De novo methylation of an embryonic globin gene during normal development is strand specific and spreads from the proximal transcribed region

Rakesh Singal and Jane M. vanWert

DNA methylation in eukaryotes involves addition of a methyl group to the carbon 5 position of the cytosine ring. This reaction is catalyzed by DNA methyltransferase in the context of the sequence 5′-CG-3′, which is also referred to as a CpG dinucleotide. Eukaryotic genomes are not methylated uniformly but contain methylated regions interspersed with unmethylated domains. Approximately 70% to 80% of the CpG residues in most vertebrates are methylated. In contrast to the rest of the genome, smaller regions of DNA called CpG islands are unmethylated and possess the expected CpG frequency. During early development a dramatic reduction in methylation levels occurs in the preimplantation embryo. This is followed by a wave of de novo methylation involving most CpG residues. However, CpG island–associated promoter regions are protected from methylation by mechanisms which remain unclear. The methylation profile of genes in the adult is stable over many cell generations. Genomic methylation patterns are conserved after DNA replication by the DNA methyltransferase Dnmt1, which is the major maintenance methyltransferase. Dnmt1 is recruited to replicating DNA to reproduce the methylation pattern of the parental strands in the daughter strands. Inactivation of the mouse Dnmt1 gene by gene targeting resulted in extensive demethylation of all sequences examined. However, ES cells completely lacking Dnmt1 were still capable of methylating retroviral DNA de novo. The search for the de novo methyltransferases led to the discovery of Dnmt3a and Dnmt3b. These were found to be essential for de novo methylation and for mouse development. However, it remains unclear how de novo methylation patterns are established during development, over what time intervals these changes occur, or if this process involves any strand- or sequence-specificity.

A recent study examined the methylation profile of the mouse skeletal α-actin promoter during development and differentiation. It was found that remethylation of the α-actin promoter after implantation occurs in a stochastic pattern, with some molecules being extensively methylated and others sparsely methylated. It also found that tissue-specific expression of the skeletal α-actin gene did not correlate with the methylation state of the promoter. In contrast, we have previously shown that every cytosine within each CpG dinucleotide in a 235–base pair (bp) region of the embryonic p-globin gene promoter is methylated in normal adult (definitive) erythroid cells in which the gene is silent and completely unmethylated in 5-day (primitive) erythroid cells in which the gene is actively transcribed. Other studies have also demonstrated a strong inverse correlation between the methylation status of DNA sequences near globin genes and the transcriptional activity of these genes in different tissues. To understand the de novo methylation of the p-globin gene, we have examined the strand-specific methylation pattern of the p-globin gene promoter, proximal transcribed region, and distal transcribed region on both DNA strands during development in chicken erythroid cells.

Materials and methods

Blood collection

Eggs were purchased from Truslow Farms (Chestertown, MD) and incubated in a Lyon Roll-X Automatic Incubator (Lyon Electric, Chula Vista, CA) according to the manufacturer’s instructions. Blood was collected with a sterile Pasteur pipette into room temperature phosphate buffered saline (PBS), washed twice with PBS, and spun at 320g for 5
minutes. Red blood cells (RBCs) were then resuspended in PBS and spun for 5 minutes at 720g to pellet cells.

**RNA/DNA purification**

Cells were resuspended in 10 volumes of RNA STAT-60 (Tel-Test, Friendswood, TX) and RNA and DNA were extracted according to the manufacturer’s protocol. Briefly, cells were homogenized then the lysate was extracted with chloroform and the RNA was precipitated with isopropanol. DNA was precipitated with ethanol from the organic phase left over from the chloroform extraction, the pellet was washed 3 times with 100 mM sodium citrate in 10% ethanol then precipitated with ethanol and solubilized in 8 mM sodium hydroxide.

**Reverse transcriptase–polymerase chain reaction**

Reverse transcriptase–polymerase chain reaction (RT-PCR) was carried out using the Titan One-Tube RT-PCR System (Roche Molecular Biochemicals, Alameda, CA) as per the manufacturer’s instructions.

**Bisulfite conversion and methylation analysis**

Bisulfite conversion and methylation analysis was carried out as previously described. The same template was used for methylation analysis of different regions of the P-globin gene on both DNA strands. Sequencing of the PCR-amplified product was performed using the forward and reverse primers. The α-32P–labeled ddNTP terminator kit (United States Biochemical, Cleveland, OH) was used for sequencing. The sequencing gel was dried and exposed to a phosphorimager screen (Packard Instrument, Meriden, CT). Methylation analysis was carried out by quantitating the intensity of C and T bands using optiquant software (Packard Instrument), and calculating the percentage of C/C + T bands. At least 4 CpG sites were analyzed in each region and standard error of mean was calculated. The same sites were used for methylation analysis on the other DNA strand.

