3-methylcytosine in cancer: an underappreciated methyl lesion?

Thomas Pataillot-Meakin[1, 2], Nischalan Pillay[2] and Stephan Beck[1]

[1]Department of Cancer Biology, UCL Cancer Institute, 72 Huntley Street, London WC1E 6DD, UK
[2]Department of Pathology, UCL Cancer Institute, 72 Huntley Street, London WC1E 6DD, UK

Correspondence to TPM (tameakin@hotmail.co.uk), NP (nischalan.pillay.09@ucl.ac.uk) and SB (s.beck@ucl.ac.uk)

Keywords: 3-methylcytosine, DNA methylation, biomarker, cancer, resistance, DNA, AlkB Homolog, temozolomide, methyl methanesulphonate, cisplatin, ifosfamide

1.1 DNA methyl lesions

DNA methylation is an important epigenetic mark modulating many developmental and potentially pathological processes. Changes in DNA methylation are often observed in cancer and, cytosine methylation in particular, can be mediated enzymatically, resulting in e.g. 5-methylcytosine (5mC) or chemically, resulting in e.g. 3-methylcytosine (3mC)[1, 2]. If base pairing is affected by methylation as it is by 3mC, it can lead to methyl lesions which, if refractory to repair, can add to the mutational burden of affected cells. Here, we explore some underappreciated properties of 3mC in cancer.

Arising from non-enzymatic methylation by endogenous or exogenous methyl donors, chemically mediated methyl lesions can have disruptive effects on local DNA functionality[3]. In non-malignant cells, it can be expected that DNA repair pathways are intact and that any lesions will be repaired or, if the damage confers no selective growth advantage and is potentially lethal to the cell, programmed cell death (apoptotic) pathways will be activated. Conversely, in cancer where DNA repair mechanisms and apoptotic pathways are compromised due to somatic mutations, these methyl lesions and their effects on DNA are likely to persist.

The best known and investigated example of chemically mediated methyl lesions is O^6-methylguanine (6mG), a major product of methylating compounds and chemotherapeutic drugs such as temozolomide[4]. Other, more "minor", methyl lesions (including 3mC), with their own specific repair enzymes can also be formed but are currently far less studied.

The repair enzymes of 3mC lesions, AlkB homologs (ABH) and in particular ABH2 and ABH3, have garnered recent interest based on observations made *in vitro* of their involvement in chemotherapeutic resistance in several cancer types, including glioblastoma, non-small cell lung cancer, bladder and prostate cancer[5, 6]. 3mC may therefore serve as a surrogate marker for ABH activity as, while in normal cells, ABH2 and ABH3 are ubiquitously expressed, the transcription levels of these enzymes can be decreased with TP53 pathway activation[5]. Considering that TP53 is often mutated in cancer, inactivating the TP53 damage response pathway and thereby affecting cellular ability to alter ABH levels, it is tempting to speculate on the downstream consequences. It is conceivable that with impaired repair activity, the level of 3mC is likely to rise, generating changes in local DNA structure, and thus increased local mutagenicity and thereby contributing to the additional mutational burden and/or increasing cellular fitness. In this context, 3mC may be an example of the epigenotype driving the genotype.

This leads us to question whether minor epigenetic lesions have been overlooked due to their lower rate of formation resulting in an underappreciation of their impact on cancer development and progression?

1.2 3mC levels

In RNA, 3mC can be catalysed both enzymatically (e.g. in tRNA) and chemically, but in DNA 3mC can only be formed chemically via non-enzyme catalysed direct methyl group transfer[7]. Direct transfer can occur either from endogenous S-adenosyl methionine (SAM), the primary methyl donor for enzymatic methylation such as for 5mC in DNA; or by exogenous methylating compounds such as temozolomide, tobacco-specific nitrosamines (a major group of carcinogens found in tobacco smoke) or methyl methanesulphonate (MMS)[2, 8]. Direct chemical formation of 3mC by SAM however, occurs at a rate about 2000 times slower than by other chemical compounds like those mentioned above[9]. It is interesting to note that when the magnesium co-factor for SAM is chelated with ETDA *in vitro*, 3mC formation is inhibited, despite an overall increase in methyl products, highlighting the specificity in 3mC generation[10].

