THE JOURNAL OF PHYSICAL CHEMISTRY B
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCB

Article

Light–Nucleotide versus Ion–Nucleotide Interactions for Single-Nucleotide Resolution

Published as part of The Journal of Physical Chemistry virtual special issue "Dave Thirumalai Festschrift".

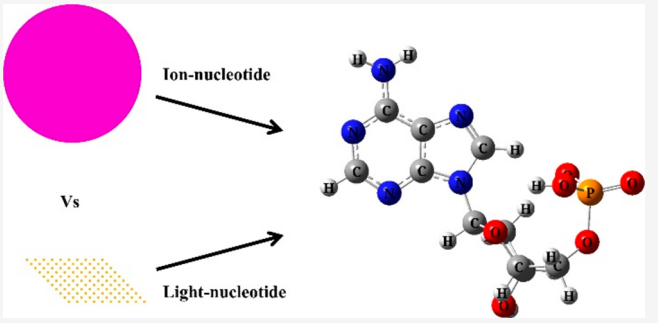
Mohsen Farshad and Jayendran C. Rasaiah*

Cite This: https://dx.doi.org/10.1021/acs.jpcb.0c10759

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Several parallel reads of ionic currents through multiple CsgG nanopores provide information about ion–nucleotide interactions for sequencing single-stranded DNA (ssDNA) using base-calling algorithms. However, the information in ion–nucleotide interactions seems insufficient for single-read nanopore DNA sequencing. Here we report discriminative light–nucleotide interactions calculated from density functional theory (DFT), which are compared with ionic currents obtained from molecular dynamics (MD) simulations. The MD simulations were performed on a system containing a transverse nanochannel and a longitudinal solid state nanopore. We show that both of the transverse and longitudinal ionic currents during the translocation of A_{16}, G_{16}, T_{16}, and C_{16} through the nanopore, overlapped widely. On the other hand, the UV–vis and Raman spectra of different types of single nucleotides, nucleosides, and nucleobases show relatively higher resolution than the ionic currents. Light–nucleotide interactions provide better information for characterizing the nucleotides in comparison to ion–nucleotide interactions for nanopore DNA sequencing. This can be realized by using optical techniques including surface-enhanced Raman spectroscopy (SERS) or tip-enhanced Raman spectroscopy (TERS), while plasmon excitation can be used to localize light and control the rate of nucleotide flow.



INTRODUCTION

Recent progress in nanopore DNA sequencing is promising in detecting adenine (A), guanine (G), thymine (T), and cytosine (C) by analyzing the ionic current patterns generated by the movement of ions during translocation of single-stranded DNA (ssDNA) through a nanopore under an applied electric field.[1−15] In nanopore DNA sequencing, the nucleotides in single-stranded DNA (ssDNA) are recognized by a sensor probing a characteristic property inside the nanopore that is affected by the presence of nucleotides.[1] The modulation of this property can reflect the characteristics of nucleotides of different types. Since 1996,[2] the potential use of ionic current sensors to distinguish between nucleotide types in nanopore DNA sequencing has been extensively explored.[1,16−19] Electrodes embedded in two chambers interconnected by a nanopore across a membrane generate ionic currents through the membrane under an applied electric field in an ionic aqueous solution, e.g., 1 M KCl.[1]

The negatively charged nucleotides enter the nanopore and are then driven through it electrophoretically under an applied electric field.[1] Motor proteins, e.g., phi29 polymerase and helicase, promote consecutive translocation of the nucleotide at a slow rate lower than 1 base per millisecond (b/ms) through the nanopore.[20,21] These two factors facilitate the gain of adequate informative measurements of ionic currents as several ions interact with a minimum number of nucleotides for a relatively long time within the nanopore.[20−25] As a result, more ions interact with each nucleotide to provide a significant amount of information for several reads during sequencing in the current Oxford nanopore technology. However, the ionic current modulations caused by the presence of nucleotides inside the nanopores are not entirely distinguishable even with advanced technologies.[1,15] Considering the length of the nanopores, e.g., MspA and CsgG,[26] there always will be a few nucleotides within the pore interacting with ions present during the translocation of ssDNA through the pore under an applied field. In this regard, it is difficult to attribute the ionic currents to individual nucleotides because the signals of nucleotides overlap with each other. Also, the motor enzyme

Received: December 1, 2020
Revised: February 10, 2021

sometimes malfunctions in pushing the DNA through the pore consecutively.[21,27]

Because of the interaction of a relatively small group of ions with each nucleotide within the nanopore, ion−nucleotide signals are hidden in random thermal Brownian motions of ions.[1] Furthermore, the number of ions as charge carriers in the nanopore is not constant and varies in time. This fluctuation in the number of ions leads to further fluctuation in the ionic current through the nanopore.[1,28] The signal-to-noise ratio increases by ensemble measurements averaging over a relatively long time as reversible stochastic thermal fluctuations cancel each other, and the ions move in the direction of the applied external field in a nonequilibrium system. On the other hand, slow structural dynamics of DNA within the nanopore do not provide an adequate amount of sampling over nucleotides conformations during ionic currents measurements, which is believed to lead to fluctuation of ionic currents.[29]

If these limitations are overcome, there would still be another fundamental challenge with conventional nanopore DNA sequencing by using ionic currents. We intuitively think that nanopore DNA sequencing via ionic current measurements is difficult, principally because of the relatively large size of ions as the entities that detect the nucleotides. In this regard, we aim to investigate the interaction of nucleotides with light as a smaller entity and compare the resolution of light−nucleotide interactions with ion−nucleotide interactions to distinguish between different nucleotides. To compare the ion−nucleotide with light−nucleotide interactions, we first performed all-atom molecular dynamics (MD) simulations on a solid-state nanopore model.[30−34] The longitudinal and transverse ionic currents are measured through the nanopore and nanochannel, respectively, during translocation of homo-oligonucleotides. Figure 1a is an image from a MD simulation during voltage-driven translocation of A_{16} through the nanopore. Figure 1b is a schematical representation of the nanochannel with a transversely applied electric field and a nanopore with a longitudinally applied electric field. The ionic currents obtained from the MD simulations of four types of homo-oligonucleotides overlap widely. The averaged longitudinal and transverse ionic currents were partially distinguishable, and their combination showed enhancement in discriminating between the homo-oligonucleotides. The accuracy of classifying homo-oligonucleotides with averaged ionic currents by using a neural network algorithm was 60.6%. The MD simulations indicated that single-read identification of the homo-oligonucleotides was not effective in our system in the given time scale. However, the optical spectra and Raman spectra calculated by using density functional theory (DFT) showed significant discrimination between single nucleotides in comparison to raw ionic currents calculated from MD simulations. We calculated the optical properties (UV and visible) of single nucleotides in the gas and aqueous phase (implicit water model) using time-dependent DFT (TDDFT). We further calculated the Raman and IR spectra of nucleotides in the gas phase using DFT. Since optical devices can be integrated with solid-state nanopores,[35−37] the integration of a spectrometer, especially as in enhanced Raman spectroscopy, in the presence of localized surface plasmon resonance of nanostructures[38−43] or near-field spectroscopy[44]—with a scanning tip near nucleotides associated with our solid-state model can potentially enhance single-read nanopore DNA sequencing techniques, as observed by Belkin et al.[27]

