Published in final edited form as:

ACS Nano. 2018 May 22; 12(5): 4194–4205. doi:10.1021/acsnano.8b01198.

Single Locked Nucleic Acid-Enhanced Nanopore Genetic Discrimination of Pathogenic Serotypes and Cancer Driver Mutations

Kai Tian^{#†}, Xiaowei Chen^{#†}, Binqun Luan^{*,‡}, Prashant Singh[‡], Zhiyu Yang[§], Kent S. Gates[§], Mengshi Lin[‡], Azlin Mustapha[‡], and Li-Qun Gu^{*,†}

[†]Department of Bioengineering and Dalton Cardiovascular Research Center, University of Missouri, Columbia, Missouri 65211, United States

[‡]Food Science Program, Division of Food Systems and Bioengineering, University of Missouri, Columbia, Missouri 65211, United States

[§]Department of Chemistry, University of Missouri, Columbia, Missouri 65211, United States

^{||}Computational Biology Center, IBM Thomas J. Watson Research, Yorktown Heights, New York 10598, United States

[#] These authors contributed equally to this work.

Abstract

Accurate and rapid detection of single-nucleotide polymorphism (SNP) in pathogenic mutants is crucial for many fields such as food safety regulation and disease diagnostics. Current detection methods involve laborious sample preparations and expensive characterizations. Here, we investigated a single locked nucleic acid (LNA) approach, facilitated by a nanopore single-molecule sensor, to accurately determine SNPs for detection of Shiga toxin producing *Escherichia coli* (STEC) serotype O157:H7, and cancer-derived *EGFR* L858R and *KRAS* G12D driver mutations. Current LNA applications that require incorporation and optimization of multiple LNA nucleotides. But we found that in the nanopore system, a single LNA introduced in the probe is sufficient to enhance the SNP discrimination capability by over 10-fold, allowing accurate detection of the pathogenic mutant DNA mixed in a large amount of the wildtype DNA. Importantly, the molecular mechanistic study suggests that such a significant improvement is due to the effect of the single-LNA that both stabilizes the fully matched base-pair and destabilizes the mismatched base-pair. This sensitive method, with a simplified, low cost, easy-to-operate LNA design, could be generalized for various applications that need rapid and accurate identification of single-nucleotide variations.

Graphical Abstract

*Corresponding Authors: bluan@us.ibm.com, gul@missouri.edu.

Notes

The authors declare no competing financial interest.

Keywords

locked nucleic acid; nanopore; single nucleotide polymorphism; driver mutation; E. coli O157:H7 foodborne pathogen serotype; cancer diagnostics; EGFR; KRAS

Single-nucleotide polymorphism (SNP) is the alteration of one nucleotide occurring at a specific site in the genome between either paired chromosomes or members of species. SNPs play crucial roles in both genetic and epigenetic levels of gene expressions and, therefore, are used as important biomarkers for diagnostics^{1–4} and standards in pathogenic species identification.^{5–8} Although some technologies, such as real-time PCR,⁹ microarray¹⁰ and sequencing,^{11,12} have been widely utilized for SNP detection in clinical settings, it is still highly demanded to develop high-resolution approaches to accurate and rapid genotyping for SNP discrimination.

Locked nucleic acids (LNAs) are a class of artificial RNA mimicking nucleotides. Due to the special “locked” ribose ring (Figure 1a), LNA can increase the double strands’ thermal stability when hybridized to a complementary DNA or RNA.¹³ This function renders LNA a high-performance probe in a variety of hybridization-based applications, from SNP discrimination¹⁴ and microRNA detection^{15,16} to gene silencing¹⁷ and DNA zyme activity enhancement.¹⁸ However, designing LNA probes remains a challenge because most applications require incorporation of multiple LNA nucleotides (at least three) and optimization of the position of each LNA in the probe. As such, the LNA design has to be a complicated, laborious, and expensive process. This challenge is partially due to the fact that the molecular mechanism for the LNA effect remains unclear. In particular, current technologies are not sensitive enough to precisely elucidate the role of a single LNA in its applications such as SNP discrimination.

Nanopore is a label-free, ultrasensitive, single-molecule-based sensing technology. It has been broadly investigated for various genetic,^{19–23} epigenetic,^{24–27} and proteomic^{28–30} detection strategies.³¹ Many excellent studies have demonstrated nanopore’s single-nucleotide sensitivity^{32–34} and ability to detect single-nucleotide polymorphism.^{33,35–37} The nanopore-based next-generation sequencing technology is being developed.^{36,38–40} Through collaboration, we have developed two different single-molecule platforms, nanolock-nanopore⁴¹ and nanocross-nanopore^{42,43} biosensors, for the detection of cancer-derived point mutations.

Motivated by the merits of LNAs and nanopore’s single-nucleotide sensitivity, here we report a combined single-LNA-nanopore approach to LNA mechanistic study and SNP

discrimination (Figure 1b). This approach verifies that a single LNA introduced in the probe is sufficient to accurately discriminate various SNPs. Combined with molecular dynamic (MD) simulation, this approach allows elucidating the singleLNA mechanism: while the LNA can stabilize a matched basepair, it, however, can surprisingly deteriorate a mismatched one as well as a neighboring matched one, yielding a dramatically magnified contrast between the pathogen and nonpathogen nanopore signatures. The approach also allows investigating how to regulate the LNA-enhanced SNP discrimination capability by various sequence factors, such as different mismatched base-pairs, the target sequence length, the SNP position, and the types of neighboring nucleotides. All these findings are useful for optimizing the performance of the singleLNA-nanopore sensors for SNP discrimination. The approach could be expanded to the mechanistic and functional study of various artificial nucleotides.

To demonstrate the universal applicability of the single-LNAnanopore sensor, we studied three important SNPs in broad fields from food science to oncology. The first target is the +93 SNP in *Escherichia coli uidA* gene. Foodborne pathogen detection is globally important for food safety and foodborne disease prevention. In addition to time-consuming culturebased approaches,⁴⁴ foodborne pathogens can be detected through molecular methods, which identify pathogen-secreted proteins⁴⁵ or specific pathogenic genes^{46–48} (e.g., nucleic acids amplification-based assays).^{49,50} *E. coli* O157:H7 is the most frequently isolated Shiga toxin producing *E. coli* (STEC) serotype. Because *E. coli* O157:H7 and non-O157 serotypes only differ by one nucleotide in the *uidA* gene (*uidA* +93), this SNP has been used as a biomarker for distinguishing the O157 pathogenic serotype from all other *E. coli* serotypes.^{49–54} The second and third targets are *EGFR* L858R and *KRAS* G12D driver mutations. Driver mutations cause malfunctions of proteins, affecting cell proliferation and eventually developing cancers. Epidermal growth factor receptor (EGFR) is a transmembrane protein family of tyrosine kinases that regulates cellular proliferation, differentiation, and survival.⁵⁵ Because the *EGFR* L858R mutation is highly associated with nonsmall-cell lung cancer (NSCLC),^{56,57} genotyping the SNP of this mutation is important for both cancer diagnostics and treatment. The membrane-tethered KRAS (a GTPase) turns on/off many signal transduction pathways. Its mutations such as G12D impair the intrinsic GTPase activity and the interaction with proteins that modulate KRAS activation, causing KRAS to remain activated and constitutively turn on the downstream signaling.⁵⁸ This mutation has been widely discovered in lung⁵⁹ and pancreatic^{60–62} cancers and is a therapy target¹ and the hallmark of prognostics.^{63,64}

RESULTS

As shown in Figure 1b, the nanopore for this study is a 2 nm protein pore assembled by α -hemolysin. It was reconstituted in a lipid membrane that insulates the solutions on both sides of the pore. A transmembrane voltage is applied to produce an ion current across the pore. Single-stranded nucleic acids can freely translocate through this nanopore, but double-stranded nucleic acids must be unzipped prior to translocation, driven by the voltage. This process can be revealed by the nanopore current signature.^{26,65,66} To detect an SNP, we first design a probe that forms a fully matched duplex with the pathogenic gene target and a single-mismatched duplex with the nonpathogenic target at the SNP site. The probe flanks a

3'-poly(dC)₁₅ tag for trapping the duplex into the nanopore from its cis entrance and unzipping the duplex.^{26,67} The unzipping/translocation process reduced the nanopore current to the level of $I/I_0 \sim 10\%$ (I and I_0 are the blocked and open pore currents).⁶⁷ The blockade duration, that is, the unzipping time (τ_{uz}), indicates the duplex stability. The fully matched duplex is more stable than the mismatched one and, thus, can be discriminated from the prolonged unzipping time. The ratio of unzipping times between the fully matched and mismatched duplexes ($\tau_{uz-FM}/\tau_{uz-MM}$) measures the SNP discrimination capability. Each target is detected by two probes: a DNA probe containing all regular nucleotides and a LNA probe with a locked nucleoside substitution at the SNP site. Due to the use of LNA, an increase in SNP discrimination capability can be observed. The fold of increase is defined as the enhancement magnitude.

Single LNA-Enhanced Genetic Discrimination of Pathogenic Serotype and Cancer Driver Mutations.

