Recognizing nucleotides by cross-tunneling currents for DNA sequencing

V. M. K. Bagci[1] and Chao-Cheng Kaun[1,2,*]
[1]*Research Center for Applied Sciences, Academia Sinica, Taipei 11529, Taiwan*
[2]*Department of Physics, National Tsing-Hua University, Hsinchu 30013, Taiwan*
(Received 30 January 2011; revised manuscript received 23 March 2011; published 22 July 2011)

Using first-principles calculations, we study electron transport through nucleotides inside a rectangular nanogap formed by two pairs of gold electrodes which are perpendicular and parallel to the nucleobase plane. We propose that this setup will enhance the nucleotide selectivity of tunneling signals to a great extent. Information from three electrical probing processes offers full nucleotide recognition, which survives the noise from neighboring nucleotides and configuration fluctuations.

DOI: 10.1103/PhysRevE.84.011917 PACS number(s): 87.15.ag, 87.14.gf, 87.15.Pc, 87.85.Qr

I. INTRODUCTION

A fast and affordable DNA sequencing technique would trigger personalized medicine and new phase of pharmaceuticals [1]. Using electron tunneling across a single-stranded (ss) DNA molecule to detect its constituent nucleobases, adenine (A), thymine (T), cytosine (C), and guanine (G), is a promising approach [1–3]. Since the tunneling current is sensitive to the molecule present in the interelectrode gap, nucleotide recognition is possible [4]. But, at the same time, fluctuations on the geometrical configuration, i.e., the molecular position and orientation within the gap, lead to a distribution of currents and thus reduce nucleotide selectivity (signal-to-noise ratio) as well as efficiency of devices [4–8]. Therefore, a key issue in progress lies in finding ways to increase the device selectivity. Holding DNA bases temporarily in the tunneling gap has shown potential in reducing the system instability [2]. Various electrodes have thus been built, including molecule-functionalized electrodes that can induce hydrogen bonds with passing nucleosides or nucleotides in the constructed nanogap [9–12], and reconfigurable gold electrodes that can form a nanogap having the similar size of a nucleotide [13]. So far, three of four single nucleotides have been statistically identified by repeating thousands of measurements, while their magnitude order of conductance values are controversial (G, C, T and C, A, G, respectively) [12,13]. A physical picture that can help to understand these results and a method that can enhance the selectivity and efficiency are therefore desired.

Reading DNA sequences by tunneling currents parallel to the base plane has been studied theoretically [3]. The currents provide signals for nucleotide identification, but they are small in magnitude and exhibit exponential dependence on geometrical configurations. On the other hand, tunneling currents perpendicular to the base plane [14,15] may increase the selectivity of the signals. In such a setup, the nucleobase would be a monolayer of atoms between two electrodes, and thus the electrode gap can be smaller than 8 Å, which would not only restrict configuration fluctuations but also enhance the conduction. Furthermore, the phosphate and sugar groups common to all nucleotides, namely the backbone, can be bypassed. Therefore, parallel, in conjunction with perpendicular tunneling signals, can be used to give better nucleotide identification, which has not been addressed to date.

Naturally, due to the miniaturization in one dimension, efforts to avoid clogging the gap may be necessary [3,16]. Yet, a recent experiment reports that a DNA molecule can indeed be driven through a nanopore smaller than its radius [17]. In addition, experimental setups with interacting reconfigurable electrodes, constructed by piezoactuators or a scanning tunneling microscope tip over substrates, were used to form variable tunneling gaps with success [13,18,19]. Based on this concept, a servo system can be built, so ss-DNA may first be allowed to pass through a larger gap, and after a nucleotide is initially detected, the gap size can be shrunk for reading. Several techniques can facilitate this process further. For example, polarized laser pulses could be applied to hold and spin nucleotides in space [20] before they enter the nanogap. A transverse-electrical field could be used to increase nucleotide stabilization inside the gap [5,13,14,21]. Applying a magnetic [22] or optical [23] tweezer to induce a controllable "reverse DNA translocation" so clearer and more stable signals are available is also an attractive means to this end.

