Solid-State Nanopore Analysis of Diverse DNA Base Modifications Using a Modular Enzymatic Labeling Process

By: Fanny Wang, Osama K. Zahid, Brandi E. Swain, Derek Parsonage, Thomas Hollis, Scott Harvey, Fred W. Perrino, Rahul M. Kohli, Ethan W. Taylor, and Adam R. Hall

Abstract:

Many regulated epigenetic elements and base lesions found in genomic DNA can both directly impact gene expression and play a role in disease processes. However, due to their noncanonical nature, they are challenging to assess with conventional technologies. Here, we present a new approach for the targeted detection of diverse modified bases in DNA. We first use enzymatic components of the DNA base excision repair pathway to install an individual affinity label at each location of a selected modified base with high yield. We then probe the resulting material with a solid-state nanopore assay capable of discriminating labeled DNA from unlabeled DNA. The technique features exceptional modularity via selection of targeting enzymes, which we establish through the detection of four DNA base elements: uracil, 8-oxoguanine, T:G mismatch, and the methyladenine analog $1,N^6$-ethenoadenine. Our results demonstrate the potential for a quantitative nanopore assessment of a broad range of base modifications.
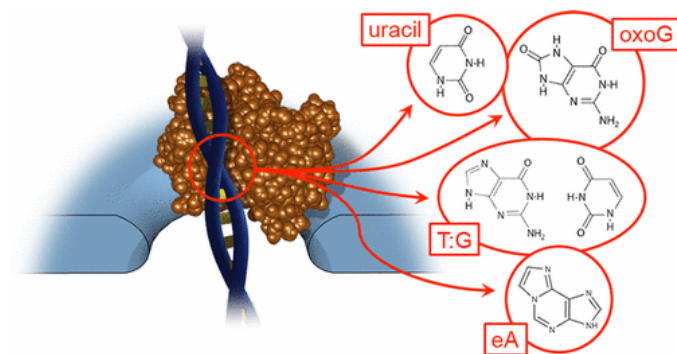
Keywords: Nanopore | base lesion | epigenetics | detection | cancer

Article:

A variety of noncanonical bases are prevalent in genomic DNA and play crucial roles in cell functions that include gene expression and suppression,[1] transposon expression,[2, 3] stem cell differentiation,[4] and chromosomal inactivation.[5] For example, the abundance and position of epigenetic modifications are tightly regulated and errors in this regulation have been linked to a wide range of diseases[6] including cancer. In addition, DNA base damage elements generated both endogenously and exogenously are a major source of point mutations if not correctly repaired by cellular processes. The locations of these elements can be random or could be linked to sequence accessibility in chromatin structures. While the impact of modified DNA bases is clear, their detection can be challenging. For example, direct sequencing approaches typically lack the ability to identify noncanonical bases that may be present in DNA,[7] with some extension to the epigenetic modification methylcytosine, specifically.[8] Conventional technologies like mass spectrometry and high-performance liquid chromatography are burdensome, expensive, and destructive to the DNA and in some cases can induce additional lesions,[9] leading to the misrepresentation of density. In addition, immunological methods like the enzyme-linked immunosorbent assay (ELISA) rely on antibodies that can suffer from cross-reactivity.[10] One approach that addresses some of these concerns has been pioneered recently by Song et al.[11] in which a single, high-affinity tag was attached enzymatically to 5-hydroxymethylcytosine bases, permitting downstream analysis, enrichment, or sequencing. While the process has been adapted to access some additional elements of the demethylation pathway,[12-14] only a limited suite of modifications are suitable for such tagging. As the ability to probe a wide variety of base modifications would be of significant value, we set out to develop a new, modular strategy employing solid-state (SS) nanopore technology that would be permissive for a host of different DNA modifications.

SS nanopores[15, 16] have been widely studied as a means to assess biological molecules like DNA,[17, 18] RNA,[19, 20] and proteins[21-23] using the principle of resistive pulse sensing. The platform consists of an insulating thin-film membrane that contains a nanometer-scale pore, positioned in an electrolyte solution. Application of an electrical bias across the membrane generates an electric field through the pore, and as charged molecules are threaded electrophoretically one-by-one, they temporarily occlude the aperture and interrupt the measured ionic current. These brief electrical disruptions are designated as "events", and their properties have been used to study molecular attributes,[24] probe intermolecular interactions,[25, 26] and determine analyte concentrations.[27, 28] Historically, a significant limitation of this measurement approach has been a lack of selectivity: all molecules of like-charge will translocate and contribute to the overall signal, thus requiring differentiation ex post facto via often subtle differences in event characteristics. We have developed a SS nanopore assay that enables nearly binary detection and quantification of DNA featuring a single biotin affinity tag.[29] Briefly, when target DNA fragments (below ~250 bp) or a key chaperone protein (monovalent streptavidin,[30] MS) are introduced individually to an SS nanopore of appropriate diameter, their rapid translocations prevent events from being resolved by conventional electronics (Figure 1a, left and center). However, when the two molecules bind, the larger nucleoprotein complex interacts with the walls of the nanopore during passage, slowing its translocation to a resolvable speed and yielding events (Figure 1a, right). Recently, we expanded this basic approach to assess hydroxymethylcytosine epigenetic modifications[28] by employing an established method for

