DOI: 10.1002/anie.200702159

Epigenetic Sequencing

Selective Detection of 5-Methylcytosine Sites in DNA**

*Sébastian Bareyt and Thomas Carell**

After the complete sequencing of the human genome, one of the next important steps is to decipher the genetic mechanisms which allow cells to switch genes on and off. These processes are fundamental for cellular development.[1] At the chemical level, gene silencing requires the selective methylation of of the C5-position of cytidines in DNA, and this established method is today known as the epigenetic control of gene function.[2] Failure of the epigenetic control of genes has been connected to many human diseases, including neurodevelopmental disorders and cancer.[3] The identification of 5-methylcytosine residues (^{5Me}dC) in genes is consequently of paramount importance for the diagnosis of modern diseases.

Today, the most common detection method uses bisulfite, which selectively hydrolyzes the amino group on cytosines to give uridines and allows the identification of nonmethylated dC nucleobases in DNA after amplification through the polymerase chain reaction (PCR) by using DNA-sequencing methods.[4] Other chemical methods exist which allow the selective transformation of either dC or ^{5Me}dC residues in DNA,[5] however, they generally do not possess the selectivity needed for efficient detection.[6] Very recently, the reaction of ^{5Me}dC-containing DNA with OsO_4 was studied in detail.[7] Although OsO_4 reacted selectively with the more electrophilic C5–C6 double bond of ^{5Me}dC, and left the same bond in dC intact, the efficient parallel conversion of thymidines (dT) is a major problem.

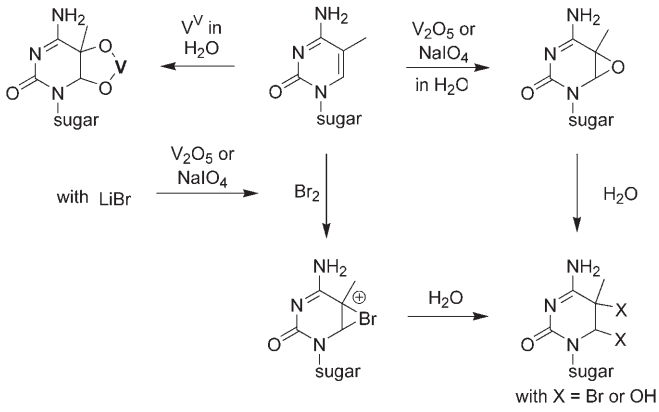
The redox potential of dC and dT residues are similar to but slightly higher than that of ^{5Me}dC. This should, in principle, allow ^{5Me}dC residues in DNA to be selectively targeted, for example, by the application of oxidation methods with finely tuned reactivity. Some vanadium(V) species are known to epoxidize[8] or *cis*-hydroxylate[9] double bonds, and their redox potential can be tuned by controlling the cluster size and composition.[10] Herein we report that mixtures of V^V species, or of $NaIO_4$ in the presence of LiBr, allow the desired conversion of the C5–C6 double bond of ^{5Me}dC to be achieved with unprecedented selectivity in mixed DNA sequences.

[*] Dr. S. Bareyt, Prof. Dr. T. Carell
Center for Integrative Protein Science
Department of Chemistry and Biochemistry
Ludwig-Maximilians-University Munich
Butenandtstrasse 5–13 (Haus F), 81377 München (Germany)
Fax: (+49) 89-2180-777-56
E-mail: thomas.carell@cup.uni-muenchen.de

[**] We are grateful to the Alexander von Humboldt Foundation for a Postdoctoral Fellowship to S.B. We thank the DFG (SFB 749) for financial support, and Prof. Dr. K. Karaghiosoff for recording the ^{51}V NMR spectra.

Supporting information for this article (including experimental details) is available on the WWW under http://www.angewandte.org or from the author.