We first investigated how LNA enhances the discrimination of *E. coli* O157 and non-O157 serotypes. The target and probesequences are given in Table 1. The sequence of the 17-nt synthetic target is truncated from the antisense strand of the *uidA* gene. The SNP site is located in the middle of the sequence, which is a cytosine (C) in the O157 target T1_C and an adenine (A) in the non-Q157 target T1_A. The DNA probe P1_G contains a regular guanosine (G), and the LNA probe P1_{LG} contains a locked guanosine (^LG) at the SNP site, such that they form a G/^LG-C pair with T1_C (P1_G·T1_C and P1_{LG}·T1_C) and a G/^LG-A mismatched base-pair with T1_A (P1_G·T1_A and P1_{LG}·T1_A). Using the DNA probe, τ_{uz} for P1_G·T1_C was 37 ± 3 ms, 1.6-fold as long as the 22 ± 4 ms for P1_G·T1_A (Figure 2a and b, Figure 3a), indicating that the O157/non-O157 discrimination capability is 1.6 (Figure 3b). Strikingly, when using the LNA probe, τ_{uz} for the fully matched P1_{LG}·T1_C was increased to 61 ± 14 ms, while τ_{uz} for mismatched P1_{LG}·T1_A was decreased to 5.0 ± 1.2 ms (Figure 2c and d, Figure 3a), therefore leading to significant increase of the discrimination capability to 12 (Figure 3b). This key result suggests that the probe's ^LG at the SNP site can enhance the O157/non-O157 discrimination capability by 7.5-fold (Figure 3c).

To study the universal applicability of the LNA mechanism, we expanded the SNP target to cancer driver mutations, EGFR L858R and KRAS G12D. Both single-nucleotide genetic alterations are cancer biomarkers and therapy targets. We found that the LNA effect on the discrimination of the two medically relevant SNPs is similar to that of the discrimination of *E. coli* O157 SNP. Similar to *E. coli* O157, EGFR L858R is also an A > C substitution in the antisense strand, so the detection of this target can verify the ^LG effect on the same SNP in a different sequence. The sequences of the 17-nt mutant (T6_C) and wild-type (T6_A) targets are truncated from the antisense strand of the EGFR gene. T6_C has a cytosine (C) and T6_A has an adenine (A) at the mutation site, thus the corresponding DNA probe (P6_G) and single-^LG probe (P6_{LG}) can form a G/^LG-C pair with T6_C and a G/^LG-A mismatched base-pair with T6_A. Using the DNA probe, τ_{uz} for P6_G·T6_C was 6.8 ± 3.4 ms, 3.6-fold as long as the 1.9 ± 0.5 ms for P6_G·T6_A (Figure 3a), indicating that the L858R mutant discrimination capability is 3.6 (Figure 3b). Using the ^LG probe greatly extended τ_{uz} of fully matched P6_{LG}·T6_C to 42.2 ± 10.4 ms, while slightly shortened τ_{uz} of mismatched P6_{LG}·T6_A to 1.9 ± 0.2 ms (Figure 3a), therefore enhancing the mutation discrimination capability to 22

(Figure 3b). Overall, a single L G enhances the L858R discrimination capability by 6.2-fold (Figure 3c), similar to the 7.5-fold for *E. coli* O157 discrimination.

KRAS G12D is a C > T substitution (sense strand). The sequences of the 17-nt mutant (T7_T) and wild-type (T7_C) targets contain a thymidine (T) and a cytosine (C) at the mutation site, respectively. The two probes, one (P7_A) with a regular adenosine (A) and the other (P7_{LA}) with a locked adenosine (^LA) at the mutation site, can form an A/^LA–T pair with T7_T (P7_A·T7_T and P7_{LA}·T7_T) and an A/^LA···C mismatched base-pair with T7_C (P7_A·T7_C and P7_{LA}·T7_C). Using the DNA probe, τ_{uz} for P7_A·T7_T was 26.3 ± 10.8 ms. This τ_{uz} is 3.42-fold as long as the 7.7 ± 2.5 ms for P7_A·T7_C (Figure 3a), showing that the KRAS G12D discrimination capability is 3.42 (Figure 3b). Using the ^LA probe extended τ_{uz} of fully matched P7_{LA}·T7_C to 69.6 ± 17.5 ms, while shortened τ_{uz} of mismatched P7_{LA}·T7_C to 6.5 ± 1.3 ms (Figure 3a). The overall effect is greatly increasing the discrimination capability to 10.7 (Figure 3b). Therefore, a single ^LA enhances the KRAS mutant discrimination capability by 3.1-fold (Figure 3c).

The results from three different SNP species indicate that the introduction of a single LNA to the SNP site in the probe is sufficient enough to enhance the nanopore's SNP discrimination capability. The enhancement is contributed by two factors: (1) LNA stabilizes a fully matched duplex, in which LNA forms a Watson–Crick base-pair with the pathogenic target DNA. This can be verified by the fact that the LNA probe significantly extends the fully match unzipping time by 1.7-fold for *E. coli* O157, 6.2-fold for EGFR L858C, and 2.7-fold for KRAS G12D (Figure 3d, $\tau_{uz-FM} (LNA) / \tau_{uz-FM} (DNA)$). (2) Strikingly, LNA destabilizes the mismatched duplex with nonpathogenic DNA, as reflected by shortening the mismatched unzipping time by 4.5-, 1.1-, and 1.2-fold for *E. coli* non-O157, wild-type EGFR, and KRAS (Figure 3d, $\tau_{uz-MM} (DNA) / \tau_{uz-MM} (LNA)$). In addition, the melting profiles (Figure S3) support that the order of T_m for the four duplexes in *E. coli* O157 detection is consistent with that of their stabilities (unzipping time) in the nanopore, but the nanopore method is more efficacious (Figure S3). Overall, the single LNA plays opposite roles in stabilizing the fully matched duplex and destabilizing the one-mismatched duplex. The two factors together amplify the difference of their unzipping times, leading to the enhancement of SNP discrimination.

Regulation of the LNA Effect by Sequence Factors.

To improve the single-LNA-nanopore for SNP discrimination in future applications, we utilized the *E. coli uidA* gene +93 SNP as the basis to investigate how the LNA effect is regulated by various sequence factors, such as the base type of the SNP site, the LNA's neighboring sequence, the LNA position in the sequence, and the target length.

Similar to the LNA role in destabilizing the G···A mismatched duplex (T1_A, Figure 3), the single ^LG in the probe also reduces the stability of the G···G mismatched duplex with the target T1_G. τ_{uz} was decreased by 2.7-fold (Figure 4d) from 7.2 ± 3.1 ms for P1_G·T1_G to 2.7 ± 1.4 ms for P1_{LG}·T1_G (Figure 4a), resulting in a significant enhancement of the G–C/G···G discrimination capability by 4.37-fold (Figure 4c). However, LNA shows a different role in another duplex containing a G···T mismatched base-pair with target T1_T. The LG in the probe slightly increased, rather than decreased τ_{uz} (Figure 4d) from 8.3 ± 4.8 ms for

$P1_G \cdot T1_T$ to 10.4 ± 1.5 ms for $P1_{LG} \cdot T1_T$ (Figure 3a). This is the only one among all targets we studied that shows stabilization of a mismatched duplex by LNA. We interpret that the $G \cdots T$ mismatched base-pair, which is generally considered as a stable noncanonical pair,⁶⁸ can be grouped with A–T and G–C and further stabilized by LNA (simulation study below). Note that the similar $G \cdots U$ noncanonical pair widely participates in RNA tertiary structures^{69,70} even though the G–C/G $\cdots$ T discrimination capability is still moderately high (5.87, Figure 4b), and the LNA results in 1.31-fold enhancement of discrimination.

Next, we substituted L G's neighbor nucleotides in the probe from AG/ L GC (purine-G/ L G-pyrimidine, $P1_{LG}$, and $P1_G$) to AG/ L GA (purine-G/ L G-purine, $P2_{LG}$, and $P2_G$) and CG/ L GC (pyrimidine-G/ L G-pyrimidine, $P3_{LG}$, and $P3_G$). Their corresponding targets are $T2_C/T2_A$ and $T3_C/T3_A$. The probes still form a G/ L G–C pair with $T2_C$ and $T3_C$ and a G/ L G $\cdots$ A mismatched base-pair with $T2_A$ and $T3_A$. For the AG/ L GA motif, L G enhanced the G–C/G $\cdots$ A discrimination capability by 3.18-fold (Figure 4c) from 1.71 to 5.44 (Figure 4b). This enhancement is mainly contributed by destabilizing the G $\cdots$ A mismatched duplex as τ_{uz} was shortened by 3.65-fold between $P2_{LG} \cdot T2_A$ vs $P2_G \cdot T2_A$ (Figure 3d), though L G did not improve the stability of the fully matched duplex ($P2_{LG} \cdot T2_C$ vs $P2_G \cdot T2_C$). For the C/ L G/GC motif, L G enhanced the G–C/G $\cdots$ A discrimination capability by 3.25-fold (Figure 4c) to 5.88 (Figure 4b). This enhancement is contributed by both the 1.93fold increase in τ_{uz} between fully matched $P3_{LG} \cdot T3_C$ and $P3_G \cdot T3_C$ and the 1.68-fold decrease in τ_{uz} between mismatched $P3_{LG} \cdot T3_A$ and $P3_G \cdot T3_A$.

Furthermore, we truncated the target sequences from the *uidA* gene that shifts the C > A SNP position from the middle ninth to 15th near the 3'-end ($T4_G$ and $T4_A$). The unzipping time data (Figure 4a) suggest that L G weakly increased the G–C/G $\cdots$ A discrimination capability to 2.55 (Figure 4b), a 1.99-fold enhancement (Figure 4c) relative to that using the DNA probe. These values are much lower than the discrimination capability of 12.2 (Figure 3b) and 7.35-fold enhancement (Figure. 3c) for the SNP in the middle of the sequence ($T1_G$ and $T1_A$).

