted into the corresponding sites of the pGL3-Basic vector. For analysis of enhancer function of M α 4-CGI three different vector constructs were generated. The M α 4-CGI was PCR amplified and inserted into the KpnI and NheI sites of the pGL3-Promoter vector, a firefly reporter vector which harbors the SV40 promoter using the following primers: Fwd: 5' GGTGGTGCTACCTATTGGCAAGATGCCCTA 3' and Rev: 5' GGTGGTGCTAGCAGTCACTGAGGTGCCCTCT 3 α . The M α 4-CGI was inserted in to the PstI and SpeI sites of the CpG free pCpGL-EF1 α vector (a kind gift from Dr. Michael Rehli, University of Rogensburg, Switzerland, Schmidl et al., 2009) using the following primers: Fwd: 5 α GGTGGTCTGCAGTATTGGCAAGATGCCCTA 3 α and Rev: 5 α GGTGGTACTAGTAGTCACTGAGGTGCCCTCT 3 α . To generate the M α 4-CGI-M α 4-Promoter vector, the M α 4-Promoter was gel purified following digestion of the pGL3-

Basic- M α 4-Promoter with BglIII and HindIII and inserted into the pGL3-M α 4-CGI-SV40-Promoter vector replacing the SV40 promoter. All DNA inserts were sequence verified before transfection.

2.5. In vitro methylation of reporter constructs

Individual reporter constructs (10 μ g) were treated with 10 units of M.SssI methylase (New England Biolabs) overnight in the presence of 160 μ M S adenosylmethionine (New England Biolabs). Mock methylation was performed in the absence of the M.SssI methylase. The M.SssI methylase enzyme was heat inactivated at 65 °C for 20 min and the DNA was ethanol precipitated before transfection into cells and assayed for luciferase activity (Hattori et al., 2004). The methylation status of the constructs was verified by digestion with the methyl-sensitive restriction enzyme HpaII (New England Biolabs).

2.6. Transient transfections and dual luciferase reporter assays

For dual luciferase reporter assays, GC-1spg cells, NIH3T3 and HEK 293 were plated at 2×10^5 cells/ml in a 24 well dish and then transfected with 1 μ g of DNA using Lipofectamine 2000 (Life technologies). 100 ng of *Renilla* luciferase reporter vector pRL-TK (Promega) was cotransfected to monitor transfection efficiencies. 48 h post transfection, the cells were assayed for both firefly and *Renilla* luciferase activity using the dual luciferase reporter assay system (Promega) using a BMG labtech Novostar microplate reader. Each assay was performed in duplicate for each of four independent transfections. Results are expressed as a ratio of the firefly to *Renilla* luciferase activities. A two-tailed student t-test was performed to calculate the p values.

2.7. Transcription elongation analysis

The intragenic M α 4-CGI was cloned in to the NotI and NheI sites present in the linker between the firefly and the *Renilla* luciferase reporters in the TAN1 reporter construct (A kind gift from Dr. Ed Grabczyk, Department of Genetics LSU Health Sciences Center New Orleans LA; Banerjee et al., 2009). To test *in vitro* methylation, the M α 4CGI was first PCR amplified using the high fidelity Takara Taq DNA polymerase. The PCR product was gel purified using a Gel/PCR DNA fragment extraction kit (IBI scientific). Following gel purification, the PCR product was digested with NotI and NheI. Subsequently, the DNA was ethanol precipitated and *in vitro* methylated or mock methylated. 16 μ g of insert was *in vitro* methylated using the M. SssI methylase (NEB) in the presence of 160 μ M S-adenosylmethionine and NEB buffer 2 for 8–10 h at 37 °C. Mock methylation was performed under the same conditions without the M. SssI methylase (Hattori et al., 2004). The reaction was terminated by heat inactivating the methylase at 65 °C for 20 min. The methylation status was verified by digestion with the methyl-sensitive restriction enzyme HpaII. The methylated and mock-methylated intragenic M α 4-CGI fragments were then ligated to the TAN1 vector construct digested with NotI and NheI. The ligation products were then linearized with SacII, run on a 1% agarose gel and the appropriately sized DNA fragments were gel purified and used for transfection of HEK 293 cells on a 96 well dish. HEK 293 cells were plated at 2×10^4 cells/ml one day before transfection. On the day of transfection, 100 ng of gel purified ligation product was transfected using turbofect (Thermo

Scientific) according to the manufacturer's instructions. 48hr post transfection the cells were assayed for both firefly and *Renilla* luciferase activity as above. Eight independent transfections were performed for *in vitro* methylation and mock methylation. Results are expressed as a ratio of the *Renilla* to firefly luciferase activities. A two-tailed-student t-test was performed to calculate the p values.

3. Results

3.1. A single differentially methylated intragenic CpG island (M α 4-CGI) is present in the *Atp1a4* gene

As a first step towards determining whether DNA methylation is involved in regulating the expression of the α 4 Na,K-ATPase gene, Methyl Primer Express[®] Software v1.0 (Applied Biosystems; Life Technologies) was used to identify a single CpG island in the *Atp1a4* locus. The mouse alpha4 CpG island (M α 4-CGI) begins in the intron between exon 8 and exon 9, and ends in the beginning of the intron between exon 10 and exon 11 corresponding to the genome coordinates Chr1:174,174,430-174,175,532 of the mm9 mouse genome assembly (Figure 1A). This region is a 1.1kb sequence with a CpG expected/CpG observed ratio of 0.7 and a C+G percentage of 51.41% and therefore fulfills the definition of a CpG island proposed by Deaton and Bird (2011). As a preliminary indicator of a potential tissue-specific role in regulating Na,KATPase α 4 gene expression, the methylation status of the M α 4-CGI was examined in DNA from mouse kidney (a tissue that does not express the Na,K-ATPase α 4 isoform) and mouse sperm (cells known to contain this protein). Note that while sperm contain the highest levels of this protein, the expression of α 4 begins in cells earlier on in the spermatogenic lineage (Blanco et al., 2000). DNA methylome analyses of germ cells and somatic tissues/cells have been recently reported (Kobayashi et al., 2012; Hon et al., 2012; Stadler et al., 2011). Analysis of the data associated with these studies suggests that the intragenic M α 4-CGI identified in our study is a hypomethylated region (HMR) in mouse sperm but is methylated in somatic tissues Therefore we wanted to verify these results and hence performed bisulfite sequencing of this region from mouse sperm (germ cells) and kidney (a somatic tissue).

Bisulfite DNA sequencing analysis revealed that the M α 4-CGI is differentially methylated in sperm and kidney: 17.3% of the CpGs in this region were methylated in DNA from sperm whereas 71.7% of the CpGs were methylated in the DNA from kidney (Figure 1B).

3.2. CpGs in the putative *Atp1a4* promoter (M α 4-Promoter) are not differentially methylated in mouse kidney and sperm

As many promoters are positioned within the region between -700bp and +200bp relative to the TSS (Weber et al., 2007) and because this region is included in those that previous studies suggest may be involved in the tissue-specific expression of this gene in human and mouse (Blanco et al., 2006; McDermott et al., 2012), we choose the -700bp to +200bp region of *Atp1a4* (corresponding to the genome coordinates Chr1:174,188,355-174,189,255 of the mm9 mouse genome assembly) for analysis as a promoter (M α 4-Promoter).

Although our analysis of *Atp1a4* did not uncover a cluster of CpG dinucleotides characteristic of a CpG island in the M α 4-Promoter, this region does contain individual CpG dinucleotides that could potentially play an important role in control of *Atp1a4* expression. Other promoters with CpG contents below those necessary to qualify as CpG islands have been identified that are able to regulate gene expression through CpG methylation. As examples, both the insulin gene and the OCT4 gene have promoters that do not contain CpG islands, yet both are regulated by DNA methylation of the individual CpGs in their promoter regions (Kuroda et al., 2009; Hattori et al., 2004). Bisulfite sequencing of the M α 4-Promoter region revealed equivalent methylation levels in kidney and sperm with 87.2% of the CpGs in this region of *Atp1a4* methylated in the kidney and 88.6% of the CpGs methylated in sperm (Figure 1C).

3.3. Upregulation of *Atp1a4* expression and demethylation of the M α 4-CGI and the M α 4-Promoter following treatment of GC-1spg cells with 5-azaC

The GC-1spg cell line is an immortalized male germ cell line (Hofmann et al., 1992) known to express low levels of α 4 mRNA (Blanco et al., 1999). We reasoned that if DNA methylation is in any way important for the regulation of expression of *Atp1a4* and underlies the low level of expression of this gene in GC-1spg cells, treatment with 5-azaC should increase *Atp1a4* expression. qPCR analysis demonstrates that treatment of GC-1spg cells with 5 μ M 5-azaC for 72 h leads to a >10 fold increase in α 4 mRNA levels when compared to vehicle-treated GC-1spg cells (Figure 2A).