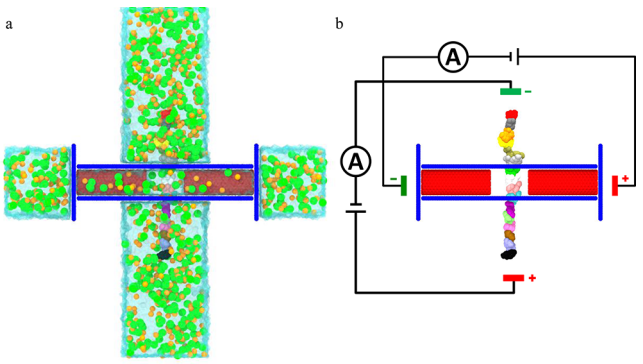


Figure 1. Schematic representation of the MD model for nanopore DNA sequencing via ionic currents. The longitudinal nanopore and transverse nanochannel are vertical and horizontal, respectively. (a) An image of the system during MD simulation. The cyan color reservoirs represent water. The green and orange spheres represent Cl^- and K^+ ions, respectively. The graphene nanosheets and CNTs are represented with blue and red colors, respectively. (b) Schematic illustration of the system with a transversely and longitudinally applied electric field of 0.5 V through the nanochannel and nanopore, respectively.

METHODS

Molecular Dynamics (MD) Simulations. We used the 3D-DART website to build homo-oligonucleotide dsDNA sequences of A_{16}, G_{16}, T_{16}, and C_{16}.[45] The dsDNA structures were converted to ssDNA with a script executed through VMD.[46] VMD was used for visualization and preparing the MD systems. As demonstrated in Figure 1, two 5 nm long single-walled CNTs of diameter 1.49 nm were surrounded by four single-layer graphene nanosheets to form two membranes. The distance between two horizontal CNTs is 2.2 nm. The space between the CNTs and a vertical 2 nm diameter nanopores between the top graphene nanosheets forms a longitudinal nanopore of a diameter of 2 nm. The two vertical graphene nanosheets at the ends contain a nanopore of diameter 1.5 nm that connects the reservoirs to the CNTs and to the 2 nm diameter vertical nanopore. The CNTs act as transverse nanochannel with a diameter of 1.49 nm in our model. After the structures of the membranes were created, the system was solvated in a cubic 1 M KCl solution with randomly distributed Cl^- and K^+ ions in a 1 g/mL water solvent. To decrease the computing time, we decreased the size of the system. In this regard, we cut the cubic reservoir to a cross-shaped system shown in Figure 1a. We employed the NAMD2 package[47] to perform all-atom MD simulations. The CHARMM36 force field[48] with TIP3P model for waters was applied to account for all interactions within the cutoff radius in the system. The particle mesh Ewald (PME)[49] was used to account for electrostatic and van der Waals long-range interactions at farther distances than the cutoff radius of 12 Å. The integrity of the reservoirs was maintained by the combination of periodic boundary conditions on the farthest edges and boundary forces on the other edges. We applied an electric force on homo-oligonucleotides in a gas phase MD simulation to straighten them up. After minimization of each system, the NVT ensemble was used for the simulations at the temperature of 300 K. The translocation of straight ssDNA through nanopore was restricted to the axial z direction by fixing the backbone atoms of the ssDNA harmonically to the x

https://dx.doi.org/10.1021/acs.jpcb.0c10759
J. Phys. Chem. B XXXX, XXX, XXX−XXX

and y directions. The graphene sheets and nanotubes were fixed in their positions. The NAMD2 program uses the Verlet algorithm to compute the time evolution of the positions and velocities of the atoms by integration over the Newtonian equation of motions every 2 fs time step. The trajectories were written to an output file every 500 time steps which is equal to 1 ps. In our MD model, the two short nanotubes with lengths of 5 nm and diameters of 1.49 nm are surrounded by four graphene nanosheets. The system contains four reservoirs next to each graphene sheet and five virtual walls imposed on the edges of the system by external boundary forces to maintain the integrity of water reservoirs in rectangular shapes shown in Figure 1. The top and bottom nanosheets resemble a hydrophobic membrane with a 2 nm diameter longitudinal nanopore across the membrane. The 16-mer homo-oligonucleotides along with ions and water translocate through the longitudinal nanopore under a longitudinally applied electric field. The left and right graphene nanosheets form a hydrophobic membrane through which the ions and water translocate under a transversely applied electric field. Consequently, the solutions in the reservoirs are interconnected with each other through the intersection of transverse nanochannel and longitudinal nanopore. Consequently, the ions entering the nanopore can exit from the nanochannel, and vice versa. The translocation of ions under transversely and longitudinally applied electric fields generate ionic currents through the nanopore and nanochannel, respectively. The voltage-driven nucleotides inside the pore disturb the movement of Cl^- and K^+ ions, and as a result, they modulate the ionic currents. The modulation of ionic current can be indicative of nucleotide types. We performed a 100 ns simulation for each nucleotide. Because the ionic currents data were collected when the homo-oligonucleotides were completely occupying the pore, the currents were measured only over part of the simulation—over 100 ns in total.

