DNA Demethylation

DOI: 10.1002/anie.201106690

Oxidized Cytosine Metabolites Offer a Fresh Perspective for Active DNA Demethylation

*Kathrin I. Ladwein and Manfred Jung**

5-carboxycytosine · 5-formylcytosine · DNA methylation · epigenetic modifications · nucleobases

Already back in 1948, five years before Watson and Crick suggested their model for the structure of DNA, Hotchkiss reported that DNA can be methylated at the 5-position of cytosine (C),[1] and 5-methylcytosine (5mC) is now considered the fifth DNA base. DNA methylation is catalyzed by DNA methyltransferases (Dnmts) and occurs mostly at clustered CpG sites. Many of these are found in gene promoters, and their hypermethylation usually leads to gene silencing. DNA methylation plays a pivotal physiological role in the epigenetic regulation of gene expression, X-chromosome inactivation, and genomic imprinting.[2] The first DNA methyltransferase, Dnmt1, was discovered in 1988.[3] Subsequently Dnmt2, Dnmt3A/B, and Dnmt3L (an inactive homologue) were described. Knockout of, for example, Dnmt3B leads to embryonic lethality,[4,5] which indicates that accurate DNA methylation is a prerequisite during early development. Hence, deregulated DNA methylation is linked to many diseases, for example, cancer.[6]

In contrast to, for example, histone acetylation as another key epigenetic mechanism, DNA methylation is very stable, and active DNA demethylation mechanisms have been subjects of debate.[5,7] This discussion displays parallels to the history of lysine methylation, which was first suggested to be an irreversible modification. One obvious means of weakening a relatively stable carbon–nitrogen or carbon–carbon bond is oxidation. In the case of lysine demethylation two such mechanisms were discovered which proceed through either dehydrogenation or oxygenation.[8]

After the discovery of the TET (ten–eleven translocation) family of proteins, a model of an active oxidative demethylation of DNA was proposed.[9] TET proteins are 2-oxoglutarate- and Fe^{II}-dependent, and catalyze the conversion of 5mC to 5-hydroxymethylcytosine (5hmC) in DNA. According to the model, 5hmC might be further oxidized to 5-formylcytosine (5fC) and 5-carboxycytosine (5caC)[10] (see Figure 1), but until recently there was no evidence for the existence of these putative intermediates. In June 2011, T. Carell and co-workers could prove for the first time the presence of 5fC in samples of digested DNA isolated from murine embryonic stem (mES) cells. They were able to show that on average 5–15% of the 5hmC bases in the mES cell DNA are oxidized to 5fC. In contrast, 5caC could not be detected, but they also suggested its existence.[11] The Carell group could also show that 5fC from oxidized DNA can be enriched using a biotinylated hydrazide, which is a valuable tool for further studies. Subsequently, two publications followed, in which the detection of 5fC and/or 5caC was described.[12] These studies further proved the involvement of TET enzymes in the formation of highly oxidized cytosine derivatives in DNA.

The Zhang group could show that incubation of a 5mC-containing DNA substrate with TET1 leads to a decrease of the 5mC level with the concomitant appearance of 5hmC and two unknown compounds.[12a] Similar results were obtained with the TET2 and TET3 subtypes. It was verified that the new compounds match independent 5fC and 5caC controls by chromatographic comparison. MS studies provided final confirmation. Treatment of the TET2-catalyzed reaction mixture with $NaBH_4$ led to reduced levels of the oxidized cytosine metabolites and an increase of 5hmC, indicating that both are generated by the oxidation of 5hmC. Furthermore, ectopic expression of the catalytic domain of TET2 fused to GFP in HEK293 cells, followed by analysis of the DNA fragments by FACS (fluorescence-activated cell sorting) and two-dimensional thin-layer chromatography, indicated an increased level of 5hmC. Again, two new products were observed that correlated indeed to 5fC and 5caC. Under physiological conditions, both 5fC and 5caC could be detected in 14-day-old mouse embryos.

