

Published in final edited form as:

*Anal Chem.* 2008 March 15; 80(6): 2212–2221. doi:10.1021/ac702265b.

# Accurate DNA Fragment Sizing by Capillary Electrophoresis with Laser-Induced Fluorescence Array for Detection of Sequence Specificity of DNA Damage

Erwin V. Fundador<sup>†</sup>, Dharamainder Choudhary<sup>‡</sup>, John B. Schenkman<sup>§</sup>, and James F. Rusling<sup>\*,†,§</sup>

Department of Chemistry, University of Connecticut, Storrs, Connecticut 06269-3060, and Departments of Cell Biology and Surgery, University of Connecticut Health Center, Farmington, Connecticut 06032

## Abstract

Cancer has been linked to mutations within specific codons in genes that code for critical biomolecules such as tumor suppressor proteins (e.g., p53). Activated metabolites like benzo[*a*]pyrenediol epoxide act on preferred nucleotide sequences of DNA, and such mutations have been identified in cancers. DNA reaction site identification depends on accurate analysis of oligonucleotide fragment sizes produced by strand breakage at the damaged sites. Herein, we report a new method for DNA fragment sizing using capillary electrophoresis with laser-induced fluorescence detection (CE-LIF). Absolute sizing accuracy and speed are achieved by utilizing a CE-LIF array with two-color fluorescence detection. Accuracy depends on correcting results with commercial standards by referring them to primary standards with the same sequences and identical labels as sample fragments. The method is demonstrated by detection of a [...GGC GCG-CAG...] *G* reaction site for styrene oxide on an oligonucleotide representing the CYP1B1 gene. This approach avoids the need for radioactive isotopes and is less labor intensive and faster than the alternative PAGE with <sup>32</sup>P end labeling.

Cytochromes P450 in oxidative liver metabolism convert foreign hydrophobic substances into hydrophilic compounds that can be more readily excreted by the kidneys.<sup>1</sup> However, some metabolites may be chemically reactive electrophiles. These reactive metabolites may damage DNA, proteins, and other biomolecules or be inactivated by metabolic bioconjugation enzymes.<sup>2</sup> DNA damage often involves covalent binding of reactive metabolites to nucleobases, which may then lead to apoptosis (programmed cell death) or to neoplastic transformations.

Cancer has been linked to specific changes within the amino acid sequence of certain proteins that control the movement of cells through the cell cycle, e.g., the p53 tumor suppressor protein.<sup>3,4</sup> These amino acid changes result from mutations in specific codons in the genes. The majority of human cancers are found to have mutations in the p53 gene.<sup>5–7</sup> These mutations do not occur at random within the gene but occur at specific positions with

© 2008 American Chemical Society

\*To whom correspondence should be addressed. James.Rusling@uconn.edu.

<sup>†</sup>University of Connecticut.

<sup>‡</sup>Department of Surgery, University of Connecticut Health Center.

<sup>§</sup>Department of Cell Biology, University of Connecticut Health Center.

## SUPPORTING INFORMATION AVAILABLE

Correlation table of estimated vs actual DNA fragment size and additional figure showing capLC–MS detection of nucleobase adducts. This material is available free of charge via the Internet at <http://pubs.acs.org>.

greater frequency. Mutations at guanine bases shown as G\* appearing at codons 248 (CG\*G\*) and 273 (CG\*T) of the p53 gene are common in many types of cancer.<sup>8,9</sup> In addition, codon 157 (G\*TC) is also frequently mutated in lung cancer.<sup>10</sup> Mutations in these codons can change the identity of amino acids inserted into an expressed protein and may result in the loss of that protein's function. For example, in DNA damage induced by the benzopyrene metabolite benzo[*a*]pyrene-7,8-diol-9,10-epoxide (BPDE), adduct formation at the first guanine of codon 248 of the p53 gene causes G→T transversions.<sup>3</sup> This results in a point mutation in the p53 tumor suppressor protein, the substitution of an arginine for leucine at residue 248. The loss of the tumor suppressive capabilities of this mutated protein may contribute to uncontrolled cell growth or cancer.<sup>11,12</sup>

Mapping the nucleobase sequence specificity for reaction of a reactive metabolite with DNA can be used as a predictive tool for carcinogenicity. There are two methods for detecting the sequence specificity of a DNA-damaging compound. The first involves isotopic <sup>15</sup>N or <sup>13</sup>C labeling of specific nucleobases within an oligonucleotide and determining the amount of damage at the site by mass spectrometry.<sup>13–16</sup> This method is rapid and also identifies the type of nucleobase adduct formed at the specific site. However, isotope-labeled oligonucleotides are up to 100-fold more expensive than fluorescent-labeled oligonucleotides. Also, with isotope labeling, it is only practical to characterize previously identified specific mutation sites, and the method is limited by the size of the labeled oligonucleotides.

The second method involves size analysis of DNA fragments that result from processing at the damaged site, which can be used to determine damage at all reactive points within a DNA strand.<sup>17–19</sup> There are three major variations, differing by how the damaged site is processed to produce DNA fragments. These are as follows: (1) ligation-mediated polymerase chain reaction,<sup>20,21</sup> (2) terminal transferase-dependent polymerase chain reaction,<sup>22,23</sup> and (3) cleavage at the damaged site of a 5'-labeled DNA fragment. Many studies have been done on the sequence specificity of DNA damage using the third variation,<sup>24</sup> which is the simplest and does not use PCR amplification but is less sensitive.<sup>25,26</sup> In this variant, <sup>32</sup>P 5'-labeled DNA is incubated with an agent that causes nucleobase damage and is cleaved at the point of damage. The smaller DNA fragments produced by cleavage are then separated based on size. Mapping of the damaged point can then be determined based on the size of the fragment (Scheme 1).

Regardless of how the damaged site is processed to generate DNA fragments, size analysis is usually performed by polyacryl-amide gel electrophoresis (PAGE) with <sup>32</sup>P detection. This mode of sizing and detection is relatively inexpensive, but is labor intensive, time-consuming (1 day), and requires certification and expertise in handling radioactivity.

To avoid the use of radioactive materials, Hardman<sup>27</sup> and Bando<sup>28</sup> replaced <sup>32</sup>P with a fluorophore and used a slab gel DNA sequencer. This instrument is still used in some biotechnology laboratories, but now faces stiff competition from capillary electrophoresis with laser-induced fluorescence (CE-LIF) detection.<sup>29</sup> This is because CE-LIF requires minimal intervention by the user, which makes it faster and more suitable for automation.<sup>29,30</sup>

The aim of this paper is to demonstrate that CE-LIF presents a more convenient, safer, and faster (40–70 min) method for accurate oligonucleotide size analysis than <sup>32</sup>P-PAGE. DNA analyses of multiple samples can be done by CE-LIF instrumentation utilizing an array of capillaries. In this method, fluorescent-tagged DNA fragment samples are loaded simultaneously onto each column along with size standards labeled with a fluorophore different from the unknowns. Using multiple detectors specific for the different fluorophores,

sample and standards are run at the same time on the same column using two-color detection. The method can resolve fragments that differ by a single base. To date, CE-LIF arrays have not been used to determine the sequence specificity of DNA damage. One problem is the inability of existing CE-LIF protocols to measure absolute fragment size. For example, commercial size standards designed for the Beckmann Coulter CEQ 2000 CE-LIF arrays usually have sequences different from those of the fragments being measured. Also, the preferred fluorophore for sample fragments (WellRedD4) must be different from that of the commercial size standards (WellRedD1) to be compatible with two-color detection. Such differences cause variations in electrophoretic mobility between fragments and standards as stated by the manufacturer of the size standards. Based on our experience, size estimates using such standards can give an error by up to six bases. This error is acceptable for the standard usage of these size standards (i.e., single nucleotide polymorphism and amplified fragment length polymorphism analysis). This is because these methods depend on differences in fragment sizes between two samples and not on absolute sizes.<sup>31</sup>

Our approach to this problem is to synthesize and use primary standards with the same sequence and fluorophore as the samples. However, these primary standards cannot be run directly with the samples in a CE-LIF array since they cannot be differentiated by two-color detection. Thus, samples and primary standards are run in separate capillaries with commercial standards. However, direct comparison of migration times cannot be used to determine size of the sample fragments because of variations from capillary to capillary. Thus, determination of fragment sizes was done using the procedure described below.

We recently reported the detection of abasic sites in DNA by fluorescent tagging and CE-LIF.<sup>32</sup> In the present paper, we report a CE-LIF array method to accurately size DNA fragments that result from cleavage at damaged bases. Like <sup>32</sup>P-PAGE,<sup>20</sup> this method can be used to determine the sequence specificity of DNA damage as long as cleavage can be done at the damage points. Accurate sizing of these fragments provides the exact location of the DNA damage. For proof of concept, we used styrene oxide to create DNA damage and cleavage points. The method, summarized in Scheme 2 involves the following: (1) generation of a 5'-fluorophore-labeled oligonucleotide (e.g., from a plasmid), (2) incubation of the tagged DNA fragment with a metabolite of a xenobiotic, (3) maximizing the depurination of alkylated N-7 guanines and N-3 adenines while minimizing the depurination of intact guanines and adenines,<sup>33</sup> (4) cleavage of the abasic sites created by depurination of the adducts by mild piperidine (1 M at 37 °C for 20 min),<sup>34,35</sup> (5) denaturation of the sample at 90 °C, (6) CE-LIF size analysis of the DNA fragments using commercial standards, and (7) correction of the estimated size using a set of "primary standards". Herein, we demonstrate the new method and its accuracy for the reaction of styrene metabolite styrene oxide with an oligonucleotide representing a CYP gene coding for an enzyme involved in reactive metabolite generation.

