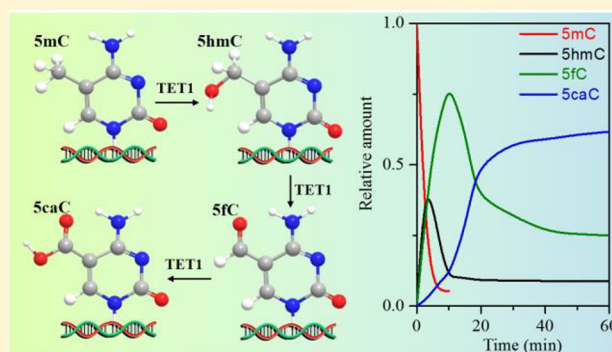


Label-Free Raman Observation of TET1 Protein-Mediated Epigenetic Alterations in DNA

Xiaojun Luo,[†] Lijuan Jiang,[†] Tuli Kang,[†] Yingfang Xing,[†] Erjin Zheng,[‡] Ping Wu,^{*,†} Chenxin Cai,^{*,†} and Qiuming Yu^{*,‡}[†]Jiangsu Key Laboratory of New Power Batteries, Jiangsu Collaborative Innovation Center of Biomedical Functional Materials, College of Chemistry and Materials Science, Nanjing Normal University, Nanjing, Jiangsu 210097, P.R. China[‡]Department of Chemical Engineering, University of Washington, Seattle, Washington 98195, United States

Supporting Information

ABSTRACT: Epigenetic modifications of DNA are known to modulate gene activity and expression and are believed to result in genetic diseases, such as cancer. Four modified cytosines were discovered in mammalian genomes: 5-methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxycytosine (5caC). They are regarded as DNA epigenetic markers and play key roles in the regulation of the dynamic balance between DNA methylation and demethylation. Although detection approaches toward 5mC are ubiquitous, few assays have reported the simultaneous determination of all four modified cytosines as well as monitoring of their dynamic alterations. Here, we developed a label-free surface enhanced Raman spectroscopy (SERS)-based method for directly sensing the four DNA modifications by using a plasmonic gold nanohole array (PGNA) with well-controlled hot spots and an open surface as the substrate. This method is based on identifying SERS spectral features resulting from DNA base modifications. Our study shows that 5mC, 5hmC, 5fC, and 5caC exhibit distinct Raman spectroscopic signatures at 785, 660, 1450, and 1680 cm^{-1} , respectively. Moreover, the developed method can be used for tracking of the dynamic alterations among these four modified cytosines in DNA mediated by the ten-eleven translocation (TET) protein. The dynamic stepwise conversion from 5mC into 5hmC, 5fC, and 5caC is further demonstrated to be a typical three-step consecutive reaction with rate constants of 0.6, 0.25, and 0.15 min^{-1} , respectively, which has not been achieved before via a SERS-based method.



Many human diseases, including cancer, have been demonstrated to be the result of a stepwise accumulation of epigenetic alterations.^{1–5} Various epigenetic marks have been described, ranging from DNA to histone modifications, which can affect the way the cell reads its genome and hence its transcriptional output.⁶ DNA methylation, occurring at the carbon 5 position of the pyrimidine ring in cytosine residues (5-methylcytosine (5mC)), is the best characterized epigenetic event and plays a significant role in the modulation of transcriptional activity and other genome functions. Aberrations of the DNA methylation can alter the expression of genes, perturb entire metabolic pathways, and even promote the carcinogenic transformation of healthy cells.^{7–10} Recently, 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxycytosine (5caC) were discovered as new DNA epigenetic markers, which are downstream derivatives of 5mC and can be enzymatically oxidized from 5mC by ten-eleven translocation (TET) family proteins.^{11,12} It is currently accepted that 5hmC, 5fC, and 5caC perform epigenetic functions distinct from 5mC. For example, 5hmC, widely accumulated in neurons and self-renewing and pluripotent stem cells, plays a significant role in

stem cell differentiation, epigenetic reprogramming, leukemia, neurodegenerative diseases, hypoxia, and angiogenesis. Extensive loss of hydroxymethylcytosine may cause cell mutation and even carcinogenesis.^{13,14} Moreover, 5hmC, 5fC, and 5caC have demonstrated direct involvement in DNA demethylation, where the demethylation is initiated by the oxidation of 5mC to 5hmC and then to 5fC and 5caC. These oxidized products can then be excised by thymine DNA glycosylase and the subsequent base-excision repair pathway.^{15,16} These findings implicated the importance of DNA epigenetic modifications in gene expression patterns and the regulation of the dynamic balance between DNA methylation and demethylation. Hence, the idea of exploring a new method to detect DNA epigenetic modifications as well as monitor their alterations is explored.

The typical methods for detecting DNA modifications include bisulfite conversion, enzymatic approaches, polymerase chain reactions, liquid chromatography, capillary electrophoresis, electrochemical measurements, and various spectroscopic

Received: February 24, 2019

Accepted: May 10, 2019

Published: May 10, 2019

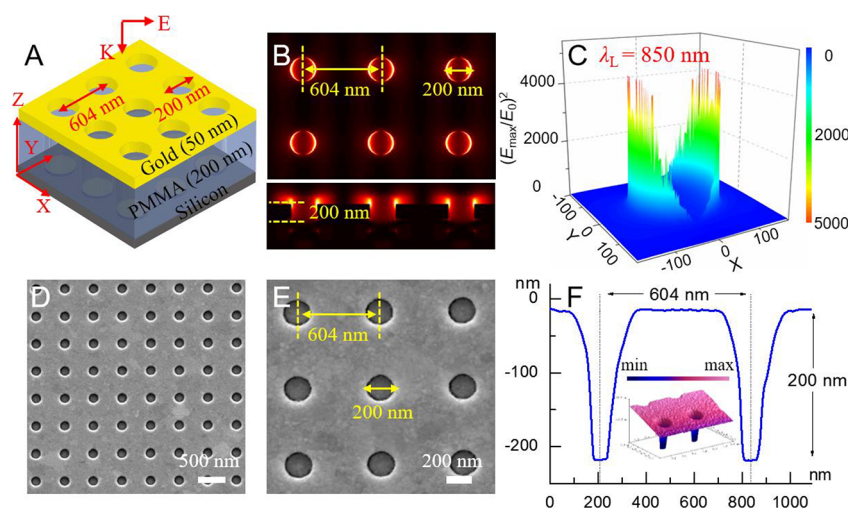


Figure 1. (A) 3D schematic illustration of the PGNA nanostructure. (B) FDTD simulated electric field distribution at the x - y (top) and x - z (bottom) plane of PGNA. (C) 3D electric field distribution displayed at the top Au/water interface of a nanohole. The scale bar represents $(E_{\max}/E_0)^2$. (D and E) SEM images at low and high magnification of the as-fabricated PGNA. (F) A line profile shows the depth as 200 nm, pitch as 604 nm, and radius as 100 nm of the as-fabricated PGNA. Inset is a $1.5\ \mu\text{m} \times 1\ \mu\text{m}$ 3D AFM image.

copy-based approaches. Among these methods, the bisulfite conversion assay is currently considered the gold standard assay technique; however, it fails to discriminate 5hmC from 5mC and even interpreted 5fC and 5caC as unmodified cytosines.^{6,11,17–19} Currently, very few works have reported detection methods available to detect all four epigenetically modified cytosines in DNA. Ito and co-workers successfully quantified the genomic content of 5mC, 5hmC, 5fC, and 5caC by collecting their liquid chromatographic peaks and subjecting them to high-resolution mass spectrometric analysis.¹¹ However, this method requires complex procedures and long detection times, which make it not conducive for rapid analysis.