Similar results were published by He and co-workers.[12b] They overexpressed a full-length TET2 protein in HEK293T cells and incubated the nuclear extract with DNA substrates. Only when 5mC or 5hmC were incorporated in the substrates was the formation of an unknown material observed. Similar enzymatic activity could also be shown for TET1 and TET3. The comparison of the novel compound to a chemically synthesized 5caC standard in HPLC experiments showed that the two samples eluted at the same time, indicating that they detected 5caC. Again, MS experiments proved the formation

[*] Dr. K. I. Ladwein, Prof. Dr. M. Jung
Institut für Pharmazeutische Wissenschaften
Albert-Ludwigs-Universität Freiburg
Albertstrasse 25, 79104 Freiburg (Germany)
E-mail: manfred.jung@pharmazie.uni-freiburg.de
Prof. Dr. M. Jung
Freiburg Institute for Advanced Studies (FRIAS)
Albert-Ludwigs-Universität Freiburg (Germany)

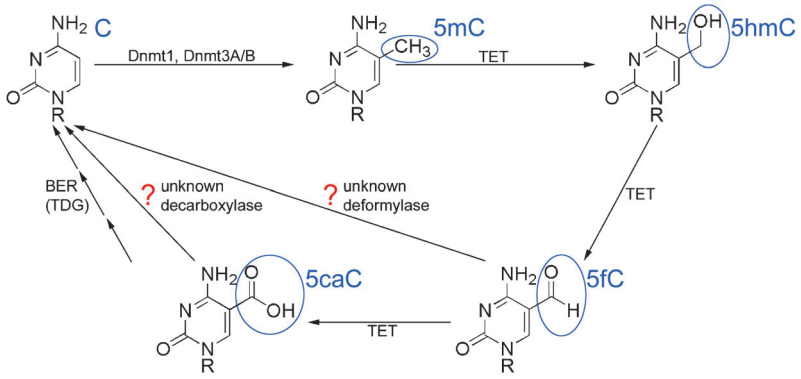

Figure 1. Cytosine (C) is methylated by DNA methyltransferases (Dnmts) to 5-methylcytosine (5mC). TET oxygenases oxidize 5mC stepwise via 5-hydroxymethylcytosine (5hmC) and 5-formylcytosine (5fC) to 5-carboxycytosine (5caC). The final step back to C might involve a base excision repair (BER), for example, by thymine-DNA glycosylase (TDG). R = DNA backbone.

of this metabolite. The authors could also show, using ^{14}C-labeled *S*-adenosyl methionine, that 5caC was formed by oxidation of 5mC and not by an independent carboxylation reaction. In contrast to Zhang and co-workers, they were not able to detect 5fC.

While the last step in the demethylation of 5caC to give C is still not understood, He et al. proposed a model based on a base-excision repair (BER) mechanism towards 5caC with 5caC as the recognition motif. To test this hypothesis, they looked at TDG (thymine-DNA glycosylase), an enzyme known to be involved in BER. It excises modified bases from DNA with formation of an abasic site, ultimately followed by the integration of an unmodified cytosine.[5,13] By incubating purified TDG with 5caC-containing DNA as the substrate, in fact, they could show TDG-excision activity towards the oxidized cytosine metabolite.

In conclusion, these three publications could prove that further oxidized cytosine derivatives besides 5mC and 5hmC exist in DNA in vitro and in vivo, and that TET enzymes do not use only 5mC as a substrate. They are also able to oxidize 5hmC and its product 5fC, resulting ultimately in 5caC. One potential possibility regarding the last step in the process of active DNA methylation could be the BER mechanism, as proposed by He and others.[5,12b] Another possibility could be a direct decarboxylation reaction, but the corresponding enzyme has not yet been discovered. Additionally, Pfaffeneder et al. have discussed a potential deformylation of 5fC via a vinyl carbanion intermediate.[11] Again, here is the question whether an enzyme that would catalyze this reaction can be found. As already mentioned, several possible DNA demethylation mechanisms in mammalian cells have been described and extensively discussed.[5] It is conceivable that several of these pathways act in concert to adjust to a changing environment. Additional experiments are needed to shed light on the active DNA demethylation process.

Detailed knowledge of the active DNA demethylation process will contribute to a better understanding of the regulation of DNA hypermethylation that causes the aberrant silencing of specific genes, for example, of tumor suppressor genes in certain cancer types.[14] Additionally, it will be important with respect to the reprogramming of somatic cells back into pluripotent cells. Upon development, certain genes are repressed by DNA methylation, whereas others, which are needed for further differentiation, are activated. This shows clearly that the process of active DNA demethylation might also have a great impact on stem cell research. It will be exciting to follow further studies concerning active DNA demethylation and to answer the question if the sixth (5hmC), seventh (5fC), and eighth (5caC) bases are only intermediates in the stepwise demethylation of 5mC to C or if each of them is an epigenetic modification with its own regulatory character.

