Capillary electrophoretic analysis of genomic DNA methylation levels

Dirk Stach, Oliver J. Schmitz¹, Stephan Stilgenbauer², Axel Benner², Hartmut Döhner², Manfred Wiessler and Frank Lyko⁴,*

Division of Molecular Toxicology, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, ¹Department of Analytical Chemistry, University of Wuppertal, Gaußstrasse 20, 42119 Wuppertal, Germany, ²Department of Internal Medicine III, University of Ulm, Robert-Koch-Strasse 8, 89081 Ulm, Germany, ³Central Unit Biostatistics and ⁴Research Group Epigenetics, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

Received September 6, 2002; Revised October 16, 2002; Accepted November 6, 2002

ABSTRACT

Changes in DNA methylation have been found in the large majority of tumors. This phenomenon includes both genome-wide hypomethylation and gene-specific hypermethylation. However, the clinical relevance of either mechanism has remained contentious. In order to determine DNA methylation levels from a large number of clinical samples, we have established a method for accurate high-throughput quantification of 5-methylcytosine in genomic DNA. Our protocol requires a small amount (<1 µg) of DNA that is enzymatically hydrolyzed to single nucleotides. Single nucleotides are then derivatized with a fluorescent marker and separated by capillary electrophoresis. After calibration of the method, we have determined cytosine methylation levels from tumor samples of 81 patients that had been diagnosed with chronic lymphocytic leukemia (CLL). These patients showed a high variability in their methylation levels with a general trend towards hypomethylation. Because of its high accuracy and throughput our method will be useful in determining the role of genomic DNA methylation levels in tumorigenesis.

INTRODUCTION

Cytosine methylation is required for mammalian development (1). Loss of methylation impairs various aspects of epigenetic gene regulation, including genomic imprinting (2), X-chromosome inactivation (3) and retroviral silencing (4). However, the precise function of DNA methylation is still unclear. It is widely believed that DNA methylation patterns represent an epigenetic programme of the genome that dictates the interpretation of genetic information. Consistent with epigenetic programming, methylation levels change dynamically during mammalian development (5,6). During preimplantation development, the mammalian genome becomes rapidly demethylated (7,8). This presumably represents a deprogramming step to initiate cellular differentiation. After implantation, de novo methylation restores DNA methylation levels, which coincides with an increasing commitment of cells to particular, specialized functions.

The genomic DNA methylation level also plays an important role in tumorigenesis (9,10). Determination of total cytosine methylation levels revealed significantly lower levels in tumor-derived genomic DNA compared with control tissue (11,12). In addition, global genomic hypomethylation has also been shown to result in chromosome instabilities (13), which is a hallmark of cancer cells. More recently, numerous studies have focused on the methylation status of single genes or single cytosine residues. This identified various loci in the genome that showed aberrant methylation in a high number of tumors (14–16). Many of these sites were hypermethylated, which is consistent with the general view of gene-specific hypermethylation in cancer (17). Mechanistically, gene-specific hypermethylation is believed to be a consequence of overexpression or ectopic activity of human DNA methyltransferases (18–20). This effect is not considered to be specific but would rather affect DNA methylation on a global level.

The analysis of global DNA methylation changes requires a method for the precise determination of cytosine methylation in a high number of samples. For this purpose, most laboratories use high performance liquid chromatography (HPLC) (21). However, HPLC requires large amounts of DNA and significant experimental experience to obtain reproducible results. Both points are impedimental to the large-scale analysis of clinical samples. In order to systematically analyze the role of the global DNA methylation level in tumorigenesis we have established a novel method based on capillary electrophoresis. This method is characterized by both high sensitivity and high reproducibility. We have analyzed tumor samples from 81 patients diagnosed with B-cell chronic lymphocytic leukemia (CLL). This revealed a significant variability in DNA methylation levels between individual patients. CLL is the most common leukemia in adults and is characterized by a highly variable clinical course (22,23). Our data can provide the foundation for a detailed statistical

*To whom correspondence should be addressed. Tel: +49 6221 423800; Fax: +49 6221 423802; Email: f.lyko@dkfz.de
analysis to address the role of the DNA methylation level in CLL and other types of cancer.