Electric Ionic Current. Equation 1 was used to calculate the electrical currents generated by displacements of ion i under applied electric fields in the MD simulations:

$$I(t) = \frac{1}{l_a \Delta t} \sum_{i=1}^{n} q_i [a_i(t + \Delta t) - a_i(t)] \quad (1)$$

where Δt, the time frame of the simulation, is 10^{-12} s. l_a is the length of the membranes in the x and z directions. q_i is the electrical charge of ion i which is −1 and +1 for Cl^- and K^+, respectively. Δa_i is the measure of the displacement of species i in time Δt.

Neural Network. We used a pattern recognition neural network (NN) in MATLAB for the classification of the longitudinal and transverse ionic currents to target classes of homo-oligonucleotides. The two-layer NN contains one hidden layer with 128 neurons. We used the Levenberg−Marquardt (LM) transfer function for training the data and the mean-square error (MSE) function for the calculation of prediction error. To prepare the ionic current data, we used a moving average function, movemean in MATLAB, with a sliding window of 5 ns. The movemean function is a simple moving average that smooths the noisy time series electrical currents calculated from MD trajectories.

DFT and TDDFT. The optimization of bases was performed at the DFT level of theory with the B3LYP functional and the 6-311++g(d,p) basis set. We employed TDDFT to calculate the optical properties of single nucleotides with optimized structures. Again, we used the B3LYP functional with 6-311++g(d,p) basis set. The TDDFT calculation was performed with different numbers of excited electronic states of 25, 50, 75, and 100. For the solvation effect, we used an integral equation formalism of the polarized continuum model (IEFPCM) with universal force field (UFF) radii for spheres of the cavity with a scaling factor of 1.1 of solute atoms.

RESULTS AND DISCUSSION

To evaluate the ionic current discrepancies that are measured with a high sampling frequency of 10^{12} s^{-1} between four homo-oligonucleotides, we created histograms of the ionic currents. Figures 2a and 2b present histograms of individual longitudinal and transverse ionic currents, respectively, during translocation of A_{16}, G_{16}, T_{16}, and C_{16} homo-oligonucleotides through the nanopore.

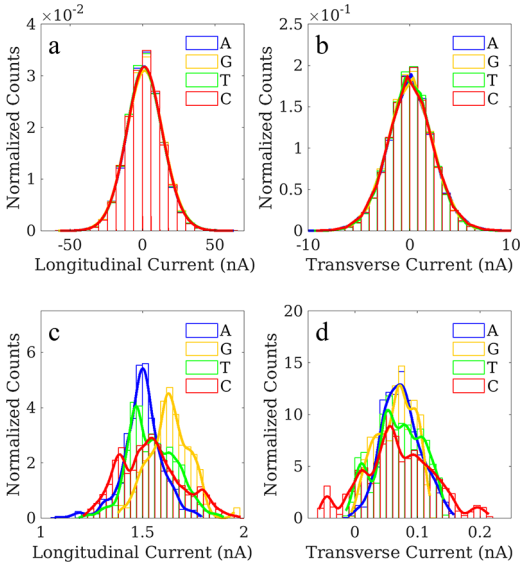


Figure 2. Histograms of ionic currents during A_{16}, G_{16}, T_{16}, and C_{16} homo-oligonucleotides through the longitudinal nanopore and transverse nanochannel generated under an applied voltage of 0.5 V in each direction. (a, b) The histograms of raw ionic currents generated through longitudinal and transverse pores, respectively. (c, d) Histogram of averaged ionic currents generated through longitudinal and transverse pores, respectively. The histograms of the ionic current of four homo-oligonucleotides are broadly overlapping.

Therefore, we averaged over a large sliding window of 5 ns containing 5000 measurements to increase the resolution of ionic currents. Figures 2c and 2d present the histogram of averaged longitudinal and transverse ionic currents, respectively. The averaged individual ionic currents of four homo-oligonucleotide over a window of 5 ns are partially distinguishable. To further increase the resolution of the averaged ionic currents, we combined the transverse and longitudinal ionic currents in a two-dimensional (2D) scatter plot presented in Figure 3a.

Initial assessment of ionic current efficiency in discriminating between the homo-oligonucleotides is achieved by observing histograms of individual ionic currents and 2D scatter plot of longitudinal and transverse ionic currents. To determine the accuracy of distinguishing between four homo-oligonucleotide with averaged ionic currents, we employed a pattern recognition neural network (NN) algorithm. With this

https://dx.doi.org/10.1021/acs.jpcb.0c10759
J. Phys. Chem. B XXXX, XXX, XXX−XXX

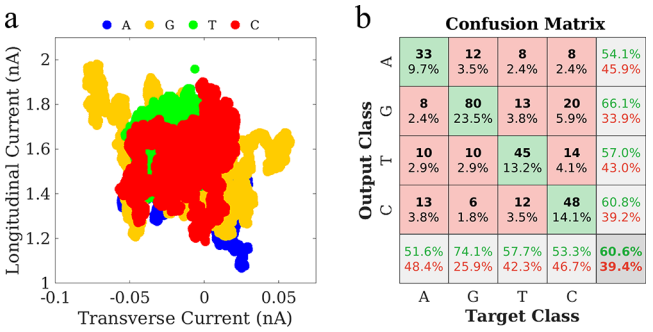


Figure 3. Combination of longitudinal and transverse ionic currents generated by applied voltage of 0.5 V in each direction. (a) A 2D scatter plot of longitudinal and transverse ionic currents. (b) Testing confusion matrix for classification of four types of homo-oligonucleotides using the pattern recognition NN model.

technique, after training the NN model, we classified the homo-oligonucleotide in their categories with an accuracy of 60.6%. The confusion matrix in Figure 3b presents the testing classification accuracy of homo-oligonucleotides by using the NN model. An increase and decrease of the time elapse (sliding window) for averaging over the ionic currents lead to higher and lower accuracy, respectively, in the classification of individual nucleotides.