specific biotin labeling of the base,[11] enabling direct assessment of a base modification with physiological relevance. However, the scope of possible targets for the labeling approach was limited intrinsically by enzymatic recognition. Here, we enhance our SS nanopore measurement scheme significantly by integrating it with an alternative, modular labeling technique that enables the targeted detection of diverse base modifications.
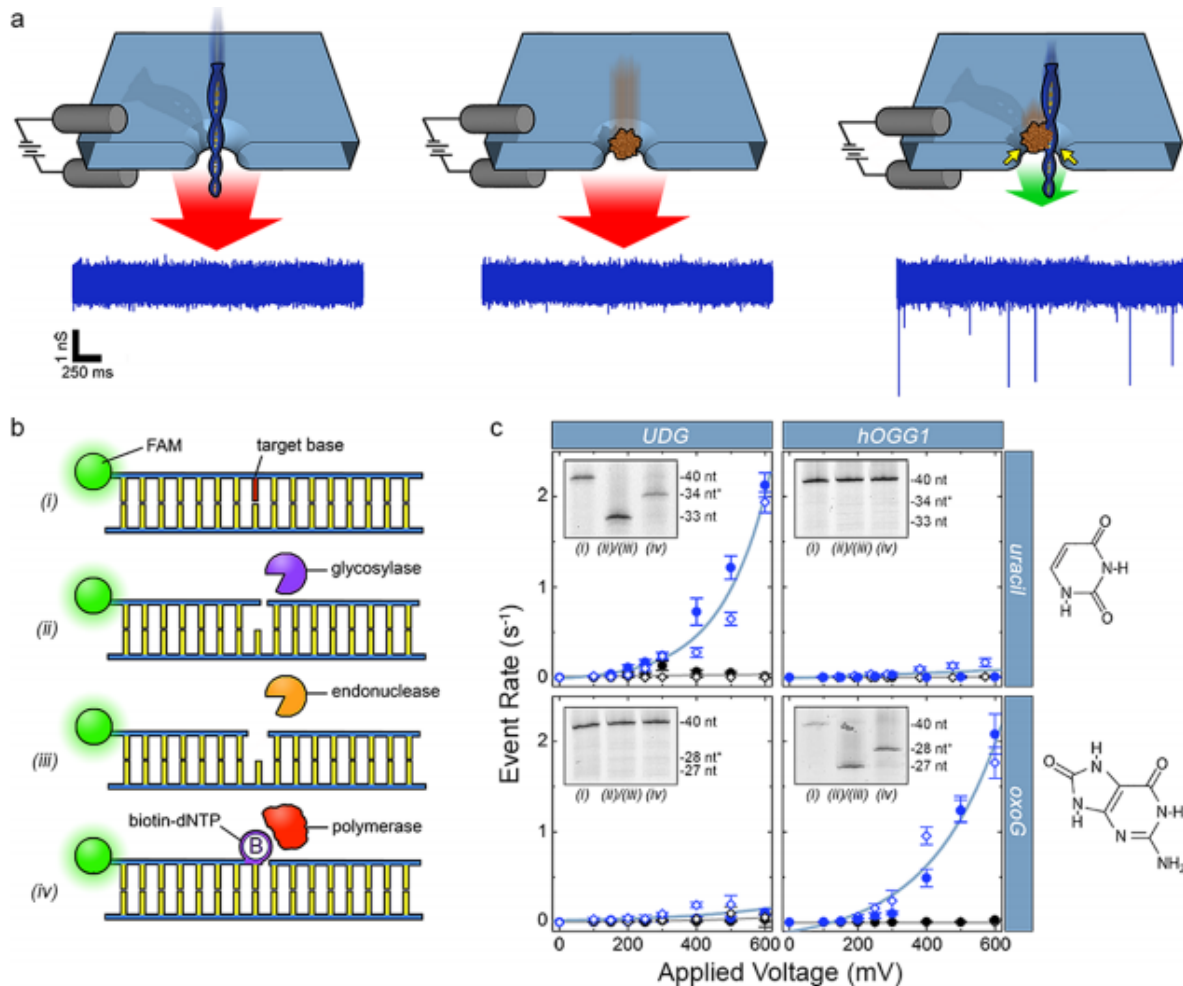
Figure 1. (a) Depiction of the selective SS nanopore assay. Individual passage of a short DNA (left) or a chaperone protein (MS, center) yields no events due to the high translocation speed (red arrows); a DNA–protein complex (right) interacts with the pore walls (yellow arrows), resulting in slower translocation speed (green arrow) and resolvable events. Sample conductance traces at bottom were measured at 300 mV using 75 bp DNA (500 nM) with a synthetic biotin. (b) Schematic representation of the general labeling approach. (i) A duplex DNA molecule featuring a target base element (red). (ii) A glycosylase recognizes and excises the base element (diagram shows the activity of a bifunctional glycosylase that nicks the phosphate backbone 3′ to the excision). (iii) An AP endonuclease cuts the backbone 5′ to the excision. (iv) A gap-filling polymerase incorporates a single biotinylated nucleotide at the modification position. (c) SS nanopore analyses of 250 nM DNA oligonucleotides featuring either a single uracil (at nucleotide position 34, top) or a single oxoG (at nucleotide position 28, bottom). Data points indicate measurements on treated DNA with (blue) and without (black) MS. Filled circles and open diamonds are independent measurements on different SS nanopore devices, and all lines are exponential fits to the data. Dramatic increases in event rate are measured for DNA-MS when a glycosylase specific for the target base is used (blue data, upper left and lower right). Almost no effect is observed for mismatched glycosylase (blue data, upper right and lower left). Insets: denaturing gel analyses of the same DNA constructs (steps numbered as in b). Lane 1: annealed oligonucleotide; lane 2: following glycosylase/endonuclease treatment; lane 3: following T4(exo-) fill-in. * indicates DNA length plus biotin tag. Right: molecular structures of the target bases.

Similar to a recent report by Riedl et al.,[31] our methodology exploits the enzymatic machinery of the DNA base excision repair (BER) pathway, which identifies and restores base lesions in vivo. In our approach (Figure 1b), a modified base element was first excised from the DNA using a DNA glycosylase, which removed the target base from the phosphate backbone, leaving an abasic (AP) site. If the glycosylase was bifunctional (i.e., had AP lyase activity), the phosphodiester bond was also cleaved 3′ to the modification, leaving a single strand nick. The ensuing steps were not affected by this activity. Next, an AP endonuclease was used to cleave the phosphodiester bond 5′ to the abasic site and remove the exposed 3′ phosphate, leaving a hydroxyl group that was amenable to the final step: treatment with a gap-filling polymerase to incorporate a biotin-conjugated nucleotide into the DNA structure. For this, we used a mutant polymerase lacking 3′–5′ exonuclease activity (T4(exo-)) and provided it with only one of the cognate biotin-dNTP, resulting in the insertion of a single affinity label at the precise location of the modified base. We have found the absence of exonuclease activity to be particularly important as processive cleavage of nucleotides from the modification site can result in the prevention or misincorporation of the biotinylated nucleotide label. We note that this methodology ultimately resulted in a nick 3′ to the inserted biotin-dNTP. While it should be possible to repair this nick through ligation, we did not include such a step because of the potential for reduced product yield. The presence of the nick did not negatively impact subsequent measurements.