Surface-enhanced Raman spectroscopy (SERS) has been proven to be a powerful analytical tool for the rapid detection and structural characterization of DNA. For example, Ren and co-workers successfully used the phosphate backbone signal as an internal standard to achieve a reliable determination of the DNA structure and the clear identification of DNA with single-base sensitivity.²⁰ He and co-workers obtained distinct and reproducible SERS spectra of DNA molecules by use of a silver nanoparticle-decorated silicon wafer (Ag NPs@Si) as the SERS chip and further set up a SERS database for artificial-intelligence-based discrimination of tumor suppressor genes, which enables a 90.28% accuracy rate for the recognition of a specific DNA target.²¹ Chen et al. developed a novel and powerful technology, nanoslit SERS, which combines SERS with plasmonic nanoslits, enabling a real-time SERS measurement of DNA with single-molecule sensitivity and subnanometer spatial resolution.²² El-Sayed and co-workers also probed the unique dehydration-induced structural modifications in DNA extracted from cancer cells using SERS.²³ Halas and co-workers demonstrated that SERS can be used to detect DNA and identify chemically modified DNA bases, including methylated adenine, 5mC, 5hmC, and oxidized guanine in purely synthesized DNA via the identification of SERS spectral variations caused by DNA base modifications.^{24,25} Recently, Ouyang et al. reported using a Raman reporter assay to detect 5mC, 5hmC, 5fC, and 5caC in DNA, in which the Raman signal reporters (SYBR Green I) were labeled to transduce the

signal.²⁶ Although an accurate and rapid detection was endowed, this method failed to obtain simultaneous detection of the different epigenetic modifications or monitor their real-time dynamical alterations in the TET protein-mediated oxidation process.

In this paper, we report a label-free SERS-based method for directly sensing 5mC, 5hmC, 5fC, and 5caC in DNA. This method is based on identifying the SERS spectral features resulting from the modified cytosine in DNA. Since the Raman signal of DNA molecules is very weak and indiscernible, a subwavelength plasmonic gold nanohole array (PGNA) in an ultrathin gold film was used as the SERS substrate (Figure 1A) to enhance the Raman signals from DNA. The enhancement of the Raman signals from the localized surface plasmon resonance (LSPR) associated with the nanoholes is approximately 6 orders of magnitude. Additionally, the periodically distributed nanoholes generate uniform “hot spots”, allowing a highly reproducible readout of the signal. Moreover, the strongest local electric field can be specifically tuned to the Au/water interface at the top surface of the PGNA, thereby enabling an open and wide surface for DNA absorption. Thus, highly sensitive and reproducible SERS signals for DNA can be obtained on PGNA. Herein, we show that 5mC, 5hmC, 5fC, and 5caC exhibit distinct Raman spectroscopic fingerprint signatures, through which a sensitive, reproducible, and direct detection of the epigenetic modifications of DNA can be achieved. This method is also used for monitoring the dynamic DNA epigenetic alterations in the TET protein-mediated oxidation process, which has not been achieved previously by using a SERS-based detection approach. This method would make the detection of DNA epigenetic modifications straightforward, in which no labeling and complex procedure are needed, thus eliminating the damage to DNA by other chemicals. More importantly, this proposed method will provide dynamic tracking of alterations in DNA epigenetic modifications with the promise of better understanding the regulation mechanism of DNA epigenetic modifications.

■ EXPERIMENTAL SECTION

Design and Fabrication of SERS-Active PGNA Substrates. The PGNA substrate was designed using Finite-Difference Time-Domain (FDTD) simulation to gain the maximum electric field on the top Au surface by adjusting the structural parameters including diameter, depth, and pitch of nanoholes (see the [Supporting Information](#) for details). The optimal PGNA was fabricated using a variety of nanofabrication techniques following our previous work.^{27–29} Typically, a 200 nm thick film of poly(methyl methacrylate) (950PMMA A6 purchased from MircoChem (Westborough, MA)) was spin-coated at a speed of 4000 rpm on the precleaned silicon substrate and then baked at 180 °C for 90 s. A nanohole pattern was generated after development in a 3:1 mixture of isopropyl alcohol/methyl isobutyl ketone (IPA/MIBK) for 70 s followed by an IPA rinse and a N₂ blow-dry. The total written area was 50 $\mu\text{m} \times 50 \mu\text{m}$ with nanoholes of 100 nm radius, 200 nm depth, and 604 nm pitch. Next, a 50 nm thick gold film was evaporated onto the nanohole array at the deposition rate of 0.1 nm s^{−1} under a background pressure of 2×10^{-6} mbar. Then, we cleaned the substrates in a UV–O₃ for 20 min, rinsed it with 18.2 M $\Omega \cdot \text{cm}^{-1}$ of deionized water, and dried it with a N₂ stream. Finally, the PGNA substrates were ready to use.

The structure of the as-prepared PGNA was confirmed by a JEOL scanning electron microscope (SEM, JSM-7600F field-emitting scanning electron microscope) and a tapping mode atomic force microscope (AFM, Nanoscope IIIa scanning probe microscope). The reflectance spectrum of the PGNA was recorded on a Nexus 670 FT-IR spectrophotometer (Nicolet Instruments).

SERS Measurements. All of the single-stranded DNA (ssDNA) sequences (listed in [Table 1](#)) were purchased from

mixtures were taken out at different incubation time intervals (0, 2.5, 5, 7.5, 10, 12.5, 15, 17.5, 20, 30, 40, 50, and 60 min) and quenched by adding 250 μL of cold buffer PN (QIAGEN). The DNA products were then purified via a QIAquick Nucleotide Removal Kit (QIAGEN) following the manufacturer's instructions, which ensures removal of primers <10 bases, enzymes, salts, and unincorporated nucleotides. The concentration and quality of the extracted DNA were evaluated by measuring the absorbance at 260 and 280 nm.

Finally, the DNA solution ($\sim 10 \mu\text{M}$, 50 μL) was dropped on the PGNA substrate at 25 °C to prepare it for the SERS measurements. All the DNA samples were measured in solution. The SERS spectra were collected using a Labram HR 800 microspectrometer (Jobin Yvon) equipped with a He–Ne laser ($\lambda = 785 \text{ nm}$; power: 0.5 mW). A low power of laser was used to prevent the evaporation of solution under irradiation. A 50 $\times$ long work distance objective (NA = 0.75) was used to focus the laser beam onto the sample and collect the scattered light. The confocal hole was set to 200 μm . The grating of 600 g mm^{−1} was used. Typically, 50 μL of the DNA solution was dropped onto the PGNA and directly measured. The collection time of each SERS spectrum was 6 s, and 5 cycles were accumulated over a spectral range from 400 to 1800 cm^{−1}. SERS spectra were randomly collected from DNA solution at different sites on the PGNA substrate. Each SERS spectrum, consisting of 539 data points, was baselined-corrected and smoothed using LabSpec 5. To obtain the SERS barcodes, the spectra of each sample were input into MATLAB to generate a 539 $\times$ 30 matrix. The matrix was visualized as a SERS barcode using the code surface in MATLAB. Principle Component Analysis (PCA) of the SERS spectra of the different DNA samples was analyzed using SPSS statistics software (SPSS Inc.). The two-dimensional Raman spectra were obtained using OMNIC software.