In this work, we analyze tunneling signals generated by the setup of Fig. 1. A stretched ss-DNA (brown cylinder depicts the backbone and colored hexagons represent the nucleobases) translocates through a rectangular gap of electrodes, and the nucleobases align on a plane in the ideal configuration. The first pair of electrodes is perpendicular to the base plane, having a gap size of 0.7 nm, while the second pair is parallel to the base plane, having a gap size of 1.4 nm (about 2 Å larger in both sides than the largest nucleotide). Due to geometrical limitation, a nucleobase would enter an interelectrode gap narrower than its width [24] in a face-to-face configuration, similar to other flat molecules, tetrathiafulvalene and tetraselenafulvalene [25]. Under a transverse-electrical field, a nucleotide inside a gap wider than its width would be stablized along the electrode [5,13,14,21]. Two gold electrodes, crossed and located at different heights, can be fabricated lithographically [26] on a polyimide-coated flexible metal substrate that can be bent mechanically [13]. When the substrate is bent, a narrow (wide) gap, at the low (high) electrode shall open up for perpendicular (parallel) tunneling detections. While the ss-DNA translocates through the gap, the position of nucleotides inside the gap varies, and the conductance does

*kauncc@gate.sinica.edu.tw

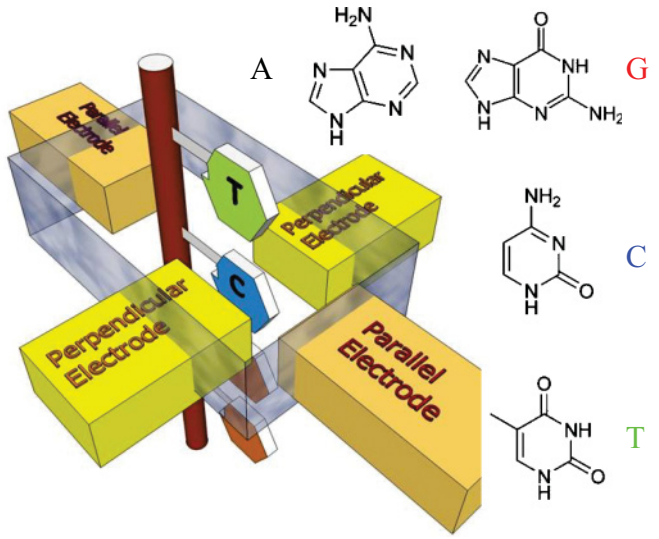

FIG. 1. (Color online) Schematic illustration of the rectangular gap of electrodes along with a stretched translocating ss-DNA. The brown cylinder represents the backbone and the colored hexagons depict the various nucleobases attached to the backbone. The gap between parallel (perpendicular) electrodes is 14.0 (7.0) Å. The perpendicular and parallel electrodes are located at different heights to avoid crosstalk. The chemical structures of the four nucleobases are also shown.

as well, producing a conductance profile. We compare the conductance profiles of nucleotides and sketch out how to use them to identify each nucleotide. We mainly consider profiles of single nucleotides that are hydrogen saturated at both ends. Water molecules and ions are not included in our calculations.

II. COMPUTATIONAL METHOD

We use first-principles quantum transport calculations, performed by ATOMISTIX TOOLKIT [27,28], based on the nonequilibrium Green's function approach on top of the density functional theory, within the generalized gradient approximations [29]. Good agreement between our calculations and experimental data on molecular junctions has been obtained previously [30,31]. The scattering region forming our simulation box [31], as indicated in Fig. 1 by a cuboid, includes a sufficiently large part of the [100] orientated electrodes and the nucleotide or nucleotides in between. The Hamiltonian was expanded in real space having *s, p, d* double ζ with a polarization atomic orbital basis set. The atomic cores were defined by the Troullier-Martins pseudopotentials [32]. After the Kohn-Sham potential is obtained self-consistently, the conductance and currents were calculated from the Landauer formula [33].

III. RESULTS AND DISCUSSION

Our calculations point out that three distinct conduction profiles are necessary to achieve differentiation of each nucleotide efficiently. The first is the perpendicular conductance profile, $G_\perp$, obtained from the conductance across the perpendicular electrodes. In the low bias regime, current is proportional to the conductance. The second is the perpendicular current profile, $I_\perp$, retrieved from the same electrodes under a bias voltage of 1.2 V, which gives the highest contrast in this bias range. And, finally, there is the parallel conductance profile, $G_\parallel$, acquired from the parallel electrodes. We also include a sample calculation for two nucleotides in tandem to test the validity of generalization of the single-nucleotide results on a translocating chain of nucleotides. Furthermore, we address the two most probable configuration fluctuations of the nucleotides, namely the rotations around the glycosidic bond and the translation of entire nucleotides along the longer span of the pore.