Lastly, we truncated the sequences of 23-nt long targets from the *uidA* gene with the middle SNP ($T5_C$ and $T5_A$). The DNA and LNA probes form a G/ L G–C pair with the O157 target $T5_C$ and a G/ L G $\cdots$ A mismatched base-pair with the non-O157 target $T5_A$. This long target demonstrates the strongest LNA effect. Based on the unzipping time (Figure 4a), the single L G enhanced the O157/non-O157 discrimination capability by 10.2-fold (Figure 4c) to 44.1 (+150 mV, Figure 4b), which is the highest among all targets we studied.

In summary, under most of the above sequence conditions, the single LNA can enhance the SNP discrimination capability (Figure 4c), and such enhancement is contributed by both stabilizing the fully matched duplex and destabilizing the one mismatched duplex (Figure 4d). Both discrimination capability and its enhancement (by using LNA) vary with these sequence factors. Understanding these factors are useful for the LNA design. For instance, to achieve optimized SNP discrimination capability, the SNP base-pair with LNA should be placed near the middle of the sequence, rather than close to a duplex terminal. Also a long target sequence is advantageous as it not only greatly enhances the SNP discrimination

capability but also improves the selectivity for binding with the probe and provides more enzyme cutting sites for target preparation in real samples.

MD Simulation Reveals the Function of the Single LNA in Duplexes

To understand the molecular mechanism underlying the different transport behaviors of the four duplexes, we carried out MD simulations on the four duplexes, containing G–C, ^LG–C, G...A, and ^LG...A base-pairs, respectively (Methods). Figure 5a illustrates the simulation system: a 6-mer dsDNA molecules solvated in a 1 M KCl electrolyte. The atomic coordinations between a matched ^LG–C and a mismatched ^LG...A are highlighted in Figure 5b and 5c, respectively. From four simulations, we extracted the structure parameters for the above four kinds of base-pairs (Table 2). For the fully matched G–C pair, the propeller twist (12.5°) is the smallest among the four kinds. In the mismatched G...A pair, the larger propeller twist (17.1°) than the one for the G–C pair, is due to the repulsion between two hydrogen atoms (instead of hydrogen-bond-forming H and O atoms as in the G–C pair) at the pairing interface (Figure 5c). When ^LG was introduced, because of the locked C3'-endo sugar-ring in ^LG and the predominately favored C2'-endo sugar-rings in other nucleotides, the propeller twist for the ^LG–C is about 5.5° larger than that of the G–C, which however had little effect on the stability of the ^LG–C pairing (Movie S1). Surprisingly, for the mismatched ^LG...A, the propeller twist (28.1°) is even much larger, suggesting that the mismatch and the locked sugar-ring conformation act synergistically to deform the ^LG...A pairing. The large propeller twist also yields that in the ^LG...A pairing frequently there is only one-hydrogen bond (i.e., N...H–N in Figure 5c) at the pairing interface, destabilizing the pairing. The change in the buckle angle is negligible between G–C and ^LG–C pairs, however, it is about 6.8° between G...A and ^LG...A. The large buckle and propeller (twisting) angles in the ^LG...A pair can yield a local instability (see below, and Movie S2). Regarding the local DNA diameter, overall, in basepairs with ^LG, the distance between two phosphorus atoms (P–P) is about 0.3–0.5 Å larger than the one without LNA (Table 2).

The time-dependent pairwise interaction energies for different base-pairs are shown in Figure 5d and 5e. We found that the time-averaged interaction energies for the P_LG·T_C are about 0.6 kcal/mol larger (or more negative) than those for the P_G·T_C duplex (Figure 5d), which is consistent with the fact the LNA can stabilize the DNA duplex formation. The larger interaction energy in P_LG·T_C could partially result from two more distant (negatively charged) phosphate groups (Table 2). The time-dependent interaction energies for the P_LG·T_A and P_G·T_A are comparable (Figure 5e), however from time to time, interaction energies for the P_LG·T_A decrease (less negative) due to the temporary breaking of the weak pairing (see structure analysis above) in the ^LG...A pair. Additionally, a neighboring matched pair, affected by the breaking of the ^LG...A pair, can break accordingly, greatly destabilizing the local DNA structure.

Besides the interaction energies, we also calculated the binding free energies of four 6-mer duplexes using the empirical approach (Methods) that takes into account the enthalpy and entropy of duplex annealing. The calculated binding free energies G (at 300 K) are –7.77, –6.83, –3.27 and –3.16 kcal/mol, respectively for duplexes P_LG·T_C, P_G·T_C, P_LG·T_A, and P_G·T_A. Therefore, consistently, with ^LG in the matched G–C base-pair, the binding free

energy of the duplex increases by 0.94 kcal/mol. Note that the binding free energies are comparable between $P_{LG} \cdot T_A$ and $P_G \cdot T_A$, likely due to the fact that structural factors discussed above are not included in the theoretical model. Nevertheless, $G(LNA) = G(P_{LG} \cdot T_A) - G(P_{LG} \cdot T_C) = 4.50$ kcal/mol and $G(DNA) = G(P_G \cdot T_A) - G(P_G \cdot T_C) = 3.67$ kcal/mol; the larger value of $G(LNA)$ suggests that when compared with DNA, the local LNA modification can enhance the mismatch discrimination.

With respect to the base-pair stability for a regular dsDNA, it was previously determined that $G \cdots A \sim G \cdots T < G \cdots G < A \cdots T < G \cdots C$ (see ref 68). Among all mismatches (non-Watson-Crick pairs), $G \cdots T$ and $G \cdots G$ are two most stable ones (with $C \cdots C$ the least stable one). To further understand how LNA affects the mismatched pair stability, we also simulated duplexes $P_{LG} \cdot T_T$, $P_G \cdot T_T$, $P_{LG} \cdot T_G$, and $P_G \cdot T_G$ (Table 1). For the $G \cdots T$ mismatched pair, the simulation result shows that the duplex $P_{LG} \cdot T_T$ is more stable than the duplexes $P_G \cdot T_T$ with the 0.23 kcal/mol larger pairing interaction energy (Figure S4a). This result is consistent with the experiment findings (Figure 4a), confirming that LNA can stabilize this specific stable mismatched base-pair. Note that $G \cdots T$ can form a stable two hydrogen-bond pairing.⁶⁸ For the $G \cdots G$ mismatched base-pair, the simulation result (Figure S4b) further shows that the pairing energy for the duplex $P_{LG} \cdot T_G$ is very close to that for the duplex $P_G \cdot T_G$, while experimentally the duplex $P_{LG} \cdot T_G$ is less stable than $P_G \cdot T_G$ (Figure 4a). This seems contradictory to the fact that LNA might further stabilize stable base-pairs. However, the most stable mismatched $G \cdots G$ results from their stronger base stacking with flanking base-pairs and is not due to the base pairing.⁶⁸ Thus, the effect of LNA on the $G \cdots G$ mismatched pair might be dwarfed by the unusual base-stacking effect. Nevertheless, except for the $G \cdots G$ mismatched pair, both experiment and simulation showed the same overall trend that LNA can not only enhance stable base pairings (e.g., $G \cdots C$ and $G \cdots T$) but also destabilize less stable mismatched pairs (e.g., $G \cdots A$).

Discrimination of *E. coli* O157 Contaminated with Non-O157 Serotype DNAs

A single LNA can significantly enhance the SNP discrimination capability. But how accurate is the method when using the enhanced discrimination capability to discern a SNP? The accuracy can be evaluated using the receiver operating characteristic (ROC) curve analysis (Figure 6a), a diagnostic tool that measures the accuracy of a test to discriminate diseased cases from normal cases. We used *E. coli* O157 DNA as the basis to illustrate this analysis. Single molecule detection can generate a long unzipping-time population for the fully matched duplex (O157 serotype) and a short unzipping-time population for the mismatched duplex (non-O157). The two populations, which are exponentially distributed, have an overlay. For every cutoff time we select to discriminate between them, there will be many events in the long-time population correctly classified as the O157 events, that is, true positive fraction ($TP = 0-1$), but some events in the short-time population could be incorrectly classified as the pathogen events, that is, false positive fraction ($FP = 0-1$). The ROC curve is the plot of TP (sensitivity) against FP ($100\% - \text{selectivity}$) at every cutoff time. The areas under the ROC curve (AUC) ($0.5-1$) represent the detection accuracy. As shown in Figure 6a, the 10-fold unzipping time difference by using LNA yields an AUC above 0.9, indicating that, with the enhanced discrimination capability, the SNP detection accuracy is “excellent”. By comparison, the 1.6-fold unzipping time difference without using LNA only

yields an ROC of 0.6, meaning that the system is not adequately accurate for SNP discrimination. If AUC = 0.9 is set as “excellent performance”, then the LNA-enhanced discrimination capability for the three SNPs we studied would all be ranked as “excellent accuracy”, due to that the unzipping time differences for *E. coli* O157 (12.2-fold), EGFR L858C (22.2-fold), and KRAF G12D (10.7fold) are all larger than 10-fold (Figure 3b).

The single LNA capability in SNP discrimination allows for accurate detection of O157:H7 DNA contaminated with nonO157:H7 DNA. We mixed $P1_{LG} \cdot T1_C$ with $P1_{LG} \cdot T1_A$ at various percentages. The long and short components identified in the distributions of the signature event duration (Figure 2b–e) can be assigned to the $P1_{LG} \cdot T1_C$ (fully matched) and $P1_{LG} \cdot T1_A$ (single-mismatched) duplexes, respectively. Analysis indicates that as the $P1_{LG} \cdot T1_C$ percentage increases from 1% to 10%, 50%, and 90%, and the fractional population of the $P1_{LG} \cdot T1_C$ signature events linearly increased from $5.2 \pm 0.8\%$ to $12 \pm 4\%$, $48 \pm 3\%$, and $76 \pm 11\%$ (Figure 6). Therefore, we demonstrated that this approach is capable of detecting the *E. coli* O157:H7 serotype accurately even at a low percentage, without interference from non-O157 serotypes.