■ RESULTS AND DISCUSSION

Design, Fabrication, and Characterization of PGNA.

We designed the PGNA ([Figure 1A](#)) using FDTD simulations. The FDTD simulated results demonstrated that the distribution and intensity of the electric field depend sensitively on the structural parameter of PGNAs, including the radius, depth, and the pitch ([Figures S1–S3](#)). The maximum local electric fields were observed around the edge of the nanoholes at the top gold film when the nanohole radius was 100 nm, the depth was 200 nm, and the pitch was 604 nm ([Figures 1B, S3D, and S4](#)). Such an electric field distribution will also greatly improve DNA Raman signals because DNA molecules will be accessibly absorbed at the top surface of the substrate and thus located in the region of enhanced electric field. Moreover, the PGNA (100–200–604, radius–depth–pitch in nm) has a strong LSPR peak at 850 nm ([Figures 1C, S1B, S2C, and S3D](#)), which is desirable to achieve the maximum SERS effect. Because the maximum Raman enhancement can be obtained for an LSPR located between the wavelengths of the excitation source and the scattered Raman photons,³⁰ it is commonly accepted that the optimal wavelength of LSPR (λ_{LSPR}) for SERS substrates exhibiting a Lorentzian-shaped LSPR peak can be described by $\lambda_{\text{LSPR}} = \lambda_{\text{Ex}} + \lambda_{\text{RS}}$.³¹ λ_{Ex} and λ_{RS} correspond to the wavelengths of the excitation source and the inelastically scattered Raman photons, respectively. When a 785 nm near-infrared laser was used as an excitation source and Raman shifted photons were collected in the range of 300–1800 cm^{−1} (please refer to the [Supporting Information](#) for details), λ_{LSPR} should be tuned

Table 1. Sequences of Oligonucleotides Used in This Work

description	sequence (5'–3')
S1, 9-mer unmethylated DNA	<u>C</u> GC GCG CGC
5mC–DNA, one of the cytosines in S1 was replaced by 5-methylcytosine	^m C GC GCG CGC
5hmC–DNA, one of the cytosines in S1 was replaced by 5-hydroxymethylcytosine	^{hm} C GC GCG CGC
5fC–DNA, one of the cytosines in S1 was replaced by 5-formylcytosine	^f C GC GCG CGC
5caC–DNA, one of the cytosines in S1 was replaced by 5-carboxylcytosine	^{ca} C GC GCG CGC
S2, 25-mer 5mC–DNA from the promoter region of the <i>Homo sapiens</i> BRCA1 gene, one of the cytosines was replaced by 5-methylcytosine	GGG G ^m CA AAA GCA AGC TGA ACC CGA A

TaKaRa Biotechnology Co. Ltd. in Dalian, China. The DNA solutions were diluted in Tris–HCl buffer (10 mM, pH 7.4). The double-stranded DNAs (dsDNA) were prepared by incubating the 1:1 mixture of the two complementary ssDNA at 95 °C for 10 min followed by cooling to room temperature. The TET1 protein active domain (residues from 1369 to 2039) was purchased from Wisegene (USA). DNA S2 (60 μM) was incubated in 50 mM of a hydroxyethyl-piperazineethane-sulfonic acid buffer (HEPES, pH 7.4) containing 100 mM NaCl, 2 mM ascorbic acid, 1 mM alpha-ketoglutarate, 105 μM Fe(NH₄)₂(SO₄)₂·6H₂O, 1.2 mM adenosine triphosphate (ATP), and 2.5 mM dithiothreitol (DTT) and, then, incubated with TET1 protein (7 μM) at 37 °C for 60 min (total volume 25 μL). Afterward, the reaction

within the spectral window of 794–850 nm for optimal SERS response. Therefore, the PGNA (100–200–604) with a strong LSPR peak at 850 nm and the maximum local electric field at the top surface of the gold nanoholes were chosen as the substrate. It shows a maximum electric field intensity (i.e., $(E_{\max}/E_0)^2$) of $\sim 5 \times 10^3$ (Figure 1C). The quality factor (Q_F) was calculated to evaluate the average SERS intensity over the entire range of the Stokes Raman shifted frequencies (300–1800 cm^{-1}) and was found to be as high as 10^6 , implying a high SERS activity of the designed PGNA.

Then, the designed optimal PGNA (100–200–604) was fabricated [citations as in the Experimental Section].^{27–29} As shown in Figure 1D, the as-fabricated PGNA shows the well-defined arrangement of the nanoholes, thereby providing uniformly distributed “hot spots”, as shown in Figure 1B, which can lead to high detection reproducibility. The magnified SEM and AFM images confirm that the structural parameter of the as-fabricated PGNA meets the FDTD designed optimal parameters (Figure 1E,F). We also compared the experimentally collected reflectance spectrum with the simulated one and found the main LSPR peaks match well (Figure S5), reconfirming the structure of the as-fabricated PGNA is the same as the simulated one. The intensities of the reflectance spectra obtained from the FDTD simulation differs from that of the experimental measurement, which is mainly because of the different dielectric media. In the FDTD simulation, the PGNA substrate was immersed in water while the experiment was carried out in air. Another reason could be due to the gold surface roughness introduced during nanofabrication compared to the perfectly smooth gold film considered in the FDTD simulation. Similar phenomena were observed in the previous studies.³² To experimentally verify the SERS performance of the as-fabricated PGNA, the SERS spectra of Rh 6G were collected on PGNA (Figure S6) where ~ 9 hot spots and 8.0×10^4 Rh 6G molecules are involved within a 2 μm diameter laser spot. An averaged enhancement factor (EF) was calculated to be 1.9×10^6 following the method reported in the literature³³ (Please refer to the Supporting Information for details). Moreover, the spectra taken at different positions were highly reproducible (Figure S7). Such strong and highly reproducible signals enable a sensitive and robust detection.