The routes investigated are depicted in Scheme 1. For the experiments, two 3′-fluorescein-labeled synthetic 30 base pair long single-stranded oligonucleotides ODN(X) containing



5'-d(AAGTGTCATGAGTT**X**GATAGGTAAAG-TATT-Fluorescein)

Scheme 1. Envisioned mechanisms and intermediates of the vanadium(V) or sodium periodate oxidation of ^{5Me}dC in the absence and presence of LiBr.

either a dC or a ^{5Me}dC (X) residue in the middle of the random DNA sequence were designed. Both DNA strands were exposed to a series of V^V-containing clusters with different redox properties and different chemical reactivity in respect to olefin dihydroxylation and olefin epoxidation.

To further modulate the redox potential of the pyrimidine bases we systematically varied the pH value of the oxidation solution. The DNA products obtained were subsequently treated with hot piperidine (85 °C, 20 min) to induce strand breaks at oxidatively damaged DNA bases. The so-treated DNA was finally analyzed by denaturing polyacrylamide gel electrophoresis. The results obtained from these experiments are depicted in Figure 1. Freshly prepared solutions of V_2O_5 showed excellent selectivity for ^{5Me}C oxidation at pH values between 3 and 5. A clearly visible cleavage product was observed for the DNA strand with X = ^{5Me}dC. No cleavage at this site occurred with X = dC, thus proving the selective conversion of the ^{5Me}dC C5–C6 double bond. No reaction at the dT sites was detected, which shows that oxidation reactions are indeed able to distinguish between the three pyrimidine bases dT, dC, and ^{5Me}dC in DNA. However, we also observed a rather efficient cleavage at dG nucleobases in both strands, thus indicating that V^V species can react with nucleobases through two distinct pathways. The first pathway is the desired electrophilic attack on the double bond. The second pathway, however, is a direct electron abstraction, which will preferentially convert the purine moieties into

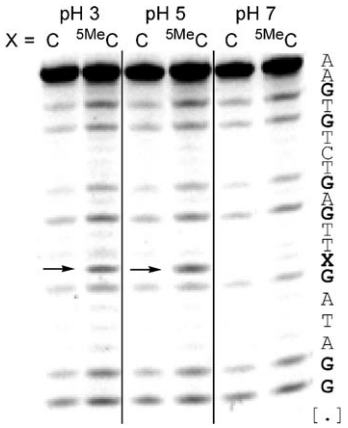


Figure 1. Sequence of ODN(X) and polyacrylamide gel electrophoresis (PAGE) of the strand cleavage after oxidation with V_2O_5 and treatment with hot piperidine. ODN(X) was treated with a solution of 5 mM V_2O_5 for 18 h at 60 °C.

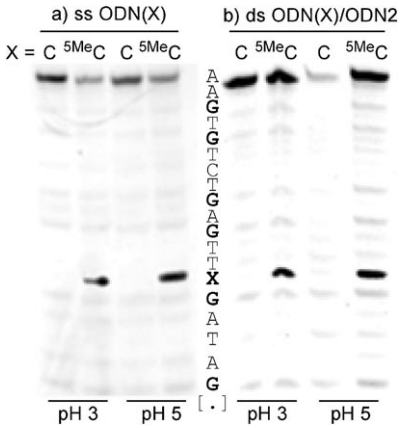


Figure 2. PAGE analysis of a) single and b) double strands after treatment with a solution of 5 mM vanadium pentoxide and 10 mM LiBr in a 100 mM sodium phosphate buffer at 60 °C for 18 h under anaerobic conditions, followed by treatment with 1 M piperidine at 85 °C for 20 min.

oxidative DNA lesions. In this latter case, dG is particularly vulnerable because of its low redox potential.[11] Special efforts were therefore taken to increase the efficiency of the ^{5Me}dC-selective reactivity and to suppress the direct electron abstraction. To this end, various other oxidizing V^V reagents were screened. The $Na[VO_2(Hhida)]\cdot 3H_2O$ ($H_3hida = N$-(2-hydroxyethyl)iminodiacetic acid) cluster was more stable at the rather low pH values, but exhibited the same reactivity as freshly prepared V_2O_5. Sodium decavanadate ($Na_6V_{10}O_{28}$), in contrast, furnished a dramatically higher amount of one-electron dG oxidation.