The raw ionic currents do not show any significant method of classification, indicating minimal information that is contained in the ionic currents. As our MD simulations imply, the ion−nucleotide interactions as a source of nucleotide recognition are not robust enough for single-read identification of nucleotides. This motivated us to search for an alternative method instead of electrical ionic currents in identifying the nucleotides.[35−37,50−54] In this connection, we investigated the light−nucleotide interactions in this study. We could not incorporate light into the solid-state model despite the schematic demonstration shown in Figure 4. However, the

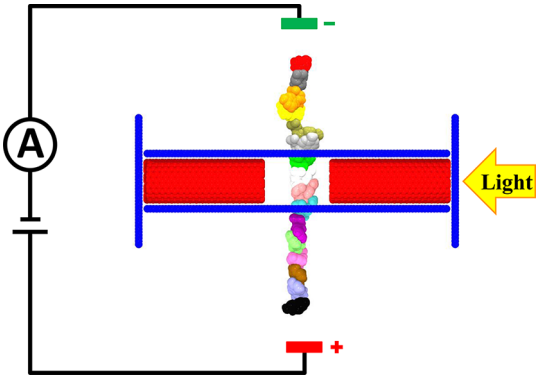


Figure 4. Schematic illustration of light incorporation in the hypothesized model.

optical device can be integrated with the solid-state nanopore. Here, we hypothesized that light can be irradiated through transverse nanochannel and interact with nucleotides translocating through the nanopore in the model (Figure 4).

As the wavelength of the light even in the UV−vis region is substantially larger than the size of a single nucleotide, the light propagated through the system would interact with many nucleotides. Therefore, in far-field spectroscopy, it is fundamentally challenging to distinguish individual nucleotides from light−nucleotide interactions if not impossible. In this regard, we assume the nucleotides are pushed through the pore individually one by one to find how single nucleotides affect the optical absorption. The single nucleotide translocation is realized experimentally by using an exonuclease enzyme coupled with the nanopore.[55] Exonuclease captures the ssDNA and cleaves off the nucleotides consecutively from it. In this scenario, the single nucleotides may show recognizable signals during passage through the nanopore. In this regard, we calculated the UV−vis, IR, and Raman spectra of individual nucleotides (Figure 5 and 6, Figures S1 and S2).

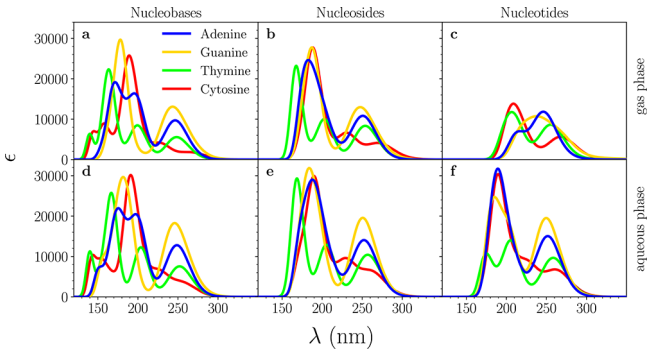


Figure 5. Interaction of light and four types of nucleotides. (a, b, c) UV−vis spectrum of optimized structure of nucleotides, nucleotides, and nucleobases in the gas phase, respectively. (d, e, f) UV−vis spectrum of optimized structure of nucleosides, nucleotides, and nucleobases in an aqueous solution, respectively.

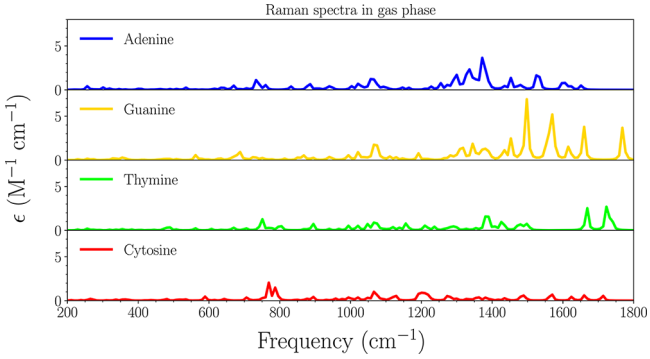


Figure 6. Interaction of light and four types of nucleotides. Raman spectra of optimized structure of nucleotides in the gas phase.

Figure 5 shows the optical spectra of stable structures of nucleotides, nucleobases, and nucleosides in gas and aqueous phases. To obtain this, we first optimized the structures of the nucleotides, nucleobases, and nucleosides in gas and aqueous phases by using the DFT level of theory. Then we employed TDDFT to calculate the excitation energies of electronic transitions for nucleotides, nucleobases, and nucleosides in the gas and aqueous phases.

Figures 5a−c show that the optical properties of the nucleotides, nucleosides, and nucleobases in the gas phase, respectively, are partially distinguishable. Also, Figures 5d−f present the corresponding spectra of nucleotides, nucleosides, and nucleobases in the aqueous phase with discriminative intensity and wavelength of the absorbed light. The intensity of absorbance is significantly different for four types of nucleotides, nucleosides, and nucleobases. Furthermore, the maximum wavelength of absorbed light varies among them.

https://dx.doi.org/10.1021/acs.jpcb.0c10759
J. Phys. Chem. B XXXX, XXX, XXX−XXX

In the above hypothetical system, although we assumed the nucleotides are passing through the pore individually, we still have to deal with the locality issue of the light as in a real experimental setup the illuminated light is not localized on the single nucleotides. This can be overcome in near-field spectroscopy with a scanning tip near the nucleotides or by using plasmons to localize light in measuring enhanced Raman signals.

Extensive research on plasmons with small structures relative to the wavelength of the illuminated light shows that the presence of plasmons enhances localized fields.[56] The collision of incident light with electrons in metallic compounds results in the collective oscillation of electrons across the plasmonic structures which develops clouds of electrons around the structures. These electron-rich areas in the vicinity of the plasmons are termed hot spots. The hot spots enhance the local electromagnetic field that amplifies the optical signals of species that are located inside the hot spot area. Furthermore, the surface plasmon resonance enables the charge transfer between substrate and absorbate species leading to resonance Raman scattering that amplifies Raman signals of the species significantly in a chemical enhancement mechanism.[57−60] Therefore, the presence of plasmons in the system increases the sensitivity of light interaction with target species. Moreover, the relatively small plasmons localize the light which promotes the possibility of single-molecule spectroscopy. Also, a mapping tip close to the species can further help with the localization of light as light cannot propagate a long distance. Near-field spectroscopy combined with plasmonic structures can lead to tip-enhanced Raman spectroscopy (TERS) techniques that can further be applied to the nanopore DNA sequencing method.[43,44] In this scenario, we can compare our calculated Raman spectra with ones observed experimentally by using surface-enhanced Raman spectroscopy (SERS) techniques with the aid of plasmons. Here, we focus on the major and characteristic signals in the region that there exist available experimental signals of four nucleotides for comparison with our results. Figure 6 shows vibrational modes of the Raman spectra for four nucleotides in wavenumber range of 200−1800 cm^{-1} taken from the full spectra that are presented in Figure S2.