SERS Detection of DNA Epigenetic Modifications. We first investigated the feasibility of using the PGNA as a SERS substrate for identifying different epigenetic modifications of DNA. To do so, we constructed unmethylated dsDNA and dsDNA sequences with 5mC, 5hmC, 5fC, and 5caC as models (see Table 1 for the sequences). These DNA samples have repeating C and G bases because DNA methylation frequently occurs at the carbon 5 position of cytosine in the CpG island, a region of DNA where a cytosine nucleotide is followed by a guanine nucleotide in the linear sequence of bases along its 5' $\rightarrow$ 3' direction. Vibrational assignment of dsDNA was based on previous publications (Table S2).^{25,34–40} As shown in curve a of Figure 2A, the SERS spectrum of the unmethylated DNA shows dominated peaks at 760, 1365, and 1680 cm^{-1} , assigned to the ring breathing mode vibration of cytosine, the guanine overlapped with CH/CH₂ deformation vibration, and the C=O stretching mode,^{34,35} respectively. When we substituted a cytosine with 5mC, the spectral variations between the unmethylated DNA and methylated DNA (5mC–dsDNA) can be distinguished. The ring breathing band of cytosine redshifts to 785 nm, as the intensity increases, in agreement

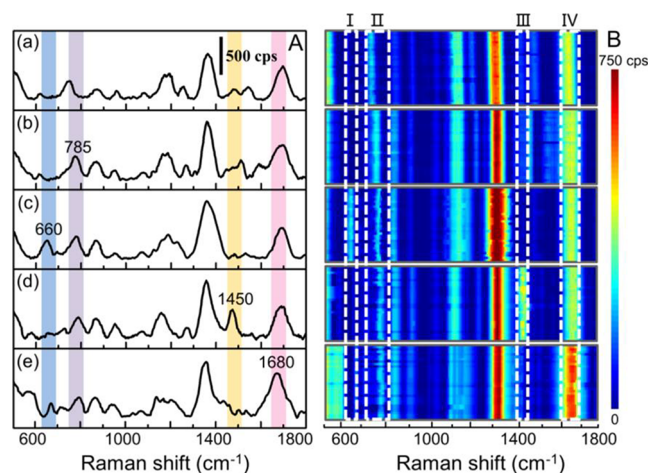


Figure 2. (A) Averaged SERS spectra and (B) barcodes of normal dsDNA (a), 5mC–dsDNA (b), 5hmC–dsDNA (c), 5fC–dsDNA (d), and 5caC–dsDNA (e). Each spectrum is an average of 30 spectra taken at different points on the PGNA substrate and normalized using the peak intensity at 1080 cm^{-1} . The concentrations of the DNA samples are 10 μM . Each barcode contains 30 spectra along the y-axis taken from different regions on the PNA substrate.

with the previous reports.²⁵ The intensity at 1365 cm^{-1} increases relatively. These band changes indicate that mC is detectable using our approach. HmC is another abundant DNA epigenetic modification. The introduction of 5hmC into DNA (5hmC–dsDNA) induces a new peak, appearing at 660 cm^{-1} assigned to C–O–H in-plane bending vibration³⁶ (curve c of Figure 2A), which had been regarded as an obvious marker for the presence of 5hmC.²⁵ The redshift and intensity increase of the pyrimidine ring breathing band at 785 cm^{-1} also appeared. However, the intensity increase at 785 cm^{-1} for 5mC is larger than the same peak increase for 5hmC, which is also in good agreement with previous results,²⁵ rendering the detection of 5mC in the presence of 5hmC possible as well. The SERS spectrum of 5fC–dsDNA is shown in curve d of Figure 2A. By comparison with the SERS spectrum of the normal dsDNA (curve a), the substitution of cytosine by 5fC also leads to a redshift and intensity increase at 785 cm^{-1} , together with a new band that strongly appears at 1450 cm^{-1} , which corresponds to the in-plane deformation mode of the C–H of the aldehyde group.³⁷ Furthermore, there is a new and moderate peak at 750 cm^{-1} , which should be assigned to the in-plane deformation mode of the CCO. We can use the 1450 cm^{-1} band as the marker to determine 5fC in DNA. Finally, the SERS spectrum of 5caC–dsDNA was collected and exhibited distinguished features from the other epigenetically modified DNAs. A broad band begins to show approximately 550 cm^{-1} , which should be assigned to in-plane deformation vibrations of the carboxylic group.^{37,38} Additionally, we found an obvious intensity increase in the peak at 1680 cm^{-1} (curve e in Figure 2A), and it was contributed from the C=O stretching mode as well as the antisymmetric stretching vibration of the carboxylic group.^{37,39,40} This peak was still present in the SERS spectra of the other epigenetically modified DNAs; however, it was less intense (curve a–d), implying it is possible to distinguish 5caC from other epigenetic modifications. Therefore, we chose the peak at 1680 cm^{-1} as a marker to indicate the substitution of cytosine by 5caC.

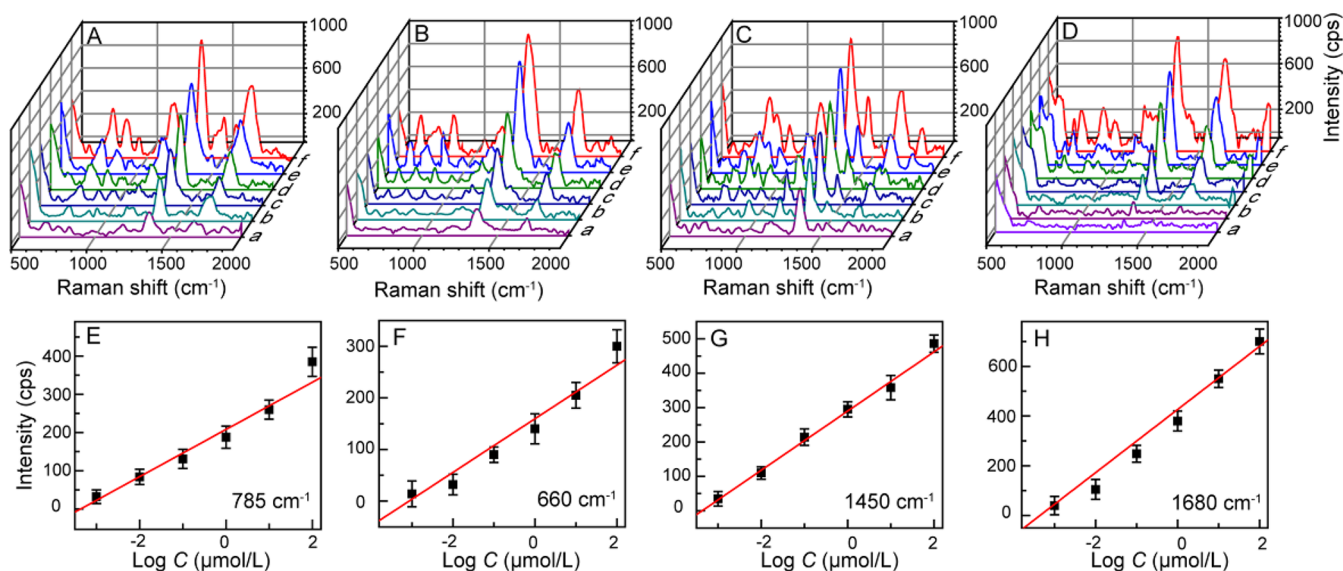


Figure 3. SERS spectra of 5mC–DNA (A), 5hmC–DNA (B), 5fC–DNA (C), and 5caC–DNA (D) with different concentrations in the range of 0.001–100 μM (curves a–f). The linear dependence of peak intensity at 785 cm^{-1} (E), 660 cm^{-1} (F), 1450 cm^{-1} (G), and 1680 cm^{-1} (H) with respect to DNA concentration in logarithmic scale. Each spectrum is an average of 30 spectra taken at different points on the PGNA substrate and normalized using the peak intensity at 1080 cm^{-1} . Error bars show the standard deviation from five replications.