Proteins that recognize 5mC in DNA, so-called methyl-binding domain (MBD) proteins, can be considered to be "readers" of the epigenetic code and have been reported to be associated with human diseases.[15] The existence of the three oxidized methyl cytosine derivatives also poses the question whether similar reader proteins might exist for these modifications and what their biological role will be. Thus, with iterative oxidation the obvious solution to the question of DNA demethylation might have been found, but its implications are still far from apparent.

Received: September 20, 2011
Published online: November 23, 2011

[1] R. D. Hotchkiss, *J. Biol. Chem.* **1948**, *175*, 315–332.
[2] a) T. B. Miranda, P. A. Jones, *J. Cell. Physiol.* **2007**, *213*, 384–390; b) J. C. Chow, Z. Yen, S. M. Ziesche, C. J. Brown, *Annu. Rev. Genomics Hum. Genet.* **2005**, *6*, 69–92; c) I. M. Morison, J. P. Ramsay, H. G. Spencer, *Trends Genet.* **2005**, *21*, 457–465.
[3] T. H. Bester, *Gene* **1988**, *74*, 9–12.
[4] M. G. Goll, T. H. Bestor, *Annu. Rev. Biochem.* **2005**, *74*, 481–514.
[5] S. C. Wu, Y. Zhang, *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 607–620.
[6] A. P. Feinberg, B. Vogelstein, *Nature* **1983**, *301*, 89–92.
[7] S. K. Ooi, T. H. Bestor, *Cell* **2008**, *133*, 1145–1148.
[8] a) Y. Shi, F. Lan, C. Matson, P. Mulligan, J. R. Whetstine, P. A. Cole, R. A. Casero, *Cell* **2004**, *119*, 941–953; b) Y. Tsukada, J. Fang, H. Erdjument-Bromage, M. E. Warren, C. H. Borchers, P. Tempst, Y. Zhang, *Nature* **2006**, *439*, 811–816.
[9] a) S. Kriaucionis, N. Heintz, *Science* **2009**, *324*, 929–930; b) M. Tahiliani et al., *Science* **2009**, *324*, 930.
[10] a) D. Globisch, M. Munzel, M. Muller, S. Michalakis, M. Wagner, S. Koch, T. Bruckl, M. Biel, T. Carell, *PLoS One* **2010**, *5*, e15367;

b) M. Münzel, D. Globisch, T. Carell, *Angew. Chem.* **2011**, *123*, 6588 – 6596; *Angew. Chem. Int. Ed.* **2011**, *50*, 6460 – 6468.
[11] T. Pfaffeneder, B. Hackner, M. Truss, M. Munzel, M. Muller, C. A. Deiml, C. Hagemeier, T. Carell, *Angew. Chem.* **2011**, *123*, 7146 – 7150; *Angew. Chem. Int. Ed.* **2011**, *50*, 7008 – 7012.
[12] a) S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, Y. Zhang, *Science* **2011**, *333*, 1300 – 1303; b) Y. F. He et al., *Science* **2011**, *333*, 1303.
[13] a) T. Lindahl, R. D. Wood, *Science* **1999**, *286*, 1897 – 1905; b) M. T. Bennett, M. T. Rodgers, A. S. Hebert, L. E. Ruslander, L. Eisele, A. C. Drohat, *J. Am. Chem. Soc.* **2006**, *128*, 12510 – 12519.
[14] a) Y. Kanai, *Pathol. Int.* **2008**, *58*, 544 – 558; b) J. G. Herman, S. B. Baylin, *N. Engl. J. Med.* **2003**, *349*, 2042 – 2054.
[15] a) M. Chahrour, S. Y. Jung, C. Shaw, X. Zhou, S. T. Wong, J. Qin, H. Y. Zoghbi, *Science* **2008**, *320*, 1224 – 1229; b) T. Clouaire, I. Stancheva, *Cell. Mol. Life Sci.* **2008**, *65*, 1509 – 1522.