## EXPERIMENTAL SECTION

### Chemicals and Materials

WellRedD4 5'-labeled forward primer (fp408), double-stranded (ds) salmon testes (ST) DNA of ~2 kilobase pairs (kbp), styrene oxide, and piperidine were from Sigma-Aldrich. Centricon filters with a cutoff mass of 3000 Da were obtained from Millipore. All enzymes used were from New England Biolabs. The DNA purification kit was from Qiagen. The separation gel (linear polyacrylamide), sample loading solution (SLS), and WellRedD1 5'-labeled DNA standard kit size 80 (13 and 88 nucleobases) were from Beckmann Coulter. Glycogen was from Roche. Human CYP1B1 gene cloned in pCWoRI was obtained using the method of Jansson et al.<sup>36</sup>

## Apparatus

DNA fragment sizing was done on a Beckman Coulter CEQ 2000 capillary array featuring eight capillaries filled with a linear polyacrylamide gel. Before loading the DNA onto the capillaries, samples were first denatured at 90 °C for 120 s. Loading of the samples onto the capillary was done by electrokinetic injection at 2 kV for 30 s. The separation voltage was 6 kV, and temperature was 60 °C. The CapLC (Waters) had a trapping column (Atlantis dC18, 0.18-mm i.d., 5- $\mu$ m particle size) and analytical column (Atlantis dC18, 150 mm, 300- $\mu$ m i.d., 5- $\mu$ m particle size) and is equipped with column switching. This allowed preconcentration of desired species in the trapping column, before switching flow to send the preconcentrated sample to the analytical column and MS detector.<sup>37</sup> A solvent gradient of 10 mM acetate buffer pH 4.5/1% acetic acid in methanol was used to elute the DNA adducts. Oligonucleotide sizes were estimated by reference to calibration curves made from data on the standards run simultaneously on the same column.

A Micromass Quattro 2 electrospray ionization mass spectrometer was interfaced to the CapLC. Multiple reaction monitoring (MRM) of daughter ions of N-7 guaninestylene oxide (GSO) and N-3 adeninestylene oxide (ASO) employed a dwell time of 1 s and an interchannel delay of 30 ms. Additional parameters were previously described.<sup>38</sup>

DNA in solution was measured by optical absorbance at 260 nm using a Hewlett-Packard 845X UV–visible system. DNA concentration was estimated based on the absorbance of double-stranded DNA at 260 nm where 1 absorbance unit equals 50  $\mu$ g mL<sup>-1</sup> in a 10-mm path length.<sup>39</sup>

## Optimization of Depurination Conditions

The protocol for thermal depurination was a modification of those of Koskinen<sup>40</sup> and Asaeda<sup>33</sup>. Incubation conditions at 80 °C were optimized for pH and time to maximize depurination of GSO and ASO while leaving unmodified bases intact. A report by Asaeda<sup>33</sup> and our previous paper<sup>32</sup> show that abasic sites are formed selectively when these alkylated bases are depurinated by heating.

## Amplification of 5'-Fluoropore-Tagged 147-bp DNA

A 5'-WellRedD4-tagged forward primer (fp408) and a nontagged reverse primer (rp553) was used to amplify a 147-bp DNA fragment from the human CYP1B1 pCWori plasmid using polymerase chain reaction (PCR). More than 40 different mutations in the CYP1B1 gene have been reported in individuals with the primary congenital glaucoma disease phenotype.<sup>41</sup> The sequences of the primers were (1) fp = WellRedD4—CTACTCGGAGCACTGCAAGG and (2) rp = CTGCCGCGCAA-CAGCAGCGCC. PCR employed 35 cycles of denaturation (30 s, 95 °C), annealing (30 s, 60 °C), and elongation (1 min, 72 °C) with an added elongation step (7 min, 72 °C) at the end of the cycle. The amplified 147-bp DNA was then purified using a Qiagen PCR purification kit.

## Creation of Primary Standards for More Accurate Determination of Fragment Sizes

The protocol used in this study is a modification of the method of Maxam and Gilbert.<sup>42</sup> A 5'-WellRedD4-tagged 147-bp DNA (10–15  $\mu$ g/mL) in 33  $\mu$ L of 10% formic acid solution was incubated for 30 min. This was followed by addition of 30  $\mu$ L of 2 mg mL<sup>-1</sup> glycogen, which acts as a carrier for minute amounts of DNA.<sup>43</sup> Ethanol precipitation was used to recover the DNA. To the final volume  $X$  of solution, 0.1 $X$  volume of 3 M sodium acetate and 2.5 $X$  volumes of cold ethanol were added and centrifuged at 16000g for 15 min. The sample was then washed 3 times with 70% cold ethanol, dried, and then reconstituted in 44  $\mu$ L of 10% piperidine. This solution was then incubated at 37 °C for 20 min.<sup>34</sup> A 14- $\mu$ L

sample of  $1.7 \text{ mg mL}^{-1}$  glycogen in 14% acetic acid was then added. DNA was ethanol precipitated as above and redispersed in  $20 \mu\text{L}$  of water.

### Identification of DNA Fragments with the CE-LIF Array

A  $2\text{-}\mu\text{L}$  sample of the reconstituted DNA after strand cleavage was mixed with  $40 \mu\text{L}$  of SLS premixed with  $0.2 \mu\text{L}$  of the DNA standard from kit size 80. The solution was then run on the capillary array using the conditions above.

### Validation of Primary Standards

In order to determine whether the estimated fragment size using our primary standards is the same as the actual value, which was 20 bases,  $2 \mu\text{L}$  of a  $2 \text{ nM}$  solution of a 5'-WellRedD4-tagged oligonucleotide (20 bases) with the same sequence as the first 20 bases of our 5'-WellRedD4-tagged 147-bp DNA was run on the capillary array as above.

### Generation of SO Nucleobase Adducts within the 5'-Fluoropore-Tagged DNA Fragment at pH 5.5 and pH 7.5

Two different  $50\text{-}\mu\text{L}$  solutions of the 5'-WellRedD4-tagged 147-bp DNA ( $10\text{--}15 \mu\text{g/mL}$ ) were made. One was in  $10 \text{ mM}$  acetate buffer +  $1 \text{ mM}$  EDTA pH 5.5, and the other one was in  $10 \text{ mM}$  phosphate buffer +  $1 \text{ mM}$  EDTA pH 7.5. Styrene oxide was then added to both solutions to a final concentration of 2%. Both solutions were then incubated at  $37^\circ\text{C}$  for 6 h. After addition of  $30 \mu\text{L}$  of  $2 \text{ mg mL}^{-1}$  glycogen, DNA was then ethanol precipitated. A  $20\text{-}\mu\text{L}$  aliquot of water was then added to reconstitute the DNA solution.

### Depurination of the DNA Base Adducts and Strand Cleavage

A  $20\text{-}\mu\text{L}$  sample of  $20 \text{ mM}$  phosphate buffer pH 7.65 +  $2 \text{ mM}$  EDTA was added to the  $20\text{-}\mu\text{L}$  reconstituted DNA solution. The resulting solution was incubated at  $80^\circ\text{C}$  for 2 h. Piperidine was then added to a final concentration of 10%, and the mixture was incubated at  $37^\circ\text{C}$  for 20 min. A  $14\text{-}\mu\text{L}$  sample of glycogen ( $1.7 \text{ mg mL}^{-1}$ ) in 14% acetic acid was then added, and the DNA was ethanol precipitated. A  $20\text{-}\mu\text{L}$  aliquot of water was then added to reconstitute the DNA.

### Determination of DNA Adducts after Incubation with Styrene Oxide

In order to determine whether the relative amounts of alkylated adenines and guanines correlated to strand breaks created at specific damaged sites, ST-DNA was dissolved to a concentration of  $25 \mu\text{g mL}^{-1}$  in (a)  $10 \text{ mM}$  acetate buffer pH 5.5 +  $1 \text{ mM}$  EDTA and (b)  $10 \text{ mM}$  phosphate buffer pH 7.5 +  $1 \text{ mM}$  EDTA. Styrene oxide was then added to each of the above solutions to a final concentration of 0.2%. The resulting solutions were incubated for 2 h at  $37^\circ\text{C}$ . Styrene oxide was then extracted 3 times with hexane, and the samples were then incubated for 2 h at  $80^\circ\text{C}$  for depurination. Released nucleobases were then filtered using Centricon filters with 3000 MW cutoff. The filtrate was dried using a vacufuge. The filtrate from incubation at pH 5.5 was dissolved in  $1000 \mu\text{L}$  of water, and that at pH 7.5 in  $400 \mu\text{L}$  of water. This was to ensure that detector signals in subsequent analyses fall in the linear range. DNA base adducts thus formed and isolated were analyzed by CapLC-MS/MS as above.