The vibrational modes in Raman spectra of four nucleotides demonstrate significant differences, and they contain characteristic signals that can be exploited to sequence the DNA.[38−43] Peaks between 600 and 800 cm^{-1} that are attributed to the ring breathing modes are distinguishable for four nucleotides. The ring breathing modes for adenine, guanine, and thymine are around 735, 685, and 750 cm^{-1}, respectively. For cytosine there are two peaks in breathing mode region around 770 and 785 cm^{-1} which are strongest peaks in the spectra of cytosine and distinguishable from other nucleotides.[41,61−66] In the structure of adenine C═O does not exist; in this regard we do not see any significant peaks above 1600 cm^{-1}. On the other hand, there are strong bands for guanine (around 1660 and 1770 cm^{-1}) and thymine (around 1670 and 1725 cm^{-1}) in that region. Also, cytosine spectra contain relatively weak bands around 1625, 1660, and 1715 cm^{-1} which can be attributed to C═O stretching modes.[61,65−67] Adenine has its strongest bands around 1300 (may be related to C−C starching or wagging), 1335, and 1370 cm^{-1} which are assigned to C−N stretching.[65−68] In this region adenine spectra contain the sharpest peaks among the nucleotides. These peaks, particularly the one at 1335 cm^{-1} along with the ring breathing mode at 735 cm^{-1}, are believed to be the important characteristic Raman signals of the adenine nucleotide. In guanine spectra, the strongest bands around 1500 and 1570 cm^{-1} are assigned to C−N stretching and N−H_2 scissoring vibrations, respectively.[61,65,66,69] These two vibrational modes have the highest intensity among all the signals of four nucleotides and therefore could be termed characteristic signals of guanine along with breathing modes at 685 cm^{-1}. For thymine, two distinguishable peaks around 1670 and 1770 cm^{-1} possibly due to C═O stretching and ring breathing mode around 750 cm^{-1} can be main characteristic signals of thymine that are distinguishable from other nucleotides.[65,66,70,71] Another significant peak in thymine spectra around 1400 cm^{-1} can be assigned to N−H bending modes.[61,65,71] In cytosine spectra, the strongest band is related to breathing mode around 770 and 785 cm^{-1}. There is another band at 1200 cm^{-1} that can be also a characteristic peak for cytosine. This vibrational frequency is related to N−H_2 rocking, C−C bending, and C−N stretching modes in different experiments.[61,65,66,72] Even though there is difficulty in assigning bands with certainty, our results confirm that the Raman spectra of nucleotides contain characteristic signals for each nucleotide that are discernible with their frequency and intensity from the signals of three other nucleotides. As the plasmons enhance the locality of light and sensitivity of Raman spectra, integration of Raman spectroscopy to the solid-state nanopore perpendicular to the DNA can lead to a robust substitute for conventional nanopore DNA sequencing methods using ionic currents.[27]

Over the past few years, the accuracy of classification of nucleotides has improved substantially by several parallel reads of ionic current signals through multiple mutant CsgG nanopores with diameters of 0.75 nm in their narrowest regions. However, to achieve single-read identification of nucleotides, investigation and development of alternative sensors with higher sensitivity are emerging.[35−37,50−54] Our investigation of ssDNA sequencing via ionic currents in different solid-state nanopore MD systems[73,74] implies the deficiency of necessary information in ionic currents for the classification of nucleotides because of high random fluctuations of ions. The enhancement in classification accuracy of averaged ionic currents through the increase of averaging time elapse over the ionic current data is interpreted in two distinct ways. It may indicate either the increase of signal-to-noise ratio and efficacy of information in averaged ionic currents or the increase in convergent of ionic currents for each homo-oligonucleotides by averaging over the ionic current data. Nevertheless, our MD simulations show that with high sampling frequency the resolution of ionic currents is insufficient to distinguish between nucleotides in single-read nanopore DNA sequencing.

The composition of nucleotides A, G, T, and C differ slightly by a few atoms. This makes it challenging to distinguish the bases from each other in nanopore DNA sequencing by using ionic currents. We speculate the ion sizes are too large for having high informative interactions with nucleotides. Analogously, this is similar to scanning tunneling microscopy (STM) with a large radius curvature of the scanning tip to image a relatively small molecule with comparable size to the curvature. In this regard, the defect of using ions for nucleotide recognition is fundamental, and searching for an alternative indicator that interacts with a smaller characteristic property of nucleotides is encouraged.

https://dx.doi.org/10.1021/acs.jpcb.0c10759
J. Phys. Chem. B XXXX, XXX, XXX−XXX

Light can interact with electrons within the molecules. The intensity and the wavelength of the light absorbance depend on the density of electrons and electronic energy states in the molecules. The DFT/TDDFT calculations indicated the spectra of different nucleotide types are distinguishable with different intensity and wavelength of the absorbed light (Figure 5 and 6, Figures S1 and S2). Our results indicate the implementation of spectroscopy in solid-state nanopores would provide optical signals with higher resolution than electrical signals obtained from ionic currents. Developing an ultrasmall solid-state nanopore requires elegant engineering techniques, but the solid-state nanopores are highly stable, chemically tunable, and compatible with the integration of electrical/optical devices.[75−77]

CONCLUSION

In summary, the averaged ionic currents obtained from the MD simulations were classified into four homo-oligonucleotide categories corresponding to the canonical bases A, G, T, and C by a NN algorithm with an accuracy of 60.6%. However, the low resolution of raw ionic currents does not permit single-read nanopore DNA sequencing in our MD simulations. This indicates the ion−nucleotide interactions do not contain adequate information for single-read identification of nucleotides in our MD simulations. On the other hand, DFT and TDDFT calculations indicate that the interaction of light and nucleotide could provide distinguishable spectra as the intensity and frequency of the absorbed light for four types of nucleotides vary significantly, but conventional optics is inadequate to stimulate and distinguish between optical spectra of different nucleotides present in a ssDNA sequence. However, we expect that a specially and specifically designed spectrometer—potentially surface-enhanced Raman spectroscopy (SERS) or near-field spectroscopy—can offer a single-read nanopore DNA sequencing technique in a well-designed solid-state nanopore with a precisely controlled translocation of nucleotides.[78] The compatibility of solid-state nanopores with the integration of optical devices owing to their robust structure permits the usage of spectroscopy or microscopy in nanopore DNA sequencing.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.0c10759.