CONCLUSIONS

We have investigated a simple yet efficient single-LNA nanopore sensor for pathogenic SNP detection. First, SNPs are widely dispersed in the genome, unlike the specific genes, allowing for a multitarget detection to increase the accuracy and to enhance the universality of the method. Second, LNA performs excellently in the discrimination of the single-base difference. Compared with a pure DNA probe, the probe with a LNA can not only elongate the unzipping time of the fully matched duplex but also shorten the unzipping time of the single-mismatched duplex. Thus, LNA significantly magnifies the difference between the target and the corresponding nontarget sequences. The molecular mechanism for LNA effect on base-pair stability was investigated numerically by simulating DNA duplexes containing a range of base-pairs with different stabilities. Overall, for stable base-pairs (such as G–C and G–T), LNA may further enhance the binding interaction energies; for less stable ones (such as G–A), LNA further reduces the binding interaction energies. Third, the single LNA introduction into the probe, which is complementary to the SNP site, is enough for the discrimination of various SNPs in the nanopore. Previously, at least three consecutive LNA nucleotides (LNA triplet) are required to generate enough difference between fully matched and single-mismatched duplexes in melting temperature analysis.^{71,72} Xi et al. recently reported that a probe with total 22 LNA nucleotides can discriminate specific microRNA in a nanopore.⁷³ By contrast, we strikingly found that just the geometrical change by a single LNA in a LNA/DNA base-pair extensively affects the stability of the whole duplex and generates a significant difference in unzipping time. This simple design makes the method more applicable and far more effective. Lastly, the simultaneous detection of the pathogenic and nonpathogenic sequences can efficiently avoid false-positive or false-negative results.

From the inter-event duration (Figure S5a–h), we evaluated the event frequency for 100 nM DNA duplexes to be $0.33 \pm 0.13 \text{ s}^{-1}$. The detection efficiency can be enhanced, for example, by 10–100-fold under a salt gradient across the membrane, with greatly increased sensitivity.^{26,74} However, for real sample detection, the analytical procedure in the current stage could

be slow (low capture frequency) due to the low DNA concentrations. One solution is using the amplification/cleavage protocol that we have used for preparing the target sample from tissue DNA.⁴¹ The role of amplification in this protocol is only for increasing the DNA amount. But unlike PCR-based SNP detections, this amplification does not participate in the nanopore sensing. The following cleavage step is for obtaining the target fragment at a desired length between two endonuclease sites. For this purpose, the CRISPR technology can also be adapted for site-specific cleavage of long sequences. During the probe/target hybridization, the probe amount is 5–10-fold excessive over the target amount to maximize the duplex formation, thus increasing the detection efficiency. As LNA can discriminate SNPs better for longer sequences (Figure 4), we may be able to detect long targets (e.g., 40–50-nt) in the future. This is advantageous because the preparation of long targets no longer needs the cleavage. Furthermore, it is useful to utilize miniaturized biomimetic nanopore platforms^{29,75,76} and nanopore chips⁷⁷ that are available for the detection of tiny samples. The miniaturized devices allow using a very small volume (1 μ L) of solution. Therefore, the detection can be performed in pre-concentrated and enriched samples with a highly elevated efficiency of single molecule detection. Ultimately, it may be possible to realize amplification-free detection in real samples. Because practical detection, such as diagnostics, often requires simultaneous detection of multiple pathogens or mutations rather than just one, we can develop methods such as “nanopore barcoding” (which we developed for multiple²⁶ and interference-free⁷⁸ microRNA detection) for simultaneous detection of multiple SNP biomarkers.

In summary, the method provides a rapid and reliable tool to detect SNPs. We expect that the outstanding ability to discriminate a SNP can also expand the usage of this method to other fields, such as human genomic SNP and epigenetic single nucleotide mutations.

METHODS

Nucleic Acids.

The DNA probes were synthesized and purified by Integrated DNA Technologies. The DNA probe with LNA was synthesized and purified by Exiqon. Before nanopore testing, the two single strands in each group are mixed with salt solution. The final concentration in the stock solution was 100 μ M duplex, 1 M KCl, 10 mM Tris, pH 7.4. The mixtures were heated to 95 °C for 5 min and then gradually cooled to room temperature and stored at 4 °C.

Nanopore Formation and Electrical Recording.

Nanopore electrical recording was conducted according to previous reports.^{79,80} The lipid bilayer membrane was formed over a 100–150 μ m orifice in the center of the Teflon film that partitioned between cis and trans recording solutions. Both solutions contained 1 M KCl and were buffered with 10 mM Tris (pH 7.4). The α -hemolysin protein was synthesized by the protein gene-carrying plasmid (T7 promoter) using coupled *in vitro* transcription and translation (IVTT) (Promega). IVTT has been described previously.⁸¹ The nanopore protein was added in the cis solution, from which it was inserted into the bilayer to form a channel. The duplexes were released to the cis solution. A transmembrane voltage was applied from the trans solution with the cis side grounded through a pair of Ag/AgCl electrodes. The ionic

flow through the pore was recorded with an Axopatch 200B amplifier (Molecular Device Inc., Sunnyvale, CA), filtered with a built-in 4-pole low-pass Bessel Filter at 5 kHz, and acquired with Clampex 9.0 software (Molecular Device Inc.) through a Digidata 1440 A/D converter (Molecular Device Inc.) at a sampling rate of 20 kHz. Singlechannel event amplitude and duration were analyzed using Clampfit 9.0 (Molecular Device Inc.), Excel (Microsoft), and SigmaPlot (SPSS) software. The nanopore measurements were conducted at 22 ± 2 °C. Data were presented as means $\pm$ SD of at least three independent experiments.

Melting Temperature Measurement.

The melting temperatures were calculated by monitoring the increase in absorbance at 260 nm as a function of temperature. The temperature was increased from 22 to 95 °C at a rate of 0.5 °C/min.

MD simulation.

We carried out MD simulations for the dsDNA fragment (Table 1 for sequences) in a 1 M KCl electrolyte that contains 134 K⁺, 124 Cl⁻, and 6376 water molecules. The G at the third position is either the locked or the normal base and is paired with either the matched C or the mismatched A. We modeled four different DNA fragments containing: G–C, ^LG–C, G...A, and ^LG...A basepairs. Here, ^LG denotes the locked nucleotide G. We used the CHARMM force field for DNA, and the one for the LNA was adopted from a previous study.⁸² We used the TIP3P force field⁸³ for water and a standard force field⁸⁴ for ions. During the equilibration, the NPT (P = 1 bar and T = 300 K) ensemble was applied, with constrained DNA backbones. In production runs (with same NPT ensemble), all constraints were removed.

We used the software package NAMD2.⁹⁸⁵ for MD simulations. The Langevin dynamics was applied to all oxygen atoms in water to keep the temperature of the system to be constant. A smooth cutoff (10–12 Å) was utilized for calculating van der Waals interactions. Electrostatic interactions were calculated using the particle mesh Ewald (PME) method (grid size ~ 1 Å). The integration time-step in a simulation was 1 fs.

Empirical Free Energy Calculations.

The empirical relations⁸⁶ that include the enthalpy gain and the entropy loss can be applied to conveniently calculate the binding free energy of each duplex. The web tools at <http://biophysics.idtdna.com> are used to analyze four DNA fragments used in simulations. The sequences of 6-mer duplexes are provided in Table 1.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We are grateful to the National Institutes of Health grants GM114204 (L.-Q.G.) and HG009338 (K.S.G. and L.-Q.G.), USDA Multistate project S1056 (A.M.) and USDA NIFA Multistate project NC1194 (M.L.) for support of this work. B.L. gratefully acknowledges the financial support from the IBM Bluegene Science Program (grant nos: W1258591, W1464125, W1464164).