## RESULTS

### Optimization of Depurination Conditions

Creation of strand breaks at damaged bases on DNA depends on the ability of piperidine to cleave an abasic site (Scheme 3). An abasic site is created by removal of a nucleobase from the DNA backbone. To create stand breaks at specific points; abasic sites have first to be



created “specifically” at the damaged nucleobase sites. N-7 alkylated guanines<sup>44</sup> and N-3 alkylated adenines<sup>45</sup> depurinate at accelerated rates compared to intact guanines and adenines. However, depurination of intact nucleobases can be significant under certain conditions.<sup>40</sup> Our preliminary studies showed that there is minimal depurination of intact bases at 80 °C in 10 mM phosphate buffer, pH 7.65. Incubation at 80 °C and pH 7.65 for 2 h depurinated almost all of the N-7 GSO and N-3 ASO as shown in Figure 1 and was optimal for depurination of styrene oxide nucleobase adducts.

### Making Primary Standards

A modified Maxam–Gilbert protocol<sup>42</sup> was used to make primary standards with the same sequences and fluorophore as the fragments to be measured. Briefly, 5′-WellRedD4-tagged 147-bp DNA of known sequence was treated with formic acid and 1 M piperidine at 37 °C for 20 min (Scheme 3).

Primary standards were produced as fragments from the 147-bp DNA and were used to calibrate the Beckmann size standards. As mentioned, samples and standards are run simultaneously on columns of the CE-LIF array, and results displayed in the same electropherogram (Figure 2). Spaces correspond to regions in the DNA sequence that do not contain adenine and guanines (e.g., CCC, CTTCTTC, TCCTC) that were not cleaved because cleavage only occurs at A and G.

Locations of these spaces helped determine the identity of each fragment peak in the electropherogram (Figure 2B,C). For example, the base sequence **CGCAA** precedes sequence **CT-TCTTC** and **ACGCGC** follows it. Therefore, peaks corresponding to fragments from cleavage at G, A, and A are seen before the “CTTCTTC space” with following peaks from cleavage at A, G, and G (Figure 2B). Since the sequence of the 147-bp DNA is known, the actual fragment size of each peak can be determined (see Scheme 4). The “estimated” size of each fragment was found based on the set of Beckmann size standards. This allowed construction of a calibration curve to correct the estimated size to the exact size (Figure 3). From the calibration curve, a correction equation (eq 1) was based on linear regression:

$$\text{estimated size} = 1.036 \times \text{corrected size} + 1.572 \quad (1)$$

### Validation of the Correction Equation

Figure 4 shows the electropherogram of a purified 5′-WellRedD4-tagged 20-base oligonucleotide. Its sequence is the same as the first 20 bases of the 5′-WellRedD4-tagged 147-bp DNA. Using the Beckmann standards, the estimated size of this oligonucleotide is  $21.7 \pm 0.4$  bases, representing a +1.7 base error. Using eq 1 to correct the estimated size gave  $19.4 \pm 0.4$  bases, indicating an error decreased to −0.6 bases. A *t*-test revealed no significant difference at the 95% confidence level between the estimated and true size of the 20-base fragment. This result demonstrates the improved accuracy in fragment size determination using the standard correction approach.

### Determination of Fragment Size after Reaction with Styrene Oxide and Strand Cleavage

5′-WellRedD4-tagged 147-bp DNA was incubated with 2% styrene oxide separately in 10 mM acetate buffer, pH 5.5 and in 10 mM phosphate buffer, pH 7.5. Figures 5 and 6 show the electropherograms of incubated samples after subsequent depurination optimized for the nucleobase adducts, and then cleavage. At both pHs, there was one predominant fragment peak. Using the Beckmann standards, this peak was estimated to be  $31.63 \pm 0.87$  nucleobases at pH 5.5 and  $31.50 \pm 0.14$  nucleobases at pH 7.5. Using eq 1 to correct the

estimated size, we determined that the corrected size of the predominant peak is  $28.91 \pm 0.87$  nucleobases at pH 5.5 and  $28.74 \pm 0.14$  nucleobases at pH 7.5. (The corrected size of the fragment can be also determined using the slightly more accurate correction Table S1, Supporting Information.)

Incubation at pH 5.5 resulted in smaller peaks that were not completely separated (Figure 5B). This may be due to the presence of a significant amount of DNA base adducts that are not N-7 guanine styrene oxide or N-3 adenine styrene oxide (e.g., N-2 guanine styrene oxide), which are difficult to depurinate.<sup>40</sup> However, incubation at pH 7.5 resulted in peaks that are distinct and sharp (Figure 6B). This is because formation of nucleobase adducts other than N-7 GSO and N-3 ASO at pH 7.5 is small.<sup>40</sup> The predominant fragment peaks are numbered 1–11 and cleavage points and fragment sizes determined (Table 1).

From the size and known sequence, we determined the cleavage point that resulted in each specific fragment (cf. Schemes 1 and 4). Mild piperidine cleaves abasic sites produced from depurination of nucleobase adducts (Scheme 3). Alkylation/depurination of C and T does not occur because alkylation of pyrimidines is very slow and only N-3 alkylated adenines and N-7 alkylated guanines are ejected by neutral thermal hydrolysis.<sup>40</sup> Correction of estimated sizes using eq 1 or the correction table gives more accurate sizes. However, fragment sizes from eq 1 suggest alkylation and cleavage at some C and T residues (Table 1), which is impossible. The correction table provides more accurate corrected sizes that all indicate G and A residues as cleavage sites (Table 1).

The alkylated base that corresponds to a particular fragment peak is identified based on the actual size of the fragment. As shown in Scheme 1, a 29-nucleobase fragment results from the depurination and cleavage at base 30. This peak corresponds to selective alkylation of guanine G in the sequence [...GGCGCG-CAG...]. Using corrected sizes (Table 1) from the correction table, we determined that the nucleobases depurinated/cleaved that result in peaks 1–11 are all G except for peak 3, which is A.

CapLC–MS/MS using MRM obtained after reaction of DNA with styrene oxide and neutral thermal hydrolysis was used to confirm formation of nucleobase adducts of styrene oxide. For samples incubated at both pHs, peaks characteristic of ASO and GSO adducts were found (Figure S1, Supporting Information). These results are consistent with the CE-LIF analysis showing cleavage at G and A sites and with earlier findings that 94% of adducts formed by incubating DNA with styrene oxide at pH 7.5 are GSO.<sup>40</sup>

## DISCUSSION

Results above show that accurate sizing of DNA fragments to determine the sequence specificity of DNA damage using CE-LIF arrays can be done (Figures 5 and 6, Table 1) by referencing commercial standards to primary standards to create a correction table (Table S1) or equation (eq 1). Commercial (Beckman Coulter) standards for CE-LIF arrays necessarily employ different fluorophores and sequences than DNA fragment samples, leading to errors from different electrophoretic mobilities. Estimated size based on commercial standards alone led to errors as large as six bases (Table 1). The correction table provides greater accuracy, but eq 1 is nevertheless useful if accuracy of  $\pm 1$  base is tolerable.

The method utilizes depurination of adducted bases followed by cleavage of the abasic site. Depurination of intact guanines and adenines is avoided to prevent abasic site formation at nonadducted sites. This is achieved by optimizing the depurination buffer pH.<sup>47</sup> Excessively high pH should be avoided because N-7 alkylated guanines can be converted into “depurination resistant” formamidopyrimidine-guanine (Fapy-guanines) at high pH.<sup>48–50</sup>

Preliminary data showed minimal depurination of intact bases in pH 7.65 buffer and no significant formation of Fapy-guanines.

Depurinations of N-7 GSO and N-3 ASO from DNA obey first-order chemical kinetics,<sup>50</sup> a property of which is that the time required to consume a certain percentage of reactants is independent of initial concentration. Thus, formation of depurinated nucleobase adducts is roughly complete within 2 h (Figure 1), independent of reactant concentration.

The predominant peak found after incubation of the 5'-WellRedD4-tagged 147-bp CYP1B1 DNA with styrene indicated adduct formation at the 30th base, which is *G* in the sequence [...GGCGCGCAG...], regardless of whether the pH was 7.5 or 5.5 (Figures 5 and 6). Coincidentally, the base sequence in this gene is found to be mutated in some families containing the primary congenital glaucoma phenotype as a 10-base pair deletion, which results in a frame shift leading to truncation of the CYP1B1 protein.<sup>51</sup>

DNA damage by BPDE<sup>52</sup> prefers guanines in [CG]<sub>n</sub> sequences, while aflatoxin diol epoxide<sup>53,54</sup> prefers the 3'-guanine in GG sequences. However, experiments by Chen<sup>55</sup> show different sequence specificities of these carcinogens. The sequence T[G-GT]G is present in the K-ras, N-ras, and H-ras gene, and only codon 12 T[GGT]G of the K-ras gene has high affinity for bulky carcinogens.<sup>56</sup> Therefore, the preference for attack at a particular base is also dependent on sequences that are three or more bases away from the target nucleobase. Alkylating agents like styrene oxide that intercalate DNA may have different sequence specificity. This is partly because these agents preferentially intercalate specific sequences before reacting with the nucleobases.<sup>52</sup>

Poorer resolution of minor peaks was found for fragment samples of DNA obtained from incubation with styrene oxide at pH 5.5 (Figure 5B). This may be due to formation and retention of significant amounts of DNA base adducts with styrene oxide other than N-7 GSO and N-3 ASO at pH 5.5.<sup>57</sup> Depurination-resistant nucleobase adducts left in the fragments may cause band broadening and decreased resolution. On the other hand, at pH 7.5, 97% of the DNA base adducts formed with styrene oxide are N-7 guanine styrene oxide (94%) plus N-3 adenine styrene oxide (3%).<sup>40</sup> In the latter case, it was possible to identify all the damage sites of the DNA by CE-LIF. This example illustrates the need for sufficient resolution in the CE-LIF for our methodology to be effective. However, it is possible that a method to depurinate all adducted bases, not just N-7 G and N-3 A adducts, can be adapted to this technology to minimize this problem.