Available IR and Raman spectra of nucleotides obtained from DFT calculations (PDF)

AUTHOR INFORMATION

Corresponding Author

Jayendran C. Rasaiah − *Department of Chemistry, University of Maine, Orono, Maine 04469, United States;* orcid.org/0000-0002-4453-7438; Phone: + 207 581 1179; Email: rasaiah@maine.edu

Author

Mohsen Farshad − *Department of Chemistry, University of Maine, Orono, Maine 04469, United States;* orcid.org/0000-0001-5095-4361

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.0c10759

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank an anonymous referee and Professor Carl Tripp for critical comments on the manuscript and the Advanced Computing Group of the University of Maine, in particular Stephen Cousins for his continuous support and provision of computer time.

REFERENCES

(1) Branton, D.; Deamer, D. W. *Nanopore Sequencing: An Introduction*; World Scientific: Hackensack, NJ, 2018.

(2) Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. Characterization of indi- vidual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 13770−13773.

(3) Loman, N. J.; Quick, J.; Simpson, J. T. A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat. Methods* **2015**, *12*, 733−735.

(4) Craig, J. M.; Laszlo, A. H.; Derrington, I. M.; Ross, B. C.; Brinkerhoff, H.; Nova, I. C.; Doering, K.; Tickman, B. I.; Svet, M. T.; Gundlach, J. H. Direct Detection of Unnatural DNA Nucleotides dNaM and d5SICS using the MspA Nanopore. *PLoS One* **2015**, *10*, e0143253.

(5) Laszlo, A. H.; Derrington, I. M.; Ross, B. C.; Brinkerhoff, H.; Adey, A.; Nova, I. C.; Craig, J. M.; Langford, K. W.; Samson, J. M.; Daza, R.; Doering, K.; Shendure, J.; Gundlach, J. H. Decoding long nanopore sequencing reads of natural DNA. *Nat. Biotechnol.* **2014**, *32*, 829−833.

(6) Laszlo, A. H.; Derrington, I. M.; Brinkerhoff, H.; Langford, K. W.; Nova, I. C.; Sam- son, J. M.; Bartlett, J. J.; Pavlenok, M.; Gundlach, J. H. Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 18904−18909.

(7) Quick, J.; et al. Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of Salmonella. *Genome Biol.* **2015**, *16*, 114.

(8) Chaisson, M. J. P.; Wilson, R. K.; Eichler, E. E. Genetic variation and the de novo assembly of human genomes. *Nat. Rev. Genet.* **2015**, *16*, 627−640.

(9) Wescoe, Z. L.; Schreiber, J.; Akeson, M. Nanopores Discriminate among Five C5- Cytosine Variants in DNA. *J. Am. Chem. Soc.* **2014**, *136*, 16582−16587.

(10) Jain, M.; Olsen, H. E.; Paten, B.; Akeson, M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol.* **2016**, *17*, 239.

(11) Quick, J.; et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature* **2016**, *530*, 228−232.

(12) Istace, B.; Friedrich, A.; d'Agata, L.; Faye, S.; Payen, E.; Beluche, O.; Ca- radec, C.; Davidas, S.; Cruaud, C.; Liti, G.; Lemainque, A.; Engelen, S.; Wincker, P.; Schacherer, J.; Aury, J.-M. de novo assembly and population genomic survey of natural yeast isolates with the Oxford Nanopore MinION sequencer. *GigaScience* **2017**, *6*, 1−13 DOI: 10.1093/gigascience/giw018.

(13) Jain, M.; et al. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nat. Biotechnol.* **2018**, *36*, 338−345.

(14) Garalde, D. R.; et al. Highly parallel direct RNA sequencing on an array of nanopores. *Nat. Methods* **2018**, *15*, 201−206.

(15) Noakes, M. T.; Brinkerhoff, H.; Laszlo, A. H.; Derrington, I. M.; Langford, K. W.; Mount, J. W.; Bowman, J. L.; Baker, K. S.; Doering, K. M.; Tickman, B. I.; Gundlach, J. H. Increasing the accuracy of nanopore DNA sequencing using a time-varying cross membrane voltage. *Nat. Biotechnol.* **2019**, *37*, 651−656.

(16) Akeson, M.; Branton, D.; Kasianowicz, J. J.; Brandin, E.; Deamer, D. W. Microsecond Time-Scale Discrimination Among Polycytidylic Acid, Polyadenylic Acid, and Polyuridylic Acid as