REFERENCES

- (1). Zorde Khvalevsky E; Gabai R; Rachmut IH; Horwitz E; Brunshwig Z; Orbach A; Shemi A; Golan T; Domb AJ; Yavin E; Giladi H; Rivkin L; Simerzin A; Eliakim R; Khalaileh A; Hubert A; Lahav M; Kopelman Y; Goldin E; Dancour A; et al. Mutant KRAS Is a Druggable Target for Pancreatic Cancer. *Proc. Natl. Acad. Sci. U. S. A* 2013, 110, 20723–20728. [PubMed: 24297898]
- (2). Halushka MK; Fan JB; Bentley K; Hsie L; Shen N; Weder A; Cooper R; Lipshutz R; Chakravarti A Patterns of SingleNucleotide Polymorphisms in Candidate Genes for Blood-Pressure Homeostasis. *Nat. Genet* 1999, 22, 239–247. [PubMed: 10391210]
- (3). Begovich AB; Carlton VE; Honigberg LA; Schrodi SJ; Chokkalingam AP; Alexander HC; Ardlie KG; Huang Q; Smith AM; Spoerke JM; Conn MT; Chang M; Chang SY; Saiki RK; Catanese JJ; Leong DU; Garcia VE; McAllister LB; Jeffery DA; Lee AT; et al. A Missense Single-Nucleotide Polymorphism in a Gene Encoding a Protein Tyrosine Phosphatase (PTPN22) Is Associated with Rheumatoid Arthritis. *Am. J. Hum. Genet* 2004, 75, 330–337. [PubMed: 15208781]
- (4). Bond GL; Hu W; Bond EE; Robins H; Lutzker SG; Arva NC; Bargonetti J; Bartel F; Taubert H; Wuerl P; Onel K; Yip L; Hwang SJ; Strong LC; Lozano G; Levine AJ A Single Nucleotide Polymorphism in the MDM2 Promoter Attenuates the P53 Tumor Suppressor Pathway and Accelerates Tumor Formation in Humans. *Cell* 2004, 119, 591–602. [PubMed: 15550242]
- (5). Filliol I; Motiwala AS; Cavatore M; Qi W; Hazbon MH; Bobadilla del Valle M; Fyfe J; Garcia-Garcia L; Rastogi N; Sola C; Zozio T; Guerrero MI; Leon CI; Crabtree J; Angiuoli S; Eisenach KD; Durmaz R; Joloba ML; Rendon A; SifuentesOsornio J; et al. Global Phylogeny of Mycobacterium Tuberculosis Based on Single Nucleotide Polymorphism (SNP) Analysis: Insights into Tuberculosis Evolution, Phylogenetic Accuracy of Other DNA Fingerprinting Systems, and Recommendations for a Minimal Standard SNP Set. *J. Bacteriol* 2006, 188, 759–772. [PubMed: 16385065]
- (6). Gopaul KK; Koylass MS; Smith CJ; Whatmore AM Rapid Identification of Brucella Isolates to the Species Level by Real Time PCR Based Single Nucleotide Polymorphism (SNP) Analysis. *BMC Microbiol* 2008, 8, 86. [PubMed: 18518958]
- (7). Gutacker MM; Mathema B; Soini H; Shashkina E; Kreiswirth BN; Graviss EA; Musser JM Single-Nucleotide Polymorphism-Based Population Genetic Analysis of Mycobacterium Tuberculosis Strains from 4 Geographic Sites. *J. Infect. Dis* 2006, 193, 121–128. [PubMed: 16323140]
- (8). Van Ert MN; Easterday WR; Simonson TS; U'Ren JM; Pearson T; Kenefic LJ; Busch JD; Huynh LY; Dukerich M; Trim CB; Beaudry J; Welty-Bernard A; Read T; Fraser CM; Ravel J; Keim P Strain-Specific Single-Nucleotide Polymorphism Assays for the Bacillus Anthracis Ames Strain. *J. Clin. Microbiol* 2007, 45, 47–53. [PubMed: 17093023]
- (9). Mhlanga MM; Malmberg L Using Molecular Beacons to Detect Single-Nucleotide Polymorphisms with Real-Time PCR. *Methods* 2001, 25, 463–471. [PubMed: 11846616]
- (10). Wang DG; Fan JB; Siao CJ; Berno A; Young P; Sapolsky R; Ghandour G; Perkins N; Winchester E; Spencer J; Kruglyak L; Stein L; Hsie L; Topaloglou T; Hubbell E; Robinson E; Mittmann M; Morris MS; Shen N; Kilburn D; et al. Large-Scale Identification, Mapping, and Genotyping of SingleNucleotide Polymorphisms in the Human Genome. *Science* 1998, 280, 1077–1082. [PubMed: 9582121]
- (11). Xu X; Hou Y; Yin X; Bao L; Tang A; Song L; Li F; Tsang S; Wu K; Wu H; He W; Zeng L; Xing M; Wu R; Jiang H; Liu X; Cao D; Guo G; Hu X; Gui Y; et al. Single-Cell Exome Sequencing Reveals Single-Nucleotide Mutation Characteristics of a Kidney Tumor. *Cell* 2012, 148, 886–895. [PubMed: 22385958]
- (12). Ahmadian A; Gharizadeh B; Gustafsson AC; Sterky F; Nyren P; Uhlen M; Lundeberg J Single-Nucleotide Polymorphism Analysis by Pyrosequencing. *Anal. Biochem* 2000, 280, 103–110. [PubMed: 10805527]
- (13). Petersen M; Bondensgaard K; Wengel J; Jacobsen JP Locked Nucleic Acid (LNA) Recognition of RNA: Nmr Solution Structures of LNA:RNA Hybrids. *J. Am. Chem. Soc* 2002, 124, 5974–5982. [PubMed: 12022830]

- (14). Simeonov A; Nikiforov TT Single Nucleotide Polymorphism Genotyping Using Short, Fluorescently Labeled Locked Nucleic Acid (LNA) Probes and Fluorescence Polarization Detection. *Nucleic Acids Res* 2002, 30, e91. [PubMed: 12202779]
- (15). Obernosterer G; Martinez J; Alenius M Locked Nucleic Acid-Based in situ Detection of Micrnas in Mouse Tissue Sections. *Nat. Protoc* 2007, 2, 1508–1514. [PubMed: 17571058]
- (16). Fiori ME; Barbini C; Haas TL; Marroncelli N; Patrizii M; Biffoni M; De Maria R Antitumor Effect of Mir-197 Targeting in P53 Wild-Type Lung Cancer. *Cell Death Differ* 2014, 21, 774–782. [PubMed: 24488097]
- (17). Jepsen JS; Wengel J LNA-Antisense Rivals SiRNA for Gene Silencing. *Curr. Opin. Drug Discovery Devel* 2004, 7, 188–194.
- (18). Vester B; Lundberg LB; Sørensen MD; Babu BR; Douthwaite S; Wengel J Lnazymes: Incorporation of LNA-Type Monomers into Dnazymes Markedly Increases RNA Cleavage. *J. Am. Chem. Soc* 2002, 124, 13682–13683. [PubMed: 12431091]
- (19). Cherf GM; Lieberman KR; Rashid H; Lam CE; Karplus K; Akeson M Automated Forward and Reverse Ratcheting of DNA in a Nanopore at 5-Å Precision. *Nat. Biotechnol* 2012, 30, 344–348. [PubMed: 22334048]
- (20). Kasianowicz JJ; Brandin E; Branton D; Deamer DW Characterization of Individual Polynucleotide Molecules Using a Membrane Channel. *Proc. Natl. Acad. Sci. U. S. A* 1996, 93, 13770–13773. [PubMed: 8943010]
- (21). Manrao EA; Derrington IM; Laszlo AH; Langford KW; Hopper MK; Gillgren N; Pavlenok M; Niederweis M; Gundlach JH Reading DNA at Single-Nucleotide Resolution with a Mutant MSPA Nanopore and Phi29 DNA Polymerase. *Nat. Biotechnol* 2012, 30, 349–353. [PubMed: 22446694]
- (22). An N; Fleming AM; Middleton EG; Burrows CJ SingleMolecule Investigation of G-Quadruplex Folds of the Human Telomere Sequence in a Protein Nanocavity. *Proc. Natl. Acad. Sci. U. S. A* 2014, 111, 14325–14331. [PubMed: 25225404]
- (23). Cao C; Ying YL; Hu ZL; Liao DF; Tian H; Long YT Discrimination of Oligonucleotides of Different Lengths with a WildType Aerolysin Nanopore. *Nat. Nanotechnol* 2016, 11, 713–718. [PubMed: 27111839]
- (24). An N; Fleming AM; White HS; Burrows CJ Crown Ether-Electrolyte Interactions Permit Nanopore Detection of Individual DNA Abasic Sites in Single Molecules. *Proc. Natl. Acad. Sci. U. S. A* 2012, 109, 11504–11509. [PubMed: 22711805]
- (25). Wallace EV; Stoddart D; Heron AJ; Mikhailova E; Maglia G; Donohoe TJ; Bayley H Identification of Epigenetic DNA Modifications with a Protein Nanopore. *Chem. Commun. (Cambridge, U. K.)* 2009, 46, 8195–8197.
- (26). Wang Y; Zheng D; Tan Q; Wang MX; Gu LQ Nanopore-Based Detection of Circulating Micrnas in Lung Cancer Patients. *Nat. Nanotechnol* 2011, 6, 668–674. [PubMed: 21892163]
- (27). Wang Y; Luan BQ; Yang Z; Zhang X; Ritzo B; Gates K; Gu LQ Single Molecule Investigation of Ag⁺ Interactions with Single Cytosine-, Methylcytosine- and Hydroxymethylcytosine-Cytosine Mismatches in a Nanopore. *Sci. Rep* 2015, 4, 5883.
- (28). Rosen CB; Rodriguez-Larrea D; Bayley H Single-Molecule Site-Specific Detection of Protein Phosphorylation with a Nanopore. *Nat. Biotechnol* 2014, 32, 179–181. [PubMed: 24441471]
- (29). Wolfe AJ; Mohammad MM; Cheley S; Bayley H; Movileanu L Catalyzing the Translocation of Polypeptides through Attractive Interactions. *J. Am. Chem. Soc* 2007, 129, 14034–14041. [PubMed: 17949000]
- (30). Wang S; Haque F; Rychahou PG; Evers BM; Guo P Engineered Nanopore of Phi29 DNA-Packaging Motor for Real-Time Detection of Single Colon Cancer Specific Antibody in Serum. *ACS Nano* 2013, 7, 9814–9822. [PubMed: 24152066]
- (31). Hornblower B; Coombs A; Whitaker RD; Kolomeisky A; Picone SJ; Meller A; Akeson M Single-Molecule Analysis of DNAProtein Complexes Using Nanopores. *Nat. Methods* 2007, 4, 315–317. [PubMed: 17339846]
- (32). Purnell RF; Schmidt JJ Discrimination of Single Base Substitutions in a DNA Strand Immobilized in a Biological Nanopore. *ACS Nano* 2009, 3, 2533–2538. [PubMed: 19694456]