Crucially, this general approach could be used to target a variety of distinct modified bases through variation of two central components: the DNA glycosylase, selected for recognition of a particular lesion, and the biotin-conjugated nucleotide, selected to match the canonical identity of the target modified base (or in the case of a mismatch target, the appropriate nucleotide for Watson–Crick base-pairing with the opposite strand). As an initial demonstration of this modularity, we first showed selective detection of uracil and oxoguanine (oxoG) bases. Uracils arise in DNA upon deamination of cytosine, resulting in a mutagenic U:G mismatch, or upon misincorporation of dUTP, resulting in a genotoxic U:A pair.[32] Meanwhile, oxoG is the major oxidative base damage associated with reactive oxygen species (ROS) due to the low redox potential of guanine and has known mutagenic potential via transversion during DNA replication.[33] For these measurements, we used synthetic 40 bp double-strand (ds-) DNA oligonucleotides, with one strand containing the target modified base at a known position and a fluorescent FAM label at the 5′ end. We utilized endonuclease IV (EndoIV) to prime the excised gap for T4(exo-) incorporation of a biotin-dNTP.

Denaturing gel analysis (Figure 1c insets) of each sequential step for the two bases using an appropriate glycosylase/nucleotide combination showed the excision of the modified base and incorporation of the biotin-dNTP. Labeling of uracil was achieved using a combination of uracil DNA glycosylase (UDG) and biotin-dUTP, while oxoG labeling employed human oxoG DNA glycosylase (hOGG1) and biotin-dGTP. Notably, we made use of a "one-pot" treatment for each of these targets (see Materials and Methods) that minimized material loss and enabled high product yields of ∼92% and ∼83%, respectively. In addition, UDG is a monofunctional glycosylase, while hOGG1 is bifunctional, showing that the approach was not affected significantly by either absence or presence of AP lyase activity in the glycosylase. Identical treatments of each base modification with nontarget components showed no significant labeling,

highlighting process selectivity that was facilitated by the low cross-recognition of each glycosylase (Supplementary Figures S1–2).

SS nanopore analyses of the same labeling products demonstrated clear specificity in the resulting electrical signal as well. For the appropriate combinations of base modification and enzymes, we observed exponential voltage-dependent event rates (Figure 1c), characteristic of the assay.[28, 34] Provided with the same total DNA concentrations (250 nM), the nearly identical event rate trends for both cases further indicated not only the similarity of the yields for the two labeling protocols, but also the reproducibility of the assay. In contrast, mismatched components yielded negligible event rates that were indistinguishable from negative controls across the entire investigated range of applied voltage (Figure 1c, black). These results suggested that nonspecific labeling of DNA was insignificant, including at the 3′ ends of the molecule (Supplementary Figure S3) and confirmed intrinsic discrimination for an intended base element.

While these data clearly demonstrated a flexible approach that could in principle be extended to a broad range of base targets,[35] glycosylases can also have additional activities that could interfere with the labeling procedure as described. For example, thymine DNA glycosylase (TDG) is a major component of the cytosine demethylation process, recognizing T:G mismatches[36] among other elements,[37] but it also recruits additional enzymes like histone acetyltransferases.[38] Because of this latter role, TDG has a high affinity for the AP site resulting from base excision, making it difficult to detach for subsequent labeling steps (Figure 2a). To address this, we sought to promote enzyme disengagement through the incorporation into the protocol of an additional endonuclease, AP endonuclease 1 (APE1). The extensive dsDNA binding surface of APE1 and the prominent kinking it induces in the DNA helix[39] have been suggested as means to promote displacement of glycosylases more efficiently than EndoIV.[40] However, the improved activity of APE1 comes at the expense of 3′–5′ exonuclease activity not found in the other enzyme, especially under key buffer conditions.[41] To partially mitigate this effect, we used the APE1 D308A mutant,[42] which features reduced 3′–5′ exonuclease activity. This inclusion improved yield significantly over wild type APE1 (Supplementary Figure S4), but the remaining nucleotide digestion activity still necessitated a supplementary purification step prior to polymerase gap filling to limit decomposition of the DNA. While this increased the number of steps and decreased overall product yield somewhat, the resulting material showed the successful incorporation of biotin-labeled nucleotides on gel (Figure 2b) at a high yield (~73%), as well as selective detection by our SS nanopore assay (Figure 2c). Indeed, we recovered the same exponential trend in the measured SS nanopore event rate and the same selectivity over a negative control as found for uracil and oxoG. The event rate dependence was slightly higher for TDG labeling than for previous examples, which could be due to minor residual enzyme binding or small differences in pore attributes (diameter, shape, etc.).