These reactive metabolites formed by cyt P450 enzymes can react with specific nucleobases within a gene and result in transitions, transversions, insertions, and deletions.<sup>58-60</sup> This results in a change in base sequence of a particular gene, which may affect codons for synthesis of specific proteins. Thus, metabolism of environmental pollutants or drugs can result in chemical reactions with specific nucleobases that ultimately alter the amino acid sequence of individual proteins and greatly influence their function. Various steps in the cell cycle are mediated by proteins that stimulate or inhibit movement through the cell cycle, and the functions can be altered by such mutations.<sup>4</sup>

With respect to metabolite-generated DNA damage, the reactivity of guanine residues in DNA has long been known.<sup>61</sup> In the 1990s, it was shown that proto-oncogenes and tumor suppressor genes have regions of higher reactivity (hot spots) to specific chemical agents.<sup>62</sup> A number of agents, such as benzo[*a*]pyrene, discussed above, generate reactive metabolites with specificity toward specific nucleotide sequences.<sup>63</sup> Within the p53 gene, these hot spots correlate to mutations within lung cancer cells.<sup>64</sup> Based on this observation, "chemical agent-specific hot spots" identified within genes should be useful biomarkers to predict the



putative carcinogenicity of a specific compound. Therefore, our method provides a valuable tool for rapid determination of sequence specificity and hot spots within specific genes.

In summary, we have shown how CE-LIF arrays can accurately size DNA fragments by using primary standards to correct size estimates derived from commercial standards. Accurate size correction is done by generating a correction table (Table S1). The method described in this paper provides a simpler, more rapid alternative to PAGE with  $^{32}\text{P}$  labeling. When coupled with metabolic enzyme activation,<sup>37,65</sup> the CE-LIF array method is very promising for predicting potential carcinogenicity of drugs and environmental chemicals and identifying potential gene mutation sites. This method is applicable for determining not only the sequence specificity of guanine and adenine damage but also cytosine and thymine damage. It would be possible to determine the sequence specificity at these points by making “primary standards” from breakpoints at cytosines and thymines.<sup>42</sup>

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

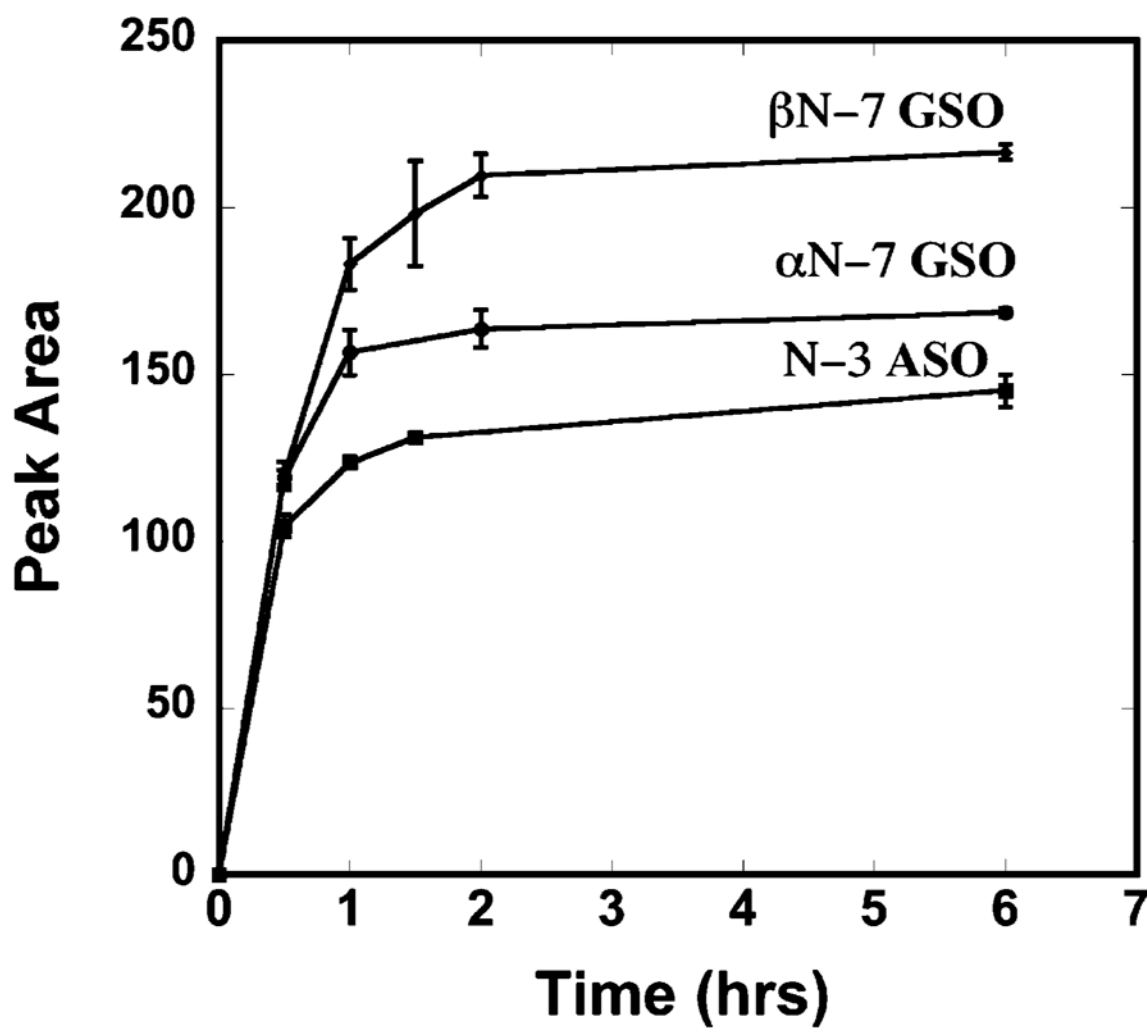
This work was supported by U.S. PHS Grant ES03154 from the National Institute of Environmental Health Services (NIEHS), NIH, and in part by Grant EY11095 from NIH and AHAF Grant G2007-036 from the American Health Assistance Program.