- (33). Kong J; Zhu J; Keyser UF Single Molecule Based SNP Detection Using Designed DNA Carriers and Solid-State Nanopores. *Chem. Commun. (Cambridge, U. K.)* 2017, 53, 436–439.
- (34). Vercoutere W; Winters-Hilt S; Olsen H; Deamer D; Haussler D; Akeson M Rapid Discrimination among Individual DNA Hairpin Molecules at Single-Nucleotide Resolution Using an Ion Channel. *Nat. Biotechnol* 2001, 19, 248–252. [PubMed: 11231558]
- (35). Ang YS; Yung LY Rapid and Label-Free Single-Nucleotide Discrimination via an Integrative Nanoparticle-Nanopore Approach. *ACS Nano* 2012, 6, 8815–8823. [PubMed: 22994459]
- (36). Cornelis S; Gansemans Y; Deleye L; Deforce D; Van Nieuwerburgh F Forensic Snp Genotyping Using Nanopore Minion Sequencing. *Sci. Rep* 2017, 7, 41759. [PubMed: 28155888]
- (37). Zhao Q; Sigalov G; Dimitrov V; Dorvel B; Mirsaidov U; Sligar S; Aksimentiev A; Timp G Detecting Snps Using a Synthetic Nanopore. *Nano Lett* 2007, 7, 1680–1685. [PubMed: 17500578]
- (38). Torigoe H; Miyakawa Y; Kozasa T; Ono A The Specific Interaction between Two T:T Mismatch Base Pairs and Mercury (II) Cation. *Nucleic Acids Symp. Ser* 2007, 51, 185–186.
- (39). Schmidt K; Mwaigwisya S; Crossman LC; Doumith M; Munroe D; Pires C; Khan AM; Woodford N; Saunders NJ; Wain J; O'Grady J; Livermore DM Identification of Bacterial Pathogens and Antimicrobial Resistance Directly from Clinical Urines by Nanopore-Based Metagenomic Sequencing. *J. Antimicrob. Chemother* 2017, 72, 104–114. [PubMed: 27667325]
- (40). Greninger AL; Naccache SN; Federman S; Yu G; Mbala P; Bres V; Stryke D; Bouquet J; Somasekar S; Linnen JM; Dodd R; Mulembakani P; Schneider BS; Muyembe-Tamfum JJ; Stramer SL; Chiu CY Rapid Metagenomic Identification of Viral Pathogens in Clinical Samples by Real-Time Nanopore Sequencing Analysis. *Genome Med* 2015, 7, 99. [PubMed: 26416663]
- (41). Wang Y; Tian K; Shi R; Gu A; Pennella M; Alberts L; Gates KS; Li G; Fan H; Wang MX; Gu LQ NanolockNanopore Facilitated Digital Diagnostics of Cancer Driver Mutation in Tumor Tissue. *ACS Sensors* 2017, 2, 975–981. [PubMed: 28750524]
- (42). Zhang X; Price NE; Fang X; Yang Z; Gu LQ; Gates KS Characterization of Interstrand DNA-DNA Cross-Links Using the Alpha-Hemolysin Protein Nanopore. *ACS Nano* 2015, 9, 11812–11819. [PubMed: 26563913]
- (43). Nejad MI; Shi R; Zhang X; Gu LQ; Gates KS Sequence-Specific Covalent Capture Coupled with High-Contrast Nanopore Detection of a Disease-Derived Nucleic Acid Sequence. *ChemBioChem* 2017, 18, 1383–1386. [PubMed: 28422400]
- (44). Hirvonen JJ; Siitonen A; Kaukoranta SS Usability and Performance of Chromagar STEC Medium in Detection of Shiga Toxin-Producing Escherichia Coli Strains. *J. Clin. Microbiol* 2012, 50, 3586–3590. [PubMed: 22933601]
- (45). Feng P Impact of Molecular Biology on the Detection of Foodborne Pathogens. *Mol. Biotechnol* 1997, 7, 267–278. [PubMed: 9219240]
- (46). Bai J; Shi X; Nagaraja TG A Multiplex PCR Procedure for the Detection of Six Major Virulence Genes in Escherichia Coli O157:H7. *J. Microbiol. Methods* 2010, 82, 85–89. [PubMed: 20472005]
- (47). Fratomico PM; DebRoy C Detection of Escherichia Coli O157:H7 in Food Using Real-Time Multiplex Pcr Assays Targeting the stx1, stx2, wzyO157, and the fliCH7 or eae Genes. *Food Analyt. Methods* 2010, 3, 330–337.
- (48). Feng P; Lampel KA Genetic Analysis of Uida Expression in Enterohaemorrhagic Escherichia Coli Serotype O157:H7. *Microbiology* 1994, 140, 2101–2107. [PubMed: 7921260]
- (49). Singh P; Mustapha A Multiplex Real-Time Pcr Assays for Detection of Eight Shiga Toxin-Producing Escherichia Coli in Food Samples by Melting Curve Analysis. *Int. J. Food Microbiol* 2015, 215, 101–108. [PubMed: 26355415]
- (50). Wang L; Li Y; Mustapha A Detection of Viable Escherichia Coli O157:H7 by Ethidium Monoazide Real-Time PCR. *J. Appl. Microbiol* 2009, 107, 1719–1728. [PubMed: 19457030]
- (51). Kaynak A; Sakalar E A Rapid and Cost-Efficient Technique for Simultaneous/Duplex Detection of Listeria Monocytogenes and Escherichia Coli O157:H7 Using Real Time PCR. *J. Food Saf* 2016, 36, 375–382.
- (52). Liu Y; Mustapha A Detection of Viable Escherichia Coli O157:H7 in Ground Beef by Propidium Monoazide Real-Time PCR. *Int. J. Food Microbiol* 2014, 170, 48–54. [PubMed: 24291180]

- (53). Anklam KS; Kanankege KS; Gonzales TK; Kaspar CW; Dopfer D Rapid and Reliable Detection of Shiga Toxin-Producing Escherichia Coli by Real-Time Multiplex PCR. *J. Food Prot* 2012, 75, 643–650. [PubMed: 22488051]
- (54). Wang LX; Li Y; Mustapha A Rapid and Simultaneous Quantitation of Escherichia Coli O157: H7, Salmonella, and Shigella in Ground Beef by Multiplex Real-Time PCR and Immunomagnetic Separation. *J. Food Prot* 2007, 70, 1366–1372. [PubMed: 17612065]
- (55). Herbst RS Review of Epidermal Growth Factor Receptor Biology. *Int. J. Radiat. Oncol., Biol., Phys* 2004, 59, S21–S26.
- (56). Gazdar AF Activating and Resistance Mutations of Egfr in Non-Small-Cell Lung Cancer: Role in Clinical Response to EGFR Tyrosine Kinase Inhibitors. *Oncogene* 2009, 28, S24–S31. [PubMed: 19680293]
- (57). Lynch TJ; Bell DW; Sordella R; Gurubhagavatula S; Okimoto RA; Brannigan BW; Harris PL; Haserlat SM; Supko JG; Haluska FG; Louis DN; Christiani DC; Settleman J; Haber DA Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non-Small-Cell Lung Cancer to Gefitinib. *N. Engl. J. Med* 2004, 350, 2129–2139. [PubMed: 15118073]
- (58). Eser S; Schnieke A; Schneider G; Saur D Oncogenic KRAS Signalling in Pancreatic Cancer. *Br. J. Cancer* 2014, 111, 817–822. [PubMed: 24755884]
- (59). Engelman JA; Chen L; Tan XH; Crosby K; Guimaraes AR; Upadhyay R; Maira M; McNamara K; Perera SA; Song YC; Chirieac LR; Kaur R; Lightbown A; Simendinger J; Li T; Padera RF; Garcia-Echeverria C; Weissleder R; Mahmood U; Cantley LC; et al. Effective Use of Pi3k and Mek Inhibitors to Treat Mutant KRAS G12d and Pik3ca H1047r Murine Lung Cancers. *Nat. Med* 2008, 14, 1351–1356. [PubMed: 19029981]
- (60). Hingorani SR; Wang L; Multani AS; Combs C; Deramaudt TB; Hruban RH; Rustgi AK; Chang S; Tuveson DA Trp53r172h and KrasG12d Cooperate to Promote Chromosomal Instability and Widely Metastatic Pancreatic Ductal Adenocarcinoma in Mice. *Cancer Cell* 2005, 7, 469–483. [PubMed: 15894267]
- (61). Ling J; Kang Y; Zhao R; Xia Q; Lee DF; Chang Z; Li J; Peng B; Fleming JB; Wang H; Liu J; Lemischka IR; Hung MC; Chiao PJ KrasG12d-Induced Ikk2/Beta/Nf-Kappab Activation by Il-1alpha and P62 Feedforward Loops Is Required for Development of Pancreatic Ductal Adenocarcinoma. *Cancer Cell* 2012, 21, 105–120. [PubMed: 22264792]
- (62). Hingorani SR; Petricoin EF; Maitra A; Rajapakse V; King C; Jacobetz MA; Ross S; Conrads TP; Veenstra TD; Hitt BA; Kawaguchi Y; Johann D; Liotta LA; Crawford HC; Putt ME; Jacks T; Wright CV; Hruban RH; Lowy AM; Tuveson DA Preinvasive and Invasive Ductal Pancreatic Cancer and Its Early Detection in the Mouse. *Cancer Cell* 2003, 4, 437–450. [PubMed: 14706336]
- (63). Lievre A; Bachet J-B; Le Corre D; Boige V; Landi B; Emile J-F; Côte J-F; Tomasic G; Penna C; Ducreux M; Rougier P; Penault-Llorca F; Laurent-Puig P KRAS Mutation Status Is Predictive of Response to Cetuximab Therapy in Colorectal Cancer. *Cancer Res* 2006, 66, 3992.
- (64). Roth AD; Tejpar S; Delorenzi M; Yan P; Fiocca R; Klingbiel D; Dietrich D; Biesmans B; Bodoky G; Barone C; Aranda E; Nordlinger B; Cisar L; Labianca R; Cunningham D; Van Cutsem E; Bosman F Prognostic Role of KRAS and BRAF in Stage II and III Resected Colon Cancer: Results of the Translational Study on the Petacc-3, Eortc 40993, Sakk 60-00 Trial. *J. Clin. Oncol* 2010, 28, 466–474. [PubMed: 20008640]
- (65). Zhang X; Xu X; Yang Z; Burcke AJ; Gates KS; Chen SJ; Gu LQ Mimicking Ribosomal Unfolding of RNA Pseudoknot in a Protein Channel. *J. Am. Chem. Soc* 2015, 137, 15742–15752. [PubMed: 26595106]
- (66). Zhang X; Zhang D; Zhao C; Tian K; Shi R; Du X; Burcke AJ; Wang J; Chen SJ; Gu LQ Nanopore Electric Snapshots of an RNA Tertiary Folding Pathway. *Nat. Commun* 2017, 8, 1458. [PubMed: 29133841]
- (67). Wang Y; Tian K; Hunter LL; Ritzo B; Gu LQ Probing Molecular Pathways for DNA Orientational Trapping, Unzipping and Translocation in Nanopores by Using a Tunable Overhang Sensor. *Nanoscale* 2014, 6, 11372–11379. [PubMed: 25144935]