Notably, this alternative method could be used to incorporate other glycosylases with similar behavior as well. As an example, we utilized human alkyladenine DNA glycosylase (hAAG), which excises alkylated bases from DNA, but has also been observed to bind tightly to its DNA template.[43] The major target of hAAG is the important epigenetic element methyladenine,[44] but this base is known to be unstable for in vitro measurements. Consequently, we instead used for our demonstration a synthetic oligonucleotide featuring the methyladenine analog 1,N^6-ethenoadenine, and employed hAAG and biotin-dATP for labeling. Subsequent analyses of the

product again indicated efficient (~74%) labeling on gel (Figure 3a) and a selective event rate increase in SS nanopore measurements (Figure 3b). We noted additional spread in the data at low voltages (<400 mV) specific to the labeled 1,N^6-ethenoadenine DNA, which we suggest may have been due to structural irregularities associated with the modified base itself (see Supplementary Figure S5). We also observed a lower maximum rate at 600 mV, which we attributed to the smaller length[28] of this DNA as compared to the other constructs described in this report. However, the selective rate difference is easily resolved, demonstrating the broad modularity of both the labeling scheme and the measurement approach.
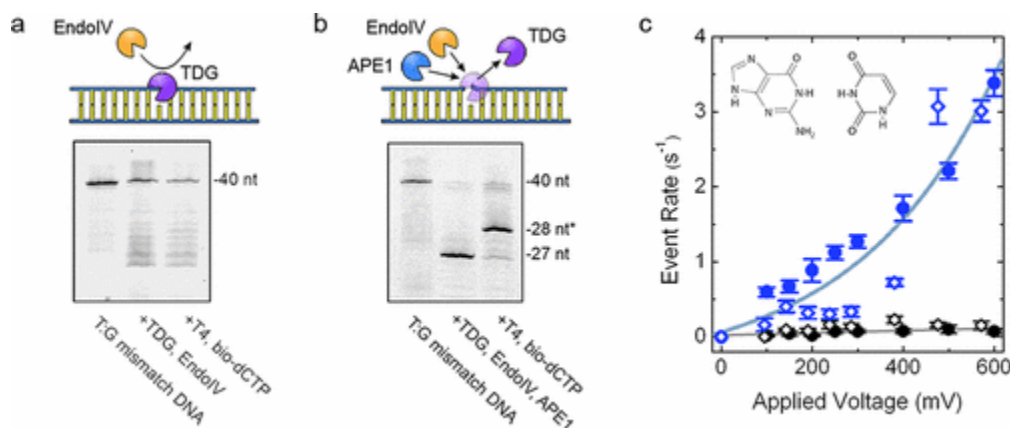
Figure 2. (a) Top: Schematic showing inaccessibility of AP site by EndoIV caused by TDG binding. Bottom: denaturing gel of labeling steps for a T:G mismatch oligonucleotide using EndoIV only. No significant labeling is observed. (b) Top: schematic showing release of TDG by APE1, leaving DNA accessible by EndoIV. Bottom: denaturing gel of labeling steps for a T:G mismatch oligonucleotide using both EndoIV and APE1 (D308A mutant), indicating recovery of high yield labeling. * indicates DNA length plus biotin tag. (c) SS nanopore analysis of 250 nM labeled construct from (b) both with (blue) and without (black) MS. Filled circles and open diamonds are independent measurements on different SS nanopore devices, and lines are exponential fits to the data. Inset: molecular structure of the T:G mismatch base element.
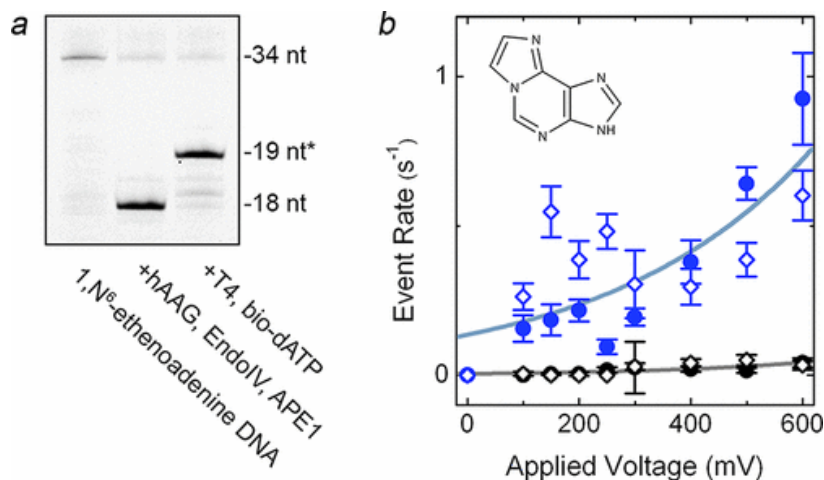
Figure 3. (a) Denaturing gel of labeling steps for a 1,N^6-ethenoadenine oligonucleotide using EndoIV and APE1 (D308A mutant), showing high yield biotin labeling. * indicates DNA length plus biotin tag. (b) SS nanopore analysis of 250 nM labeled construct from (a) both with (blue) and without (black) MS. Filled circles and open diamonds are independent measurements on different SS nanopore devices and lines are exponential fits to the data. Inset: molecular structure of 1,N^6-ethenoadenine.

In conclusion, we have shown that a variety of single-base modifications can be assessed with a selective SS nanopore assay. This was achieved by incorporating an efficient and targeted affinity-labeling technique that exploited the physiological activities of enzymes involved in the BER pathway to install a single biotin tag at the precise location of a given base element. We first showed selective recognition of uracil and oxoG bases with the glycoslyases UDG and hOGG1, respectively. Next, we sought to utilize other glycosylases by integrating a mixture of endonucleases designed to promote enzyme release and limit DNA digestion. While the alternative procedure entailed some loss of material due to increased exonuclease activity, it enabled the use of glycosylases that are specifically challenging to incorporate in the labeling approach due to strong AP binding capacity. As a demonstration, we showed that this approach could be used for the study of T:G mismatch bases with TDG and the methyladenine analog 1,N^6-ethenoadenine with hAAG. Therefore, with the flexible protocols established here, nearly any glycosylase could be integrated, facilitating the labeling and analysis of a broad range of bases that they target, including the widely studied methylcytosine.[45] The central limiting factor for this capacity is in the specificities of the glycosylases themselves, since many have recognition for multiple elements. However, the affinity for specific targets can vary wildly, offering a potential pathway to high selectivity and we expect that the use of point mutations in the glycosylases may also be able to tailor their specificity and enable high certainty in recognition.