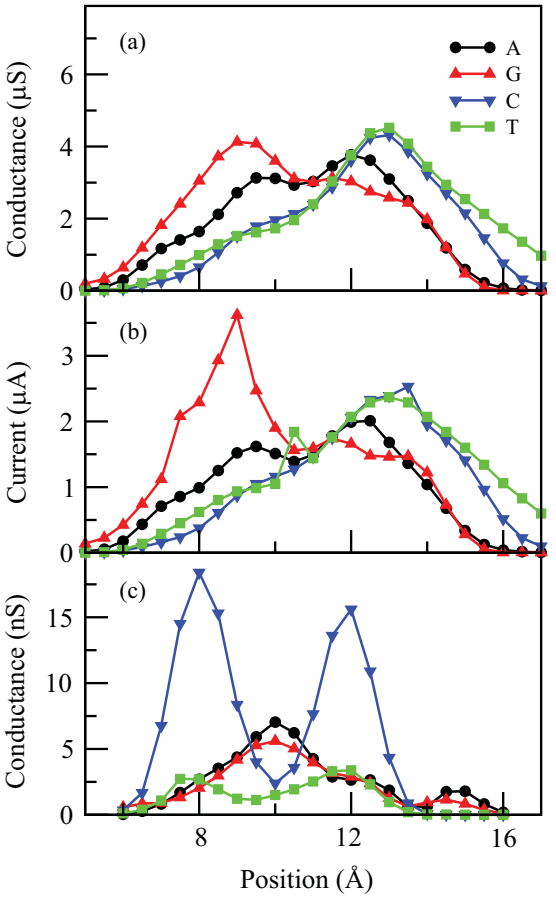

FIG. 2. (Color online) (a) Conductance vs. position profiles of the four nucleotides from perpendicular electrodes; (b) current vs. position profiles from the same electrodes ($V_b = 1.2$ V); (c) conductance vs. position profiles from parallel electrodes.

We first present the $G_\perp$ for the four nucleotides in Fig. 2(a). The conductance values are in the range of μS, which would facilitate detection. The conductance is strongly position dependent. The positions are coordinated relative to the center of the electrode. The bases A and G have two rings (purine) exhibiting two conductance peaks, while the bases C and T have single rings (pyrimidine) displaying single peaks. The peak values of conductances are too close to each other to distinguish the nucleotides. However, if we focus on the conductance values above 2 μS, A and G present a broader profile than do C and T. Therefore, $G_\perp$ may suffice to distinguish the A and G pair from the C and T pair.

To further identify between A and G, current profiles $I_\perp$ are used (under bias voltage of 1.2 V). This voltage can probe the deeper lying electronic states of the nucleotides below the Fermi level of the combined electrode-nucleotide system, while not disrupting the measurement conditions. Figure 2(b) shows that while $I_\perp$ of the other three nucleotides are similar in the shape of $G_\perp$, nucleotide G owns a prominent current peak around position 9 Å and thus can be distinguished from A. In addition, the current peak values follow the order of G, C, and T, while the difference between the latter two are smaller. Whereas a 1-nm gap (just the size of a nucleotide) is used in the experiment [13], the backbone may prefer mechanically to be outside the tunneling path through the nucleobase. Our calculated trend agrees well with the experimental observation [13], suggesting that perpendicular tunneling occurs.

The $G_\parallel$ of the four nucleotides are presented in Fig. 2(c). The position 10 Å roughly corresponds to the positions studied at Refs. [4,6]. If one focuses at that point, A has the highest conductance, followed by G, C, and, finally, T. This is consistent with the previous studies [4]. However, the entire profile shows that C has two very high peaks at different positions, while A and G present single lower peaks, and T has the lowest peak values. At points where C has very high conductance, the other nucleotides have significantly lower conductance. Therefore, $G_\parallel$ is able to distinguish C from the rest of the nucleotides. Whereas the functional molecules lead to parallel tunneling across the base in the experiment [12], our computational results, giving C (T) the maximum (minimum) conductance value, agree with their data (although the tunneling orientations in the base plane are different). At the end, as diagrammatically represented in Fig. 3, under ideal conditions the four nucleotides can be identified as follows: $G_\perp$ differentiates purines (A and G) from pyrimidines (C and T), $I_\perp$ separates G from A, and $G_\parallel$ differentiates C from T.

When the flow of ss-DNA through the pore is controlled, the gap can be occupied by successive nucleotides with equal time intervals. That is, while one nucleotide is followed by another, the two of them share the gap for a certain time, since there is only a 7-Å separation between bases along a stretched ss-DNA (center to center). We investigate the conductance profile for A followed by C, a case where overlap between two bases is the highest. The results shown in Fig. 4(a) indicate that the conductance profile of the A+C combined system is simply a superposition of profiles A and C. Following the logic in reverse, an overall current measurement can be decomposed into individual nucleotide signals.