Large amounts of LiBr were added to the reaction mixture to increase the efficiency of the electrophilic attack on the ^{5Me}dC C5–C6 double bond; this was performed on the basis of the observation of Sudalai and co-workers that this would accelerate the dihydroxylation reaction.[12] Indeed, the addition of 10 mM LiBr (see the Supporting Information) to the reaction mixture immediately increased the selectivity for reaction at the ^{5Me}dC residue by up to a factor of 4. This result was determined by quantification of the fluorescent markers on the cleaved and uncleaved strands. In the presence of LiBr, only the ODN with X = ^{5Me}dC furnishes a cleavage product at the expected ^{5Me}dC position after treatment with piperidine.

To further reduce the amount of cleavage at dG sites we next performed the DNA oxidation in the absence of oxygen. Single-electron oxidation of dG initially forms the 8-oxodG lesion, which should not be cleavable with hot piperidine. Further oxidation of 8-oxodG occurs in the presence of air, which leads to piperidine-sensitive cleavage sites.[13] Indeed, as depicted in Figure 2a, reaction of DNA with V_2O_5 in the presence of LiBr and in the absence of oxygen resulted in a highly specific reaction at the C5–C6 double bond of the ^{5Me}dC base. No reaction at any other nucleobase was detected, even at the rather high temperature of 80 °C required for full conversion. The requirement of high temperature enabled us to perform the analysis even directly on double strands, as these melt to a significant extent at these high temperatures. The counterstrand 5′-d(AATACTTTAC-CTAT^{5Me}CG-AACTCATGACACTT)-3′ (ODN2) was hybrized with each ODN(X). Treatment of the corresponding double strand under the same oxidation conditions allowed selective cleavage at the ^{5Me}dC position (Figure 2b).

With the purpose of establishing a robust system, which would also permit selective conversion of ^{5Me}dC residues at a lower reaction temperature and in the presence of oxygen, a number of stronger oxidizing agents and halogenide sources were scanned. The type of halogenide used played a surprisingly dramatic role, with bromides giving the best results.

This finding is probably due to the ability of Br^- ions to form the dihalogen most rapidly, and indicates that the addition of the salt might have changed the reaction mechanism from *cis*-hydroxylation/epoxidation to electrophilic bromination with Br^+ ions. The use of 1,2-dibromoethylene or *N*-bromosuccinimide as the halogen source, however, caused degradation of the fluorescence label, probably because of the formation of the bromine radical. The use of 5,5-dibromo-2,2-dimethyl-4,6-dioxo-1,3-dioxane,[14] as a formal source of Br^+ ions, gave similar results.

Numerous oxidants were subsequently screened together with LiBr. $Na_2S_2O_8$ produced large amounts of dG oxidation products. Sodium tungstate showed an analogous reactivity as V_2O_5 but also resulted in higher amounts of dG oxidation. Significantly better results were obtained when we switched to iodine-containing oxidizing reagents such as *o*-iodoxybenzoic acid (IBX), Dess–Martin periodinane (DMP), or sodium periodate. The three oxidants gave excellent conversion of ^{5Me}dC without any detectable dG oxidation, even in the presence of oxygen, which shows that abstraction of a single electron from the dG base can be fully suppressed in these systems. The best results were finally obtained with a combination of $NaIO_4$/LiBr (1 mM $NaIO_4$, 4 mM LiBr, 100 mM Na_3PO_4 buffer, pH 5, 40 °C, 1 h; Figure 3). This system allowed selective conversion of the C5–C6 double bond of ^{5Me}dC, at a highly reduced reaction temperature of just 40 °C. Furthermore, the reaction could even be performed in the presence of oxygen. To investigate the mechanism which underlays the surprisingly efficient conversion of ^{5Me}dC