## References

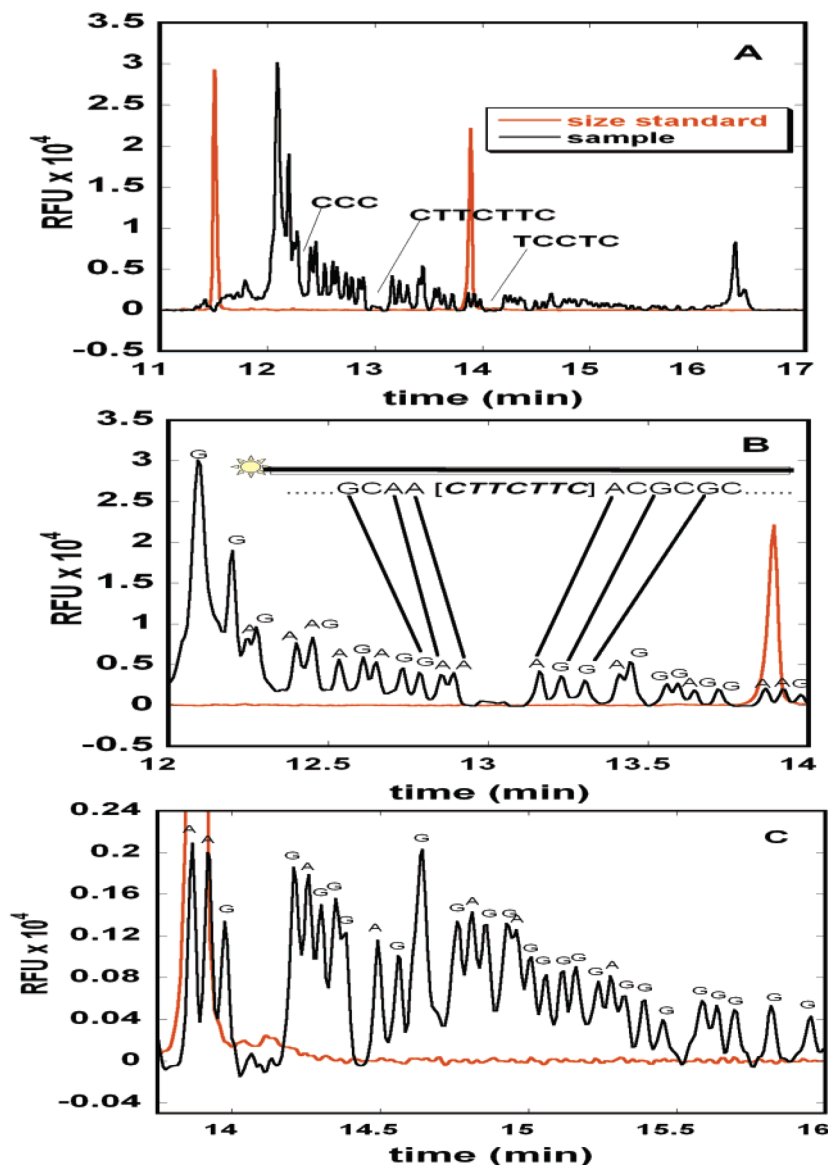
1. Rodriguez C, Ingelman-Sundberg A. *Oncogenes*. 2006; 25:1679–1691.
2. Conney A. *Cancer Res*. 2003; 63:7001–7031.
3. Rodin S, Rodin A. *Semin Cancer Biol*. 2005; 15:103–112. [PubMed: 15652455]
4. Soussi T. *Oncogene*. 2007; 26:2145–2156. [PubMed: 17401423]
5. Viktorsson K, Petris L, Lewensohn R. *Biochem Biophys Res Commun*. 2005; 331:868–880. [PubMed: 15865943]
6. Hollstein M, Sidransky D, Vogelstein B, Harris C. *Science*. 1991; 253:49–53. [PubMed: 1905840]
7. Greenblatt M, Bennett W, Hollstein M, Harris C. *Cancer Res*. 1994; 54:4855–4878. [PubMed: 8069852]
8. Pfeifer G, Denissenko M, Olivier M, Tretyakova N, Hecht S, Hainaut P. *Oncogene*. 2002; 21:7435–7451. [PubMed: 12379884]
9. Paschke T. *Mutagenesis*. 2000; 15:457–458. [PubMed: 11076995]
10. Pfeifer G, Denissenko M, Tang M. *Mutagenesis*. 1998; 13:537–538. [PubMed: 9862184]
11. Levine A. *Cell*. 1997; 88:323–331. [PubMed: 9039259]
12. Hahn WC, Weinberg RA. *Nat Rev Cancer*. 2002; 2:331–341. [PubMed: 12044009]
13. Ziegel A, Shallop A, Jones R, Tretyakova N. *Chem Res Toxicol*. 2003; 16:541–550. [PubMed: 12703972]
14. Tretyakova N, Matter B, Jones R, Shallop A. *Biochemistry*. 2002; 41:9535–9544. [PubMed: 12135376]
15. Matter B, Guza R, Zhao J, Li Z, Jones R, Tretyakova N. *Chem Res Toxicol*. 2007; 20:1379–1387. [PubMed: 17867647]
16. Rajesh M, Wang G, Jones R, Tretyakova N. *Biochemistry*. 2005; 44:2197–2207. [PubMed: 15697245]
17. Pereira T, Webb R, Reilly P, Seawright A, Prakash A. *Nucleic Acids Res*. 1998; 26:5441–5447. [PubMed: 9826770]
18. McHugh P, Knowland J. *Nucleic Acids Res*. 1995; 23:1664–1670. [PubMed: 7784169]
19. Komura J, Riggs A. *Nucleic Acids Res*. 1998; 26:1807–1811. [PubMed: 9512556]

20. Pfeifer G, Singer J, Riggs A. *Methods Enzymol.* 1993; 225:567–583. [PubMed: 7694044]
21. Angers M, Cloutier JF, Castonguay A, Drouin R. *Nucleic Acids Res.* 2001; 29:1–11. [PubMed: 11125037]
22. Pfeifer G, Chen H, Komura J, Riggs A. *Methods Enzymol.* 1999; 304:548–571. [PubMed: 10372381]
23. Cloutier J, Drouin R, Weinfeld M, O'Connor T, Castonguay A. *J Mol Biol.* 2001; 313:539–557. [PubMed: 11676538]
24. Kawanishi S, Hiraku Y, Murata M, Oikawa S. *Free Radical Biol Med.* 2002; 32:822–832. [PubMed: 11978484]
25. Meggers E, Michel-Beyerle M, Giese B. *J Am Chem Soc.* 1998; 120:12950–12955.
26. Saluz H, Wiebauer K, Wallace A. *Trends Genet.* 1991; 7:207–211. [PubMed: 1887501]
27. Hardman L, Murray V. *Biochem Mol Biol Int.* 1997; 42:349–359. [PubMed: 9238534]
28. Bando T, Iida H, Tao Z, Narita L, Fukuda N, Yamori N, Sugiyama H. *Chem Biol.* 2003; 10:751–758. [PubMed: 12954334]
29. Mardis E. *J Biomol Technol.* 1999; 10:137–143.
30. Pieris K, Green H, Hartley S, Jordan D, Lahec S, Livett R, Tsang K, Ward D. *Electrophoresis.* 2004; 25:2227–2241. [PubMed: 15274007]
31. Meudt H, Clarke A. *Trends Plant Sci.* 2007; 13:106–117. [PubMed: 17303467]
32. Fundador E, Rusling JF. *Anal Bioanal Chem.* 2007; 387:1883–1890. [PubMed: 17206410]
33. Asaeda A, Ide H, Terato H, Takamori Y, Kubo K. *Anal Chim Acta.* 1998; 365:35–41.
34. Ye N, Holmquist G, O'Connor T. *J Mol Biol.* 1998; 284:269–285. [PubMed: 9813117]
35. Cloutier J, Drouin R, Weinfeld M, O'Connor T, Castonguay A. *J Mol Biol.* 2001; 313:539–557. [PubMed: 11676538]
36. Jansson I, Stoilov B, Sarfarazi M, Schenkman J. *Pharmacogenetics.* 2001; 11:793–801. [PubMed: 11740343]
37. Tarun M, Bajrami B, Rusling JF. *Anal Chem.* 2006; 78:624–627. [PubMed: 16408950]
38. (a) Tarun M, Rusling JF. *Anal Chem.* 2005; 77:2056–2062. [PubMed: 15801738] (b) Tarun M, Rusling JF. *Crit Rev Eukaryotic Gene Expression.* 2005; 15:295–315.
39. Mey M, Legueux G, Maertens J, Maeseneire S, Soetaert W, Vandamme E. *Anal Biochem.* 2006; 353:198–203. [PubMed: 16545766]
40. Koskinen M, Vodickova L, Vodicka P, Warner S, Hemminki K. *Chem Biol Interact.* 2000; 124:13–27. [PubMed: 10658899]
41. Sarfarazi M, Stoilov I, Schenkman J. *Ophthalmol Clin N Am.* 2003; 16:543–554.
42. Maxam A, Gilbert W. *Methods Enzymol.* 1980; 65:499–560. [PubMed: 6246368]
43. Hengen P. *Trends Biochem Sci.* 1996; 21:224–225. [PubMed: 8744357]
44. Gates K, Nooner T, Dutta S. *Chem Res Toxicol.* 2004; 17:839–853. [PubMed: 15257608]
45. Bejjani B, Stockton D, Lewis R, Tomey K, Dueker D, Jabak M, Astle W, Lupski J. *Hum Mol Genet.* 2000; 9:367–374. [PubMed: 10655546]
46. Stuart G, Chamber R. *Nucleic Acids Res.* 1987; 15:7451–7462. [PubMed: 3658699]
47. Greer S, Zamenhof S. *J Mol Biol.* 1962; 4:123–141. [PubMed: 13901601]
48. Chetsanga C, Bearie B, Makaroff C. *Chem Biol Interact.* 1982; 41:217–233. [PubMed: 7049418]
49. Hemminki K, Peltonen K, Vodicka P. *Chem Biol Interact.* 1989; 70:289–303. [PubMed: 2743474]
50. Vodicka P, Hemminki K. *Chem Biol Interact.* 1988; 68:117–126. [PubMed: 3203402]
51. Bejjani B, Stockton D, Lewis R, Tomey K, Dueker D, Jabak M, Astle W, Lupski J. *Hum Mol Genet.* 2000; 9:367–374. [PubMed: 10655546]
52. Geacintov N, Shahbaz M, Ibanez V, Moussaoui K, Harvey R. *Biochemistry.* 1988; 27:8380–8387. [PubMed: 3149504]
53. Misra R, Muench K, Humayn Z. *Biochemistry.* 1983; 22:3351–3359. [PubMed: 6412744]
54. Ross M, Mathison B, Said B, Shank R. *Biochem Biophys Res Commun.* 1999; 254:114–119. [PubMed: 9920742]
55. Chen J, Zheng Y, West M, Tang M. *Cancer Res.* 1998; 58:2070–2075. [PubMed: 9605744]

56. Feng Z, Wenwei H, Chen J, Pao A, Li H, Rom W, Hung M, Tang M. J Nat Cancer Inst. 2002; 94:1527–1536. [PubMed: 12381705]
57. Latif F, Moschel R, Hemminki K, Dipple A. Chem Res Toxicol. 1988; 1:364–369. [PubMed: 2979753]
58. Jernstrom B, Graslund A. Biophys Chem. 1994; 49:185–199. [PubMed: 8018817]
59. Johnson W, He Q, Tomasz M. Bioorg Med Chem. 1995; 3:851–860. [PubMed: 7582961]
60. Aguilar F, Hussain S, Cerutti P. Proc Natl Acad Sci USA. 1993; 90:8586–8590. [PubMed: 8397412]
61. Benzer S. Proc Natl Acad Sci USA. 1961; 47:403–415. [PubMed: 16590840]
62. Puisieux A, Lim S, Groopman J, Ozturk M. Cancer Res. 1991; 51:6185–6189. [PubMed: 1933877]
63. Denissenko M, Chen J, Tang M, Pfeifer G. Proc Natl Acad Sci USA. 1997; 94:3893–3898. [PubMed: 9108075]
64. Denissenko M, Pao A, Tang M, Pfeifer G. Science. 1996; 274:430–432. [PubMed: 8832894]
65. Hvastkovs EG, So M, Krishnan S, Bajrami B, Tarun M, Jansson I, Schenkman JB, Rusling JF. Anal Chem. 2007; 79:1897–1906. [PubMed: 17261025]



**Figure 1.** Effect of incubation time on the amount of DNA base adducts depurinated. Depurination of adducts reaches completion by ~ 2 h. Incubation was done in 10 mM phosphate buffer, pH 7.65 at 80 °C.