Chemically formed 3mC preferentially occurs in single-stranded DNA (ssDNA) and RNA. Direct methyl donation occurs via "backside attack" requiring that the lone electron pair of N3 in cytosine to attack the methyl group directly opposite to the adjacent atom in the methyl-donating molecule leaving group. In double-stranded DNA (dsDNA), the base pairing guanine sterically hinders the rate of 3mC formation as the guanine occupies the space where the methylating compound needs to be[11]. Therefore, in cells with higher transcription rates, the opportunities for 3mC formation are higher due to a higher proportion of DNA existing in a single stranded state. This transcription-dependant formation rate possibly explains ABH2 and ABH3's DNA transcription-based activity regulation.

1.3 3mC effects on transcription

The effect of 3mC on transcription is dependent on polymerase fidelity. Higher fidelity polymerases, such as DNA polymerase delta are halted upon encountering 3mC. Lower fidelity polymerases such as bypass polymerases nel, kappa or RNA polymerases tend to misincorporate bases. *E. coli* RNA polymerase was shown to misincorporate bases at a ten-fold higher rate than DNA polymerase I which is more often halted when encountering 3mC[11-13]. It was shown that changing the co-factor of DNA polymerase I *in vitro* from magnesium to manganese caused an increase in misincorporation of bases and a switch in the downstream effect of 3mC from cytostatic to mutagenic, an effect similarly observed between the X and Y families of human DNA polymerases[2, 14].

$AlkB^+$ SOS^- *E. coli* treated with MMS showed little (<2%) appreciable mutations due to 3mC, compared to 30% mutagenicity in $AlkB^-$ SOS^- *E. coli*. In $AlkB^-$ SOS^+ cells, mutagenicity rose to 70%, with a concordant decrease in cytotoxicity due a lack of direct repair and increased lesion bypass using DNA polymerase V. The replicative *E. coli* DNA polymerase III holoenzyme showed a higher mutation generation rate compared to DNA polymerase I, likely due to the proofreading capability of DNA polymerase I causing replicative arrest and allowing for removal of 3mC as the rate became nearly equal when the proofreading capability was removed[10].

1.4 3mC as a biomarker

Chemotherapeutic regimes do not necessarily include methylating compounds however inducers of double-stranded DNA breaks such as platinum compounds (e.g. cisplatin) and topoisomerase inhibitors (e.g. ifosfamide) are often included. As previously seen in *E. coli,* the generation of strand breaks and pausing of replication for repair allows DNA to be in single stranded form for longer and thus may indirectly allow increased generation of 3mC and additional time to enact effects within the cell, particularly when the repair pathways have been impaired[15, 16].

Knockdown studies of ABH2 and ABH3 suggested that decreased levels of ABH increased sensitivity to chemotherapeutics[5, 6]. If the levels of ABH2 and ABH3 are linked to resistance as previously suggested then a higher than normal level of ABH2 or ABH3 would provide increased chemotherapeutic resistance. As such, due to ABH being the main method of repair for 3mC, assuming a basal level of 3mC, a lower than expected level of 3mC would indicate presence of increased resistance. If observed in untreated patients, this may predict a decreased response to chemotherapy and therefore could act as a surrogate predictive biomarker for chemotherapeutic response allowing a more personalised therapeutic regimen.

1.5 Outlook

There are many open questions that will determine the future importance and interest in 3mC over and above its current potential as cancer-specific biomarker. For example, kataegis mutation signatures (hypermutability localised to small genomic regions) have recently gained interest for understanding cancer development and

potentially helping to guide prevention and therapy but the underlying mechanisms are poorly understood. Such signatures have been shown to be significantly associated with patients treated with alkylating agents in glioblastoma multiforme, mutagens within tobacco smoke in lung cancer and altered DNA maintenance and repair in several cancers[17], which are all hallmarks of increased 3mC formation. While DNA strand breaks have been shown to be a major driver of kataegis signatures[18], it is possible that 3mC may contribute as well and thus be an example of an epigenotype driven genotype. Analysis of kataegis signature samples for 3mC should clarify this hypothesis.

In addition, 3mC may even have broader implications. As chemically mediated methylation predates enzymatic methylation, it will be interesting to explore if during evolution, any organism has ever harnessed chemical DNA methylation such as 3mC for regulating biological processes akin to processes regulated by 5mC. Prime candidates would be extremophiles particularly those that have adapted to life under high levels of alkylating compounds, the largest group of chemicals interacting with DNA and frequently part of chemotherapy in the treatment of cancer. An important first step towards this exploration will be to develop quantitative sequencing-based chemistries for the detection of 3mC in DNA and to identify proteins that specifically bind to 3mC besides AlkB and its homologs.

1.6 Financial disclosure/acknowledgements

The authors have no conflicting interests to declare.