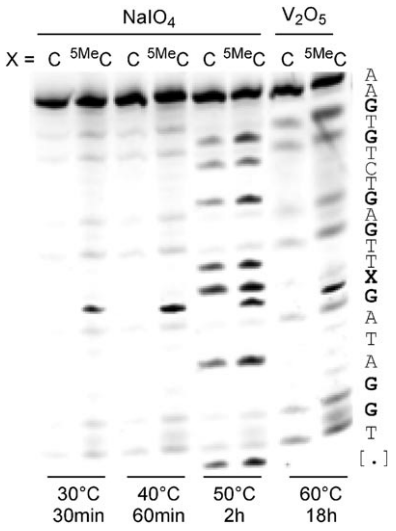


Figure 3. Sequence of ODN(X) and PAGE analysis after treatment with piperidine. The ODN was treated with a solution of 1 mM $NaIO_4$ and 4 mM LiBr in 100 mM sodium phosphate buffer (pH 5) at 30 °C to 50 °C for 30 min to 2 h or in solution. The results obtained upon changing the oxidizing conditions to 5 mM vanadium pentoxide and 10 mM LiBr at 60 °C for 18 h are also shown. Both experiments were performed in an ambient atmosphere, followed by treatment with 1 M piperidine at 85 °C for 20 min.

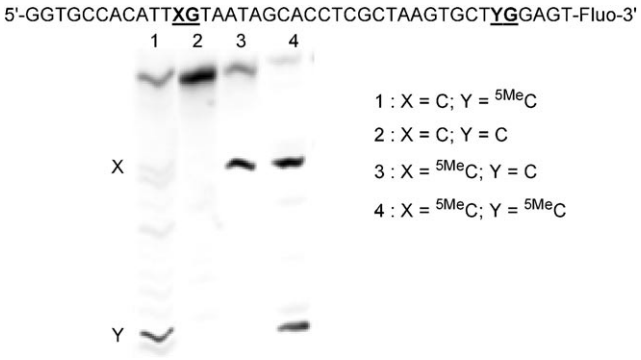


Figure 4. PAGE analysis: ODN(X,Y) was incubated in a solution of 1 mM $NaIO_4$, 4 mM LiBr, and 50 mM H_2NCHO in a 100 mM sodium phosphate buffer (pH 5) at 40 °C for 1 h in an ambient atmosphere, followed by treatment with 1 M piperidine at 85 °C for 20 min.

in more detail, the reaction was repeated on the unlabeled counterstrand ODN2 carrying a single ^{5Me}dC residue in the middle of its sequence. The resulting DNA strands were then analyzed by MALDI-TOF mass spectrometry. The mass spectra showed a new strong signal (see the Supporting Information) with an increased molecular weight of m/z 96, which is indicative of the formation of the bromohydroxylated C5–C6 double bond, as depicted in Scheme 1.

With the new method for ^{5Me}dC detection in hand, we investigated the possibility of detecting ^{5Me}dC residues in d(CpG) sites, which are the epigenetic control elements. The p16 gene is described as a regulator of the invasive proliferation of basal cell carcinoma cells. The activity of its promoter is reduced by a diminished level of methylation at certain d(CpG) sites.[15] We selected a short 40 base pair long fragment of the promoter (1751–1791) containing two of the critical d(CpG) sequences (Figure 4). Four different methylation patterns are consequently possible when X and Y are either dC or ^{5Me}dC. Treatment of the oligonucleotide under the described conditions indeed resulted in selective reaction at the ^{5Me}dC sites. In addition, however, we also observed some reaction at dT sites, which we believe can be explained by an accessibility problem. Longer single strands of DNA adopt compact structures in solution which may shield the most reactive sites. This problem could be efficiently overcome by the addition of formamide (50 mM) to the oxidizing solution; formamide unfolds these structures without modifying the reactivity. Figure 4 depicts the results of these studies and shows the extremely efficient detection of ^{5Me}dC nucleobases in mixed DNA sequences. Note that side reactions occur at almost no other nucleobases.

