# Evolution of the pathways leading to synthesis and modification of DNA

T. Duranti[1], A. La Teana[1], T. Cacciamani[1], P. Volpe[2]

[1] Institute of Biochemistry, University of Ancona
[2] Department of Biology, University of Rome "Tor Vergata"
volpe@bio.uniroma2.it

KEY WORDS: Evolution, DNA polymerases, DNA methyltransferases, restriction-modification, regulation of transcription

## Abstract

*A previous study showed that in eukaryotes the pathways leading to synthesis and post-synthetic modification of DNA employ methionine as common donor of atoms: the carbon coming from the methyl group of this amino acid is needed for replication; its entire methyl group is needed to build 5-methylcytosine ($m^5C$) on semiconservatively newly replicating chains. This work showed that the two pathways originate in prokaryotes, where an enzymic machinery yields on DNA 6-methylaminopurine ($m^6A$) in addition to $m^5C$. The formation rate of $m^6A$ gradually decreased during the bacterial culture growth cycle (CGC), while that of $m^5C$ reached an optimum in its middle. This shift suggested that the deoxycytidylate (dcm) and deoxyadenylate (dam) methylasic activities, as well as the activities of the methylasic moieties of the restriction-modification (RM) enzymes, are uncoupled.*

## Introduction

Evolution of living matter is accompanied by that of post-synthetic modification of DNA, implying the formation on it of two classes of methylated bases: $m^6A$ and $m^5C$ in prokaryotes and $m^5C$ alone in mammals (Vanyushin et al., 1970). In prokaryotic DNA, while the presence of $m^6A$ was shown to be due partly to the activity of the *dam* DNA methylases (DNAmets) and partly to the activity of the type I, type II and type III methylating RM endonucleases involved in the digestion of the infecting $m^6A$-free phage DNA, the presence of $m^5C$ was shown to be due in part to the activity of the *dcm* DNAmets and in part to the

methylasic activity of the sole type II RM system involved in the digestion of the infecting $m^5C$-free phage DNA (Marinus, 1987; Arber, 1974). In higher cells, the well documented concentration of $m^5C$ in the promoters and introns (Volpe et al., 1974; Cacciamani et al., 2002) appeared to be involved in a mechanism modulating gene expression (Volpe et al., 1974; Cacciamani et al., 2002; Liau et al., 1981; Tentravahi et al., 1981). Thus, in concomitance with the disappearance of the 'archaic' enzyme systems combining *dcm* and *dam* DNAmets and RM endonucleases, characterizing the bacterial world (Marinus, 1987), molecular selection led to the loss, in eukaryotes, of $m^6A$ not only in nuclear but also in mitochondrial and chloroplastic DNAs (Volpe, 2005). The $m^5C$ was preserved in all these three DNA species by a novel methylasic protein family likely generated by an alternate splicing (Volpe et al., 1994; Franchina et al., 2001). The present research had two main purposes. The first was an attempt to verify whether bacteria are able to use methionine as a sole common donor of atoms for a convergence of DNA synthesis and methylation, as the eukaryotic cells do (Eremenko et al., 1985; Geraci et al., 1974). The second was an effort to verify whether, in bacteria, the timing for optimal yield of $m^5C$ coincides with that for optimal yield of $m^6A$.

## Methods

*E. coli* MRE 600 cultures were grown on *Petri* dishes by inoculating 0.5-1 ml of minimal (M9) medium, with $1 \times 10^{11}$ cells/ml, into 100 ml of Luria-Bertani (LB) medium. At intervals of time during the CGC, roughly lasting 10 hrs, DNA was labelled for 40 min at 37 °C with 2 $\mu$Ci/ml [$^{14}$C]methyl-L-methionine ([$^{14}$C]Met). The $^{14}$C-labelled DNA was extracted with minor modification of the method described by (Marmur, 1961). A 30-$\mu$g aliquot of purified DNA was introduced into a glass vial containing 300 $\mu$l of 88% formic acid (75% final concentration). The vial was closed with flame and left at 160 °C for 1 hr (Eremenko et al., 1985). The hydrolysed residues were dried, suspended in 30$\mu$l of 20 mM trichloroacetic acid at pH 2.2 and run at 13,000 rev./min for 5-10 min in an Eppendorf centrifuge. A 20-$\mu$l aliquot was percolated through the 131051 L ODS-55 Biorad Bio-Sil HPLC

column (250 x 4 mm) in reverse phase condition (Gehrke et al., 1978). Fractions of 0.4 ml were collected every 30 sec, while their content was resuspended in 50 μl of water and tested for radioactivity in a Packard Tri-Carb TRβ counter.