- (68). Granzhan A; Kotera N; Teulade-Fichou MP Finding Needles in a Basestack: Recognition of Mismatched Base Pairs in DNA by Small Molecules. *Chem. Soc. Rev* 2014, 43, 3630–3665. [PubMed: 24634921]
- (69). Ben-Shem A; Jenner L; Yusupova G; Yusupov M Crystal Structure of the Eukaryotic Ribosome. *Science* 2010, 330, 1203–1209. [PubMed: 21109664]
- (70). Yusupov MM; Yusupova GZ; Baucom A; Lieberman K; Earnest TN; Cate JH; Noller HF Crystal Structure of the Ribosome at 5.5 Å Resolution. *Science* 2001, 292, 883–896. [PubMed: 11283358]
- (71). You Y; Moreira BG; Behlke MA; Owczarzy R Design of LNA Probes That Improve Mismatch Discrimination. *Nucleic Acids Res* 2006, 34, e60. [PubMed: 16670427]
- (72). Owczarzy R; You Y; Groth CL; Tataurov AV Stability and Mismatch Discrimination of Locked Nucleic Acid-DNA Duplexes. *Biochemistry* 2011, 50, 9352–9367. [PubMed: 21928795]
- (73). Xi D; Shang J; Fan E; You J; Zhang S; Wang H Nanopore-Based Selective Discrimination of Micrnas with SingleNucleotide Difference Using Locked Nucleic Acid-Modified Probes. *Anal. Chem* 2016, 88, 10540–10546. [PubMed: 27734673]
- (74). Wanunu M; Morrison W; Rabin Y; Grosberg AY; Meller A Electrostatic Focusing of Unlabelled DNA into Nanoscale Pores Using a Salt Gradient. *Nat. Nanotechnol* 2010, 5, 160–165. [PubMed: 20023645]
- (75). Czekalska MA; Kaminski TS; Jakiela S; Tanuj Sapra K; Bayley H; Garstecki P A Droplet Microfluidic System for Sequential Generation of Lipid Bilayers and Transmembrane Electrical Recordings. *Lab Chip* 2015, 15, 541–548. [PubMed: 25412368]
- (76). White RJ; Ervin EN; Yang T; Chen X; Daniel S; Cremer PS; White HS Single Ion-Channel Recordings Using Glass Nanopore Membranes. *J. Am. Chem. Soc* 2007, 129, 11766–11775. [PubMed: 17784758]
- (77). Jain M; Olsen HE; Paten B; Akeson M The Oxford Nanopore Minion: Delivery of Nanopore Sequencing to the Genomics Community. *Genome Biol* 2016, 17, 239. [PubMed: 27887629]
- (78). Tian K; Decker K; Aksimentiev A; Gu LQ InterferenceFree Detection of Genetic Biomarkers Using Synthetic DipoleFacilitated Nanopore Dielectrophoresis. *ACS Nano* 2017, 11, 1204–1213. [PubMed: 28036167]
- (79). Wang Y; Zheng D; Tan Q; Wang MX; Gu L-Q Nanopore-Based Detection of Circulating Micrnas in Lung Cancer Patients. *Nat. Nanotechnol* 2011, 6, 668–674. [PubMed: 21892163]
- (80). Shim JW; Tan Q; Gu LQ Single-Molecule Detection of Folding and Unfolding of the G-Quadruplex Aptamer in a Nanopore Nanocavity. *Nucleic Acids Res* 2009, 37, 972–982. [PubMed: 19112078]
- (81). Shim JW; Yang M; Gu LQ In Vitro Synthesis, Tetramerization and Single Channel Characterization of VirusEncoded Potassium Channel Kcv. *FEBS Lett* 2007, 581, 1027–1034. [PubMed: 17316630]
- (82). Pande V; Nilsson L Insights into Structure, Dynamics and Hydration of Locked Nucleic Acid (LNA) Strand-Based Duplexes from Molecular Dynamics Simulations. *Nucleic Acids Res* 2008, 36, 1508–1516. [PubMed: 18203740]
- (83). Jorgensen WL; Chandrasekhar J; Madura JD; Impey RW; Klein ML Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys* 1983, 79, 926–935.
- (84). Beglov D; Roux B Finite Representation of an Infinite Bulk System: Solvent Boundary Potential for Computer Simulations. *J. Chem. Phys* 1994, 100, 9050–9063.
- (85). Phillips JC; Braun R; Wang W; Gumbart J; Tajkhorshid E; Villa E; Chipot C; Skeel RD; Kale L; Schulten K Scalable Molecular Dynamics with Namd. *J. Comput. Chem* 2005, 26, 1781–1802. [PubMed: 16222654]
- (86). Owczarzy R; You Y; Groth CL; Tataurov AV Stability and Mismatch Discrimination of Locked Nucleic Acid-DNA Duplexes. *Biochemistry* 2011, 50, 9352–9367. [PubMed: 21928795]

Figure 1.

Single-LNA-nanopore sensor for enhanced SNP discrimination. (a) Structures of locked and regular nucleotides in a nucleic acid sequence. (b) LNA-enhanced SNP discrimination by detecting the unzipping time difference in the nanopore between fully matched probe-target duplex for pathogen DNA and one-mismatched duplex for the nonpathogen DNA. The details of the nanopore setup and method are described in Methods.

Figure 2.

Current traces showing single LNA-enhanced single-molecule discrimination of *E. coli* O157 and non-O157 serotype DNAs in the nanopore. (a–d) Representative sequential series of nanopore current blockade (left) generated by the P1_G·T1_C (a), P1_G·T1_A (b), P1_L_G·T1_C (c), and P1_L_G·T1_A (d) DNA duplexes and corresponding histograms of unzipping time (blockade duration, right). The fold of increase in the unzipping time between fully matched and single-mismatched duplexes is marked. Each probe-target duplex was loaded in 1 M KCl solution to 100 nM on the *cis* side of the α -hemolysin protein pore. Current traces were recorded at +120 mV.

Figure 3.

Single LNA-enhanced SNP discrimination for detection of *E. coli* O157 pathogenic serotype (T1_C vs T1_A), *EGFR* L858C (T6_C vs T6_A), and *KRAS* G12D (T7_C vs T7_A) cancer driver mutations. Sequences of all probes and targets are provided in Table 1. (a) Unzipping times (τ_{uz}) for fully matched and mismatched duplexes, by using DNA and LNA probes. (b) SNP discrimination capability for the DNA and LNA probes, calculated as the unzipping time ratio of between fully matched versus one mismatched probe-target duplexes ($\tau_{uz-FM} / \tau_{uz-MM}$). (c) Enhancement magnitude of SNP discrimination by using the LNA probe, which is the fold of increase in SNP discrimination capability. (d) Fold of increase in the unzipping time for fully matched duplexes ($\tau_{uz-FM} (LNA) / \tau_{uz-FM} (DNA)$) and fold of decrease in the unzipping time for mismatched duplexes ($\tau_{uz-MM} (DNA) / \tau_{uz-MM} (LNA)$) by using a LNA probe. The experiment conditions were the same as those in Figure 2. Histograms for obtaining τ_{uz} and inter-event interval (τ_{on}) are shown in Figure S5a,g,h.

Figure 4. Regulation of LNA-enhanced SNP discrimination capability by sequence factors. Sequence factors include the mismatch nucleotide at the SNP site ($T1_G$ and $T1_T$), LNA's neighbor nucleotides ($T2_C$ vs $T2_A$ and $T3_C$ vs $T3_A$), LNA position ($T4_C$ vs $T4_A$), and target length ($T5_C$ vs $T5_A$). Sequences of all probes and targets are listed in Table 1. (a) Unzipping times (τ_{uz}) for fully matched and mismatched duplexes, by using DNA and LNA probes. (b) SNP discrimination capability for the DNA and LNA probes, calculated as the unzipping time ratio of between the fully matched versus one-mismatched probe-target duplexes ($\tau_{uz-FM}/\tau_{uz-MM}$). (c) Enhancement magnitude of SNP discrimination by using the LNA probe, which is the fold of increase in SNP discrimination capability. (d) Fold of increase in the unzipping time for fully matched duplexes ($\tau_{uz-FM}(\text{LNA})/\tau_{uz-FM}(\text{DNA})$) and fold of decrease in the unzipping time for mismatched duplexes ($\tau_{uz-MM}(\text{DNA})/\tau_{uz-MM}(\text{LNA})$) by using a LNA probe. The experiment conditions were the same as those in Figure 3, except for the long target ($T5_C$ vs $T5_A$, 23-nt) which was recorded at 150 mV. Histograms for obtaining τ_{uz} and inter-event interval (τ_{on}) are shown in Figure S5b–f.