These validatory measurements were performed using concentrations of 250 nM, corresponding to ~65 ng of DNA per run in our current protocols. Even taking into account sample loss during the labeling procedure and the general low physiological abundance of modified bases, amounts of genomic DNA capable of supporting this assessment could in principle be obtained from microliter volumes of whole blood using commercial kits (e.g., QIAamp DNA Mini Blood Mini Kit). We also note that our approach has been verified[34] to resolve concentrations down to at least 10 nM without adjustment, making pertinent, clinically derived materials still more easily attainable. Coupled with the quantitative nature of the technique[28] and its viability among a background of nontarget components,[29, 34] our results establish a highly selective and translational SS nanopore assay. The physiologically relevant base modifications that it targets may have important impacts on biology and disease but are challenging to probe through conventional means. Furthermore, the modular labeling approach itself could also be employed independently in applications like affinity enrichment and genomic analyses[11-13] and is amenable to the integration of any label that can be incorporated by polymerase activity, including fluorescent tags.

SS Nanopore Measurements

Fabricated silicon chips (4 × 4 mm), each supporting a 10–20 μm thin film silicon nitride window (20 nm thickness), were obtained commercially (Norcada, Inc., Alberta, Canada). A single SS nanopore (diameter 7.5–9.0 nm, as determined from resistance measurement[27]) was produced in each membrane using a helium ion milling technique described elsewhere.[46] Prior to measurement, a chip was rinsed with deionized water and ethanol, dried under filtered air flow, and then exposed to air plasma (30 W) for 2 min on each side before being placed into a custom Ultem 1000 flow cell that enabled introduction of measurement buffer (1 M NaCl, 10 mM PBS buffer) to independent reservoirs on each side of the device. Ionic current

measurements were performed with a patch clamp amplifier (Axopatch 200B) through Ag/AgCl electrodes and used to verify pore diameter. After introducing biomolecules in measurement buffer to the cathode chamber, the current was recorded at a bandwidth of 200 kHz with a 100 kHz four-pole Bessel filter. Analysis was performed with custom software and an additional low-pass filter of 25 kHz. The event threshold for analysis was set at 4.5 standard deviations above the RMS noise level, and only events with durations between 12.5 μs and 2.5 ms were considered. Each rate measurement was determined by considering at least 3.5 min of uninterrupted trace recording, broken into segments of 3.2 s. The standard deviation between segments was taken as the measurement error.

Gel Electrophoresis

The denaturing gel was prepared by mixing thoroughly 70 mL of 23% gel matrix (22% acrylamide, 1% bis-acrylamide, 7 M urea in 1× tris/borate/EDTA (3:1:1) (TBE) buffer), 240 μL of 25% ammonium persulfate, and 42 μL of tetramethylethylenediamine. The gel mixture was cast and allowed to polymerize for 30 min before running samples with dye in 1× TBE (3:1:1) at 55 W for 90 min. Yields were approximated by measuring product band intensity relative to intermediates in the final lane using ImageJ analysis software.[47] For electromobility shift assays (EMSA) gels (see Supplementary Figure S3), 3.5% agarose gels were prepared in 1× TBE buffer with GelRed nucleic acid stain (Phenix Research Products, Candler, NC). Gel images were acquired using a Gel Doc system (Bio-Rad, Hercules, CA).

APE1 D308A Protein Expression

APE1 D308A plasmid (provided by the Demple Lab, Stony Brook University) was transformed into BL21*(DE3) cells and grown in 1 L LB broth at 37 °C. After bacterial cell cultures reached $OD_{600} = 0.6$, expression was induced with 0.5 mM isopropyl β-d-thiogalactopyranoside (IPTG). The cultures were then incubated for another 90 min before being harvested by centrifugation, resuspended in 50 mM HEPES-KOH (pH 7.5), 100 mM KCl, 1 mM EDTA, 0.1 mM DTT, and 10% (v/v) glycerol, and lysed by two passes through an EmulsiFlex-C5 (Avestin, Ottawa, Canada). The lysate was cleared by centrifugation at 20 000 g for 20 min, loaded onto a 15 mL SP Sepharose column (GE Healthcare, Pittsburgh, PA), and eluted with a linear gradient of 100–750 mM KCl. Elutions were analyzed by SDS-PAGE and fractions containing the protein were pooled and dialyzed overnight at 4 °C against APE1 storage buffer (50 mM HEPES-KOH (pH 7.5), 200 mM KCl, 1 mM EDTA, 0.1 mM DTT, 10% (v/v) glycerol) and concentrated using 10 kDa MWCO centrifugal spin filter columns (EMD Millipore, Billerica, MA). The final protein concentration was determined with the Bio-Rad Protein Assay (Bio-Rad), and aliquots were stored at −20 °C prior to use.

TDG Protein Expression from *E. coli*

We followed a protocol adapted by Liu et al.[48] from earlier work[49] with minor modifications. An expression plasmid for human TDG based on pET28 was transformed into BL21(DE3) cells and grown in 1 L LB broth at 37 °C. Once the cultures reached $OD_{600} = 0.6$, they were gradually cooled to 16 °C, induced with 0.25 mM IPTG and grown overnight. Cells were harvested by centrifugation, resuspended in 20 mL of TDG lysis buffer (50 mM sodium phosphate, pH 8.0,

300 mM NaCl, 25 mM imidazole) with protease inhibitors, and lysed by two passes through an EmulsiFlex-C5. The lysate was cleared by centrifugation at 20 000 *g* for 20 min, loaded onto a 1 mL column of HisPur cobalt resin (Fisher Scientific) equilibrated with TDG lysis buffer, and bound by two applications of the lysate to the column under gravity flow. The column was washed with 20 mL of TDG lysis buffer and subsequently eluted in 1 mL aliquots by a linear gradient of 100–500 mM imidazole. Elutions were analyzed by SDS-PAGE, and fractions containing the protein were pooled and dialyzed overnight at 4 °C against TDG storage buffer (20 mM HEPES, pH 7.5, 100 mM NaCl, 1 mM DTT, 0.5 mM EDTA, 1% v/v glycerol). Dialyzed proteins were concentrated using 10 kDa MWCO centrifugal spin filter columns. The final TDG concentration was determined with the Bio-Rad Protein Assay, and aliquots were stored at −80 °C prior to use.