Having demonstrated that 5-azaC upregulates α 4 mRNA levels, we next determined the methylation status of the M α 4-CGI and the M α 4-Promoter in these cells to assess whether reduced methylation of either of these regions corresponds with increased α 4 expression levels. To facilitate analysis, the M α 4-CGI sub-region containing the largest number of individual CpGs (Figure 1B) was selected for this and all subsequent methylation sequence analyses. Bisulfite DNA sequencing analysis of the M α 4-CGI demonstrates that 94.0% of the CpGs are methylated in the vehicle-treated GC-1spg cells compared to only 40.6% following treatment of GC-1spg cells with 5 μ M 5-azaC (Figure 2B). Analysis of the M α 4-Promoter region revealed that 64.7% of the CpGs are methylated in this region following treatment of GC-1spg cells with 5-azaC compared to 93.0% in the vehicle-treated GC-1spg cells (Figure 2C). In particular, sub-region I of the M α 4-Promoter showed a marked decrease in CpG methylation in the 5-azaC-treated cells (40.0%) compared to control cells (98.0%) whereas the other two promoter sub-regions showed little to no differences in methylation (sub-region II: 94.0% vs. 90.0%; sub-region III: 77.0% vs. 60.0% CpG methylation in treated vs. control cells). These results demonstrate that general inhibition of Dnmts increases *Atp1a4* expression and reduces methylation in both the M α 4-CGI and the M α 4-Promoter (in particular sub-region I), demonstrating that DNA methylation is important for regulating *Atp1a4* expression and leaves open the possibility that hypomethylation of either of these regions could underlie the upregulation of the α 4 mRNA transcript under these conditions (Figure 2A).

GC-1spg cells were also treated with 50 nM trichostatin A (TSA), a known histone deacetylase inhibitor (Godmann et al., 2010), alone or in combination with 5-azaC for 72 h.

No significant differences in $\alpha 4$ transcript levels were detected following treatment of GC-1spg cells with 50 nM TSA alone compared to vehicle-treated cells. Upregulation of the $\alpha 4$ transcript within cells treated with both 5 μ M 5-azaC and 50 nM TSA was observed however the increased expression was not significantly different when compared to cells treated with 5 μ M 5-azaC alone (Figure 2A) suggesting that histone deacetylation might not play a major role in regulating *Atp1a4* expression in GC-1spg cells.

3.4. Loss of DNA methyltransferase (Dnmt) enzymes results in upregulation of *Atp1a4* expression and demethylation of the M α 4-CGI and the M α 4-Promoter in mouse ES cells

Having demonstrated that general inhibition of Dnmts with 5-azaC in GC-1spg cells leads to upregulation of *Atp1a4* expression, it was of interest to determine the roles that individual Dnmt enzymes play in this process. To do this, *Atp1a4* expression was determined in mouse embryonic stem (ES) cells lacking the maintenance methyltransferase Dnmt1 [Dnmt knockout (Dnmt1 KO)], the individual *de novo* methyltransferase Dnmt3a (Dnmt3a KO) and Dnmt3b (Dnmt3b KO) or both Dnmt3a/3b [Dnmt3 double knockout (DKO)]. Mouse J1 wild type (wt) ES cells, derived from the 129S4/SvJae strain of mice, were used as controls as all of the Dnmt knockout cell lines used were generated using this ES cell line (Okano et al., 1999).

qPCR analysis revealed that loss of any individual Dnmt in mouse ES cells results in increased *Atp1a4* expression when compared to wt J1 ES cells: loss of Dnmt1 results in a 4.6 fold increase (Figure 3A), loss of Dnmt3a results in a 5.9 fold increase and loss of Dnmt3b results in a 3.1 fold increase in the $\alpha 4$ transcript (Figure 4A). However, loss of both Dnmt3a and 3b did not alter the expression of *Atp1a4* transcript when compared to the wt J1 cells (Figure 4A).

Next, to determine the contribution of Dnmt1, Dnmt3a and Dnmt3b to the methylation of the M α 4-CGI and the M α 4-Promoter, methylation patterns of the *Atp1a4* promoter and intragenic CGI were analyzed in the Dnmt KO ES cells using bisulfite DNA sequencing. In the wt J1 ES cells, 37.0% of the CpGs were methylated in the intragenic M α 4-CGI analyzed whereas 21.0% were methylated in this region of the Dnmt1 KO ES cells genome (Figure 3B). In comparison, varying degrees of methylation in this CpG island were displayed by the Dnmt3 KO ES cells: 3.0% of the CpGs were methylated in the M α 4-CGI in Dnmt3a KO, 20.0% in Dnmt3b KO and 19.0% in Dnmt3a/3b DKO ES cells (Figure 4B).

Analysis of the methylation status of the promoter region in ES cells revealed that 87.0% of the CpGs in the M α 4-Promoter were methylated in the wt J1 ES cells, 42.2% were methylated in the Dnmt1 KO ES cells (Figure 3C), 4.7% were methylated in Dnmt3a KO cells, 44.1% were methylated in Dnmt3b KO cells and 77.0% were methylated in the Dnmt3a/3b DKO cells (Figure 4C).

These analyses revealed an inverse relationship between the expression levels of the $\alpha 4$ transcript and the methylation state of both the M α 4-CGI and the M α 4-Promoter for the wt and single KO ES cells (compare Figures 3&4). For the Dnmt3a/3b DKO cells, this inverse relationship between methylation status and $\alpha 4$ transcript levels appears to hold for the M α 4-Promoter but not for the intragenic M α 4-CGI. The M α 4-Promoter region in the DKO

cells displayed near wild-types levels of the M α 4-Promoter methylation and α 4 transcript levels whereas these same cells displayed a moderate reduction of methylation in the intragenic M α 4-CGI (compare Figures 4A, 4B and 4C).

3.5. The M α 4-Promoter is a methylation-dependent, germ cell-specific promoter

In wt J1 and single Dnmt KO ES cells and in GC-1spg cells treated with 5-azaC the decreased methylation in the promoter corresponds to higher α 4 expression again leaving open the possibility that methylation of this region could be participating in regulating gene expression. Therefore, we next wanted to determine whether DNA methylation of the promoter could alter gene expression. As a first step, the ability of the M α 4-Promoter to drive gene expression was examined in cultured cells. GC-1spg cells but not NIH 3T3 cells (an immortalized embryonic fibroblast somatic cell line) transfected with the M α 4-Promoter-containing construct display a 2.0 fold greater normalized firefly luciferase activity compared to cells transfected with the construct lacking a promoter. These results demonstrate that the region -700bp to +200bp region of *Atp1a4* possesses promoter activity that is specific to the GC-1spg germ cell line (Figure 5A; lower panel).

Next, in order to test the methylation dependence of the M α 4-Promoter activity, the pGL3-Basic-M α 4-Promoter vector was *in vitro* methylated using the CpG methyltransferase enzyme (M.SssI) (New England Biolabs). As a control, the pGL3-Basic-M α 4-Promoter construct was mock methylated under the same conditions in the absence of the methyltransferase enzyme (Figure 5B; upper panel). GC-1spg cells containing the *in vitro* methylated expression construct displayed decreased (by 56.6%) normalized luciferase activity compared to cells containing the mock-methylated construct demonstrating that the promoter is methylation-sensitive in this germ cell line (Figure 5B; lower panel).

3.6. The M α 4-CGI does not function as an alternative promoter

As a first step towards determining if the M α 4-CGI functions as an alternative promoter, this region was assayed for promoter activity in GC-1spg and NIH3T3 cells. GC-1spg and NIH3T3 cells transfected with the pGL3-Basic-M α 4-CGI construct each displayed equivalent normalized luciferase activity as the cells transfected with the promoter-less pGL3-Basic vector (Figure 6; lower panel) demonstrating that the intragenic M α 4-CGI does not function as a promoter in either of these cell lines.

3.7. Methylation of the intragenic M α 4-CGI inhibits transcription elongation efficiency

In order to determine if DNA methylation in the intragenic M α 4-CGI inhibits transcriptional elongation, a tandem reporter construct (TAN1) that has been previously used to quantify transcription elongation through polyadenylation sequences was employed (Banerjee et al., 2009). In the TAN1 construct, transcription driven by the human cytomegalovirus (CMV) immediate early gene promoter initiates and proceeds through the firefly luciferase reporter (the upstream control reporter) sequence, through the putative transcriptional elongation-altering sequence (intervening sequence), and then through the *Renilla* luciferase reporter (downstream reporter) sequence (Figure 7A). Successful elongation through the intervening sequence is determined by assaying the ratio of expression of the downstream reporter (*Renilla* luciferase) to that of the upstream control reporter (firefly luciferase). If the

intragenic *Atp1a4* CpG island alters transcriptional elongation when methylated, it is expected that cells transfected with the reporter construct containing the methylated CpG island as the intervening sequence will display a *Renilla*/firefly luciferase activity ratio lower than cells transfected with the reporter construct containing an unmethylated CGI as the intervening sequence (Figure 7A).