Homopolymers or as Segments Within Single RNA Molecules. *Biophys. J.* **1999**, *77*, 3227−3233.
(17) Meller, A.; Nivon, L.; Brandin, E.; Golovchenko, J.; Branton, D. Rapid nanopore discrimination between single polynucleotide molecules. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 1079−1084.
(18) Deamer, D. W.; Branton, D. Characterization of Nucleic Acids by Nanopore Analysis. *Acc. Chem. Res.* **2002**, *35*, 817−825.
(19) Muthukumar, M. Polymer Translocation, 2016; OCLC: 992448436.
(20) Cherf, G. M.; Lieberman, K. R.; Rashid, H.; Lam, C. E.; Karplus, K.; Akeson, M. Automated forward and reverse ratcheting of DNA in a nanopore at 5°A precision. *Nat. Biotechnol.* **2012**, *30*, 344−348.
(21) Manrao, E. A.; Derrington, I. M.; Laszlo, A. H.; Langford, K. W.; Hopper, M. K.; Gillgren, N.; Pavlenok, M.; Niederweis, M.; Gundlach, J. H. Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat. Biotechnol.* **2012**, *30*, 349−353.
(22) Branton, D.; et al. The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* **2008**, *26*, 1146−1153.
(23) Stoddart, D.; Heron, A. J.; Mikhailova, E.; Maglia, G.; Bayley, H. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 7702−7707.
(24) Astier, Y.; Kainov, D. E.; Bayley, H.; Tuma, R.; Howorka, S. Stochastic Detection of Motor Protein−RNA Complexes by Single-Channel Current Recording. *ChemPhysChem* **2007**, *8*, 2189−2194.
(25) Derrington, I. M.; Craig, J. M.; Stava, E.; Laszlo, A. H.; Ross, B. C.; Brinkerhoff, H.; Nova, I. C.; Doering, K.; Tickman, B. I.; Ronaghi, M.; Mandell, J. G.; Gunderson, K. L.; Gundlach, J. H. Subangstrom single-molecule measurements of motor proteins using a nanopore. *Nat. Biotechnol.* **2015**, *33*, 1073−1075.
(26) Wang, S.; Zhao, Z.; Haque, F.; Guo, P. Engineering of protein nanopores for sequencing, chemical or protein sensing and disease diagnosis. *Curr. Opin. Biotechnol.* **2018**, *51*, 80−89.
(27) Belkin, M.; Chao, S.-H.; Jonsson, M. P.; Dekker, C.; Aksimentiev, A. Plasmonic Nanopores for Trapping, Controlling Displacement, and Sequencing of DNA. *ACS Nano* **2015**, *9*, 10598−10611.
(28) Mak, D.; Webb, W. Conductivity noise in transmembrane ion channels due to ion concentration fluctuations via diffusion. *Biophys. J.* **1997**, *72*, 1153−1164.
(29) Bhattacharya, S.; Yoo, J.; Aksimentiev, A. Water Mediates Recognition of DNA Sequence *via* Ionic Current Blockade in a Biological Nanopore. *ACS Nano* **2016**, *10*, 4644−4651.
(30) Aksimentiev, A.; Heng, J. B.; Timp, G.; Schulten, K. Microscopic Kinetics of DNA Translocation through Synthetic Nanopores. *Biophys. J.* **2004**, *87*, 2086−2097.
(31) Wells, D. B.; Belkin, M.; Comer, J.; Aksimentiev, A. Assessing Graphene Nanopores for Sequencing DNA. *Nano Lett.* **2012**, *12*, 4117−4123.
(32) Aksimentiev, A. Deciphering ionic current signatures of DNA transport through a nanopore. *Nanoscale* **2010**, *2*, 468.
(33) Wilson, J.; Di Ventra, M. Single-base DNA discrimination via transverse ionic trans- port. *Nanotechnology* **2013**, *24*, 415101.
(34) Boynton, P.; Di Ventra, M. Sequencing proteins with transverse ionic transport in nanochannels. *Sci. Rep.* **2016**, *6*, 25232.
(35) McNally, B.; Singer, A.; Yu, Z.; Sun, Y.; Weng, Z.; Meller, A. Optical Recognition of Converted DNA Nucleotides for Single-Molecule DNA Sequencing Using Nanopore Arrays. *Nano Lett.* **2010**, *10*, 2237−2244.
(36) Anderson, B. N.; Assad, O. N.; Gilboa, T.; Squires, A. H.; Bar, D.; Meller, A. Probing Solid-State Nanopores with Light for the Detection of Unlabeled Analytes. *ACS Nano* **2014**, *8*, 11836−11845.
(37) Spitzberg, J. D.; Zrehen, A.; van Kooten, X. F.; Meller, A. Plasmonic-Nanopore Biosensors for Superior Single-Molecule Detection. *Adv. Mater.* **2019**, *31*, 1900422.
(38) Kneipp, K.; Kneipp, H.; Kartha, V. B.; Manoharan, R.; Deinum, G.; Itzkan, I.; Dasari, R. R.; Feld, M. S. Detection and identification of a single DNA base molecule using surface-enhanced Raman scattering (SERS). *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1998**, *57*, R6281−R6284.
(39) Kneipp, K.; Kneipp, H. Single Molecule Raman Scattering. *Appl. Spectrosc.* **2006**, *60*, 322A−334A.
(40) Chen, C.; Ye, J.; Li, Yi; Lagae, L.; Stakenborg, T.; Van Dorpe, P. Detection of DNA Bases and Oligonucleotides in Plasmonic Nanoslits Using Fluidic SERS. *IEEE J. Sel. Top. Quantum Electron.* **2013**, *19*, 4600707−4600707.
(41) Barhoumi, A.; Zhang, D.; Tam, F.; Halas, N. J. Surface-Enhanced Raman Spectroscopy of DNA. *J. Am. Chem. Soc.* **2008**, *130*, 5523−5529.
(42) Pyrak, E.; Krajczewski, J.; Kowalik, A.; Kudelski, A.; Jaworska, A. Surface Enhanced Raman Spectroscopy for DNA Biosensors—How Far Are We? *Molecules* **2019**, *24*, 4423.
(43) Huang, J.-A.; Mousavi, M. Z.; Zhao, Y.; Hubarevich, A.; Omeis, F.; Giovannini, G.; Schütte, M.; Garoli, D.; De Angelis, F. SERS discrimination of single DNA bases in single oligonucleotides by electro-plasmonic trapping. *Nat. Commun.* **2019**, *10*, 5321.
(44) He, Z.; Han, Z.; Kizer, M.; Linhardt, R. J.; Wang, X.; Sinyukov, A. M.; Wang, J.; Deckert, V.; Sokolov, A. V.; Hu, J.; Scully, M. O. Tip-Enhanced Raman Imaging of Single-Stranded DNA with Single Base Resolution. *J. Am. Chem. Soc.* **2019**, *141*, 753−757.
(45) van Dijk, M.; Bonvin, A. M. J. J. 3D-DART: a DNA structure modelling server. *Nucleic Acids Res.* **2009**, *37*, W235−W239.
(46) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(47) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(48) MacKerell, A. D.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins †. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(49) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An *N* log(*N*) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(50) Zwolak, M.; Di Ventra, M. Electronic Signature of DNA Nucleotides via Transverse Transport. *Nano Lett.* **2005**, *5*, 421−424.
(51) Tsutsui, M.; Taniguchi, M.; Yokota, K.; Kawai, T. Identifying single nucleotides by tunnelling current. *Nat. Nanotechnol.* **2010**, *5*, 286−290.
(52) Yokota, K.; Tsutsui, M.; Taniguchi, M. Electrode-embedded nanopores for label-free single-molecule sequencing by electric currents. *RSC Adv.* **2014**, *4*, 15886−15899.
(53) Di Ventra, M.; Taniguchi, M. Decoding DNA, RNA and peptides with quantum tun- nelling. *Nat. Nanotechnol.* **2016**, *11*, 117−126.
(54) Verschueren, D. V.; Pud, S.; Shi, X.; De Angelis, L.; Kuipers, L.; Dekker, C. Label-Free Optical Detection of DNA Translocations through Plasmonic Nanopores. *ACS Nano* **2019**, *13*, 61−70.
(55) Clarke, J.; Wu, H.-C.; Jayasinghe, L.; Patel, A.; Reid, S.; Bayley, H. Continuous base identification for single-molecule nanopore DNA sequencing. *Nat. Nanotechnol.* **2009**, *4*, 265−270.
(56) Li, M.; Cushing, S. K.; Wu, N. Plasmon-enhanced optical sensors: a review. *Analyst* **2015**, *140*, 386−406.
(57) Stiles, P. L.; Dieringer, J. A.; Shah, N. C.; Van Duyne, R. P. Surface-Enhanced Raman Spectroscopy. *Annu. Rev. Anal. Chem.* **2008**, *1*, 601−626.
(58) Haynes, C. L.; McFarland, A. D.; Van Duyne, R. P. Surface-Enhanced Raman Spectroscopy. *Anal. Chem.* **2005**, *77*, 338 A−346 A.
(59) Willets, K. A.; Van Duyne, R. P. Localized Surface Plasmon Resonance Spectroscopy and Sensing. *Annu. Rev. Phys. Chem.* **2007**, *58*, 267−297.
(60) Schlücker, S. Surface-Enhanced Raman Spectroscopy: Concepts and Chemical Applications. *Angew. Chem., Int. Ed.* **2014**, *53*, 4756−4795.