MATERIALS AND METHODS

**In vitro methylation**

Ten micrograms of lambda DNA (Sigma) was methylated for 2 h with 50 U M.HhaI methylase (New England Biolabs), according to the manufacturer’s instructions. In a parallel experiment, 10 μg of pBluescript II KS (+) plasmid DNA (from our own laboratory stocks) was methylated for 2 h with 50 U M.HpaII methylase (New England Biolabs). Complete methylation was subsequently confirmed by a restriction protection assay. One microgram of methylated and unmethylated DNA, respectively, was incubated for 1 h with 20 U HpaII and HhaI restriction endonuclease (New England Biolabs), respectively, or mock-incubated with buffer. Reaction products were separated by an agarose gel and visualized by ethidium bromide staining.

**Isolation, hydrolysis of and derivatization of genomic DNA samples**

For initial validation experiments we used standard calf thymus DNA (Sigma). Patient DNA was prepared from peripheral blood mononuclear cells using Trizol reagent (Gibco BRL) according to the manufacturer’s instructions. Five micrograms of isolated DNA was dissolved in 5 μl of bidistilled water. Five microliters of an enzyme mixture consisting of micrococcal nuclease (150 μM/μl; Sigma) and calf spleen phosphodiesterase (2.5 μM/μl; Calbiochem-Novabiochem) and 1 μl of buffer [250 mM N-(2-hydroxyethyl)piperazine-N’-(2-ethanesulfonic acid) (HEPES), 100 mM CaCl₂, pH 6.0] were added and incubated for 3 h at 37°C (24). After enzymatic hydrolysis, 2’-deoxynucleoside-3’-phosphates were obtained. In the next step, 20 μl of 800 mM HEPES, pH 6.5, 30 μl of 1.8 M N-ethyl-N’-(3-dimethylaminopropyl)carbodiimide (EDC; Sigma) solution in 800 mM HEPES and 30 μl of 25 mM 4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene-3-propionyl ethylenediamine hydrochloride (Bodipy FL EDA; Molecular Probes) in 800 mM HEPES were added to the digest and incubated for 25 h at 25°C in the dark. After a 1:1000 (or higher) dilution with bidistilled water, 20 μl of the reaction mixture was directly injected into the capillary.

**Capillary electrophoresis**

Capillary electrophoresis was performed on a BioFocus 3000TC LIF² system (BioRad). Derivatized DNA samples were analyzed in an electrolyte consisting of 75 mM SDS in 17 mM sodium phosphate buffer (pH 9.0) containing 15% (v/v) methanol. A fused-silica capillary was used with a total length of 64 cm, corresponding to a detection window of 59.4 cm and an inner diameter of 50 μm. Separation was performed with a sample injection at 20 p.s.i. at 25°C with an applied voltage of 20 kV. Outlet was the cathode for all runs. Analyses were detected by laser-induced fluorescence with an excitation wavelength of 488 nm. Each sample was analyzed up to 12 times in the beginning of the study. Due to the high reproducibility of the method, the number of repeated analyses was gradually reduced to five. Between each run the capillary was rinsed for 60 s with 200 mM SDS, for 90 s with 1 M sodium hydroxide, for 60 s with bidistilled water and for 120 s with the electrolyte.

**Patients**

Between October 1990 and August 1998, 325 consecutive CLL patients from a single institution (Med. Klinik und Poliklinik V, University of Heidelberg) were enrolled in the study and followed with regard to survival. Detailed clinical characteristics have been described previously (25,26). Eighty-one cases were selected randomly from the various subgroups defined by VH (variable region heavy chain gene) mutation status and genomic aberration and based on the availability of DNA.