Figure 5.

MD simulations of LNA/DNA. (a) The simulation system. A dsDNA fragment (with the sequence GAGCAG in one strand) from E.coli O157 uidA gene is solvated in a 1 M KCl electrolyte. Water is shown transparently; K^+ and Cl^- are shown as tan and cyan spheres. (b and c) A close view of the $^L\text{G}-\text{C}$ (b) and $^L\text{G}\cdots\text{A}$ (c) base-pairs. (d and e) Time-dependent cumulative average of pairing energies for the $^L\text{G}-\text{C}$ and $\text{G}-\text{C}$ base-pairs (d) and pairing energies for the $^L\text{G}\cdots\text{A}$ and the $\text{G}\cdots\text{A}$ base-pairs (e) in their respective duplex.

Figure 6.

Accuracy in discrimination of *E. coli* O157 in the presence of non-O157 serotype DNAs. (a) ROC analysis of accuracy in SNP discrimination. AUC measures the accuracy. ROC was analyzed using web-based calculator (<http://www.rad.jhmi.edu/jeng/javarad/roc/JROCFITi.html>). Red solid circles are the experimental ROC curve with AUC = 0.91 for separating $P1_{LG} \cdot T1_C$ ($n = 683$) and $P1_{LG} \cdot T1_A$ duplexes ($n = 792$). Red open circles are the simulated ROC curve with AUC = 0.93 for separating $P1_{LG} \cdot T1_C$ ($n = 200$) and $P1_{LG} \cdot T1_A$ duplexes ($n = 200$) with $\tau_{uz-FM}/\tau_{uz-MM} = 10$. Blue solid circles are experimental ROC curve with AUC = 0.61 for separating $P1_G \cdot P1_C$ ($n = 587$) and $P1_G \cdot P1_A$ ($n = 641$). Blue open circles are simulated ROC curve separating $P1_G \cdot P1_C$ ($n = 200$) and $P1_G \cdot P1_A$ ($n = 200$) with $\tau_{uz-FM}/\tau_{uz-MM} = 1.6$. Black open circles are simulated reference ROC curve with AUC = 0.5 for separating $P1_G \cdot P1_C$ ($n = 200$) and $P1_G \cdot P1_A$ ($n = 200$) with $\tau_{uz-FM}/\tau_{uz-MM} = 1$. The separation performance is “perfect” for AUC = 1.0, “excellent” for AUC = 0.9–0.99, “good” for AUC = 0.8–0.89, “fair” for AUC = 0.7–0.79, “poor” for AUC = 0.51–0.69, and “worthless” for AUC = 0.5. Accordingly, the accuracy for discriminating the O157 from non-O157 serotype DNAs by using the single-LNA probe (AUC = 0.91) is “excellent”, whereas that by using the DNA probe is “poor” (AUC = 0.61). (b) Quantitative detection of *E. coli* O157 in the presence of non-O157 serotype DNAs. The O157 DNA fractional population was measured at various original O157 DNA percentages from 1% to 90%. Inset is a representative histogram of blockade duration for the mixture of the $P1_{LG} \cdot T1_C$ (O157:H7) and $P1_{LG} \cdot T1_A$ (non-O157:H7) duplexes at $P1_{LG} \cdot T1_C$ molecular percentages of 50%. The duration distribution in each histogram was fit into two components by the exponential function in log probability. Histograms and fitting for all percentages are shown in Figure S6.

Table 1.

Sequences of Probe and Target DNAs

No	Name ^a	Probe or target	Sequence ^b
1	P1 _G	O157 <i>uidA</i> G-probe	5'-GGAATTGA.GCAGCGTTGCCCCCCCCCCCCCCC-3'
	P1 _{LG}	O157 <i>uidA</i> ^L G-probe	5'-GGAATTGA ^L GCAGCGTTGCCCCCCCCCCCCCCC-3'
	T1 _C	O157 <i>uidA</i> target	3'-CCTTAACT CGTCGCAAC-5'
	T1 _A	Non-O157 <i>uidA</i> target	3'-CCTTAACT AGTCGCAAC-5'
	T1 _G	Non-O157 <i>uidA</i> target	3'-CCTTAACT GGTCGCAAC-5'
	T1 _T	Non-O157 <i>uidA</i> target	3'-CCTTAACT TGTCGCAAC-5'
2	P2 _G	O157 <i>uidA</i> G-probe	5'-GGATTGA GAAGCGTTGCCCCCCCCCCCCCCC-3'
	P2 _{LG}	O157 <i>uidA</i> LG-probe	5'-GGATTGA ^L GAAGCGTTGCCCCCCCCCCCCCCC-3'
	T2 _C	O157 <i>uidA</i> target	3'-CCTTAACT CTTGCAAC-5'
	T2 _A	Non-O157 <i>uidA</i> target	3'-CCTTAACT ATTGCAAC-5'
3	P3 _G	O157 <i>uidA</i> G-probe	5'-GGAATTGC GCGGCGTTGCCCCCCCCCCCCCCC-3'
	P3 _{LG}	O157 <i>uidA</i> LG-probe	5'-GGAATTGC ^L GC GCGGCGTTGCCCCCCCCCCCCCCC-3'
	T3 _C	O157 <i>uidA</i> target	3'-CCTTAACG CTTGCAAC-5'
	T3 _A	Non-O157 <i>uidA</i> target	3'-CCTTAACG ATTGCAAC-5'
4	P4 _G	O157 <i>uidA</i> G-probe	5'-GA GCGGCGTTGCCCCCCCCCCCCCCC-3'
	P4 _{LG}	O157 <i>uidA</i> LG-probe	5'-GA ^L GC GCGGCGTTGCCCCCCCCCCCCCCC-3'
	T4 _C	O157 <i>uidA</i> target	3'-CT CTTGCAAC-5'
	T4 _A	Non-O157 <i>uidA</i> target	3'-CT ATTGCAAC-5'
5	P5 _G	O157 <i>uidA</i> G-probe	5'-TGTGGAATTGA GCAGCGTTGGTGCCCCCCCCCCCCCCC-3'
	P5 _{LG}	O157 <i>uidA</i> LG-probe	5'-TGTGGAATTGA ^L GCAGCGTTGGTGCCCCCCCCCCCCCCC-3'
	T5 _C	O157 <i>uidA</i> target	3'-ACACCTTAACT CGTCGCAACCAC-5'
	T5 _A	Non-O157 <i>uidA</i> target	3'-ACACCTTAACT AGTCGCAACCAC-5'
6	P6 _G	<i>EGFR</i> L585R G-probe	5'-TTTGGGC GGGCCAAACCCCCCCCCCCC
	P6 _{LG}	<i>EGFR</i> L585R ^L G-Probe	5'-TTTGGGC ^L GGGCCAAACCCCCCCCCCCC
	T6 _C	<i>EGFR</i> L585R target	3'-AAAACCCG CCCGGTTG-5'
	T6 _A	<i>EGFR</i> wild-type target	3'-AAAACCCG ACCGGTTG-5'
7	P7 _A	<i>KRAS</i> G12D A-probe	5'-TGGAGTG ATGGCGTAGCCCCCCCCCCC
	P7 _{LA}	<i>KRAS</i> G12D ^L A-probe	5'-TGGAGTG ^L ATGGCGTAGCCCCCCCCCCC
	T7 _T	<i>KRAS</i> G12D target	3'-ACCTCGAC TACGCATC-5'
	T7 _C	<i>KRAS</i> wild-type target	3'-ACCTCGAC CACCGCATC-5'
8	P _G	Simulation G-probe	5'-GA GCAG-3'
	P _{LG}	Simulation ^L G-probe	5'-GA ^L GCAG-3'
	T _C	Simulation target	3'-CT CGTC-5'

No	Name ^a	Probe or target	Sequence ^b
	T _A	Simulation target	3'-CT AGTC-5'
	T _G	Simulation target	3'-CT GGTC-5'
	T _T	Simulation target	3'-CT TGTC-5'

^aThe DNA probe and LNA probe form a G/LG–C pair with the *E. coli* O157 uidA-based targets T1C, T2C, T3C, T4C, and T5C and a G/LG...A, G/LG...G, or G/LG...T mismatch with non-O157 T1A, T1G, T1T, T2A, T3A, T4A, and T5A at the +93 SNP site. The DNA probe and LNA probe form a G/LG–C pair with the EGFR L585R target T6C and a G/LG...A mismatch with the EGFR wild-type target T6A. The DNA probe and LNA probe form a A/LA–T pair with the KRAS G12D target T7T and A/LA...C mismatch with the KRAS wildtype target T7C. ^LG and ^LA in the probes denote locked guanosine and adenosine.

^bLong sequences of *E. coli* uidA and human EGFR and KRAS genes containing target SNPs are shown in Table S1.

Table 2.Structure Parameters for G–C, ^LG–C, G...A, and ^LG...A Base-Pairs in Simulated DNA Fragments

base-pair	propeller (deg)	buckle (deg)	P–P (Å)
G–C	12.5	11.3	19.4
^L G–C	18.0	12.5	19.7
G...A	17.1	8.9	19.2
^L G...A	28.1	15.7	19.8