Then, we used the color-contrast spectral barcode to reconfirm the marked bands for the DNA epigenetic modifications. The 30 SERS spectra of each dsDNA sample were displayed in a color-contrast spectral barcode, as shown in Figure 2B, where the Raman shift is represented by the x-axis and the 30 spectra are stacked along the y-axis. On each barcode, the lines represent the Raman shift of the vibrational peaks and the color contrast and the line width represent the relative intensity and the width of the peaks, respectively. The characteristic peaks of 5hmC (660 cm^{-1}), 5mC (785 cm^{-1}), 5fC (1450 cm^{-1}), and 5caC (1680 cm^{-1}) were marked in a white dashed box I–IV, respectively. We found that the SERS spectra were distinguished well among the different types of DNA epigenetic modification. For example, in box I, we found the color in the 5hmC strain is more intense than the other three types of epigenetic modifications. We also found the color in the 5fC and 5caC strains is the most intense in boxes III and IV, respectively. These results indicate that all identified markers are unique, and these results are evidence for the discrimination of the four modified cytosines. Our approach can thus enable a straightforward and simultaneous identification of different epigenetic modifications of DNA.

The different SERS characteristics of the various DNA epigenetic modifications can be further highlighted by PCA. Figure S8 shows the score plots and loading of the derived principal components (PCs). Three PC scores were obtained. The first PC accounts for 83.7%, and the second PC accounts for 6.7%, with the third PC accounting for 4.6%, revealing a clear segregation among these four DNA epigenetic modifications. Collectively, the above results demonstrate a striking distinction in the SERS characteristics of 5mC–, 5hmC–, 5fC–, and 5caC–DNA, suggesting label-free SERS-based detection of DNA epigenetic modifications is feasible.

The PGNA substrate has uniformly distributed, enhanced electric fields around the rims of the nanoholes, where DNA can easily access and adsorb to these hot spots, leading to high reproducibility. As shown in Figure 2B, the color contrast is very uniform along each line, indicating high reproducibility of

detection. The values of relative standard deviation (RSD) were calculated to be 5.6%, 8.8%, 4.8%, and 5.0% for 30 replicas collected at different spots on the substrate when detecting 5mC–, 5hmC–, 5fC–, and 5caC–DNA (Figure S9), respectively, which further confirms the high detection reproducibility of the proposed approach.

Detection sensitivity is very critical for the potential application of the approach in the diagnosis of diseases. We evaluated the sensitivity of our detection method by monitoring the characteristic peak intensity variance with various DNA concentrations. A 50 μL drop of the DNA sample in Tris-HCl (pH 7.4) was dropped onto the PGNA substrate, and the SERS spectra were collected subsequently. We used characteristic peaks at 785, 660, 1450, and 1680 cm^{-1} as the markers for 5mC, 5hmC, 5fC, and 5caC, respectively. The SERS signal of PO_2^- (1080 cm^{-1}) was used as the internal standard, and all spectra were normalized by the intensity of this peak. We plotted the intensity of peaks at 785, 660, 1450, and 1680 cm^{-1} as a function of the logarithm of DNA concentrations in the range of 0.001–100 μM . The Raman peaks show incremental intensity increases with the increase of the concentration of DNA samples (Figure 3A–D). Good linear relationships are present between the relative intensities of the peaks and the DNA concentrations ranging from 1 nM to 100 μM (Figure 3E–H). The limits of detection of 5mC–, 5hmC–, 5fC–, and 5caC–DNA are 0.1, 0.05, 0.02, and 0.05 nM, respectively, when the signal-to-noise ratio is 3, indicating the high sensitivity of our DNA epigenetic modifications analysis method.

We also tested the feasibility of using our method to detect DNA epigenetic modifications in real DNA samples. A total of 3 cancer cell lines, including human breast carcinoma cells (MCF-7 and SK-BR-3) and human lung adenocarcinoma cell (A549), were chosen to investigate the genome-wide DNA methylation. A normal human umbilical vein endothelial cell (HUVEC) was used as control. The real DNA samples were extracted from the living cells and quantified using a UV–vis spectrophotometer for the subsequent analysis. As shown in

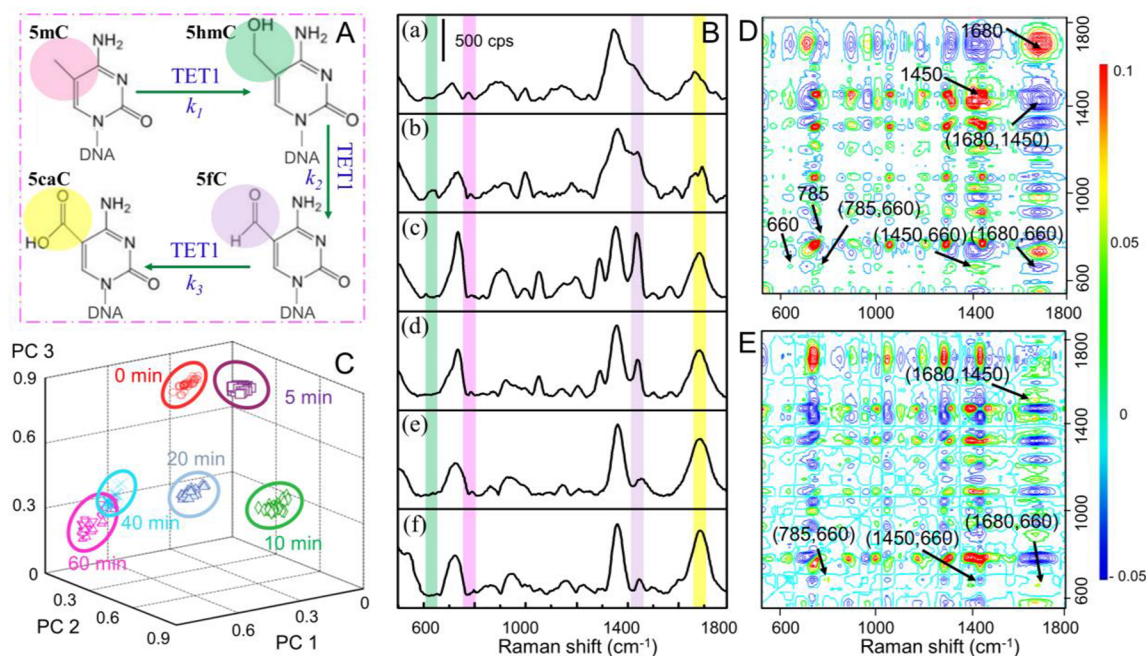


Figure 4. (A) Illustration of the TET1 protein-mediated oxidation process. (B) Averaged SERS spectra of S2, 25-mer unmethylated DNA, collected at different time intervals during the TET1 protein-mediated oxidation process. Spectra a–f represent the reaction times at 0, 5, 10, 20, 40, and 60 min, respectively. The concentrations of the DNA samples are $10 \mu\text{M}$. Each spectrum is an average of 30 spectra taken at different points on the PGNA substrate and normalized using the peak intensity at 1080 cm^{-1} . (C) PCA analysis of SERS spectra collected at different time intervals during the oxidation process. (D) 2D synchronous and (E) 2D asynchronous Raman correlation maps in the region of $500\text{--}1800 \text{ cm}^{-1}$. The color bar shows the different intensities of the 2D correlation peaks, which indicates the extent correlation. The interval time for taking the 2D Raman correlation data set is 0, 5, 10, 20, 40, and 60 min, respectively.