Uracil Labeling

A custom 40 nt oligonucleotide with a 5′ FAM (sequence: TCA CGA CTA GTG TTA ACA TGT GCA CCT GCA GAA <u>U</u>GA GAA T) was annealed to a complementary sequence by mixing both at an equimolar ratio, incubating in deionized water at 95 °C for 10 min, and cooling to room temperature over 1 h. To cap the 3′ ends of the DNA, a 100 µL aliquot was prepared containing 385 pmol of duplex DNA, 30 nmol of 2′,3′-dideoxyadenosine 5′-triphosphate (ddATP) (GE Healthcare), 500 U of terminal transferase (New England Biolabs, Ipswich, MA), and 25 mmol $CoCl_2$ (New England Biolabs) and incubated in 1× terminal transferase reaction buffer (New England Biolabs) at 37 °C for 1.5 h. The resulting material was purified with the QIAquick PCR purification kit (Qiagen, Valencia, CA) to allow for buffer exchange. To excise uracil, a 30 µL aliquot was prepared containing 100 pmol of capped duplex DNA, 20 U of *E. coli* UDG (New England Biolabs), 40 U of EndoIV (New England Biolabs), 3 µg of bovine serum albumin (BSA, New England Biolabs), and incubated in 1× NEB2 buffer (New England Biolabs) at 37 °C for 1 h. Next, 1.5 nmol of biotinylated dUTP (PerkinElmer, Waltham, MA) and 0.12 U of T4(exo-) (Lucigen, Middleton, WI) was added to a final volume of 40 µL in 1× NEB2 buffer, and the mixture was incubated at 37 °C for 30 min. Finally, the mixture was subjected to purification by QIAquick PCR purification kit to remove proteins and excess nucleotides.

OxoG Labeling

A custom 40 nt oligonucleotide with a 5′ FAM (sequence: TCA CGA CTA GTG TTA ACA TGT GCA CCT <u>G</u>CA GAA TGA GAA T, where <u>G°</u> is oxoG) was annealed to a complementary sequence by mixing both at an equimolar ratio, incubating in deionized water at 95 °C for 10 min, and cooling to room temperature over 1 h. To excise oxoG, a 30 µL aliquot was prepared containing 100 pmol of duplex, 6.5 U hOGG1 (New England Biolabs), 40 U EndoIV, 3 µg BSA, and incubated in 1X NEB2 buffer at 37 °C for 1 h. Next, 1.5 nmol of biotinylated dGTP (PerkinElmer) and 0.12 U T4(exo-) were added to a final volume of 40 µL in 1X NEB2 buffer and the mixture was incubated at 37 °C for 30 min. Finally, the mixture was subjected to QIAquick PCR purification kit purification to remove proteins and excess nucleotides.

T:G Mismatch Labeling

A custom 40 nt oligonucleotide with a 5′ FAM (sequence: TCA CGA CTA GTG TTA ACA TGT CGA CCT <u>T</u>GA GAA TGA GAA T) was annealed to a complementary sequence (except with a guanine opposite the indicated thymine) by mixing both at an equimolar ratio, incubating in deionized water at 95 °C for 10 min, and cooling to room temperature over 1 h. To excise target thymine, a 30 μL aliquot was prepared containing 100 pmol of duplex, 7.5 mg of human TDG,[49] 40 fg of APE1 (D308A mutant[42]), 3 μg of BSA, and incubated in 1× HEMN.1 buffer (20 mM HEPES (pH 7.3), 100 mM NaCl, 2.5 mM $MgCl_2$, 0.2 mM EDTA) at 37 °C for 1 h. The mixture was purified with a QIAquick PCR purification kit. Then, 40 U of EndoIV and 3 μg of BSA were added to a total volume of 30 μL in 1× NEB2 buffer and incubated at 37 °C for 30 min. 1.5 nmol of biotinylated dCTP (PerkinElmer, Waltham, MA) and 0.12 U of T4(exo-) were added to a final volume of 40 μL in 1× NEB2 buffer, and the mixture was further incubated at 37 °C for 30 min. Finally, the mixture was subjected to a second purification to remove proteins and excess nucleotides.

$1,N^6$-Ethenoadenine Labeling

A custom 34 nt oligonucleotide with a 5′ FAM (sequence: CAG TTG AGG ATC CCC ATA <u>A^e</u>TG CGG CTG TTT TCT G, where $\underline{A^e}$ is $1,N^6$-ethenoadenine) was annealed to a complementary sequence by mixing both at an equimolar ratio, incubating in deionized water at 95 °C for 10 min, and cooling to room temperature over 1 h. To cap the 3′ ends of the DNA, a 100 μL aliquot containing 385 pmol of duplex DNA, 30 nmol of ddATP, 500 U of terminal transferase, and 25 mmol of $CoCl_2$ was incubated in 1× terminal transferase reaction buffer at 37 °C for 1.5 h. The resulting material was purified with a QIAquick PCR purification kit to allow for buffer exchange. To excise target $1,N^6$-ethenoadenine, a 80 μL aliquot was prepared containing 100 pmol of duplex, 425 U of hAAG (New England Biolabs), 200 fg of APE1 D308A mutant, 8 μg of BSA, and 1× Thermopol buffer (New England Biolabs) and incubated at 37 °C for 1 h. The mixture was purified with a QIAquick PCR purification kit. Then, 40 U of EndoIV and 3 μg of BSA were added to a total volume of 30 μL in 1× NEB2 buffer and incubated at 37 °C for 30 min. 1.5 nmol of biotinylated dATP (PerkinElmer, Waltham, MA) and 0.12 U of T4(exo-) were added to a final volume of 40 μL in 1× NEB2 buffer, and the mixture was further incubated at 37 °C for 30 min. Finally, the mixture was subjected to a second QIAquick purification to remove proteins and excess nucleotides.