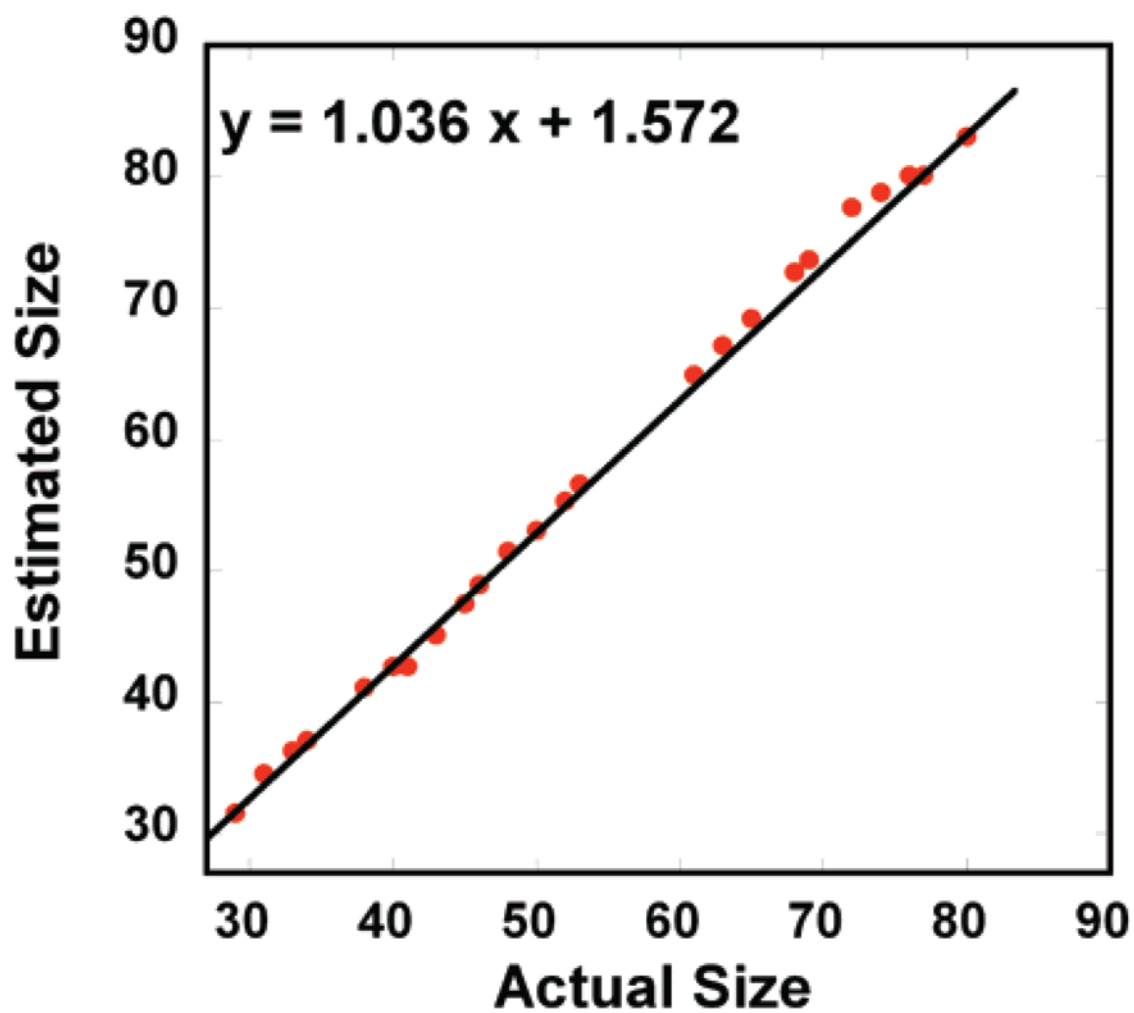
**Figure 2.**

DNA fragment peaks produced after incubation of 5'-WellRedD4-tagged 147-bp DNA in 10% formic acid to eject G and A, followed by mild piperidine treatment for cleavage at abasic sites. Beckman standards are in red. Peaks elute in the order of increasing size, i.e., larger fragments appear at longer times. RFU is relative fluorescence unit. The DNA sequence was ([WellRedD4]5'-CTACTCGGAGCACTGGAAGGTGCAGCGGCGCGCAG**CCC**ACAGCATGATGCGCA**ACTTCTTC**ACGCGCCAGCCGCGCAGCCGCCAAG**TCCTC**GAGGGCCACGTGCTGAGCGAGGCGCGCAGCTGGTGGCGCTGCTGGTGC

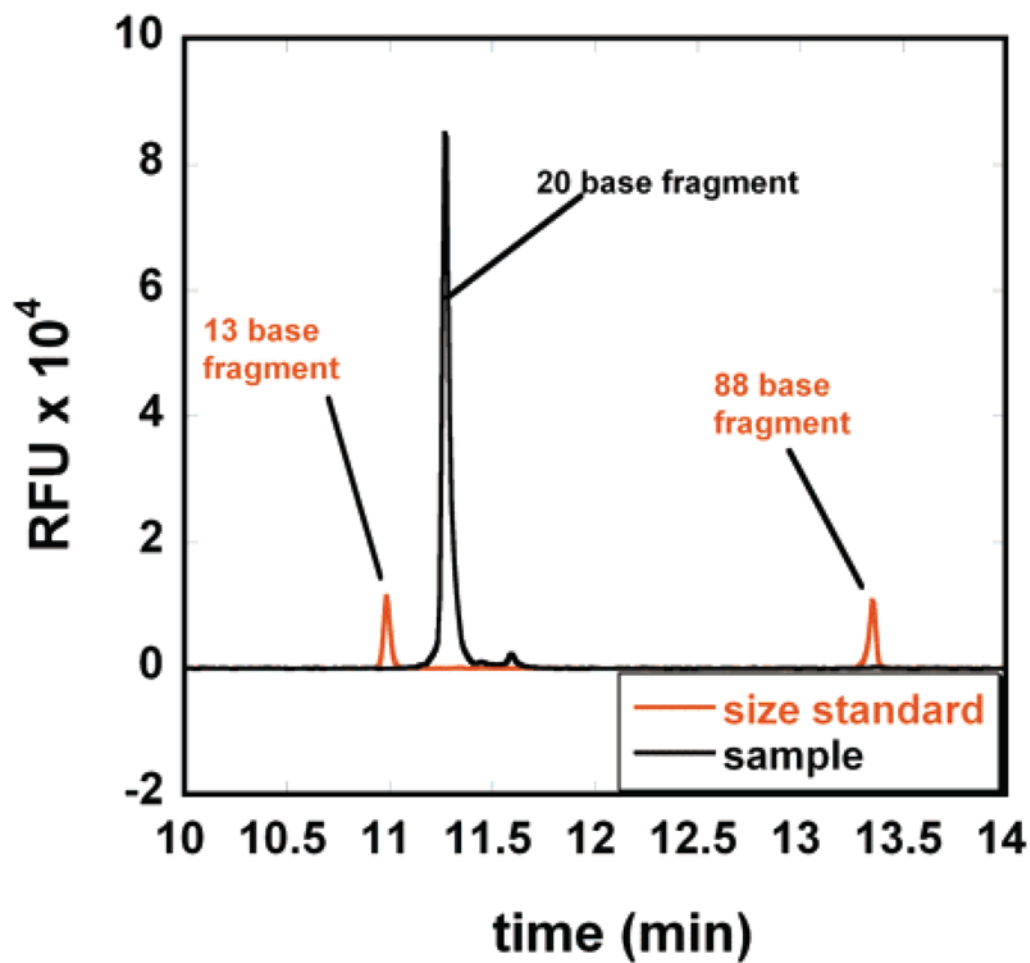
GCGGCAG). (B) and (C) are the expanded regions of (A), and peak labels denote original ejected A or G. Multiletter labels in (A) and (B) represent consecutive sequences without an A and G. If cytosines (C) and thymines (T) were cleavable, fragments should appear in this spaces. Spaces in (A) reflect regions **CCC**, **CTTCTTC**, and **TCCTC** respectively denoted in boldface type within the sequence (Scheme 4). In (B), based on the original sequence, the **CTTCTTC** space is preceded by peaks for fragments cleaved at G, A, and A and followed by A, G, and G.