**RESULTS**

Capillary electrophoresis represents a powerful tool for the detection of genomic DNA methylation (24,27). In order to determine DNA methylation levels in clinical samples we established a capillary electrophoretic method. To this end, we modified a procedure where DNA is digested to single nucleotides followed by chemical derivatization with a fluorescent marker and separation by micellar electrokinetic chromatography (24) (Fig. 1A). Derivatized nucleotides are detected by laser-induced fluorescence at 510 nm, thus permitting a well defined experimental readout with very little baseline fluctuation (Fig. 1B). Under the derivatization conditions used, the fluorescent marker does not react with ribonucleotides, which is probably due to the formation of a cyclic phosphate between the 3’-phosphate group and the 2’-hydroxy group of the ribose (28). The presence of RNA would therefore not affect the quantification of deoxyribonucleotides.

As a first step towards the calibration of the method we determined the quantum yield factor (QYF) of 5-methyl-2’-deoxycytosine-3’-monophosphate (5m-dCMP). The QYF is an important parameter in quantitative fluorescence detection and accounts for base-specific quenching effects (29–31). This results in a disproportionate detection of fluorescent signals. 5m-dCMP was synthesized by ammonolysis of O4-ethyl-thymidine (32) and spectroscopically identified by NMR and MS (a kind gift from H. C. Kliem, DKFZ Heidelberg). After labeling with the fluorescence marker and purification by preparative HPLC the product was used as a standard. A comparison with equal amounts of labeled dAMP as a reference set to 100% or 1.00, yielded a QYF for labeled 5m-dCMP of 98% or 0.98, respectively (data not shown). In the next step, we sought to address the kinetic aspects of the chemical derivatization. For a precise determination of cytosine methylation levels in genomic DNA samples we analyzed DNA from premethylated standards that had a methylation level similar to human genomic DNA (3–4%). In order to calibrate the method with a complex, yet well defined standard we used bacteriophage lambda DNA that had been methylated in vitro with the bacterial methylase M.HpaII. The lambda genome has been sequenced to completion and contains 24182 cytosine residues. It contains 328 M.HpaII target sites that can be methylated on both strands. Thus, M.HpaII-methylated lambda DNA would contain 656 5-methylcytosine residues, corresponding to a 2.7% cytosine
methylation level. Complete methylation was ensured by the experimental conditions of the in vitro methylation reaction and validated by a protection assay using the \(Hpa\) II restriction endonuclease (Fig. 2). Methylated (and unmethylated) lambda DNA was then hydrolyzed, derivatized and analyzed by capillary electrophoresis and nucleotide-specific derivatization factors were determined (Table 1). For this, the calculated fraction of each nucleotide was compared with the respective fluorescence signal. In addition, the fluorescence signal was corrected by the QYFs as well. The median derivatization factors obtained by this procedure were: \(d\)AMP 0.70 ± 0.02, \(d\)GMP 1.50 ± 0.08, \(d\)TMP 1.07 ± 0.04, \(d\)CMP 1.01 ± 0.04 and 5m-\(d\)CMP 0.96 ± 0.05. The differences between the nucleotide-specific factors reflect the different reactivities for each nucleotide towards the fluorescence marker. Derivatization factors were confirmed by a parallel analysis of pBluescript plasmid DNA that had been methylated with the bacterial methylase \(M. Hha\) I. pBluescript DNA contains 1493 cytosine residues, 48 of which can be methylated by \(M. Hha\) I. This corresponds to a cytosine methylation level of 3.2%. The resulting derivatization factors were highly similar to those determined with \(M. Hpa\) II-methylated lambda DNA.