Supporting Information

The Supporting Information is available at https://doi.org/10.1021/acs.nanolett.7b03911.

Author Contributions

designed and oversaw the project, analyzed data, and wrote the manuscript. All authors reviewed the manuscript.

The authors declare the following competing financial interest(s): A.R.H. is listed as inventor on patents covering the presented labeling approach and the SS nanopore assay it enhances. The other authors declare no competing interests.

Acknowledgment

References

1. Jaenisch, R.; Bird, A. Nat. Genet. **2003**, 33, 245– 254 DOI: 10.1038/ng1089 Google Scholar

2. Slotkin, R. K.; Martienssen, R. Nat. Rev. Genet. **2007**, 8, 272– 285 DOI: 10.1038/nrg2072 Google Scholar

3. Zhang, G.; Huang, H.; Liu, D.; Cheng, Y.; Liu, X.; Zhang, W.; Yin, R.; Zhang, D.; Zhang, P.; Liu, J.; Li, C.; Liu, B.; Luo, Y.; Zhu, Y.; Zhang, N.; He, S.; He, C.; Wang, H.; Chen, D. Cell **2015**, 161, 893– 906 DOI: 10.1016/j.cell.2015.04.018 Google Scholar

4. Reik, W.; Dean, W.; Walter, J. Science **2001**, 293, 1089– 1093 DOI: 10.1126/science.1063443 Google Scholar

5. Okamoto, I.; Otte, A. P.; Allis, C. D.; Reinberg, D.; Heard, E. Science **2004**, 303, 644– 649 DOI: 10.1126/science.1092727 Google Scholar

6. Egger, G.; Liang, G. N.; Aparicio, A.; Jones, P. A. Nature **2004**, 429, 457– 463 DOI: 10.1038/nature02625 Google Scholar

7. Korlach, J.; Turner, S. W. Curr. Opin. Struct. Biol. **2012**, 22, 251– 261 DOI: 10.1016/j.sbi.2012.04.002 Google Scholar

8. Grunau, C.; Clark, S. J.; Rosenthal, A. Nucleic Acids Res. **2001**, 29, e65– e72 Google Scholar

9. Cadet, J.; Douki, T.; Ravanat, J. L. Environ. Health Perspect. **1997**, 105, 1034– 1039 DOI: 10.2307/3433836 Google Scholar

10. Baker, M. Nature **2015**, 521, 274– 276 DOI: 10.1038/521274a Google Scholar

11. Song, C.-X.; Clark, T. A.; Lu, X.-Y.; Kislyuk, A.; Dai, Q.; Turner, S. W.; He, C.; Korlach, J. Nat. Methods **2011**, 9, 75– 77 DOI: 10.1038/nmeth.1779 [Google Scholar](#)

12. 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 DOI: 10.1016/j.cell.2013.04.001 [Google Scholar](#)

13. Zhang, L.; Szulwach, K. E.; Hon, G. C.; Song, C. X.; Park, B.; Yu, M.; Lu, X. Y.; Dai, Q.; Wang, X.; Street, C. R.; Tan, H. P.; Min, J. H.; Ren, B.; Jin, P.; He, C. Nat. Commun. **2013**, 4, 1517 DOI: 10.1038/ncomms2527 [Google Scholar](#)

14. Lu, X.; Han, D.; Zhao, B. S.; Song, C.-X.; Zhang, L.-S.; Doré, L. C.; He, C. Cell Res. **2015**, 25, 386– 389 DOI: 10.1038/cr.2015.5 [Google Scholar](#)

15. Dekker, C. Nat. Nanotechnol. **2007**, 2, 209– 215 DOI: 10.1038/nnano.2007.27 [Google Scholar](#)

16. Wanunu, M. Phys. Life Rev. **2012**, 9, 125– 158 DOI: 10.1016/j.plrev.2012.05.010 [Google Scholar](#)

17. Li, J. L.; Gershow, M.; Stein, D.; Brandin, E.; Golovchenko, J. A. Nat. Mater. **2003**, 2, 611– 615 DOI: 10.1038/nmat965 [Google Scholar](#)

18. Storm, A. J.; Chen, J. H.; Zandbergen, H. W.; Dekker, C. Phys. Rev. E **2005**, 71, 051903 DOI: 10.1103/PhysRevE.71.051903 [Google Scholar](#)

19. Skinner, G. M.; van den Hout, M.; Broekmans, O.; Dekker, C.; Dekker, N. H. Nano Lett. **2009**, 9, 2953– 2960 DOI: 10.1021/nl901370w [Google Scholar](#)

20. van den Hout, M.; Skinner, G. M.; Klijnhout, S.; Krudde, V.; Dekker, N. H. Small **2011**, 7, 2217– 2224 DOI: 10.1002/smll.201100265 [Google Scholar](#)

21. Firnkes, M.; Pedone, D.; Knezevic, J.; Doeblinger, M.; Rant, U. Nano Lett. **2010**, 10, 2162– 2167 DOI: 10.1021/nl100861c [Google Scholar](#)

22. Plesa, C.; Kowalczyk, S. W.; Zinsmeester, R.; Grosberg, A. Y.; Rabin, Y.; Dekker, C. Nano Lett. **2013**, 13, 658– 663 DOI: 10.1021/nl3042678 [Google Scholar](#)

23. Larkin, J.; Henley, R. Y.; Muthukumar, M.; Rosenstein, J. K.; Wanunu, M. Biophys. J. **2014**, 106, 696– 704 DOI: 10.1016/j.bpj.2013.12.025 [Google Scholar](#)

24. Storm, A. J.; Storm, C.; Chen, J. H.; Zandbergen, H.; Joanny, J. F.; Dekker, C. Nano Lett. **2005**, 5, 1193– 1197 DOI: 10.1021/nl048030d [Google Scholar](#)