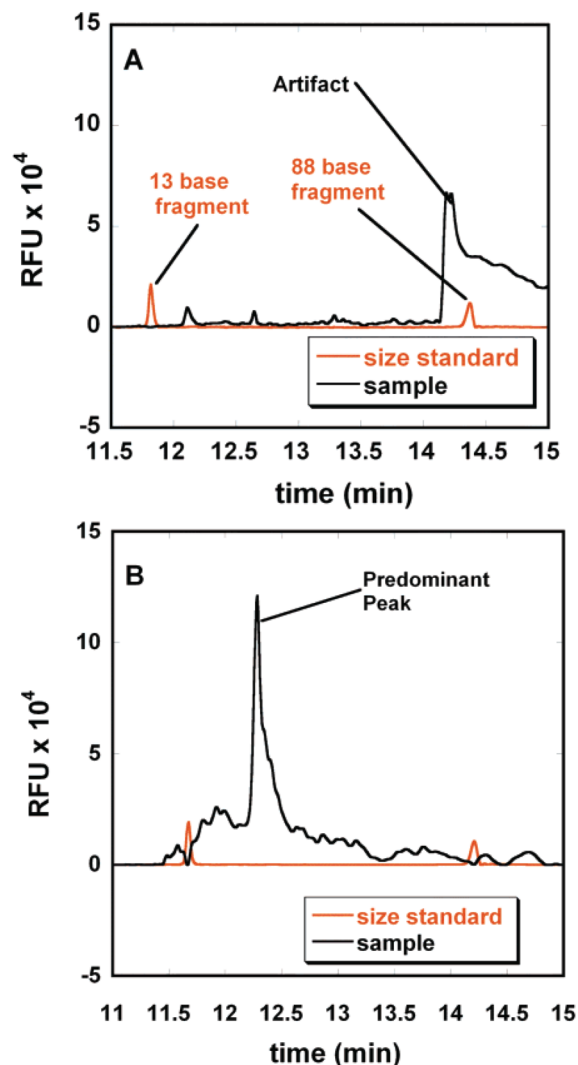
Identities of all peaks were determined in this manner. It is important that capillary temperature be at maximum value (e.g., 60 °C) in order to prevent formation of secondary structures, which lead to anomalous migration.



**Figure 3.** Calibration curve relating estimated fragment size to actual size using the standard correction procedure described in the text.

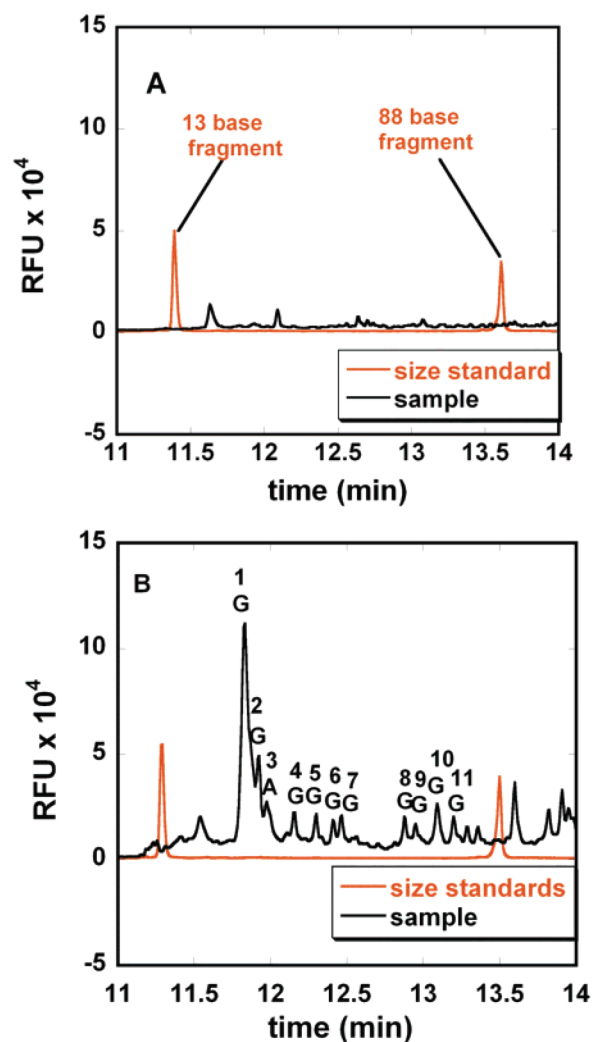


**Figure 4.** Electropherogram of the WellRedD4-tagged 20 -ase oligonucleotide and two standards. Using the Beckmann Coulter standards, the estimated oligonucleotide size is 21.7. Correction of the estimated size using eq 1 or Figure 3 gave a corrected size of 19.4.



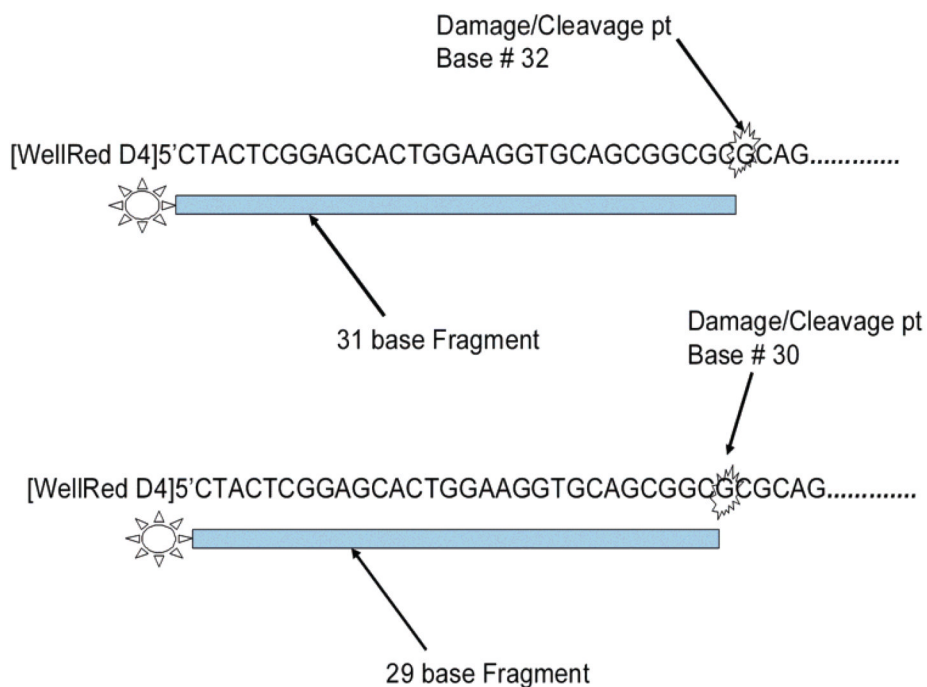
**Figure 5.**

CE-LIF of 147-bp DNA after incubation at pH 5.5 with (A) 0 and (B) 2% styrene oxide, followed by depurination and cleavage with mild piperidine. Using the Beckmann Coulter standards, the estimated size of the peak is 31. Using the correction eq 1 derived from primary standards, the size of the peak is 29. This corresponds to the attack of styrene oxide on the 30th base, which is *G* in the sequence [...GGCGCGCAG...]. The artifact peak in (A) at 14.3 min was identified as part of the 147-bp DNA that was not properly sieved due to overloading. It was not observed in (B) because most of the 147-bp DNA was cleaved.



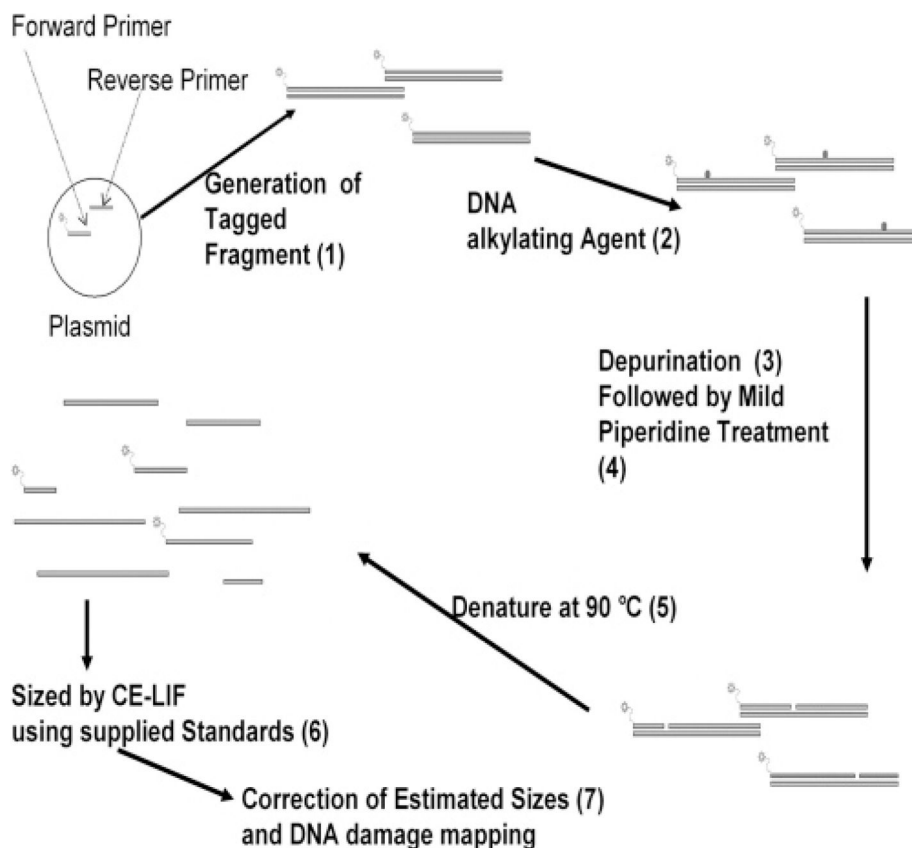
**Figure 6.** CE-LIF of 147-bp DNA after incubation at pH 7.5 with (A) 0 and (B) 2% styrene oxide, followed by optimal depurination and cleavage with mild piperidine. Using the Beckmann Coulter standards, the estimated size of the predominant peak was 31. Using the correction equation, the size of the peak is 29. This corresponds to the attack of styrene oxide on the 30th base, which is *G* in the sequence [...GGCGCGCAG...]. The predominant peaks within the range of the size standards are labeled 1–11. Using the correction table, these correspond to DNA fragments that are 29, 31, 33, 40, 45, 48, 50, 63, 65, 69, and 72 bases, respectively.