**Primer sequences**

See Table 1 for primer sequences for RT-PCR and bisulfite genomic sequencing.

**Results**

We examined the expression of avian β-type embryonic P-globin and adult β4-globin during development in primary erythroid cells using RT-PCR. P-globin mRNA is easily detected in 4-day to 5-day primitive embryonic erythroid cells but is barely detectable by day 11 of embryonic development (Figure 1). In contrast, the β4-globin starts expressing day 5 to day 6 and continues through adult life. These results are consistent with the earlier studies that examined the expression of avian β-type globin genes during development in erythroid cells.

We have previously shown that the P-globin gene promoter is completely unmethylated in primitive erythroid cells and completely methylated in erythroid cells from adult chickens. To elucidate the strand- and sequence-specificity of DNA methylation of the P-globin gene during development, we employed the bisulfite genomic sequencing method. This technique is based on bisulfite-induced oxidative deamination of genomic DNA under conditions in which cytosine is converted to uracil and 5 mC remains unchanged. The target sequence is amplified by PCR using strand-specific primers. Upon sequencing of the amplified DNA, all uracil and thymine residues become detectable as thymine and only 5 mC residues amplify as cytosines. Unlike restriction enzyme–based techniques, bisulfite genomic sequencing also permits an independent methylation analysis of the 2 strands of DNA in a given sequence. DNA strands are no longer complementary after bisulfite treatment because of the conversion of unmethylated cytosines to uracils. The 2 strands were amplified separately with strand-specific primers. These primers amplify bisulfite-treated DNA irrespective of the methylation pattern of the CpG dinucleotides. Methylation analysis was carried out independently for the 2 strands using the same bisulfite-treated genomic DNA as a template. The percentage methylation refers to the overall methylation pattern of the CpG dinucleotides in the genomic DNA derived from red cells of several dozen chicken embryos.

We determined the temporal methylation pattern on both DNA strands of 3 different regions of the P-globin gene (ie, promoter, proximal transcribed region or exon 1, and distal transcribed region or exon 3) during development in primary erythroid cells. The promoter and proximal transcribed regions are CpG dense and constitute a CpG island, whereas the distal transcribed region has normal CpG density (Figure 2).

![Figure 1. Expression of P- and β4-globin genes during development in chicken erythroid cells.](image-url)
The CpG dinucleotides in the distal transcribed region are completely unmethylated on both DNA strands until day 5 (Figure 3 and Figure 6C). Methylation progresses on both strands in a similar fashion and is complete in DNA from adult erythroid cells.

We next examined the methylation pattern of the proximal transcribed region (exon 1). Interestingly, methylation starts in this region as early as day 5, but methylation of the template strand lags behind that of the coding strand by almost 48 hours (Figure 4 and Figure 6B). As a control, fully unmethylated DNA from 4-day-old chicken embryonic erythroid cells was mixed with fully methylated DNA from adult erythroid cells in varying proportions, and methylation analysis of the proximal transcribed region was carried out. Methylation percentages were similar on 2 DNA strands in these samples (data not shown). Next, we examined the methylation pattern of the \( \beta \)-globin gene promoter region. The template and coding strands become methylated in a similar fashion (Figure 6A). However, methylation of the promoter lags as compared with the pattern observed in the proximal transcribed and distal transcribed regions (Figure 6D). To exclude PCR bias, the \( \beta \)-globin promoter and exon 1 regions were amplified with a single set of primers using bisulfite-treated DNA from day 8 erythroid cells as a template. As shown in Figure 5, the proximal transcribed region is methylated to a greater degree than the promoter region and these unmethylated regions indicated by letters a (promoter), b (proximal transcribed region or exon 1) and c (distal transcribed region or exon 3).

Promoter region CpG islands are usually unmethylated in all normal tissues, regardless of the transcriptional activity of the gene.\(^1,5\) The main exceptions include nontranscribed genes on the inactive X-chromosome and silenced alleles of imprinted genes.\(^1,5\) Recently, methylation has been proposed as the primary control mechanism for certain germ-line–specific genes with CpG-rich promoters.\(^2,3\) Here we show that, in the case of the developmentally regulated \( \beta \)-globin gene, methylation of both the \( \beta \)-globin gene promoter and proximal transcribed region and \( \beta \)-globin-poor (distal transcribed region) regions correlate inversely with the stage-specific expression in avian erythroid cells.