25. Smeets, R. M. M.; Kowalczyk, S. W.; Hall, A. R.; Dekker, N. H.; Dekker, C. Nano Lett. **2009**, 9, 3089– 3095 DOI: 10.1021/nl803189k [Google Scholar](#)

26. Marshall, M. M.; Ruzicka, J.; Zahid, O. K.; Henrich, V. C.; Taylor, E. W.; Hall, A. R. Langmuir **2015**, 31, 4582– 4588 DOI: 10.1021/acs.langmuir.5b00457 [Google Scholar](#)

27. Wanunu, M.; Dadosh, T.; Ray, V.; Jin, J.; McReynolds, L.; Drndic, M. Nat. Nanotechnol. **2010**, 5, 807– 814 DOI: 10.1038/nnano.2010.202 [Google Scholar](#)

28. Zahid, O. K.; Zhao, B. S.; He, C.; Hall, A. R. Sci. Rep. **2016**, 6, 29565 DOI: 10.1038/srep29565 Google Scholar

29. Carlsen, A. T.; Zahid, O. K.; Ruzicka, J. A.; Taylor, E. W.; Hall, A. R. Nano Lett. **2014**, 14, 5488– 5492 DOI: 10.1021/nl501340d Google Scholar

30. Howarth, M.; Chinnapen, D. J. F.; Gerrow, K.; Dorrestein, P. C.; Grandy, M. R.; Kelleher, N. L.; El-Husseini, A.; Ting, A. Y. Nat. Methods **2006**, 3, 267– 273 DOI: 10.1038/nmeth861 Google Scholar

31. Riedl, J.; Ding, Y.; Fleming, A. M.; Burrows, C. J. Nat. Commun. **2015**, 6, 8807 DOI: 10.1038/ncomms9807 Google Scholar

32. Kavli, B.; Otterlei, M.; Slupphaug, G.; Krokan, H. E. DNA Repair **2007**, 6, 505– 516 DOI: 10.1016/j.dnarep.2006.10.014 Google Scholar

33. Grollman, A. P.; Moriya, M. Trends Genet. **1993**, 9, 246– 249 DOI: 10.1016/0168-9525(93)90089-Z Google Scholar

34. Zahid, O. K.; Wang, F.; Ruzicka, J. A.; Taylor, E. W.; Hall, A. R. Nano Lett. **2016**, 16, 2033– 2039 DOI: 10.1021/acs.nanolett.6b00001 Google Scholar

35. Jacobs, A. L.; Schär, P. Chromosoma **2012**, 121, 1– 20 DOI: 10.1007/s00412-011-0347-4 Google Scholar

36. Wiebauer, K.; Jiricny, J. Proc. Natl. Acad. Sci. U. S. A. **1990**, 87, 5842– 5845 DOI: 10.1073/pnas.87.15.5842 Google Scholar

37. Maiti, A.; Drohat, A. C. J. Biol. Chem. **2011**, 286, 35334– 35338 DOI: 10.1074/jbc.C111.284620 Google Scholar

38. Tini, M.; Benecke, A.; Evans, R. M.; Chambon, P. Mol. Cell **2002**, 9, 265– 277 DOI: 10.1016/S1097-2765(02)00453-7 Google Scholar

39. Mol, C. D.; Izumi, T.; Mitra, S.; Tainer, J. A. Nature **2000**, 403, 451– 456 DOI: 10.1038/35000249 Google Scholar

40. Waters, T. R.; Gallinari, P.; Jiricny, J.; Swann, P. F. J. Biol. Chem. **1999**, 274, 67– 74 DOI: 10.1074/jbc.274.1.67 Google Scholar

41. Chou, K.-M.; Cheng, Y.-C. J. Biol. Chem. **2003**, 278, 18289– 18296 DOI: 10.1074/jbc.M212143200 Google Scholar

42. Masuda, Y.; Bennett, R. A. O.; Demple, B. J. Biol. Chem. **1998**, 273, 30360– 30365 DOI: 10.1074/jbc.273.46.30360 Google Scholar

43. Abner, C. W.; Lau, A. Y.; Ellenberger, T.; Bloom, L. B. J. Biol. Chem. **2001**, 276, 13379– 13387 DOI: 10.1074/jbc.M010641200 Google Scholar

44. Wyatt, M. D.; Allan, J. M.; Lau, A. Y.; Ellenberger, T. E.; Samson, L. D. BioEssays **1999**, 21, 668– 676 DOI: 10.1002/(SICI)1521-1878(199908)21:8<668::AID-BIES6>3.3.CO;2-4 Google Scholar

45. Morales-Ruiz, T.; Ortega-Galisteo, A. P.; Ponferrada-Marín, M. I.; Martínez-Macías, M. I.; Ariza, R. R.; Roldán-Arjona, T. Proc. Natl. Acad. Sci. U. S. A. **2006**, 103, 6853– 6858 DOI: 10.1073/pnas.0601109103 [Google Scholar](#)

46. Yang, J.; Ferranti, D. C.; Stern, L. A.; Sanford, C. A.; Huang, J.; Ren, Z.; Qin, L.-C.; Hall, A. R. Nanotechnology **2011**, 22, 285310 DOI: 10.1088/0957-4484/22/28/285310 [Google Scholar](#)

47. Collins, T. J. BioTechniques **2007**, 43, 25– 30 DOI: 10.2144/000112517 [Google Scholar](#)

48. Liu, M. Y.; Torabifard, H.; Crawford, D. J.; DeNizio, J. E.; Cao, X.-J.; Garcia, B. A.; Cisneros, G. A.; Kohli, R. M. Nat. Chem. Biol. **2016**, 13, 181– 187 DOI: 10.1038/nchembio.2250 [Google Scholar](#)

49. Morgan, M. T.; Bennett, M. T.; Drohat, A. C. J. Biol. Chem. **2007**, 282, 27578 DOI: 10.1074/jbc.M704253200 [Google Scholar](#)