Figure S10A,B, we can observe redshift and the intensities increase at 785 cm^{-1} (Raman characteristic of 5mC) in the order of MCF-7, A549, SK-BR-3, and HUVEC cells, indicating the expression level of 5mC increases in the order of MCF-7, A549, SK-BR-3, and HUVEC cells. This result agrees well with the previous reports,^{41–43} in which cancer cells have been demonstrated to have global hypomethylation related to normal cells. Moreover, results show the methylation level displayed in a cell type-dependent manner. SK-BR-3 cells show a higher methylation level than A549 and MCF-7 cells, which is also in good agreement with the results determined by high-performance capillary electrophoresis.⁴³ These results confirmed the feasibility of our method for detecting real DNA samples. We also noted there are weak Raman peaks at 1450 cm^{-1} (Raman characteristic of 5fC) in cancer cells (curves b–d in Figure S10A,C), while this peak is almost negligible in HUVEC cells (curve a), indicating that the expression level of 5fC is higher in cancer cells than in normal cells. However, the Raman signals for 5hmC and 5caC were hardly detected, which could be due to their low expression levels in living cells.

Monitoring of Dynamic DNA Epigenetic Alterations in the TET1 Protein-Mediated Oxidation Process. Next, we used the developed method to monitor the dynamic alterations in the DNA epigenetic modifications. They are closely involved in the regulation of DNA methylation and demethylation processes, which affects the cells fate and their aberrated performance in the initiation and progression of cancers.^{44–47}

The methylated DNA fragment in the human BRCA1 gene (S2, shown in Table 1) was selected as a study model because the BRCA1 gene shows up-regulated expression of methylation. In TET1 protein-mediated oxidation, 5mC can be

oxidized stepwise to 5hmC, 5fC, and 5caC (Figure 4A). We collected the SERS spectra of the mixture of the DNA oxidation products at different time intervals during the TET1 protein-mediated oxidation process. As shown in Figure 4B, the methylated DNA S2 has variable spectroscopic profiles during the oxidation process, especially at the marked bands. For example, the intensity at 785 cm^{-1} declines upon the oxidation of 5mC. The intensified Raman peaks at 660 cm^{-1} appear at 5 min and are quickly consumed. The intensity at 1450 cm^{-1} first increases, reaches the maximum at 10 min, and then decreases. Additionally, the incremental intensity at 1680 cm^{-1} is observed throughout the entire TET1-mediated oxidation. These peak variations are caused by stepwise oxidation of 5mC into 5hmC, 5fC, and 5caC, which is mediated by TET1 protein (the detailed analysis is provided later).

The characteristic spectroscopic changes can be highlighted by PCA. Three-dimensional plots were constructed for the first three principal components (PC1, PC2, and PC3). As depicted in Figure 4C, the corresponding plots reveal a clear segregation of the DNA epigenetic modifications under various TET1 protein-oxidized times. The three PCs account for 80.9%, 10.0%, and 4.9%, respectively. The PC1 score value indicates that 80.9% of the total spectral variance differs significantly. This result demonstrates the alteration in the spectroscopic characteristics of the DNA modifications experienced in the TET1 protein-mediated oxidation can be distinctly discriminated. Along with findings in the SERS spectra, the results suggest the proposed method can be used to track the dynamic alterations of DNA epigenetic modifications.

The 2D synchronous/asynchronous correlation analysis was applied to analyze the dynamic fluctuations of the SERS

spectra of the DNA epigenetic modifications in the TET1 protein-mediated oxidation process. As shown in Figure 4D, the synchronous 2D correlation map in the region of 500–1800 cm^{-1} shows auto peaks at 660, 750, 1300, 1450, and 1680 cm^{-1} at the diagonal, suggesting these bands dramatically change during the TET1 protein-mediated oxidation process. Because the peaks at 785, 660, 1450, and 1680 cm^{-1} were chosen as the Raman markers of 5mC, 5hmC, 5fC, and 5caC, respectively, the cross peaks (785, 660), (1450, 660), (1680, 660), and (1680, 1450) cm^{-1} were the focus. Figure 4D shows the positive synchronous cross peaks at (785, 660) and (1450, 660) cm^{-1} and the negative synchronous cross peaks at (1450, 660) and (1680, 1450) cm^{-1} , and Figure 4E shows the positive asynchronous cross peaks at (785, 660), (1680, 660), and (1680, 1450) cm^{-1} and the negative asynchronous cross peak at (1450, 660) cm^{-1} . According to Noda's rules,⁴⁸ the signs of the 2D Raman synchronous/asynchronous map suggest the following sequence of the spectral changes: 785 $\rightarrow$ 660 $\rightarrow$ 1450 $\rightarrow$ 1680 cm^{-1} ; i.e., Raman bands at 785 cm^{-1} first begin to show up, 660 cm^{-1} changes second, followed by 1450 cm^{-1} , and finally 1680 cm^{-1} changes (the details are summarized in Table 2), thus indicating the sequence of the alterations in the epigenetic modifications of DNA is 5mC $\rightarrow$ 5hmC $\rightarrow$ 5fC $\rightarrow$ 5caC.

Table 2. Order of Band Changes During TET1 the Protein-Mediated Oxidation Process

synchronous at $\Phi(\nu_1, \nu_2)$	asynchronous at $\Psi(\nu_1, \nu_2)$	ν_1^a
(785, 660) $\text{cm}^{-1} > 0$	(785, 660) $\text{cm}^{-1} > 0$	lead
(1450, 660) $\text{cm}^{-1} > 0$	(1450, 660) $\text{cm}^{-1} < 0$	lag
(1680, 660) $\text{cm}^{-1} < 0$	(1680, 660) $\text{cm}^{-1} > 0$	lag
(1680, 1450) $\text{cm}^{-1} < 0$	(1680, 1450) $\text{cm}^{-1} > 0$	lag

^aThe term lead means that the intensity change of the band at ν_1 occurs before ν_2 . The term lag means that the intensity change of the band at ν_1 occurs after ν_2 .

Furthermore, the kinetics of the DNA epigenetic alterations in the TET1 protein-mediated oxidation process were investigated by time-resolved SERS with PGNA. We plotted the relative intensities at 785, 660, 1450, and 1680 cm^{-1} versus time (Figure 5). The curve a in Figure 5 shows the dependence

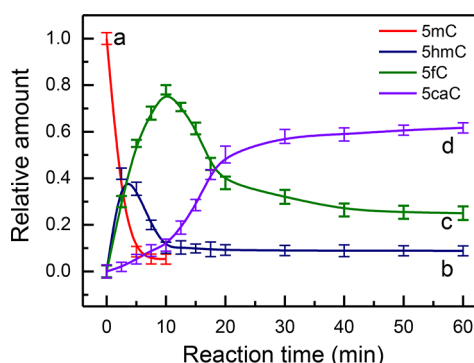


Figure 5. Dependence of the relative amount of 5mC-modified (a), 5hmC-modified (b), 5fC-modified (c), and 5caC-modified (d) DNA as a function of reaction time during the TET1 protein-mediated oxidation process. The amount of DNA was calculated by $\Delta I_B / \Delta I_{\text{total}}$, where $\Delta I_{\text{total}} = \sum_B \Delta I_B$, ΔI_B refers to the relative intensity at 785, 660, 1450, and 1680 cm^{-1} band, respectively. Error bars are based on five measurements.