https://dx.doi.org/10.1021/acs.jpcb.0c10759
J. Phys. Chem. B XXXX, XXX, XXX−XXX

(61) Otto, C.; van den Tweel, T. J. J.; de Mul, F. F. M.; Greve, J. Surface-enhanced Raman spectroscopy of DNA bases. *J. Raman Spectrosc.* **1986**, *17*, 289−298.
(62) Suh, J. S.; Moskovits, M. Surface-enhanced Raman spectroscopy of amino acids and nucleotide bases adsorbed on silver. *J. Am. Chem. Soc.* **1986**, *108*, 4711−4718.
(63) Bell, S. E. J.; Sirimuthu, N. M. S. Surface-Enhanced Raman Spectroscopy (SERS) for Sub-Micromolar Detection of DNA/RNA Mononucleotides. *J. Am. Chem. Soc.* **2006**, *128*, 15580−15581.
(64) Green, M.; Liu, F.-M.; Cohen, L.; Köllensperger, P.; Cass, T. SERS platforms for high density DNA arrays. *Faraday Discuss.* **2006**, *132*, 269−280.
(65) Madzharova, F.; Heiner, Z.; Gühlke, M.; Kneipp, J. Surface-Enhanced Hyper-Raman Spectra of Adenine, Guanine, Cytosine, Thymine, and Uracil. *J. Phys. Chem. C* **2016**, *120*, 15415−15423.
(66) Pyrak; Jaworska; Kudelski. SERS Studies of Adsorption on Gold Surfaces of Mononucleotides with Attached Hexanethiol Moiety: Comparison with Selected Single- Stranded Thiolated DNA Fragments. *Molecules* **2019**, *24*, 3921.
(67) Papadopoulou, E.; Bell, S. E. J. Label-Free Detection of Single-Base Mismatches in DNA by Surface-Enhanced Raman Spectroscopy. *Angew. Chem., Int. Ed.* **2011**, *50*, 9058−9061.
(68) Giese, B.; McNaughton, D. Surface-Enhanced Raman Spectroscopic and Density Functional Theory Study of Adenine Adsorption to Silver Surfaces. *J. Phys. Chem. B* **2002**, *106*, 101−112.
(69) Giese, B.; McNaughton, D. Density functional theoretical (DFT) and surface-enhanced Raman spectroscopic study of guanine and its alkylated derivatives. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5161−5170.
(70) Aroca, R.; Bujalski, R. Surface enhanced vibrational spectra of thymine. *Vib. Spectrosc.* **1999**, *19*, 11−21.
(71) Cho, K.-H.; Choo, J.; Joo, S.-W. Tautomerism of thymine on gold and silver nanoparticle surfaces: surface-enhanced Raman scattering and density functional theory calculation study. *J. Mol. Struct.* **2005**, *738*, 9−14.
(72) Śanchez Cortés, S.; Garcia-Ramos, J. V. SERS of cytosine and its methylated derivatives on metal colloids. *J. Raman Spectrosc.* **1992**, *23*, 61−66.
(73) Farshad, M.; Rasaiah, J. C. Molecular Dynamics Simulation Study of Transverse and Longitudinal Ionic Currents in Solid-State Nanopore DNA Sequencing. *ACS Applied Nano Materials* **2020**, *3*, 1438−1447.
(74) Farshad, M.; Rasaiah, J. C. Reverse Translocation of Nucleotides through a Carbon Nanotube. *J. Phys. Chem. B* **2020**, *124*, 937−943.
(75) Dekker, C. Solid-state nanopores. *Nat. Nanotechnol.* **2007**, *2*, 209−215.
(76) Chen, Q.; Liu, Z. Fabrication and Applications of Solid-State Nanopores. *Sensors* **2019**, *19*, 1886.
(77) Goto, Y.; Akahori, R.; Yanagi, I.; Takeda, K.-i. Solid-state nanopores towards single-molecule DNA sequencing. *J. Hum. Genet.* **2020**, *65*, 69−77.
(78) Fragasso, A.; Schmid, S.; Dekker, C. Comparing Current Noise in Biological and Solid-State Nanopores. *ACS Nano* **2020**, *14*, 1338−1349.

https://dx.doi.org/10.1021/acs.jpcb.0c10759
J. Phys. Chem. B XXXX, XXX, XXX−XXX