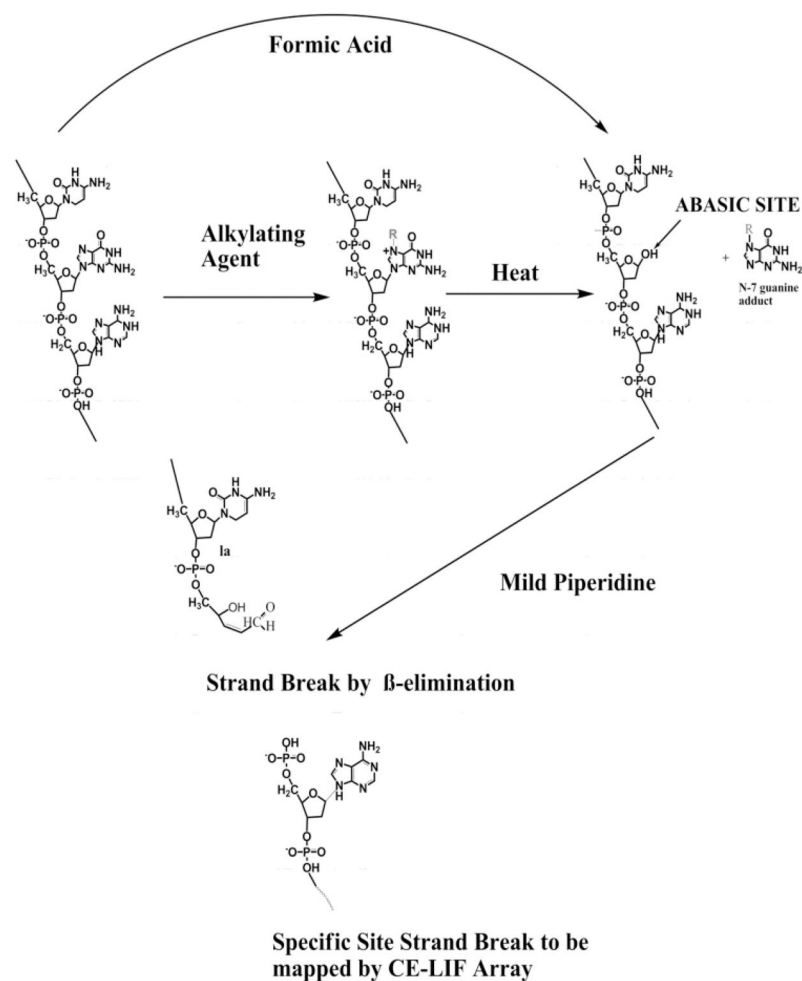
**Scheme 1. Observation of a 31-Base Fragment Means Damage at Base 32 and a 29-Base Fragment Means Damage at Base 30 As Illustrated<sup>a</sup>**

<sup>a</sup> The dotted line implies that the sequence extends to 147 bases.



### Scheme 2. Mapping DNA Damage by CE-LIF<sup>a</sup>

<sup>a</sup> In this method (1) a 147-bp WellRedD4 5'-tagged DNA is (2) incubated with an alkylating agent, e.g., styrene oxide. The (3) alkylated N-7 guanines and N-3 adenines are then depurinated at 80 °C. The (4) abasic sites at specific points, which result from depurination of DNA base adducts are treated with 1 M piperidine at 37 °C for 20 min resulting in strand breaks. The (5) resulting fragments are then denatured at 90 °C and (6) sized by the CE-LIF array using size standards. (7) Estimated sizes are then corrected using primary standards. The corrected size is then used to map the sequence specific DNA damage.

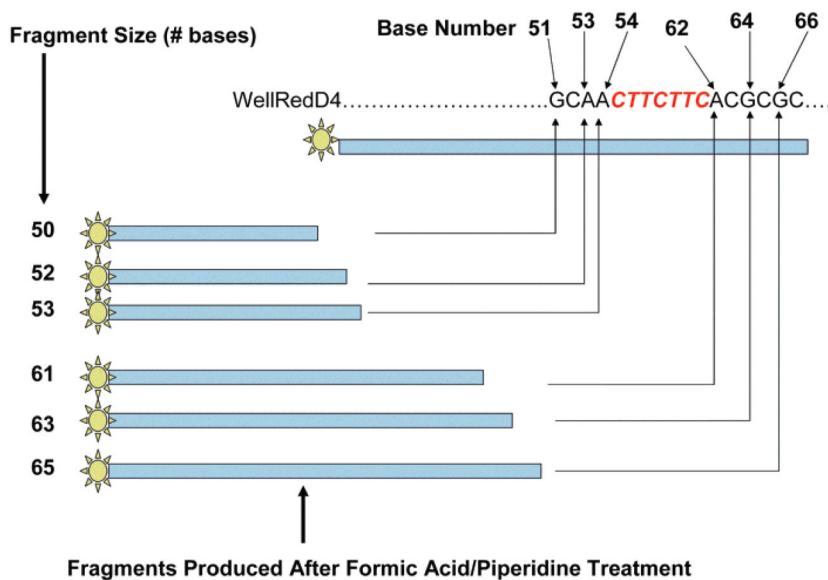


### Scheme 3. Alkylation or Protonation at the N-7 of Guanine<sup>a</sup>

<sup>a</sup> This creates a positive charge on this nucleobase thus making it a good leaving group.<sup>44</sup>

Abasic sites are formed by DNA base alkylation/heating or formic acid treatment.

Subsequent mild piperidine treatment cleaves abasic site via  $\beta$  elimination.<sup>46</sup>



**Scheme 4. Illustration of a Section of 5'-WellRedD4 147-bp DNA Correlated with Results in Figure 2<sup>a</sup>**

<sup>a</sup> At top, dotted lines indicate bases present before and after the sequences shown. Fragments produced from cleavage have sizes equal to  $N-1$  bases, where  $N$  is the base number of the adenine or guanine cleaved. The fragments separated by CE-LIF are seen as peaks in the electropherogram with the first peak being the smallest (Figure 2b). There are no fragments produced between 53 and 61 bases in length because there is no G or A at which to cleave between base numbers 54 and 62, resulting in a space in the electropherogram (e.g., CTTCTTC space) as seen in Figure 2B.

**Table 1**

Actual Sizes of Predominant DNA Fragment Peaks from 147 5'-WellRedD4-Tagged bp DNA Incubated with Styrene Oxide at a pH 7.5 Followed by Depurination and Mild Piperidine Cleavage<sup>a</sup>

peak no.	size estimated from Beckmann Coulter stds	corrected size based on Table S1	corrected size based on eq 1	nucleobase alkylated and depurinated from Table S1 vs eq 1 <sup>a</sup>
1	31.49 ± 0.14	29	28.44 ≈ 28	G vs C
2	34.75 ± 0.18	31	31.89 ≈ 32	G vs C
3	36.43 ± 0.19	33	33.52 ≈ 34	A vs G
4	42.26 ± 0.18	40	39.16 ≈ 39	G vs A
5	47.39 ± 0.17	45	44.05 ≈ 44	G vs T
6	51.17 ± 0.09	48	47.84 ≈ 48	G vs G
7	52.86 ± 0.40	50	49.47 ≈ 49	G vs C
8	66.92 ± 0.41	63	63.04 ≈ 63	G vs G
9	69.38 ± 0.29	65	65.42 ≈ 65	G vs G
10	74.10 ± 0.34	69	69.90 ≈ 70	G vs C
11	77.86 ± 0.45	72	73.13 ≈ 73	G vs C

<sup>a</sup>The nucleobase site of cleavage is also determined. Break points are shown as the superscripted numbers: ([WellRedD4]5'...CG<sup>1</sup>CG<sup>2</sup>-CA<sup>3</sup>GCCCACAG<sup>4</sup>CATG<sup>5</sup>ATG<sup>6</sup>CG<sup>7</sup>CAACTTCTTCACG<sup>8</sup>CG<sup>9</sup>CCAG<sup>10</sup>CCG<sup>11</sup>CGCAGCC....