The DNA fragment containing the M α 4-CGI was PCR amplified and gel purified, methylated *in vitro* (or mock methylated) and then ligated with the NotI and NheI-digested TAN1 reporter construct. This strategy positioned the CpG island-containing fragment as the intervening sequence between the firefly and *Renilla* reporter sequences (Figure 7A) and limited the methylation of the construct to only the CGI. The ligated DNA in the reaction mix was then digested with SacII to linearize the construct and the appropriate size band was gel purified and transfected into HEK 293 cells and assayed for firefly and *Renilla* luciferase activity 48 h later. Due to the technical difficulties associated with generating large quantities of these constructs, HEK 293 cells were utilized in these assays as this cell line can be transfected with very high efficiency (Thomas and Smart, 2005) and thus requires less DNA for each transfection.

Compared to mock methylation, *in vitro* methylation of the intragenic M α 4-CGI resulted in decreased expression of the downstream *Renilla* luciferase reporter (Figure 7B) demonstrating a reduction in the efficiency of transcriptional elongation through the methylated CGI. These results suggest that methylation of the intragenic CpG island could decrease *Atp1a4* gene expression by hindering RNA polymerase II-dependent transcriptional elongation of the α 4 transcript.

3.8. The M α 4-CGI functions as a silencer element in GC-1spg cells

Besides serving as alternative promoters and as regulators of transcriptional elongation, a role for intragenic CpG islands as enhancers or silencers of transcription is beginning to be uncovered (Yu et al., 2013; Schmidl et al., 2009). To complete the functional characterization, by testing all three proposed mechanisms by which intragenic CGI are thought to regulate gene expression (alternate promoters, transcriptional elongation, and silencer/enhancer), the ability of this CGI to modulate the activity of multiple promoters (SV40, EF1 α , and M α 4-Promoter) was determined.

GC-1spg cells transfected with pGL3-M α 4-CGI-SV40-Promoter and pGL3-M α 4-CGIM α 4-Promoter displayed reduced normalized firefly luciferase activity compared to GC-1spg cells transfected with their promoter-only control vectors (Figure 8; lower panel) demonstrating that in the unmethylated state the intragenic M α 4-CGI possesses silencer-like activity when combined with the SV40 promoter or the mouse *Atp1a4* promoter. In contrast, GC-1spg cells transfected with pCpGL-EF1 α -Promoter-M α 4-CGI displayed similar normalized firefly luciferase activity when compared to cells transfected with the promoter-only control (Figure 8; lower panel) suggesting that the silencer-like activity of the unmethylated intragenic CpG island is not as effective with all promoters.

It is known that silencer/enhancer cis elements are able to exert their function irrespective of position (upstream or downstream of the promoter) or orientation (Griffiths et al., 2000).

Therefore, since the M α 4-CGI was present downstream of the promoter in the constructs used to analyze its ability to alter transcriptional elongation (see above; Figure 7B) in HEK 293 cells, the silencer activity of this CGI was also analyzed in these cells. HEK 293 transfected with the pGL3-M α 4-CGI-SV40-Promoter or the promoter-only control displayed equivalent normalized firefly luciferase activities (Figure S1) suggesting that the silencer activity of the unmethylated M α 4-CGI may be cell type-specific and demonstrates that silencer activity was unlikely to have influenced the assays for transcriptional elongation regulation (Figure 7B).

Having identified a silencer-like activity for the unmethylated M α 4-CGI in GC1-spg cells, it was of interest to determine whether methylation could alter this activity. Towards this end, the pGL3-M α 4-CGI-SV40-Promoter, pCpGL-M α 4-CGI-EF1 α -Promoter and pGL3-M α 4-CGI-M α 4-Promoter along with their respective promoter-only vector controls were *in vitro* methylated or mock-methylated (Figure S2; upper panel and Figure 9; upper panel) and then transfected into GC-1spg cells along with pRL-TK vector and the cells assayed for luciferase activity 48 h later. *In vitro* methylation of the M α 4-CGI eliminates the silencer function of M α 4-CGI with respect to the SV40 promoter and uncovered apparent enhancer activity with respect to the EF1 α promoter (Figure S2; lower panel). *In vitro* methylation of the vector containing the M α 4-Promoter alone significantly decreased the normalized luciferase activity compared to the mock-methylated vector (compare bars 1&3 in Figure 9B) as previously shown (Figure 5B). In addition, *in vitro* methylation of the vector containing both the promoter and the CGI resulted in a decrease of normalized luciferase activity compared to the vector containing the methylated M α 4-Promoter alone (compare bars 3&4 in Figure 9B). These data indicate that methylation of the *Atp1a4* CGI allows increased expression from non-self promoters (SV40 and EF1 α) compared to the unmethylated state but the methylated CGI retains its repressor activity towards its own promoter in these cultured germ cells.

4. Discussion

The overall goal of this study was to begin to dissect how the expression of the α 4 Na,K-ATPase is regulated. Our data suggest that DNA methylation of the promoter and the intragenic CGI could combine to provide multiple modes of regulation for optimizing the *Atp1a4* expression level in somatic and male germ cells.

DNA methylome analysis using whole genome shotgun bisulfite sequencing in germ cells (sperm and oocytes) has recently been reported and shows that the sperm genome is generally hypermethylated (with average methylation levels of 89.4%) outside of regions rich in CpG dinucleotides (Kobayashi et al., 2012). Mining of the data associated with these whole genome methylation studies revealed that the intragenic M α 4-CGI identified and characterized in our study is a hypomethylated region (HMR) in mouse sperm. Analysis of databases containing information from other recent DNA methylome studies of mouse cells and somatic tissues reveals that the intragenic M α 4-CGI is methylated in 17 different somatic tissues (Hon et al., 2012), mouse ES cells and neural progenitor cells (Stadler et al., 2011). These genome wide methylation studies (publically available in UCSC database) are

in line with our findings (Figure 1B) and suggest that DNA methylation in the M α 4-CGI could be a potential mechanism for regulation of *Atp1a4* expression.

In our studies, methylation analysis via bisulfite sequencing demonstrated that the intragenic CpG island, but not the individual CpGs in the promoter, is differentially methylated in mouse kidney compared to mature sperm (Figure 1B & 1C). In addition, the mouse *Atp1a4* promoter was not identified as a HMR in the whole genome studies discussed above. While these data could be interpreted to indicate that methylation of the promoter region is not important in regulating *Atp1a4* expression, it has been documented that some testis-specific genes such as *Pgk-2*, *ApoA1* and *Alf*, known to be regulated by methylation of CpGs in their promoter region, undergo remethylation of the CpGs present in their promoter region in the mature spermatozoa as part of the sperm maturation process that takes place in the epididymis (Ariel et al., 1994; Xie et al., 2002). Thus it remains possible that hypomethylation of the *Atp1a4* promoter is involved in increasing *Atp1a4* expression earlier in sperm development and that the individual CpGs in the promoter had been remethylated in the mature sperm as translation is shut down in these cells. Our dual luciferase reporter assay confirms that the activity of the M α 4- Promoter is methylation dependent in GC1-spg cells (Figure 5B). In all strongly suggesting that methylation of the *Atp1a4* promoter may play a role in regulating the expression of this gene in the mouse. The analysis of the methylation state of the *Atp1a4* promoter in cells at earlier stages of spermatogenesis awaits further study.

Expression and methylation analysis in Dnmt KO ES cells suggests that Dnmt3a plays a major role in regulating the expression of *Atp1a4* (Figure 4A & 4B). Interestingly, an essential role for Dnmt3a in male germ cell development has been reported, with specific deletion of Dnmt3a, but not Dnmt3b, in male germ cells resulting in azoospermia and loss of methylation at imprinted loci (Kaneda et al., 2004). Thus Dnmt3a may be the important DNA methyltransferase directing the normal development of male germ cells, where at least part of its function is to ensure the proper level of expression of certain germ cell-specific genes like *Atp1a4*.