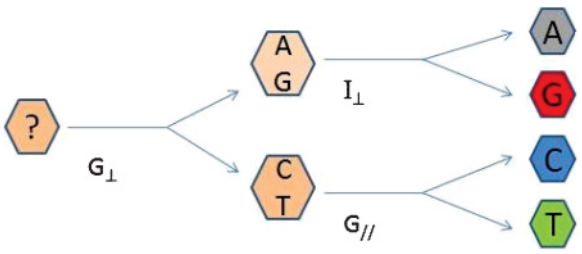

FIG. 3. (Color online) Flow diagram illustration of the three probing processes for nucleotide identification.

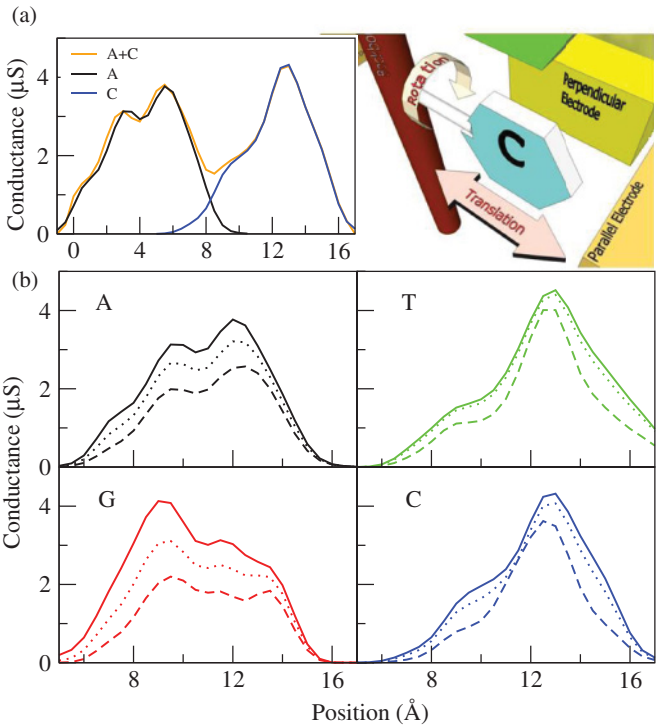

FIG. 4. (Color online) (a) Conductance vs. position profiles, $G_\perp$, for two consecutive nucleotides (A+C) along with two individual nucleotides. (b) The comparative conductance vs. position profiles of the four nucleotides for translational (dotted line) and rotational (dashed line) deviations. Schematic illustration is shown in the upper right panel.

We now investigate the effects of configuration fluctuations on conductance for the nucleotides within the pore: the noise problem. Rotation of the base around the glycosidic bond and translation of the whole nucleotide within the pore (see the upper right panel of Fig. 4) are considered. The $G_\parallel$ calculation is not included because the difference between the conductance of nucleotide C and the others is large enough to survive under the small variations in angle or position. Figure 4(b) includes the $G_\perp$ for the four nucleotides with ideal configurations (solid line), rotation of 10° in glycosidic angle configurations (dashed line), and translation of 0.5 Å configurations (dotted line). These values are chosen to represent the maximum allowed variations in angle and position within the pore volume for the largest nucleotide. The nucleotide G and A have larger conductance variations than C and T, consistent with the experimental observation [13]. Nucleotides have the highest conductance when they are located at the ideal position. While nucleotides inside the pore could have an ensemble of configurations, the highest conductance values recorded would converge to the ideal case (the reading rate could exceed the translocation rate largely). The dependence of conductance on configuration variations is not very strong for all nucleotides in this asymmetric gap and the overall shape of the conductance profiles is preserved.

In conclusion, we suggest that the use of two pairs of gold electrodes forming a rectangular gap can enhance the electrical signature of each nucleotide and thus the efficiency of DNA sequencing. The electrodes perpendicular to the base

planes have a smaller gap than parallel ones, which can reduce the configuration fluctuation, bypass the phosphate and sugar groups, and enhance the transverse electronic transport. With the help of parallel electrodes, each nucleotide can be identified rapidly. The noise, from neighboring nucleotides and configuration fluctuations, will not wash out the signal by this approach.

ACKNOWLEDGMENTS

We thank Chia-Fu Chou and Yi-Der Lin for insightful discussions. This work was partially supported by the National Science Council, the National Center for Theoretical Sciences, and the National Center for High-performance Computing, Republic of China.