of the amount of 5mC versus the reaction time. The amount of 5mC drastically decreases in 5 min, demonstrating a rapid consumption of the initially present 5mC, which implies a large rate constant of k_1 as shown in the first oxidation step in Figure 4A. The analysis of the amount of 5hmC can be divided into two periods of time between 0 and 3 min and between 3 and 60 min (curve b in Figure 5). The amount of 5hmC dramatically increased in the first 3 min and then declined slowly and remained constant after 10 min. The initial increase of 5hmC was a result of the conversion of 5mC into 5hmC. 5hmC was produced rapidly and, then, slowly oxidized to 5fC. Thus, we observed the following radioactive decay caused by the gradual consumption of 5mC and a subsequent conversion of 5hmC into 5fC. After the decrease, the amount of 5hmC remained constant, indicating the end of the 5hmC oxidation. The slow conversion from 5hmC to 5fC was observed by a similar profile for the amount of 5fC as a function of time (curve c in Figure 5). A peak value of 5fC appearing at 10 min is longer than 3 min for 5hmC, suggesting a smaller value of k_2 than k_1 . Differing from the changes of 5hmC and 5fC, the amount of 5caC displayed a continuous increase along the reaction time (curve d in Figure 5). This is because 5caC is the final oxidative product of stepwise oxidation of 5mC by the TET1 protein, which has never been consumed during the oxidation process. These are typical characteristics of the consecutive reaction, demonstrating that the TET1 protein-mediated oxidation of 5mC is a consecutive reaction consisting of three consecutive elementary steps. In other words, 5mC is oxidized stepwise to 5hmC, 5fC, and 5caC. We found more than 95% 5mC was oxidized into 5hmC (9%), 5fC (24%), and 5caC (62%). The time to consume 5mC is 5 min, and the times to reach the maximum amount of 5hmC and 5fC are 3 and 10 min, respectively. These results agree with the results obtained by HPLC measurement,⁴⁹ indicating the feasibility of our label-free SERS-based method for monitoring the level of DNA epigenetic modifications in a TET1 protein-mediated oxidation process. The rate constants of k_1 , k_2 , and k_3 were calculated to be 0.6, 0.25, and 0.15 min^{-1} , respectively (the details are provided in the Supporting Information). The reaction rates for 5hmC- and 5fC-DNA are 2.4- and 4.0-fold lower than that of 5mC-DNA, respectively, strongly suggesting that the TET1 protein-catalyzed oxidation is a kinetically relevant pathway.

CONCLUSION

In summary, we developed a SERS-based detection method for DNA epigenetic modifications using PGNA as the SERS-active substrate, which enables a highly sensitive and reproducible SERS signal of the analyte due to the uniform distribution of nanoholes on the surface. This method is based on identifying SERS spectral features resulting from DNA base modifications. Raman bands at 785, 660, 1450, and 1680 cm^{-1} were selected as the detection markers for 5mC, 5hmC, 5fC, and 5caC, respectively, because all identified markers are unique and evident for the discrimination of the four modified cytosines. We demonstrated that the proposed method can sensitively detect 5mC-, 5hmC-, 5fC-, and 5caC-dsDNA with the limits of detection of 0.1, 0.05, 0.02, and 0.05 nM, respectively. More importantly, our method can monitor the dynamic epigenetic alterations in a TET1 protein-mediated oxidation process on the basis of the relative intensity changes at the assigned marked bands with respect to time. The reaction kinetic parameters at each step can be quantitatively

determined. The proposed method is direct, label-free, sensitive, and reproducible. It provides a novel assay approach for determining DNA epigenetic modifications as well as monitoring their dynamic alterations, which is of interest in DNA analysis, especially DNA modification.

■ ASSOCIATED CONTENT

● Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.analchem.9b01004](https://doi.org/10.1021/acs.analchem.9b01004).

FDTD simulations of the dependence of the electromagnetic field distribution on the PGNA structure, calculation details of the SERS enhancement factor and rate constants of TET1-protein mediated oxidation of 5mC in DNA, extraction of DNA from the living cells, table of typical vibrational modes of DNA, distributions of the electric field on the x - z plane of the nanohole array, electric field enhancement, FDTD reflectance spectrum, SERS spectra, principle component analysis, and peak intensities (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: wuping@njnu.edu.cn (P.W.).

*E-mail: cxcai@njnu.edu.cn (C.C.).

*E-mail: qyu@uw.edu (Q.Y.).

ORCID

Erjin Zheng: 0000-0002-0576-5614

Ping Wu: 0000-0003-2366-5390

Qiuming Yu: 0000-0002-2401-4664

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the NSFC (21405083 and 21675088), JSFC (BK20181383), and Priority Academic Program Development of Jiangsu Higher Education Institutions. Q.Y. acknowledges the support by the National Science Foundation (CMMI-1661660).