It is not completely clear why loss of either Dnmt3 isoform alone reduces promoter and CGI methylation and increases *Atp1a4* expression whereas loss of both *de novo* methyltransferases has less effect on methylation and no effect on gene expression (Figure 4A, 4B & 4C). However, it has been reported that, in addition to *de novo* methylation, Dnmt3a and Dnmt3b may possess deaminase activity and may therefore also be involved in the process of active demethylation (Metivier et al., 2008). Thus, it is possible that the loss of both Dnmt3a and Dnmt3b actually reduces demethylation which results in increased methylation of the *Atp1a4* promoter and the CGI in the Dnmt3a/3b DKO ES cells compared to the Dnmt3a KO or Dnmt3b KO ES cells.

Studies have suggested that intragenic CGIs are able to alter transcription by at least three fundamentally different mechanisms: 1) by functioning as an alternative promoters that drives the expression of potential regulatory RNA molecules (Illingworth et al., 2010; Maunakea et al., 2010; Sleutels et al., 2002; Herzing et al., 1997; Rinn et al., 2007), 2) by serving as cis regulatory elements that can either enhance or repress gene expression (Yu et

al., 2013) or 3) by altering elongation of their transcripts (Lorincz et al., 2004). To provide a comprehensive analysis of whether the $\alpha 4$ intragenic CGI impacts expression of *Atp1a4* and if its function is regulated by methylation, we tested all three of these possibilities.

First, the potential role of the M $\alpha 4$ -CGI in regulating gene expression was assessed. Towards that end, we tested the role of the intragenic M $\alpha 4$ -CGI as an alternative promoter. Our data demonstrates that the M $\alpha 4$ -CGI is unable to drive reporter gene expression in GC-1spg cells or NIH3T3 cells (Figure 6) and is therefore not a functional promoter in this context.

Further analysis of the M $\alpha 4$ -CGI as a cis regulatory element that can either enhance or repress gene expression demonstrates that the CGI functions as a silencer with respect to the *Atp1a4* and SV40 promoters in GC-1spg cells (Figure 8). However the intragenic CGI was unable to repress all promoters, as it did not significantly inhibit the activity of the EF1 α promoter in GC-1spg cells (Figure 8). Interestingly, methylation of this intragenic CGI resulted in loss of its ability to silence non-native promoters (SV40 and EF1 α). Upon methylation, the CGI's silencer activity with respect to the activity of the SV40 promoter was eliminated and an apparent enhancer activity was uncovered with respect to the EF1 α promoter activity (Figure S2). However, the M $\alpha 4$ -CGI functions as a silencer element of its own promoter regardless of its methylation status in GC-1spg cells (Figure 9). The existence of a CGI with both enhancer and repressor functions is not unprecedented. For example, the CGI found in the 3' end of the human *APOE* gene was shown to decrease the activity of the promoter (i.e. functions as a silencer) of the gene located upstream of the *APOE* gene locus while it increased the activity of the *APOE* promoter (i.e. functions as an enhancer) in luciferase reporter assays (Yu et al., 2013).

Finally, we explored the possibility that methylation of the intragenic CGI might regulate *Atp1a4* expression by altering transcriptional elongation using a tandem reporter construct (TAN1)(Banerjee et al., 2009). The efficacy of this reporter has been previously determined (Banerjee et al., 2009; Dominguez-Sanchez et al., 2011) providing confidence that the TAN1 construct could be used for our experiment to determine the effect of intragenic CpG island methylation on transcriptional elongation. In the methylated state the M $\alpha 4$ -CGI significantly reduced the *Renilla* to firefly luciferase activity ratio compared to the unmethylated state in HEK 293 (Figure 7B). Therefore, as both reporters are encoded for by the same transcript, methylation of the intragenic CGI reduces the availability of the *Renilla* luciferase portion of the mRNA by inhibiting elongation of the bivalent transcript through the CGI (Figure 7B).

While it is not yet clear exactly how methylation of a CGI could inhibit transcriptional elongation, at least a couple of mechanisms are possible. First, it may be that methylation in the CGI prevents recruitment of factors required for transcription elongation by inhibiting their binding to the DNA. Second, methylation of the CGI could alter local chromatin structure to reduce transcriptional elongation. Lorincz et al. (2004) have shown that a region of dense DNA methylation inside an actively transcribed gene could trigger a pattern of histone modifications that results in a compact or repressed chromatin state, which in turn reduces the RNA polymerase II elongation efficiency and results in premature dissociation

of the polymerase. It is not completely obvious however how compact chromatin, mediated by histone modifications, could account for a decrease in transcriptional elongation observed in our reporter constructs. Evidence does exist suggesting that transiently transfected DNA interacts with all four core histones (Hebbar et al. 2008) leaving open the possibility that methylation of the CGI triggers chromatin modifications that could cause interference with RNA polymerase procession even in our reporter assays.

It is evident that no single mechanism of gene regulation functions totally independently; transcription factors, DNA methylation, and histone modifications function interdependently to regulate gene expression and these gene regulatory mechanisms can even influence the presence/absence of the others (for example DNA methylation can stimulate histone modifications to stabilize gene repression; Cedar and Bergman, 2009). With regards to histone modifications, although we have shown that acetylation/deacetylation may not be involved in regulating *Atp1a4* expression, there are at least two other histone modifications that can alter chromatin structure: methylation and phosphorylation. Depending upon the site of modification, histone methylation can either activate or repress gene expression. H3K4 and H3K36 methylation are associated with transcriptional activation while H3K4 demethylation and H3K9 and H3K27 methylation all mediate transcriptional repression (Kouzarides et al., 2007). In addition, histone phosphorylation/dephosphorylation are generally associated with transcriptional activation and repression respectively (Lo et al., 2001; Nowak et al., 2000). Therefore, methylation and/or dephosphorylation of histones mediated by M α 4-CGI methylation could underlie the reduction in transcriptional elongation that regulates *Atp1a4* expression. However, the interaction and interdependence of these individual histone modifications with respect to gene expression make it premature to speculate as to the exact mechanism by which methylation of the M α 4-CGI influences transcriptional elongation and will therefore be the subject of future studies.

Overall, our results suggest that together the *Atp1a4* promoter and intragenic CGI could have multiple effects on gene expression that may depend on cellular context and methylation state, therefore we propose the model depicted in Figure 10 to unify our findings and suggest how these two DNA elements may regulate *Atp1a4* expression. In somatic cells, methylation of both the promoter and the CGI would ensure the lowest level of *Atp1a4* expression (Figure 10A). The *Atp1a4* promoter is cell type-specific (Figure 5A) and will be inactive irrespective of its methylation state in somatic cells. In addition, since the methylated or unmethylated intragenic *Atp1a4* CGI represses the activity of the α 4 Na,K-ATPase gene promoter in germ cells, the CGI (irrespective of methylation) could also repress any aberrant expression driven by this same promoter in somatic cells. The methylated CGI could 1) repress any aberrant expression from the methylated promoter (indicated by the “X” at the TSS in Figure 10A) and/or prevent the complete synthesis of any aberrantly expressed transcripts by inhibiting transcriptional elongation (indicated by the truncated light grey transcript in Figure 10A) keeping the transcript levels of *Atp1a4* lowest in somatic cells. These expectations are supported by our findings that both the CGI and the promoter are extensively methylated in DNA isolated from mouse kidney (Figure 1B&1C) and that these regions have not been identified as HMRs in multiple somatic tissues and cells (Stadler et al., 2011, Hon et al., 2012).

In germ cells it is expected that both the *Atp1a4* promoter and the intragenic CGI will be found in hypomethylated states (see Figure 10B), as methylation of the promoter inhibits its activity (Figure 5B) and methylation of the CGI inhibits transcriptional elongation (Figure 7B). If so, the unmethylated *Atp1a4* promoter will be active in germ cells (indicated by the green arrow at the TSS in Figure 10B.I) and transcriptional elongation will proceed uninterrupted through the unmethylated intragenic CGI resulting in the production of the full-length $\alpha 4$ transcript (indicated by the black transcript in Figure 10B). However, our data suggests that the unmethylated CGI also functions to repress the activity of the *Atp1a4* promoter in germ cells (Figure 8). To align these seemingly conflicting findings, we postulate the existence of a mechanism that could reverse the repressor function of the intragenic CGI specifically in germ cells. If such a mechanism did exist, this would allow for maximal activity of the hypomethylated *Atp1a4* promoter activity combined with uninhibited transcriptional elongation through the hypomethylated CGI.

Although currently unsupported with respect to *Atp1a4* expression, mechanisms that derepress the expression of other genes are known. For example, the intracellular C-terminal fragment of heparin-binding EGF-like growth factor (HB-EGF-C) can inhibit promyelocytic leukemia zinc finger's (PLZF's) repressor function towards *cyclin A gene* expression (Shaknovich et al., 1998; Yeyati et al., 1999). As another example, Kdm2b is an H3K36me2- specific demethylase that preferentially binds to long unmethylated CGIs. Kdm2b has been shown to interact with PRC1 (part of the polycomb repressor complex) thereby recruiting the repressor complex to unmethylated CGIs of lineage-specific genes resulting in repression of these genes in mouse ES cells. Depletion in the levels of Kdm2b upon differentiation of mouse ES cells results in a de-repression of these repressed lineage-specific genes (He et al., 2013).