1.7 References

1) Meng H, Cao Y, Qin J, Song X, Zhang Q, Shi Y *et al.* DNA methylation, its mediators and genome integrity. *Int J Biol Sci.* 11(5), 604-617 (2015).

2) Barrows LR, Magee PN. Nonenzymatic methylation of DNA by S-adenosylmethionine in vitro. *Carcinogenesis*. 3(3), 349-351 (1982).

3) *****Furrer A, Van LB. Handling the 3-methylcytosine lesion by six human DNA polymerases members of the B-, X- and Y-families. *Nucleic Acids Res.* 42(1), 553-566 (2014). *3mC was shown to impair B family DNA polymerase progression and halt X family DNA polymerases but Y family DNA polymerases were able to bypass the lesion. DNA polymerase nel was shown to be critical for efficient 3mC bypass.*

4) Zhang J, Stevens MFG, Bradshaw TD. Temozolomide: Mechanisms of action, repair and resistance. *C Mol Pharm.* 5, 102-114 (2012).

5) Fujii T, Shimada K, Anai S, Fujimoto K, Konishi N. ALKBH2, a novel AlkB homologue, contributes to human bladder cancer progression by regulating MUC1 expression. *Cancer Sci.* 104(3), 321-327 (2013).

6) **Johannessen TC, Prestegarden L, Grudic A, Hegi ME, Tysnes BB, Bjerkvig R. The DNA repair protein ALKBH2 mediates temozolomide resistance in human glioblastoma cells. *Neuro.Oncol.* 15(3), 269-278 (2013). *Glioblastoma overexpressing ALKBH2 showed increased temozolomide resistance while siRNA knockdown of ALKBH2 increased sensitivity. ALKBH2 levels a partially mediated by TP53.*

7) Friedman S. In vitro methylation of rat liver tRNA-synthesis of 3-methylcytosine and 1-methylhypoxantheine. *Biochem.Biophys.Res.Commun.* 74(4), 1505-1510 (1977).

8) Rydberg B, Lindahl T. Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J*. 1(2), 211-216 (1982).

9) Delaney JC, Essigmann JM. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. *Proc.Natl.Acad.Sci.U.S.A*. 101(39), 14051-14056 (2004).

10) Ludlum DB. Methylated polydeoxyribocytidylic acid templates for RNA polymerase. *Biochim.Biophys.Acta*. 247(3), 412-418 (1971).

11) Singer B, Kusmierek JT. Chemical mutagenesis. *Annu.Rev.Biochem*. 51, 655-693 (1982).

12) Ludlum DB. Magee PN. Reaction of nitrosoureas with polycytidylate templates for ribonucleic acid polymerase. *Biochem J*. 128(3), 729-731 (1972).

13) *Saffhill R. Differences in the promutagenic nature of 3-methylcytosine as revealed by DNA and RNA polymerising enzymes. *Carcinogenesis*. 5(5), 691-693 (1984). *Using isolated polymerases and oligonucleotides, 3mC was shown to halt DNA polymerase progression while generating mutations when RNA polymerase encounters it.*

14) Boiteux S, Laval J. Mutagenesis by alkylating agents: coding properties for DNA polymerase of poly (dC) template containing 3-methylcytosine. *Biochimie*, 64(8-9), 637-641 (1982).

15) *Yang Y, Gordenin DA, Resnick MA. A single-strand specific lesion drives MMS-induced hyper-mutability at a double-strand break in yeast. *DNA Repair (Amst)*. 9(8), 914-921 (2010). *Local hypermutability is increased twety thousand fold at alkylating agent-induced double strand breaks. Based on observed mutation bias, 3mC was indentified as the likely main cause.*

16) Sikora A, Mielecki D, Chojnacka A, Nieminuszczy J, Wrzesinski M, Grzesiuk E. Lethal and mutagenic properties of MMS-generated DNA lesions in Escherichia coli cells deficient in BER and AlkB-directed DNA repair. *Mutagenesis*. 25(2), 139-147 (2010).

17) **Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV *et al.* Signature of mutational processes in human cancer. *Nature*. 500, 415-421 (2013). *Mutation signatures were identified from over four millions somatic mutations and mapped to several underlying causes including APOBECs but many remain cryptic. Kataegsis was observed in multiple cancer types.*

18) Sakofsky CJ, Roberts SA, Malc E, Mieczkowski PA, Resnik MA, Gordenin DA *et al.* Break-induced replication is a source of mutation clusters underlying kataegis. *Cell Rep.* 7(5), 1640-1649 (2014).