As a consequence, the method could now be used to determine the cytosine methylation levels from genomic DNA samples. In order to determine the reproducibility of our method we first analyzed commercially available calf thymus DNA. Aliquots from an identical source were hydrolyzed and derivatized in four parallel, independent reactions. The nucleotide mixture from each reaction was then analyzed four to five times by capillary electrophoresis and the cytosine methylation levels were determined. This revealed a very high degree of reproducibility between individual derivatizations as well as between individual measurements (Fig. 3A). In all cases, the standard deviation was below 5% (Fig. 3A). Similarly, variations in the amount of injected nucleotides had no influence on the determined cytosine methylation level (Fig. 3B). The relationship between the quantity of nucleotide and the fluorescence was linear over a broad range of dilutions (data not shown). Consequently, the analysis of various dilutions of the labeled nucleotide mixture resulted in a highly reproducible level of cytosine methylation (Fig. 3B). In conclusion, these results suggest a significant robustness of our method and demonstrate a high degree of reproducibility.

After the establishment of reproducible experimental conditions we sought to determine the genomic DNA methylation from patient samples. To this end we analyzed genomic DNA methylation levels in tumor samples from a thoroughly characterized collective of patients diagnosed with CLL (25). As for numerous other types of cancer, changes in DNA methylation have also been found to be associated with CLL (33) but little is known about their causes or consequences. In order to confirm the reproducibility of our method with standard clinical samples, we analyzed genomic DNA from four different patients seven to 10 times. This confirmed the excellent reproducibility of our analytical conditions and a standard deviation below 5% of the individual mean value (Fig. 4). Thus, the mean value from four to five measurements adequately reflects the genomic cytosine methylation levels of a patient sample.
Seventy-seven additional patients were then selected randomly from various genetic subgroups and their genomic cytosine methylation levels were determined by capillary electrophoresis. Again, all samples were measured at least five times and their standard deviation was always below 5% of the individual mean value. A comparison of mean cytosine methylation levels from all samples revealed a significant degree of variation between individual patients (Fig. 5). The highest methylation level was 4.46% and the lowest level was 2.92%. The median cytosine methylation level of all 81 patients was 3.69%. This value is in good agreement with the results from a previous study (34). We also determined DNA methylation levels in FACS-sorted CD19-positive B cells from seven healthy individuals. These samples revealed an average methylation level of 4.02% (data not shown). While this result is consistent with general hypomethylation in tumor cells, the control values need to be interpreted with caution since CLL can originate from various stages of B-cell differentiation and precisely matched controls cannot be obtained. More importantly, our results demonstrate a significant variation in genomic cytosine methylation levels between individual patients. This variation can now be used to establish

<table>
<thead>
<tr>
<th></th>
<th>Methylation</th>
<th>Unmethylation</th>
<th>d-factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Calc. (%)</td>
<td>Meas. (%)</td>
</tr>
<tr>
<td>Lambda</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>24 320</td>
<td>25.07</td>
<td>35.3</td>
</tr>
<tr>
<td>T</td>
<td>24 182</td>
<td>24.93</td>
<td>17.4</td>
</tr>
<tr>
<td>C</td>
<td>23 526</td>
<td>24.25</td>
<td>24.3</td>
</tr>
<tr>
<td>5mC</td>
<td>656</td>
<td>0.68</td>
<td>0.71</td>
</tr>
<tr>
<td>pBluescript II KS (+)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>1468</td>
<td>24.79</td>
<td>35.9</td>
</tr>
<tr>
<td>T</td>
<td>1468</td>
<td>24.79</td>
<td>22.7</td>
</tr>
<tr>
<td>G</td>
<td>1493</td>
<td>25.21</td>
<td>16.2</td>
</tr>
<tr>
<td>C</td>
<td>1445</td>
<td>24.40</td>
<td>24.4</td>
</tr>
<tr>
<td>5mC</td>
<td>48</td>
<td>0.81</td>
<td>0.84</td>
</tr>
</tbody>
</table>

DNA standards were hydrolyzed, derivatized, and analyzed by capillary electrophoresis. Nucleotide-specific derivatization factors (d-factor) were determined by comparing the calculated fraction of each nucleotide (Calc.) with the measured fluorescence signal (Meas.); 5mC indicates 5-methylcytosine; n indicates the number of bases. The measured values represent mean values from 10 to 12 analytical runs. Thus, the sum of all nucleotide fractions might be slightly above or below 100%.