Therefore, since the cis-acting intragenic CGI likely represses *Atp1a4* expression via recruitment of trans-acting factors, it is possible that loss or inhibition of these repressive factors, specifically in germ cells, would result in de-repression and maximal *Atp1a4* expression when needed as the cells progress through the process of spermatogenesis. Furthermore, when it is necessary to tune *Atp1a4* expression lower at any point during spermatogenesis, varying the activity/availability of the postulated repressive factors, combined with methylation at the CGI and/or the promoter, could provide for an adjustable scale of *Atp1a4* expression levels. Methylation of the CGI would reduce transcriptional elongation. Methylation of the promoter would reduce its ability to drive *Atp1a4* expression irrespective of methylation status of the CGI. Interestingly, our findings that the CGI, but not the promoter, is hypomethylated in mature sperm compared to kidney (Figure 1B&1C) may represent one of these intermediate states of *Atp1a4* expression.

5. Conclusions

In conclusion, two DNA elements in mouse *Atp1a4*, the promoter and an intragenic CGI, were identified that have both methylation-dependent and methylation-independent effects on gene expression in reporter assays. These DNA elements have the potential to combine to produce varying levels of expression to provide precise levels of $\alpha 4$ Na,K-ATPase levels in somatic cells and in developing spermatogenic cells (Figure 10). Future experiments will

entail determining the methylation state of these two DNA elements in multiple different stages of developing spermatogenic cells and making use of emerging technologies, like zinc finger effectors (ZFEs) or transcription activator like effectors (TALEs) to direct targeted demethylation of the CpGs (Chen et al., 2014; Maeder et al., 2013) in the *Atp1a4* promoter and CGI in cells and/or whole organisms in an attempt to understand the biological consequence of site-specific *Atp1a4* demethylation of individual CpGs *in vivo*.

Beyond providing basic information regarding the expression of sperm-specific genes, understanding the mechanisms regulating the expression of *Atp1a4* could provide insights into human male fertility regulation, causes of male infertility and identify sites for manipulations by potential male factor contraceptive agents.

Acknowledgments

The authors wish to thank Dr. Taiping Chen, MD Anderson cancer center, TX for providing the wild-type and Dnmt KO ES cells, Dr. Michael Rehli, University of Rogensburg, Switzerland for the CpG free pCpGL EF1 α Vector and Dr. Ed Grabczyk, Department of Genetics LSU Health Sciences Center New Orleans LA for the TAN1 reporter construct. We would also like to thank Dr. David Pennock, Miami University for his insightful suggestions regarding our model. This work was supported by funding from Miami University and the National Institutes of Health (1R15HD050283-01A1 and 1R15HD065633-01 to PFJ).

References

- Ariel M, McCarrey J, Cedar H. Methylation Patterns of testis-specific genes. Proceedings at the National Academy of Sciences of the United States of America. 1991; 88(6):2317–2321.
- Ariel M, Cedar H, McCarrey J. Developmental-changes in spermatogenic specific genes include reprogramming in the epididymis. Nature Genetics. 1994; 7(1):59–63. [PubMed: 8075642]
- Banerjee A, Sammarco M, Ditch S, Wang J, Grabczyk E. A Novel Tandem Reporter Quantifies RNA Polymerase II Termination in Mammalian Cells. Plos One. 2009; 4(7)
- Blanco G, Melton R, Sanchez G, Mercer R. Functional characterization of a testes-specific alpha-subunit isoform of the sodium/potassium adenosinetriphosphatase. Biochemistry. 1999; 38(41): 13661–13669. [PubMed: 10521273]
- Blanco G, Sanchez G, Melton RJ, Tourtellotte WG, Mercer RW. The alpha4 isoform of the Na,K-ATPase is expressed in the germ cells of the testis. Journal of Histochemistry Cytochemistry. 2000; 48:1023–1032. [PubMed: 10898797]
- Blanco G, Nguyen A, Rodova M. The transcription factor CREM tau and cAMP regulate the activity of the Na,K-ATPase alpha 4 isoform promoter. FASEB Journal. 2006; 20(5):A1238–A1238.
- Bonny C, Goldberg E. The CpG-rich promoter of human LDH-C is differentially methylated in expressing and non-expressing tissues. Developmental Genetics. 1995; 16(2):210–217. [PubMed: 7736669]
- Cedar H, Bergman Y. Linking DNA methylation and histone modifications: patterns and paradigms. Nature Reviews Genetics. 2009; 10(5):295–304.
- Chen H, Kazemier HG, de Groote, M L, Ruiters MHJ, Xu G, Rots MG. Induced DNA demethylation by targeting Ten-Eleven Translocation 2 to the human ICAM-1 promoter. Nucleic Acids Research. 2014; (3):1563–157442. [PubMed: 24194590]
- Choi Y, Chae C. Demethylation of somatic and testis-specific histone H2A and H2B genes in F9 embryonal carcinoma cells. Molecular and Cellular Biology. 1993; 13(9):5538–5548. [PubMed: 8355699]
- Deaton A, Bird A. CpG islands and the regulation of transcription. Genes and Development. 2011; 25(10):1010–1022. [PubMed: 21576262]
- De Smet C, Lurquin C, Lethe B, Martelange V, Boon T. DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. Molecular and Cellular Biology. 1999; 19(11):7327–7335. [PubMed: 10523621]

- Domingue-Sanchez M, Barroso S, Gomer-Gonzalez B, Luna R, Aguilera A. Genomic Instability and Transcription Elongation Impaired in Human cells Depleted of THO/TREX. *Plos Genetics*. 2011; 7(12):1–17.
- Feng S, Cokus S, Zhang X, Chen P, Bostick M, Goll M, Hetzel J, Jain J, Strauss S, Halpern M, Ukomadu C, Sadler K, Pradhan S, Pellegrini M, Jacobsen S. Conservation and divergence of methylation patterning in plants and animals. *The Proceedings of the National Academy of Science USA*. 2010; 107(19):8689–8694.
- Gardiner M, Frommer M. Transcripts and CpG islands associated with the Proopiomelanocortin gene and other neutrally expressed genes. *Journal of Molecular Endocrinology*. 1994; 12(3):365–382. [PubMed: 7916974]
- Geering K. Subunit assembly and functional maturation of Na,K ATPase. *Journal of Membrane Biology*. 1990; 115(2):109–121. [PubMed: 2162391]
- Godmann M, May E, Kimmins S. Epigenetic Mechanisms Regulate Stem Cell Expressed Genes Pou5f1 and Gfra1 in a Male Germ Cell Line. *Plos one*. 2010; 5(9):1–17.
- Griffiths, AJF.; Miller, JH.; Suzuki, DT.; Lewontin, RC.; Gelbart, WM. *An Introduction to Genetic Analysis*. 7th edition. New York: W H Freeman and Company; 2000. Transcription: An overview of gene regulation in eukaryotes.
- Hattori N, Nishino K, Ko Y, Ohgane J, Tanaka S, Shiota K. Epigenetic control of mouse Oct-4 gene expression in embryonic stem cells and trophoblast stem cells. *Journal of Biological Chemistry*. 2004; 279(17):17063–17069. [PubMed: 14761969]
- Hebbar PR, Archer TK. Altered histone H1 stoichiometry and absence of nucleosome positioning on transfected DNA. *Journal of Biological Chemistry*. 2008; 283(8):4595–4601. [PubMed: 18156629]
- He J, Shen L, Wan M, Taranova O, Wu H, Zhang Y. Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. *Nature Cell Biology*. 2013; 15:373–384. [PubMed: 23502314]
- Herzing L, Romer J, Horn J, Ashworth A. Xist has properties of the X-chromosome inactivation center. *Nature*. 1997; 386(6622):272–275. [PubMed: 9069284]
- Hlivko J, Chakraborty S, Hlivko T, Sengupta A, James P. The human Na,KATPase alpha4 isoform is a ouabain-sensitive alpha isoform that is expressed in sperm. *Molecular Reproduction and Development*. 2006; 73(1):101–115. [PubMed: 16175638]
- Hofmann M, Narisawa S, Hess R, Millan J. immortalization of germ-cells and somatic testicular cells using SV40 Large T-antigen. *Experimental Cell Research*. 1992; 201(2):417–435. [PubMed: 1322317]
- Hon GC, Hawkins R D, Caballero OL, Lo C, Lister R, Pelizzola M, Valsesia A, Ye Z, Kuan S, Edsall LE, Camargo AA, Stevenson BJ, Ecker JR, Bafna V, Strausberg RL, Simpson AJ, Ren B. Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. *Genome Research*. 2012; 22:246–258. [PubMed: 22156296]
- Hong S, Kang M, Oh J, Jung Y, Kim Y, Kim S, Choi S, Seo E, Choi S, Rhyu M. DNA Methylation and Expression Patterns of Key Tissue-specific Genes in Adult Stem Cells and Stomach Tissues. *Journal of Korean Medical Sciences*. 2009; 24:918–929.
- Iannello R, Gould J, Young J, Giudice A, Medcalf R, Kola I. Methylation-dependent silencing of the testis-specific Pdha-2 basal promoter occurs through selective targeting of an activating transcription factor/cAMP-responsive element-binding site. *Journal of Biological Chemistry*. 2000; 275(26):19603–19608. [PubMed: 10766751]
- Illingworth R, Gruenewald-Schneider U, Webb S, Kerr A, James K, Turner D, Smith C, Harrison D, Andrews R, Bird A. Orphan CpG Islands Identify Numerous Conserved Promoters in the Mammalian Genome. *Plos Genetics*. 2010; 6(9)
- Illingworth R, Kerr A, DeSousa D, Jorgensen H, Ellis P, Stalker J, Jackson D, Clee C, Plumb R, Rogers J, Humphray S, Cox T, Langford C, Bird A. A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *Plos Biology*. 2008; 6(1):37–51.
- Jimenez T, McDermott J, Sanchez G, Blanco G. Na,K-ATPase alpha 4 isoform is essential for sperm fertility. *Proceedings of the National Academy of Science USA*. 2011; 108(2):644–649.