In conclusion, we have developed a novel chemical method which allows the targeting of ^{5Me}dC in DNA. The selective detection of ^{5Me}dC in the presence of all other canonical nucleobases (dC, dT, dA, and dG) was achieved using either a combination of V_2O_5/LiBr or $NaIO_4$/LiBr. Whereas oxidation of dG to 8-oxo-dG occurs with V_2O_5/LiBr, the $NaIO_4$/LiBr system is able to fully suppress the modification of all other nucleobases. We believe that the mechanism proceeds through the formation of a bromonium cation intermediate at the C5–C6 double bond of the pyrimidine rings. The next challenge will be to increase the sensitivity of the method. A novel strategy is being developed which could enable us to establish a rapid and easy to use analysis of epigenetic information in genes of real samples.

Received: May 16, 2007
Published online: November 26, 2007

Keywords: DNA cleavage · DNA · epigenetic information · methylcytosines · sequence determination

[1] A. Goren, H. Cedar, *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 25–32.
[2] A. Razin in *DNA Methylation and Cancer Therapy*, Landes Bioscience, Georgetown, **2005**, p. 1–12.
[3] a) K. D. Robertson, A. P. Wolffe, *Nat. Rev. Genet.* **2000**, *1*, 11–19; b) D. N. Cooper, M. Krawczak, *Hum. Genet.* **1989**, *83*, 181–188; c) A. Bird, M. Taggart, M. Frommer, O. J. Miller, D. Macleod, *Cell* **1985**, *40*, 91–99.
[4] a) M. Sasaki, J. Anast, W. Bassett, T. Kawakamia, N. Sakuragi, R. Dahiya, *Biochem. Biophys. Res. Commun.* **2003**, *309*, 305–309; b) J. G. Herman, J. R. Graff, S. Myohanen, B. D. Nelkin, S. B. Baylin, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 9821–9826.
[5] a) R. G. H. Cotton, N. R. Rodrigues, R. D. Campbell, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 4397–4401; b) C. S. Serva, E. Weinhold, R. J. Roberts, S. Klimasauskas, *Nucleic Acids Res.* **1998**, *26*, 3473–3479.
[6] T. Rein, M. L. DePanphilis, H. Zorbas, *Nucleic Acids Res.* **1998**, *26*, 2255–2264.
[7] a) A. Okamoto, K. Tainaka, T. Kamei, *Org. Biomol. Chem.* **2006**, *4*, 1638–1640; b) K. Tanaka, K. Tainaka, A. Okamoto, *Bioorg. Med. Chem.* **2007**, *15*, 1615–1621; c) K. Tanaka, K. Tainaka, T. Kamei, A. Okamoto, *J. Am. Chem. Soc.* **2007**, ASAP.

[8] M. Noritaka, N. Yoshinao, Y. Kazuya, *J. Mol. Catal. A* **2006**, *251*, 286 – 290.
[9] H. Mimoun, L. Saussine, E. Daire, M. Postel, J. Fischer, R. Weiss, Raymond, *J. Am. Chem. Soc.* **1983**, *105*, 3101 – 3110.
[10] D. C. Crans, J. J. Smee, *Comprehensive Coordination Chemistry II*, *Vol. 4*, Pergamon, Oxford, **2004**, pp. 175 – 239.
[11] A. J. Stemmler, C. J. Burrows, *J. Biol. Inorg. Chem.* **2001**, *6*, 100 – 106.
[12] L. Emmanuvel, T. M. A. Shaikh, A. Sudalai, *Org. Lett.* **2005**, *7*, 5071 – 5074.
[13] C. Vialas, C. Calaparols, G. Pratviel, B. Meunier, *J. Am. Chem. Soc.* **2000**, *122*, 2157 – 2167.
[14] R. Bloch, *Synthesis* **1978**, 140 – 142.
[15] S. S. Månsson, J. Reis-Filho, G. Landberg, *J. Pathol.* **2007**, *212*, 102 – 111.