**Discussion**

This is the first study to report the strand-specific methylation pattern of a tissue-specific gene during development in primary erythroid cells. We report several interesting observations. The methylation pattern of the CpG-dense 5’ region and the CpG-poor 3’ region of the \( \beta \)-globin gene correlate inversely with stage-specific expression in primary avian erythroid cells. De novo methylation targets the CpG-dense proximal transcribed region on the coding (top) strand initially, followed by spread into the 3’ transcribed region and into the promoter region. Methylation of the template (bottom) strand lags behind that of the coding strand, and complete methylation of both strands occurs after the gene has been silenced. Our results indicate that establishment of the de novo methylation pattern involves strand-specificity and methylation spreading.

The distribution of methylated and unmethylated CpG dinucleotides in vertebrates conforms to a generalized pattern. About 70% to 80% of CpG sites contain methylated cytosines.\(^3\) The main exceptions include nontranscribed genes on the inactive X-chromosome and silenced alleles of imprinted genes.\(^1,5\) Hemi-methylation in the DNA of eukaryotic cells has been reported for maintenance methylation, where a tight coordination of DNA methylation and replication has been shown.\(^26\) Synthesis of embryonic globin chains in definitive erythroid cells has been shown to decrease with ontogeny.\(^27\) It is possible that in early definitive erythroid cell progenitors transcription through the \( \beta \)-globin gene prevents methylation of the template strand.
maturation, changes in the balance of positive and negative trans-acting factors results in silencing of \( \beta \)-globin gene transcription. This would then allow the methyltransferases to gain access to the template strand. With successive cell divisions, CpG dinucleotides in the \( \beta \)-globin gene may then become completely methylated. In the case of an integrated adenovirus in a mammalian cell line, a few 5'-CG-3' sequences can remain hemi-methylated for several cell generations before methylation involves both strands. Our results suggest that a similar phenomenon may exist for de novo methylation of a tissue-specific gene during normal embryonic development.

Recent studies have suggested that strand-specific methylation could be important in the understanding of molecular mechanisms targeting DNA methylation. A zinc finger protein, CCCTC-binding factor (CTCF) mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus.\(^{28,29}\) Interestingly, methylation of the top but not the bottom strand of radiolabeled oligonucleotide probes derived from mouse and human imprint control regions inhibited CTCF binding, indicating that CTCF makes important contacts with some of the cytosine residues on the top DNA strand.\(^{29}\) Four (MeCP2, MBD1, MBD2, and MBD3) of the 5 known proteins with methyl-CpG–binding domain are implicated in transcriptional repression.\(^{8}\) MBD2 and MBD3 form homo- and hetero-dimers (or multimers) in vitro and in vivo. Significantly, the MBD2-MBD3 complex showed an affinity to hemi-methylated DNA.\(^{30}\)

The endogenous targets for de novo methylation in the
vertebrate genome remain to be determined. A previous study with transgenic mice has shown that the bacterial lacI gene acts as a target for de novo methylation. When the GC content of this transgene was decreased to more closely resemble the GC content in mammalian cells, without altering the encoded amino acid sequence, methylation of the transgene was significantly reduced. It has been proposed that clustering of CpG sites may act as a target for de novo methylation. In the methylated at high levels in somatic cells. -deoxycytidine (5-Aza-CdR) treatment. T24 cells after 5-aza-

The spreading of methylation from the foci of methylated CpG sites has been demonstrated in vitro in cultured cells and in transgenic mice. Methylation spreading has also been observed to occur as a function of aging and neoplasia. In the vertebrate genome remain to be determined. A previous study with transgenic mice has shown that the bacterial lacI gene acts as a target for de novo methylation. When the GC content of this transgene was decreased to more closely resemble the GC content in mammalian cells, without altering the encoded amino acid sequence, methylation of the transgene was significantly reduced. It has been proposed that clustering of CpG sites may act as a target for de novo methylation. Our results are consistent with this hypothesis since methylation initiates in the CpG-dense proximal transcribed region of the p-globin gene. However, other target sequences for de novo methylation also exist. The tandem B1 repetitive elements located in a methylating center upstream of the mouse adenine phosphoribosyltransferase (Aprt) gene were shown to be targets for de novo methylation. Further, the human Ali repetitive elements are methylated at high levels in somatic cells.

References

De novo methylation of an embryonic globin gene during normal development is strand specific and spreads from the proximal transcribed region

Rakesh Singal and Jane M. vanWert