Figure 3. Reproducibility of experimental conditions. (A) Reproducibility of the derivatization reaction. Calf thymus DNA from a single source was derivatized in four independent parallel reactions and then analyzed repeatedly by capillary electrophoresis. The resulting DNA methylation levels showed very little variation with a standard deviation well below 5%. (B) The DNA methylation level is not influenced by the concentration of analytes. Different dilutions of the same derivatized calf thymus DNA sample were analyzed by capillary electrophoresis and the cytosine methylation level was determined. The following concentrations were used: 1:20 000 (filled circles), 1:10 000 (open squares), 1:5000 (filled diamonds), 1:1000 (open circles), 1:500 (filled squares). The results indicate a minimal variation with a standard deviation below 5%.

Figure 4. Determination of DNA methylation levels in individual CLL patients. Genomic DNA was isolated from tumor samples of CLL patients and the cytosine methylation level was determined by capillary electrophoretic analysis. The figure shows repeated analyses of four independent representative samples. In all cases, the standard deviation from the mean value was below 5%.
correlations between the DNA methylation and clinicopathological parameters. Data analysis from a sufficiently large number of patients might help to define the role of DNA methylation in tumorigenesis.

**DISCUSSION**

Changes in DNA methylation are a prominent feature of tumor cells. However, reliable methods for the quantification of genomic cytosine methylation levels from small DNA samples are still scarce. In this respect, capillary electrophoresis offers a number of advantages including high reproducibility, brief analysis time and a potential for automation. The sensitivity of capillary electrophoretic analyses can be significantly increased by the use of a fluorescent label. However, this raises some problems in the quantification and interpretation of experimental data. Thus, a method based on the detection of fluorescently labeled nucleotides needs to be calibrated for the determination of absolute cytosine methylation levels.

In a first step, we established precise quantum yield factors for each nucleotide. Due to base-dependent quenching effects for fluorescent signals (29–31) the different values of the integrated peak areas do not reflect a correct quantitative proportion of each nucleotide. Consistent with our expectations, the factor for 5m-dCMP was very similar to dCMP. In a second step, the kinetics of the derivatization was also taken into consideration. This was done by the analysis of complex DNA (lambda phage and plasmid) that had been methylated in a defined pattern by bacterial DNA methylases. The resulting derivatization factors revealed a different reactivity for each nucleotide towards the fluorescence marker. With the help of these factors it is now possible to precisely determine cytosine methylation in a wide variety of samples. Our method has the sensitivity to detect cytosine methylation levels of <0.2% (Dirk Stach, unpublished data). While we used 5 μg of genomic DNA for analysis, more recent experiments have shown that reproducible results can be obtained with as little as 100 ng (Oliver J. Schmitz, unpublished data). Thus, we have established a method for high-throughput analysis of DNA methylation levels from small clinical samples.

Altered DNA methylation patterns and overexpression of DNA methyltransferases have been documented for various forms of leukemia (20,33). In the present study, we have determined the level of cytosine methylation in 81 well characterized CLL patients. This revealed considerable differences between individual patients. Inter-individual differences in tissue-specific DNA methylation levels have been shown to be negligible in normal human tissue (35). Therefore, our results demonstrate a significant heterogeneity in DNA methylation levels among patients. This variability also provides the foundation for a statistical analysis of DNA methylation. The patients are part of a highly characterized collective (25,26) and we will be able to analyze various clinical as well as biological parameters for their dependency on DNA methylation. This kind of analysis will help to determine the oncogenic potential of DNA methylation and determine the health risk associated with particularly high or low cytosine methylation levels.

**ACKNOWLEDGEMENTS**

This paper is dedicated to Harald zur Hausen on the occasion of his retirement as head of the German Cancer Research Center with gratitude and appreciation for 20 years of leadership. The work was supported by the European Commission (QLG2-1999-00876), the BMBF (01K09938), and an Emmy Noether fellowship from the Deutsche Forschungsgemeinschaft to F.L.

**REFERENCES**