- Jones P. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*. 2012; 13(7):484–492.
- Kanber D, Berulava T, Ammerpohl O, Mitter D, Richter J, Siebert R, Horsthemke B, Lohmann D, Buiting K. The Human Retinoblastoma Gene Is Imprinted. *Plos Genetics*. 2009; 5(12)
- Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, Sasaki H. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature*. 2004; 429(6994):900–903. [PubMed: 15215868]
- Kobayashi H, Sakurai T, Imai M, Takahashi N, Fukuda A, Yayoi O, Sato S, Nakabayashi K, Hata K, Sotomaru Y, Suzuki Y, Kono Y. Contribution of Intragenic DNA Methylation in Mouse Gametic DNA Methylomes to Establish Oocyte-Specific Heritable Marks. *Plos Genetics*. 2012; 8(1)
- Kouzarides T. Chromatin modifications and their function. *Cell*. 2007; 128:693–705. [PubMed: 17320507]
- Kuroda A, Rauch T, Todorov I, Ku H, Al-Abdullah I, Kandeel F, Mullen Y, Pfeifer G, Ferreri K. Insulin Gene Expression Is Regulated by DNA Methylation. *Plos One*. 2009; 4(9)
- Lingrel J, Kuntzweiler T. Na⁺K⁺-ATPase. *Journal of Biological Chemistry*. 1994; 269(31):19659–19662. [PubMed: 8051040]
- Lingrel J, Moseley A, Dostanic I, Coughon M, He S, James P, Woo A, O'Connor K, Neumann J, Jorgensen P, Karlsh S, Maunsbach A. Functional roles of the alpha isoforms of the Na,K-ATPase. Na,K-ATPase and Related Cation Pumps. *Annals of the New York Academy of Sciences*. 2003; 986:354–359. [PubMed: 12763850]
- Lister R, Pelizzola M, Dowen R, Hawkins R, Hon G, Tonti-Filippini J, Nery J, Lee L, Ye Z, Ngo Q, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar A, Thomson J, Ren B, Ecker J. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009; 462(7271):315–322. [PubMed: 19829295]
- Livak K, Schmittgen T. Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods*. 2001; 25(4):402–408. [PubMed: 11846609]
- Lorincz M, Dickerson D, Schmitt M, Groudine M. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nature Structural and Molecular Biology*. 2004; 11(11):1068–1075.
- Lo WS, Duggan L, Tolga Emere NC, Belotserkoyskya R, Lane WS, Shiekhatar R, Berger SL. A histone kinase that works in concert with the histone acetyltransferase Gen5 to regulate transcription. *Science*. 2001; 293:1142–1146. [PubMed: 11498592]
- Macleod D, Ali R, Bird A. An alternative promoter in the mouse major histocompatibility complex class II I-A beta gene: Implications for the origin of CpG islands. *Molecular and Cellular Biology*. 1998; 18(8):4433–4443. [PubMed: 9671453]
- Maeder ML, Angstman JF, Richardson ME, Linder SJ, Cascio VM, Tsai SQ, Ho QH, Sander JD, Reyon D, Bernstein BE, Costello JF, Wilkinson MF, Joung JK. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nature Biotechnology*. 2013; 31(12):1137–1144.
- Mancini-DiNardo D, Steele S, Ingram R, Tilghman S. A differentially methylated region within the gene Kcnq1 functions as an imprinted promoter and silencer. *Human Molecular Genetics*. 2003; 12(3):283–294. [PubMed: 12554682]
- Maunakea A, Nagarajan R, Bilenky M, Ballinger T, D'Souza C, Fouse S, Johnson B, Hong C, Nielsen C, Zhao Y, Turecki G, Delaney A, Varhol R, Thiessen N, Shchors K, Heine V, Rowitch D, Xing X, Fiore C, Schillebeeckx M, Jones S, Haussler D, Marra M, Hirst M, Wang T, Costello J. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature*. 2010; 466(7303):253–U131. [PubMed: 20613842]
- McDermott J, Sanchez G, Chennathukuzhi V, Blanco G. Green fluorescence protein driven by the Na,K-ATPase alpha 4 isoform promoter is expressed only in male germ cells of mouse testis. *Journal of Assisted Reproduction and Genetics*. 2012; 29(12):1313–1325. [PubMed: 23229519]
- McDonough A, Geering K, Farley R. The Sodium-pump needs its beta-subunit. *FASEB Journal*. 1990; 4(6):1598–1605. [PubMed: 2156741]

- Metivier R, Gallais R, Tiffocche C, Le Peron C, Jurkowska RZ, Carmouche RP, Ibberson D, Barath P, Demay F, Reid G, Benes V, Jeltsch A, Gannon F, Salbert G. Cyclic DNA methylation of a transcriptionally active promoter. *Nature*. 2008; 452(7183):45–50. [PubMed: 18322525]
- Muller C, Readhead C, Diederichs S, Idos G, Yang R, Tidow N, Serve H, Berdel W, Koeffler H. Methylation of the cyclin A1 promoter correlates with gene silencing in somatic cell lines, while tissue-specific expression of cyclin A1 is methylation independent. *Molecular and Cellular Biology*. 2000; 20(9):3316–3329. [PubMed: 10757815]
- Newton D, Kastelic P, Wong B, van der Hoorn F, Thundathil J. Elevated testicular temperature modulates expression patterns of sperm proteins in Holstein bulls. *Molecular Reproduction and Development*. 2009; 76(1):109–118. [PubMed: 18459118]
- Nowak S, Corces V. Phosphorylation of histone H3 correlated with transcriptionally active loci. *Genes & Development*. 2000; 14:3003–3013. [PubMed: 11114889]
- 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(3):247–257. [PubMed: 10555141]
- Pinheiro A, Nunes M, Milagre I, Rodrigues E, Silva M, de Almeida I, Rivera I. Demethylation of the Coding Region Triggers the Activation of the Human Testis-Specific PDHA2 Gene in Somatic Tissues. *Plos One*. 2012; 7(6)
- Rinn J, Kertesz M, Wang J, Squazzo S, Xu X, Brugmann S, Goodnough L, Helms J, Farnham P, Segal E, Chang H. Functional demarcation of active and silent chromatin domains in human HOX loci by Noncoding RNAs. *Cell*. 2007; 129(7):1311–1323. [PubMed: 17604720]
- Rountree MR, Selker EU. DNA methylation inhibits elongation but not initiation of elongation in *Neurospora crassa*. *Genes & Development*. 1997; 11(18):2383–2395. [PubMed: 9308966]
- Sato S, Maeda C, Hattori N, Yagi S, Tanaka S, Shiota K. DNA Methylationdependent Modulator of Gsg2/Haspin Gene Expression. *Journal of Reproduction and Development*. 2011; 57(4):526–533. [PubMed: 21606629]
- Schmidl C, Klug M, Boeld T, Andreesen R, Hoffmann P, Edinger M, Rehli M. Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. *Genome Research*. 2009; 19(7):1165–1174. [PubMed: 19494038]
- Shaknovich R, Yeyati S, Ivins A, Melnick C, Lempert S, Waxman A, Ze-lent, Licht J. The promyelocytic leukemia zinc finger protein affects myeloid cell growth, differentiation, and apoptosis. *Molecular and Cellular Biology*. 1998; 18:5533–5545. [PubMed: 9710637]
- Skou J. The energy coupled exchange of Na⁺ for K⁺ across the cell membrane - The Na⁺,K⁺-pump. *FEBS Letters*. 1990; 268(2):314–324. [PubMed: 2166689]
- Sleutels F, Zwart R, Barlow D. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature*. 2002; 415(6873):810–813. [PubMed: 11845212]
- Southwood C, Lipovich L, Gow A. Tissue-Restricted Transcription from a Conserved Intragenic CpG Island in the Klf1 Gene in Mice. *Biology of Reproduction*. 2012; 87(5)
- Stadler MB, Murr R, Burger B, Ivanek R, Lienert F, Scholer A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D, Tiwari VK, Schubeler D. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature*. 2011; 480:490–495. [PubMed: 22170606]
- Sweadner K. Isozymes of the Na⁺/K⁺-ATPase. *Biochimica Et Biophysica Acta*. 1989; 988(2):185–220. [PubMed: 2541792]
- Thomas P, Smart TG. HEK293 cell line: a vehicle for expression of recombinant protein. *Journal of Pharmacological and Toxicological Methods*. 2005; 51(3):187–200. [PubMed: 15862464]
- Trasler J, Hake L, Johnson P, Alcivar A, Millette C, Hecht N. DNA methylation and demethylation events during meiotic prophase in the mouse testis. *Molecular and Cellular Biology*. 1990; 10(4): 1828–1834. [PubMed: 2320009]
- Weber M, Hellmann I, Stadler M, Ramos L, Paabo S, Rebhan M, Schubeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genetics*. 2007; 39(4):457–466. [PubMed: 17334365]
- Woo A, James P, Lingrel J. Characterization of the fourth alpha isoform of the Na,KATPase. *Journal of Membrane Biology*. 1999; 169(1):39–44. [PubMed: 10227850]
- Woo A, James P, Lingrel J. Sperm motility is dependent on a unique isoform of the Na,K-ATPase. *Journal of Biological Chemistry*. 2000; 275(27):20693–20699. [PubMed: 10764792]