## Results

When percolated through the column, control bases, mixed in equimolar proportions, were eluted in the following time-order: C, m5C, G, T, A and m6A (Fig. 1a). There was however a striking quantitative difference between the
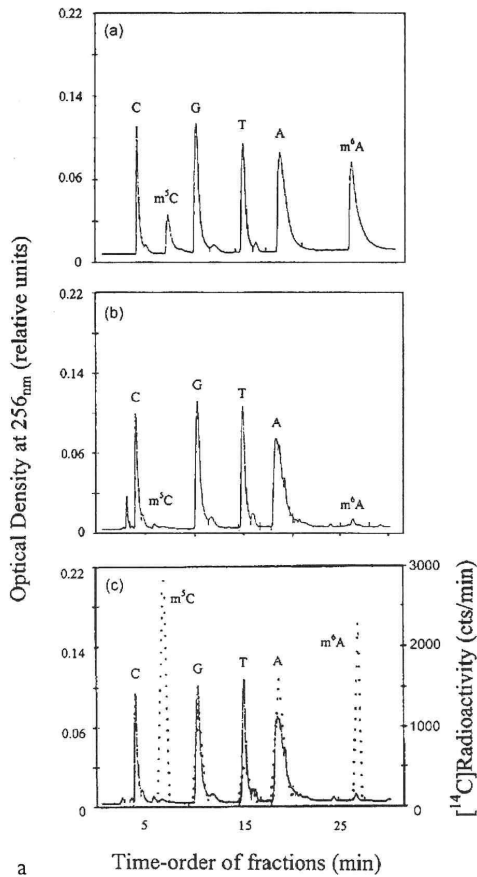


Fig. 1 - HPLC of the E. coli DNA hydrolysate. (a) Free bases from a standard solution. (b) Bases present in 30 μg of an unlabelled DNA hydrolysate extracted from a 2-ml suspension of 1x10⁸-1x10¹⁰ bacteria maintained in M9. (c) Bases present in 30 μg of a labelled DNA hydrolysate obtained from a 2-ml suspension of 1x10⁸-1x10¹⁰ bacteria maintained in LB. The left ordinate shows the concentration of bases detected in arbitrary units by a computerized instrument (continuous line); the right ordinate shows their radioactivity (dotted line). The analysis was repeated several times.
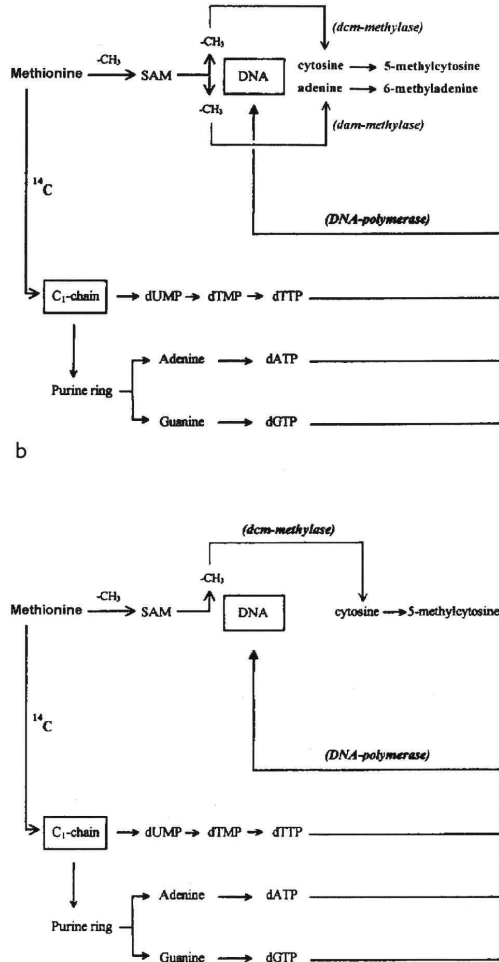