[1] D. Branton *et al*., Nat. Biotechnol. **26**, 1146 (2008).
[2] T. Thundat, Nat. Nanotechnol. **5**, 246 (2010).
[3] M. Zwolak and M. Di Ventra, Rev. Mod. Phys. **80**, 141 (2008).
[4] M. Zwolak and M. Di Ventra, Nano Lett. **5**, 421 (2005).
[5] J. Lagerqvist, M. Zwolak, and M. Di Ventra, Nano Lett. **6**, 779 (2006).
[6] R. Zikic, P. S. Krstic, X. G. Zhang, M. Fuentes-Cabrera, J. C. Wells, and X. Zhao, Phys. Rev. E **74**, 011919 (2006).
[7] X. G. Zhang, P. S. Krstic, R. Zikic, J. C. Wells, and M. Fuentes-Cabrera, Biophys. J. **91**, L04 (2006).
[8] M. Krems, M. Zwolak, Y. V. Pershin, and M. Di Ventra, Biophys. J. **97**, 1990 (2009).
[9] M. H. Lee and O. F. Sankey, J. Phys. Condens. Matter **21**, 035110 (2009).
[10] S. Chang, S. Huang, J. He, F. Liang, P. Zhang, S. Li, X. Chen, O. F. Sankey, and S. Lindsay, Nano Lett. **10**, 1070 (2010).
[11] H. He, R. H. Scheicher, R. Pandey, A. R. Rocha, S. Sanvito, A. Grigoriev, R. Ahuja, and S. P. Karna, J. Phys. Chem. C **112**, 3456 (2008).
[12] S. Huang *et al*., Nat. Nanotechnol. **5**, 868 (2010).
[13] M. Tsutsui, M. Taniguchi, K. Yokota, and T. Kawai, Nat. Nanotechnol. **5**, 286 (2010).
[14] J. W. Lee and A. Meller, in *Perspectives in Bioanalysis*, edited by K. Mitchelson (Elseiver, Amsterdam, 2007), pp. 245–263.
[15] M. Xu, R. G. Endres, and Y. Arakawa, Small **3**, 1539 (2007).
[16] J. Lagerqvist, M. Zwolak, and M. Di Ventra, Phys. Rev. E **76**, 013901 (2007).
[17] U. Mirsaidov, J. Comer, V. Dimitrov, A. Aksimentiev, and G. Timp, Nanotechnology **21**, 395501 (2010).
[18] H. Tanaka and T. Kawai, Nat. Nanotechnol. **4**, 518 (2009).
[19] D. A. Ryndyk, E. Shapir, D. Porath, A. Calzolari, R. Di Felice, and G. Cuniberti, ACS Nano **3**, 1651 (2009).
[20] S. S. Viftrup, V. Kumarappan, S. Trippel, H. Stapelfeldt, E. Hamilton, and T. Seideman, Phys. Rev. Lett. **99**, 143602 (2007).
[21] J. Lagerqvist, M. Zwolak, and M. Di Ventra, Biophys. J. **93**, 2384 (2007).
[22] H. Peng and X. S. Ling, Nanotechnology **20**, 185101 (2009).
[23] U. F. Keyser, B. N. Koeleman, S. van Dorp, D. Krapf, R. M. M. Smeets, S. G. Lemay, N. H. Dekker, and C. Dekker, Nat. Phys. **2**, 473 (2006).
[24] M. Tsutsui, M. Taniguchi, and T. Kawai, Appl. Phys. Lett. **93**, 163115 (2008).
[25] M. Taniguchi, M. Tsutsui, K. Shoji, H. Fujiwara, and T. Kawai, J. Am. Chem. Soc. **131**, 14146 (2009).
[26] H. Song, Y. Kim, Y. H. Jang, H. Jeong, M. A. Reed, and T. Lee, Nature **462**, 1039 (2009).
[27] J. Taylor, H. Guo, and J. Wang, Phys. Rev. B **63**, 245407 (2001).
[28] M. Brandbyge, J.-L. Mozos, P. Ordejón, J. Taylor, and K. Stokbro, Phys. Rev. B **65**, 165401 (2002).
[29] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[30] C.-C. Kaun and T. Seideman, Phys. Rev. B **77**, 033414 (2008).
[31] A. Sen and C.-C. Kaun, ACS Nano **4**, 6404 (2010).
[32] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
[33] S. Datta, *Electrical Transport in Mesoscopic Systems* (Cambridge University Press, Cambridge, UK, 1997).