■ REFERENCES

- (1) Hanahan, D.; Weinberg, R. A. *Cell* **2000**, *100*, 57–70.
- (2) Hanahan, D.; Weinberg, R. A. *Cell* **2011**, *144*, 646–674.
- (3) Jenuwein, T.; Allis, C. D. *Science* **2001**, *293*, 1074–1080.
- (4) Geiman, T. M.; Robertson, K. D. *J. Cell. Biochem.* **2002**, *87*, 117–125.
- (5) Rizzi, G.; Lee, J. R.; Dahl, C.; Guldberg, P.; Dufva, M.; Wang, S. X.; Hansen, M. F. *ACS Nano* **2017**, *11*, 8864–8870.
- (6) Kelsey, G.; Stegle, O.; Reik, W. *Science* **2017**, *358*, 69–75.
- (7) Shen, L.; Song, C.-X.; He, C.; Zhang, Y. *Annu. Rev. Biochem.* **2014**, *83*, 585–614.
- (8) Li, W.; Wu, P.; Zhang, H.; Cai, C. *Anal. Chem.* **2012**, *84*, 7583–7590.
- (9) Chen, R. Z.; Pettersson, U.; Beard, C.; Jackson-Grusby, L.; Jaenisch, R. *Nature* **1998**, *395*, 89–93.
- (10) Liu, S.; Wu, P.; Li, W.; Zhang, H.; Cai, C. *Chem. Commun.* **2011**, *47*, 2844–2846.
- (11) Ito, S.; Shen, L.; Dai, Q.; Wu, S. C.; Collins, L. B.; Swenberg, J. A.; He, C.; Zhang, Y. *Science* **2011**, *333*, 1300–1303.
- (12) He, Y.-F.; Li, B.-Z.; Li, Z.; Liu, P.; Wang, Y.; Tang, Q.; Ding, J.; Jia, Y.; Chen, Z.; Li, L.; Sun, Y.; Li, X.; Dai, Q.; Song, C.-X.; Zhang, K.; He, C.; Xu, G.-L. *Science* **2011**, *333*, 1303–1307.
- (13) Yu, M.; Hon, C. C.; Szulwach, K. E.; Song, C.-X.; Zhang, L.; Kim, A.; Li, X.-K.; Dai, Q.; Shen, Y.; Park, B.; Min, J. H.; Jin, P.; Ren, B.; He, C. *Cell* **2012**, *149*, 1368–1380.
- (14) Gabrieli, T.; Sharim, H.; Nifker, G.; Jeffet, J.; Shahal, T.; Arieli, R.; Levi-Sakin, M.; Hoch, L.; Arbib, N.; Michaeli, Y.; Ebenstein, Y. *ACS Nano* **2018**, *12*, 7148–7158.
- (15) Weber, A. R.; Krawczyk, C.; Robertson, A. B.; Kusnierczyk, A.; Vagbo, C. B.; Schuermann, D.; Klungland, A.; Schar, P. *Nat. Commun.* **2016**, *7* (9), 10806.
- (16) Raiber, E.-A.; Murat, P.; Chirgadze, D. Y.; Beraldi, D.; Luisi, B. F.; Balasubramanian, S. *Nat. Struct. Mol. Biol.* **2015**, *22*, 44–49.
- (17) Song, C.-X.; Szulwach, K. E.; Dai, Q.; Fu, Y.; Mao, S.-Q.; Lin, L.; Street, C.; Li, Y.; Poidevin, M.; Wu, H.; Gao, J.; Liu, P.; Li, L.; Xu, G.-L.; Jin, P.; He, C. *Cell* **2013**, *153*, 678–691.
- (18) Szulwach, K. E.; Jin, P. *BioEssays* **2014**, *36*, 107–117.
- (19) Nestor, C.; Ruzov, A.; Meehan, R. R.; Dunican, D. S. *BioTechniques* **2010**, *48*, 317–319.
- (20) Xu, L.-J.; Lei, Z.-C.; Li, J.; Zong, C.; Yang, C. J.; Ren, B. *J. Am. Chem. Soc.* **2015**, *137*, 5149–5154.
- (21) Shi, H.; Wang, H.; Meng, X.; Chen, R.; Zhang, Y.; Su, Y.; He, Y. *Anal. Chem.* **2018**, *90*, 14216–14221.
- (22) Chen, C.; Li, Y.; Kerman, S.; Neutens, P.; Willems, K.; Cornelissen, S.; Lagae, L.; Stakenborg, T.; Van Dorpe, P. *Nat. Commun.* **2018**, *9*, 1733.
- (23) Panikkanvalappil, S. R.; Mackey, M. A.; El-Sayed, M. A. *J. Am. Chem. Soc.* **2013**, *135*, 4815–4821.
- (24) Barhoumi, A.; Zhang, D.; Tam, F.; Halas, N. J. *J. Am. Chem. Soc.* **2008**, *130*, 5523–5529.
- (25) Barhoumi, A.; Halas, N. J. *J. Phys. Chem. Lett.* **2011**, *2*, 3118–3123.
- (26) Ouyang, L.; Hu, Y.-W.; Zhu, L.-H.; Cheng, G. J.; Irudayaraj, J. *Biosens. Bioelectron.* **2017**, *92*, 755–762.
- (27) Sun, F.; Bai, T.; Zhang, L.; Ella-Menye, J. R.; Liu, S.-J.; Nowinski, A. K.; Jiang, S.; Yu, Q. *Anal. Chem.* **2014**, *86*, 2387–2394.
- (28) Xu, J.; Turner, J. W.; Idso, M.; Biryukov, S. V.; Rognstad, L.; Gong, H.; Trainer, V. L.; Wells, M. L.; Strom, M. S.; Yu, Q. *Anal. Chem.* **2013**, *85*, 2630–2637.
- (29) Hou, C.; Galvan, D. D.; Meng, G.; Yu, Q. *Phys. Chem. Chem. Phys.* **2017**, *19*, 24126–24134.
- (30) Yu, Q.; Guan, P.; Qin, D.; Golden, G.; Wallace, P. M. *Nano Lett.* **2008**, *8*, 1923–1928.
- (31) Haynes, C. L.; Van Duyne, R. P. *J. Phys. Chem. B* **2003**, *107*, 7426–7433.
- (32) Wang, D.; Yu, X.; Yu, Q. *Nanotechnology* **2012**, *23* (9), 405201.
- (33) Li, J.; Chen, C.; Jans, H.; Xu, X.; Verellen, N.; Vos, I.; Okumura, Y.; Moshchalkov, V. V.; Lagae, L.; Van Dorpe, P. *Nanoscale* **2014**, *6*, 12391–12396.
- (34) Chandra, G. K.; Eklouh-Molinier, C.; Fere, M.; Angiboust, J. F.; Gobinet, C.; Van-Gulick, L.; Jeannesson, P.; Piot, O. *Anal. Chem.* **2015**, *87*, 2655–2664.
- (35) Guerrini, L.; Krpetic, Ž.; van Lierop, D.; Alvarez-Puebla, R. A.; Graham, D. *Angew. Chem., Int. Ed.* **2015**, *54*, 1144–1148.
- (36) Hobro, A. J.; Abdali, S.; Blanch, E. W. *J. Raman Spectrosc.* **2012**, *43*, 187–195.
- (37) Machado, N. F. L.; Marques, M. P. M.; Batista de Carvalho, L. A. E.; Castro, J. L.; Otero, J. C. *J. Raman Spectrosc.* **2017**, *48*, 413–417.
- (38) Camafeita, L. E.; Sánchez-Cortés, S.; García-Ramos, J. V. *J. Raman Spectrosc.* **1995**, *26*, 149–154.
- (39) Sanchez-Cortes, S.; Garcia-Ramos, J. V. *J. Raman Spectrosc.* **1992**, *23*, 61–66.
- (40) Pagannone, M.; Fornari, B.; Mattei, G. *Spectrochim. Acta* **1987**, *43A*, 621–625.
- (41) Yoshida, W.; Baba, Y.; Karube, I. *Anal. Chem.* **2016**, *88*, 9264–9268.
- (42) Szyf, M. *Ageing Res. Rev.* **2003**, *2*, 299–328.
- (43) Paz, M. F.; Fraga, M. F.; Avila, S.; Guo, M.; Pollan, M.; Herman, J. G.; Esteller, M. *Cancer Res.* **2003**, *63*, 1114–1121.

- (44) Ito, S.; D' Alessio, A. C.; Taranova, O. V.; Hong, K.; Sowers, L. C.; Zhang, Y. *Nature* **2010**, *466*, 1129–1133.
- (45) Hu, L.; Lu, J.; Cheng, J.; Rao, Q.; Li, Z.; Hou, H.; Lou, Z.; Zhang, L.; Li, W.; Gong, W.; Liu, M.; Sun, C.; Yin, X.; Li, J.; Tan, X.; Wang, P.; Wang, Y.; Fang, D.; Cui, Q.; Yang, P.; He, C.; Jiang, H.; Luo, C.; Xu, Y. *Nature* **2015**, *527*, 118–122.
- (46) Yang, R.; Yu, T.; Kou, X.; Gao, X.; Chen, C.; Liu, D.; Zhou, Y.; Shi, S. *Nat. Commun.* **2018**, *9* (14), 2143.
- (47) Guo, J.; Su, Y.; Zhong, C.; Ming, G.; Song, H. *Cell* **2011**, *145*, 423–434.
- (48) Noda, I. *Appl. Spectrosc.* **1993**, *47*, 1329–1336.
- (49) Kizaki, S.; Sugiyama, H. *Org. Biomol. Chem.* **2014**, *12*, 104–107.