Fig. 2 - Pathways of synthesis and methylation of DNA. (a) In prokaryotes, to accomplish DNA replication, the DNA polymerase (DNApol) system (Kornberg et al., 1992) exploits the dATP, dGTP and dTTP resulting from the C1-chain oxidation which uses, via formic acid, the carbon coming from the -CH3 of methionine to build the -CH3 of T and the purine heterocycle of A and G (this carbon cannot enter the pyrimidine ring of C). In addition, the bacterial cell, via SAM, exhibits two parallel pathways for DNA post-synthetic modification: on the one hand, the dcm methylasic system (including the methylation function exerted by the type II RM endonucleases) constructs m5C along the semiconservatively newly replicating DNA chains; on the other, the dam methylasic system (including the methylation function exerted by the type I, II and III RM endonucleases) along these chains constructs m6A. (b) In eukaryotes. as in bacteria, methionine serves as a common donor of atoms: in cytosol, through the C1-chain oxidation, the carbon coming from its -CH3 enters the purine heterocycle of A and G and the -CH3 of T. After the construction of the corresponding dATP, dGTP and dTTP, in the same cytosol, the dNMPs (dAMP, dGMP and dTMP) are inserted, in the nucleus, into semiconservatively newly replicating DNA chains by the DNApol α system (Delfini et al., 1985). In the nucleus, the dcm methylasic system also attaches to given Cs located along these newly formed chains the whole -CH3 of methionine brought via SAM.

eluted unmethylated vs. methylated bases, coming from a bacterial DNA hydrolysate: C, G, T and A essentially resulted in equimolar proportions, as expected; $m^5C$ and $m^6A$ were present in very small amounts (Fig. 1b). Figure 1c showed the profile of a bacterial DNA hydrolysate treated with [$^{14}$C]Met in mid-CGC. This profile revealed that C was the only base not incorporating the labelled carbon coming from the precursor. By contrast, $m^5C$ was labelled to a notable extent, comparable to that of $m^6A$. The radioactivity found in the two modified bases meant that post-synthetic DNA modification took place, as for eukaryotes (Volpe et al., 1974; Eremenko et al., 1985; Geraci et al., 1974), by virtue of transfer, via S-adenosyl-L-methionine (SAM), of the whole radioactive -$CH_3$ from [$^{14}$C]Met to the DNA Cs and As. In addition, in figure 1c, the radioactivity found in the A, G and T bases revealed that, once again as for eukaryotes (Volpe et al., 1974; Eremenko et al., 1985; Geraci et al., 1974), DNA synthesis occurred because there was, via $C_1$-chain oxidation, an insertion of the labelled carbon coming from the -$CH_3$ of [$^{14}$C]Met into the purine heterocycle of A and G and into the -$CH_3$ of T. The quantitative analysis of the experiments performed through the CGC in the same conditions of figure 1c showed that, while cell proliferation is sigmoidal, the synthetic and methylasic pathways of DNA develop in a differential fashion. (i) The specific labelling accounting for synthesis of A, G and T appeared to be much lower than that accounting for construction of $m^5C$ and $m^6A$. This meant that the methylasic pathway evolves faster than the synthetic one. (ii) The specific labelling of T was always lower than that of G and A, since — one thought — the exit of the $C_1$-chain, oriented towards synthesis of G and A, is highly facilitated in comparison with that oriented towards synthesis of T. (iii) The specific labelling of $m^5C$ was much lower than that of $m^6A$, probably because the dam enzyme system works faster than the dcm enzyme one. (iv) In harmony with the development of the specific labelling of all bases taken together and of the specific labelling of single A, G and T, the specific labelling of $m^6A$ gradually decreased during the CGC, while that of $m^5C$ rapidly decreased after the inoculum to re-emerge again, with a sharp peak, in mid-CGC. An interpretation of these patterns was that the dcm and dam machineries should be uncoupled.