- Xie W, Han S, Khan M, DeJong J. Regulation of ALF gene expression in somatic and male germ line tissues involves partial and site-specific patterns of methylation. *Journal of Biological Chemistry*. 2002; 277(20):17765–17774. [PubMed: 11889132]
- Yeyati PL, Shaknovich R, Boterashvili S, Li J, Ball JH, Waxman S, Nason-Burchenal K, Dmitrovsky E, Zelent A, Licht JD. Leukemia trans- location protein PLZF inhibits cell growth and expression of cyclin A. *Oncogene*. 1999; 18:925–934. [PubMed: 10023668]
- Yu C, Cudaback E, Foraker J, Thomson Z, Leong L, Lutz F, Gill J, Saxton A, Kraemer B, Navas P, Keene C, Montine T, Bekris L. Epigenetic signature and enhancer activity of the human APOE gene. *Human Molecular Genetics*. 2013; 22(24):5036–5047. [PubMed: 23892237]

List of abbreviations

bp	base pair
CGI	CpG island
CMV	human cytomegalovirus
Dnmt	DNA methyltransferase
DMEM	Dulbecco's Modified Eagles media
DKO	double knockout
ES cell	mouse embryonic stem cell
FBS	fetal bovine serum
HB-EGF-C	C-terminal fragment of heparin-binding EGF-like growth factor
HEK 293	human embryonic kidney 293
IACUC	University Institutional Animal Care and Use Committee
IV	<i>in vitro</i> methylated
kDa	kilodalton
KO	knockout
LIF	leukemia inhibitory factor
Mα4-CGI	intragenic CpG island
Mα4-Promoter	<i>Atp1a4</i> promoter
MM	mock methylated
PBS	phosphate buffered saline
PLZF	promyelocytic leukemia zinc finger
qPCR	quantitative real time polymerase chain reaction
TALEs	transcription activator like effectors
TSA	trichostatin A
TAN1	tandem reporter construct
TF	transcription factor

Tk	thymidine kinase
TSS	transcription start site
wt	wild type
ZFEs	zinc finger effectors
5-azaC	5-aza2-Deoxycytidine

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

- A differentially methylated intragenic CpG island is present in the *Atp1a4* gene.
- Treatment of GC-1spg cells with 5-azaC leads to upregulation of *Atp1a4* gene expression.
- Maintenance and *de novo* methylation are involved in regulation of *Atp1a4* expression in ES cells.
- Methylation of the M α 4-Promoter regulates its activity.
- Methylation of the intragenic M α 4-CGI alters transcription elongation efficiency.
- The M α 4-CGI displays silencer activity which may serve to fine tune *Atp1a4* expression in germ cells.

Figure 1. Location and methylation pattern of the intragenic CpG island and promoter region in the *Atp1a4* locus

A: Upper panel: The position of the Ma4-CGI relative to the transcription start site (TSS) of *Atp1a4* is indicated by the black box (corresponding to the genome coordinates Chr1:174,166,639-174,167,742, of the mm9 mouse genome assembly). In addition, the region analyzed as a putative promoter (-700bp to +200bp relative to the TSS; corresponding to the genome coordinates Chr1:174,152,939-174,153,139, of the mm9 mouse genome assembly) for the gene is indicated by the line under the region surrounding transcriptional start site (indicated by the arrow located at +1). **B:** Methylation analysis of Ma4- CGI (displayed as four non-overlapping sub-regions I, II, III and IV with total of 43 CpGs) in mouse sperm and kidney. **C:** Methylation analysis of Ma4-Promoter (displayed as three nonoverlapping sub-regions I, II and III with total of 18 CpGs) in mouse sperm and kidney. Open circles represent unmethylated CpGs and closed circles represent methylated CpGs. Columns represent individual CpGs and rows represent individual clones. For each sample (kidney and sperm) a total of twenty clones, from two biological replicates, were analyzed by bisulfite sequencing.

Figure 2. 5-Aza2-Deoxycytidine (5-azaC) treatment but not Trichostatin A (TSA) treatment upregulates *Atp1a4* expression and reduces DNA methylation in GC-1spg cells

A: Quantitative real time PCR analysis of mouse $\alpha 4$ mRNA expression levels in GC-1spg cells treated with 5-azaC, TSA and both 5-azaC and TSA relative to the respective vehicle-treated (0 μ M 5-azaC, 0 nM TSA and 0 μ M 5-azaC + 0 nM TSA) GC-1spg control cells. The C_t method was used to calculate the fold change in the mRNA levels. The data are presented as the mean $\pm$ SD of three different biological replicates (n=3). (*) indicates a statistically significant difference with a p value <0.0005. **B:** Methylation analysis of the M α 4-CGI region containing the largest number of CpGs (M α 4-CGI sub-region IV, refer to Figure 1) in 5 μ M 5-azaC-treated GC-1spg cells and 0 μ M 5-azaC-treated GC-1spg control cells. **C:** Methylation analysis of the M α 4- Promoter in 5 μ M 5-azaC-treated GC-1spg cells and 0 μ M 5-azaC treated GC-1spg control cells. The numbers in parentheses represent the % change in methylated CpGs relative to the 0 μ M 5-azaC-treated GC-1spg control cells. For each condition (5-azaC-treated and vehicle-treated) a total of ten clones, from one biological replicate, were analyzed by bisulfite sequencing.

Figure 3. Loss of DNA methyltransferase enzyme 1 (Dnmt1) results in upregulation of *Atp1a4* expression and reduced DNA methylation in mouse ES cells

A: Quantitative real time PCR showing expression of the Mα4 mRNA transcript in Dnmt1 KO ES cells compared to J1 wt ES cells. (*) represents a statistically significant difference with a p value of < 0.005. The data are presented as the mean $\pm$ SD of three different biological replicates (n=3). **B:** Methylation analysis of Mα4-CGI in the J1 wt ES cells and Dnmt1KO ES cells. **C:** Methylation analysis of the Mα4-Promoter in J1 wt ES cells and Dnmt1 KO ES cells. The numbers in parentheses represent the % change in methylated CpGs relative to the J1 wt ES cells. For J1 wt and Dnmt1KO ES cells a total of ten clones, from one biological replicate, were analyzed by bisulfite sequencing.