## Discussion

The data shown here demonstrated that the two pathways, leading to DNA synthesis and methylation, originated in the prokaryotic cell by exploiting methionine as an universal donor of a carbon atom and a -$CH_3$ (Fig. 2a). This model was crucial in the evolution of the genome functions. It was preserved by the eukaryotic cell, although at the level of the DNA modifying enzyme bifurcation one of the two branches, that yielding $m^6A$, was suppressed (Fig. 2b). The abandonment of $m^6A$ by the higher cells was a consequence of the substitution of the bacterial antiphagic RM defense (Arber, 1974) with the innovative immunitary reaction. In eukaryotes, the maintained $m^5C$ exhibited an increasing role in the genetic regulation of transcription (Volpe et al., 1974; Cacciamani et al., 2002; Volpe, 2005). In this framework, the question concerning the speed of the reactions in the pathway leading to DNA synthesis vs. that of the reactions in the pathway leading to post-synthetic DNA modification assumed great interest. During the bacterial CGC, the rate of formation of T was always lower than that of formation of A and G, i. e. the time required by the carbon from the -$CH_3$ of methionine to enter the -$CH_3$ of T was longer when compared to the time required for its insertion into the purine heterocycle of A and G. This differential rate was in harmony with the fact that, in HeLa cells, even if the synthesis of T was, quantitatively, of the same order of that of A, a notable delay of its optimum with respect to that of synthesis of A and G occurred (Eremenko et al., 1985; Geraci et al., 1974). Such a correspondence was expected, since in both pro- and eukaryotes the process leading to synthesis of T had to join in a pyrimidine ring carbamylphosphate and aspartate before accepting, at the position 5 of U, the -$CH_3$ constructed with the carbon coming from formic acid. This atom entered directly the A and G purine heterocycle. Anyway, the synthesis of A, G and T occurred in comparable compartments (in the prokaryotic liquid matrix and in the eukaryotic cytosol), whereas the post-synthetic modification process occurred in different places (in E. coli it yielded $m^5C$ and $m^6A$ in the same liquid matrix, in HeLa it yielded the sole $m^5C$ in the nucleus). The longer time required by SAM to enter the nucleus justified the advantage of the bacterial cell over the eukaryotic one in exploiting the -$CH_3$ of methionine: if in E. coli the specific building of A, G and T was much lower than that of $m^5C$ and $m^6A$, in HeLa the specific building of A, G and T turned out to be either similar or only a little lower than that of $m^5C$ (Eremenko et al., 1985; Geraci et al., 1974). The large amount of DNA methylation, which served for antiphagic defense in bacteria, became unnecessary in eukaryotes. In conclusion, the bacterial world used methionine to promote the pathways for synthesis and post-synthetic modification of DNA at the same time (Fig. 2a). This capability was inherited without major changes by the eukaryotic world, with the exception that the reactions yielding $m^6A$ were suppressed (Fig. 2b). The diversification of the speed of the single reactions occurring in the pathways leading to synthesis and post-synthetic modification of DNA seemed particularly relevant, since it showed the biochemical evolution as a function of formation of the eukaryotic cell nucleus from the bacterial liquid matrix.

## References

Arber W. 1974. DNA modification and restriction. Progr Nucl Acids Res Mol Biol 14: 1-37.

Cacciamani T, Virgili S, Centurelli M, Bertoli E, Eremenko T and Volpe P. 2002. Specific methylation of the CpG-rich domains in the promoter of the human tissue transglutaminase gene. Gene 297: 103-12.

Delfini C, Alfani E, De Venezia V, Oberholtzer G, Tomasello C, Eremenko T and Volpe P. 1985. Cell-cycle dependence and properties of the HeLa cell DNA polymerase system. Proc Natl Acad Sci USA 82: 2220-4.

Eremenko T, Palitti F, Morelli F, Whitehead EP and Volpe P. 1985. Hypomethylation of repair patches in HeLa cells. Mol Biol Rep 10: 177-82.

Franchina M, Hooper J and Kay PH. 2001. Five novel alternatively spliced transcripts of DNA (cytosine-5) methyltransferase 2 in human peripheral blood leukocytes. Int J Biochem Cell Biol 33: 1104-15.

Gehrke CW, Kuo KC, Davis GE and Suits RD. 1978. Quantitative high-performance liquid chromatography of nucleosides in biological materials. J Chromatogr 150: 455-76.

Geraci D, Eremenko T, Cocchiara R, Granieri A, Scarano E and Volpe P. 1974. Correlation between synthesis and methylation of DNA in HeLa cells. Biochim Biophys Res Commun 57: 353-8.

Kornberg A and Baker TA. 1992. DNA Replication. Freeman WH and Co, New York, 113-94.

Liau MC, Chang CF, Saunders GF and Tsai YH. 1981. S-adenosylhomocysteine hydrolases as the primary target enzymes in androgen regulation of methylated complexes. Arch Biochem Biophys 208: 261-72.

Marinus MG. 1987. DNA methylation in Escherichia coli. Ann Rev Biochem 21: 113-31.

Marmur J. 1961. A procedure for the isolation of deoxiribonucleic acid from microorganisms. J Mol Biol 3: 208-18.

Tentravahi V, Guntaka RV, Erlanger BF and Miller OJ. 1981. Amplified rDNA genes in rat hepatoma cell line enriched in $m^5C$. Proc Natl Acad Sci USA 78: 489-93.

Vanyushin BF. Tkacheva SG and Belozersky AN. 1970. Rare bases in animal DNANature 225: 948-9.

Volpe P and Eremenko T. 1974. Preferential methylation of regulatory gene sequences in HeLa cells. FEBS-Lett 44: 121-6.

Volpe P and Cascio O. 1994. Evolution of the DNA methylase protein family in Vertebrates. Phys Proc Acad Lincei 5: 79-87.

Volpe P. 2005. The Language of Methylation in Genomics of Eukaryotes. Biochemistry. (Ru) 70: 584-95.