Figure 4. Loss of *de novo* methyltransferase enzymes (Dnmt3a and Dnmt3b) results in increased *Atp1a4* expression and reduced DNA methylation in mouse ES cells

A: Quantitative real time PCR showing expression of the Mα4 mRNA transcript in Dnmt3a KO, Dnmt3b KO and Dnmt3a/3b DKO ES cells. (*) and (**) represent statistically significant differences with p values of < 0.0005 and < 0.005 respectively. The data are presented as the mean ± SD of three different biological replicates (n=3). The C_t method was used to calculate the fold change in the mRNA levels. **B:** Methylation analysis of Mα4-CGI in Dnmt3a KO, Dnmt3b KO, Dnmt3a/3b DKO and J1 wt ES cells. **C:** Methylation analysis of the Mα4-Promoter region in Dnmt3a KO, Dnmt3b KO, Dnmt3a/3b - DKO and J1 wt ES cells. The numbers in parentheses represent the % change in methylated CpGs relative to the J1 wt ES cells. For J1 wt and all the KO ES cells, a total of ten clones from one biological replicate, were analyzed by bisulfite sequencing.

Figure 5. The Mα4-Promoter is a cell type-specific promoter that displays methylation dependent activity

A: Normalized firefly luciferase activity was determined using the dual luciferase reporter assay following transfection of the pGL3-Basic-Mα4-Promoter vector (refer to construct diagram above the graph; open lollipops represent unmethylated CpGs) in GC-1spg cells and NIH3T3 cells. The normalized firefly luciferase reporter activity in cells transfected with the pGL3-Mα4-Promoter vector is represented as fold change over the same activity determined in cells transfected with the promoter-less pGL3-Basic-vector (set at 1.0). **B:**

The effect of *in vitro* methylation on activity of the M α 4-Promoter was determined following mock methylation (MM) and *in vitro* methylation (IV) of the pGL3-M α 4-Promoter-vector (refer to construct diagram above the graph; open lollipops represent unmethylated CpGs and closed lollipops represent methylated CpGs). The normalized firefly luciferase activity from cells transfected with pGL3-M α 4-Promoter-IV is represented as fold change compared to the same activity determined in cells transfected with pGL3-M α 4-Promoter-MM (which is set at 1.0). The data are represented as mean $\pm$ SD of four independent transfections (n=4) performed in duplicates. (*) represents a statistically significant difference with a p value of < 0.05.

Figure 6. The Mα4-CGI does not function as an alternative promoter in GC-1spg cells or NIH3T3 cells

Normalized firefly luciferase activity was determined using the dual luciferase reporter assay following transfection of the pGL3-Basic-Mα4-CGI vector (refer to construct diagram above the graph; open lollipops represent unmethylated CpGs) in GC-1spg cells and NIH3T3 cells. The normalized firefly luciferase activity from cells transfected with pGL3-Basic-Mα4-CGI vector is represented as fold change compared to the same activity determined in cells transfected with the promoter-less pGL3-Basic vector (which is set at 1.0). The data are presented as mean $\pm$ SD of four independent transfections (n=4) performed in duplicates. A twotailed t-test was performed to calculate p values and the data was found not to be statistically different.

Figure 7. Ma4-CGI methylation decreases transcription elongation efficiency in somatic cells
A: TAN1 reporter construct employed to determine the effect of Ma4-CGI methylation on transcriptional elongation efficiency. **B:** The dual luciferase assay was utilized to determine the effect of *in vitro* methylation of the Ma4-CGI on transcription elongation efficiency in HEK 293 cells transfected with the mock methylated (MM) or *in vitro* methylated (IV) TAN1 reporter construct. The *Renilla*/firefly activity ratio of TAN1-Ma4-CGI-IV was normalized to that of TAN1-Ma4-CGI-MM (which is set at 1.0). The data are represented as mean $\pm$ SD of eight independent transfections (n=8) performed in duplicates. (*) represents a statistically significant difference with a p value < 0.0005.

Figure 8. The Mα4-CGI functions as a silencer in GC-1spg cells

The three constructs, pGL3-Mα4-Promoter-Mα4-CGI, pGL3-SV40-Promoter-Mα4-CGI and the CpG-free pCpGL-EF1α-Promoter-Mα4-CGI and their respective promoter-only control constructs were transfected along with pRL-TK vector into GC-1spg cells (refer to construct diagram above the graph; open lollipops represent unmethylated CpGs). The normalized firefly luciferase activities were determined 48 hr later and each was normalized to its promoter-only control (which were each set at 1.0). The data are presented as mean $\pm$ SD of four independent transfections (n=4) performed in duplicates. (*) and (**) represent statistically significant differences with p values of <0.05 and <0.005 respectively.

Figure 9. Methylation of Ma4-CGI does not alter its silencer activity with respect to the Ma4-Promoter in GC-1spg cells

A: Schematic representation of the luciferase constructs. The lollipops represent individual CpGs. Open lollipops represent unmethylated CpGs and closed lollipops represent methylated CpGs. **B:** The pGL3-Ma4-Promoter and the pGL3-Ma4-Promoter-Ma4-CGI vectors were either mock-methylated (MM) or *in vitro* methylated (IV) and transfected in the GC-1spg cells along with pRL-TK *Renilla* vector. Normalized firefly luciferase activities were determined and normalized to the same activity found in cells transfected with the pGL3-Ma4-Promoter MM (which is set at 1.0). The data are presented as mean $\pm$ SD of four independent transfections (n=4) performed in duplicates. (*) represents a statistically significant difference with a p value < 0.05.

A: Somatic cells **B: Germ cells**

Figure 10. A model for the potential regulation of *Atp1a4* expression by the Ma4-Promoter and Ma4-CGI

A: Somatic cells: Ma4-Promoter is a cell type-specific promoter and is not active in somatic cells. In addition, since the Ma4-CGI represses the activity of the Ma4-Promoter in germ cells, it is possible that this CGI will also repress any aberrant activity of this promoter in somatic cells. Methylation at the intragenic Ma4-CGI serves to inhibit transcriptional elongation in HEK 293 cells. Hence any $\alpha 4$ transcript that might be initiating as result of aberrant activity of the methylated Ma4-Promoter (indicated by the “X” on the arrow at the TSS) will be blocked at the transcription elongation step thus providing a mechanism for completely shutting “Off” $\alpha 4$ expression in somatic cells. **B: Germ cells:** The unmethylated Ma4-Promoter is an active promoter in GC-1spg cells. Also the activity of the Ma4-Promoter is regulated by methylation. Hence $\alpha 4$ transcript initiation driven by the unmethylated Ma4-Promoter in germ cells is active. Following initiation, transcription elongation proceeds uninterrupted through the unmethylated Ma4-CGI resulting in “High” $\alpha 4$ expression (indicated by the black full length primary transcript). The unmethylated Ma4-CGI also functions to silence the activity of the unmethylated Ma4-Promoter in GC-1spg cells. To explain this seemingly contradictory finding, we postulate the existence of a mechanism that would specifically inhibit the repressor function of the CGI in germ cells (not illustrated in figure, see text for details) and would allow for “optimal” $\alpha 4$ expression.

Table1.1

Bisulfite sequencing primer list.

Bisulfite M4-CGI Forward 1	5'tgtgggggttataggaaatatt3'
Bisulfite M4-CGI Reverse 1	5' ttgaagtggaaagaggcacct3'
Bisulfite M4-CGI Forward 2	5'aattgtttggtgaagaatttgg3'
Bisulfite M4-CGI Reverse 2	5' agtgtgggggttataggaaatatt 3'
Bisulfite M4-CGI Forward 3	5' gaaaagaggtttttgaatgaaaga 3'
Bisulfite M4-CGI Reverse 3	5' aattgtttggtgaagaatttgg 3'
Bisulfite M4-CGI Forward 4	5' ggggggttttttagtattgtgat 3'
Bisulfite M4-CGI Reverse 4	5' ctatcacaacacaaactcctt 3'
Bisulfite M4-Promoter Forward 1	5' ttgtatagatagaataggagtggg 3'
Bisulfite M4-Promoter Reverse 1	5'ttccaacttactaaaaccaa3'
Bisulfite M4-Promoter Forward 2	5' tgggttttagtagagttggaaa 3'
Bisulfite M4-Promoter Reverse 2	5' atactccaaccctaaatcct 3'
Bisulfite M4-Promoter Forward 3	5' tagggttggagtagtaggttt 3'
Bisulfite M4-Promoter Reverse 3	5' aaactctacctcctttctttt3'

Table 1.2

qPCR primer list.

qPCR α4 forward	5' gcaagacacgcagaaattca 3'
qPCR α4 reverse	5' cctggaacgtaggacaggaa 3'
qPCR Gapdh forward	5 ' gacatgccgcctggagaaac 3'
qPCR Gapdh reverse	5' agccaggatgcccttagt 3'