On Nanopore DNA Sequencing by Signal and Noise Analysis of Ionic Current

Chenyu Wen,[1] Shuangshuang Zeng,[1] Zhen Zhang,[1] Klas Hjort,[1] Ralph Scheicher,[2] and Shi-Li Zhang[1,*]

[1] Department of Engineering Sciences, Uppsala University, 751 21 Uppsala, Sweden

[2] Department of Physics and Astronomy, Uppsala University, 751 20 Uppsala, Sweden

[*] Contact person. E-mail address: shili.zhang@angstrom.uu.se

Abstract

DNA sequencing, *i.e.*, the process of determining the succession of nucleotides on a DNA strand, has become a standard aid in biomedical research and is expected to revolutionize medicine. With the capability of handling single DNA molecules, nanopore technology holds high promises to become speedier in sequencing at lower cost than what are achievable with the commercially available optics- or semiconductor-based massively parallelized technologies. Despite tremendous progress made with biological and solid-state nanopores, high error rates and large uncertainties persist with the sequencing results. Here, we employ a nano-disk model to quantitatively analyze the sequencing process by examining the variations of ionic current when a DNA strand translocates a nanopore. Our focus is placed on signal-boosting and noise-suppressing strategies in order to attain the single-nucleotide resolution. Apart from decreasing pore diameter and thickness, it is crucial to also reduce the translocation speed and facilitate a stepwise translocation. Our best-case scenario analysis points to severe challenges with employing plain nanopore technology, *i.e.*, without recourse to any signal amplification strategy, in achieving sequencing with the desired single-nucleotide resolution. A conceptual approach based on strand synthesis in the nanopore of the translocating DNA from single-stranded to double-stranded is shown to yield a 10-fold signal amplification. Although it involves no advanced physics and is very simple in mathematics, this simple model captures the essence of nanopore sequencing and is useful in guiding the design and operation of nanopore sequencing.

Keywords: nanopore, DNA sequencing, ionic current, model, series resistance, noise, signal

1. Introduction

Precision Medicine (PM) promises a forthcoming revolution in healthcare by taking into account individual variabilities in genes, environment and lifestyle for each person in order to dramatically improve on disease treatment and prevention.[1] A backbone technology for PM applications is DNA sequencing, *i.e.*, the process to determine the succession of nucleotides (nts) on a DNA strand. Enabled by the vast development of massively parallelized sequencing (MPS) technologies,[2,3] sequencing has already become a standard aid in basic biomedical research and begins to revolutionize the medical practice.[1,4] Sequencing also finds other applications such as in anthropology,[5] agriculture,[6] and criminology.[7],[8-11] The development of sequencing technologies has led to a sharp cost reduction by more than four orders of magnitude in just 15 years.[8] Existing MPS technologies are either optics[11] or semiconductor[13] based, but they all rely on the use of large quantities of short DNA copies usually produced by means of polymerase chain reaction (PCR). However, the PCR reproduction not only adds cost but also causes sequence errors. Moreover, the *de novo* assembly using sequence data of short DNA strands increases both cost and error frequencies. Hence, DNA sequencing has been continuously searching for technologies that provide low error rates and handle long DNA strands at high throughput and low costs.

Sequencing based on handling single DNA strands is intuitively advantageous since no PCR is necessary and very long DNA strands are possible. In this work, nanopore technology for sequencing single DNA strands by means of monitoring variations in ionic current is theoretically scrutinized in order to provide guidelines for design and operation with the technology. It should be noted that other technologies, prominently that based on zero-mode waveguides,[14] for sequencing single DNA strands also are available. Detecting tunneling current in various nanostructures [15-21] or Fano resonance [22-24] in two-dimensional (2D) materials including graphene [25,26] has also been exploited as alternative means for single-DNA sequencing. Use of nanopores for analyzing DNA strands started in the 1990's.[27,28] The analysis is accomplished by associating the minute changes in ionic current through the nanopore to the nts on the DNA strand translocating the pore. This innovative concept has motivated numerous follow-up efforts with impressive achievements and breakthroughs.[29,30] Especially noteworthy is the realization of sequencing single-stranded DNA (ssDNA) by employing α-hemolysin (α-HL) bio-nanopores,[30,31] which has led to a prototype product MinION currently with an error rate around 15%.[32] Sequencing ssDNA has also been achieved using Mycobacterium smegmatis prion A (MspA), instead of α-HL, with a remarkable detection length of 4500 nts although the error rate is high at 23%.[33] However, bio-nanopores with lipid membrane are in general not robust and durable.[29] They may denature at high temperature and high salt concentrations. Furthermore, the size of α-HL and MspA pores is fixed.[29,34,35] Attention has, therefore, been turned to solid-state nanopores that can be manufactured using standard silicon nanofabrication technology for micro/nanoelectronics.[29] Apart from a potentially easily tunable pore size, high-density solid-state nanopores of controlled properties can be fabricated by exploiting the large vocabulary of materials and materials combinations available with silicon technology. The nanopore arrays can be naturally integrated with

electronic circuitry for on-chip signal processing. Nanopores in ultrathin SiN_x and SiO_2 membranes were the first choice to explore,[36-41] owing to their good controllability[42] and processability.[43,44] However, membranes of these dielectric materials are hard to reach a thickness matching the distance of two adjacent nts on an ssDNA, 0.8 nm,[45] with good homogeneity and without pinholes. Two-dimensional materials such as graphene[25,46,47] and transition metal dichalcogenide MoS_2[48,49] have caught increasing attention due to the excellent stability and processability in their natural single-layer form of 0.5-0.7 nm thickness. But achieving low background noise and single-nt resolution still remains a formidable challenge.

The nanopore sequencing primarily relies on monitoring the ionic current through the pore. A fairly good understanding of the mechanism pertaining to variations in ionic current caused by DNA translocation has been established and it gives valuable insights for general nanopore design and operation.[29] But much remains to be confirmed with respect to the influence on ionic current of pore size, size of nts, translocation speed and manner, sampling rate and bandwidth, noise, morphology (position and orientation) of nts inside and nearby the pore, ionic strength and nature of ions in the two electrolyte reservoirs, membrane properties, etc. The nanopore sequencing builds on the assumption that the variations in ionic current are predominantly, if not solely, determined by the differences in nts. Theoretical studies, having recourse to comprehensive computational approaches such as molecular dynamics, have made tremendous progress in elucidating the mechanistic details about how the ionic current can be affected by the various aforementioned parameters.[50-52] However, such numerical approaches are extremely resource-demanding if all possibilities enlisted above are included. As a complementary approach, we focus in this work on a simple nano-disk model to be detailed momentarily below. This treatment, though very simple, captures the most important features of nanopore sequencing. The model treats each nt on the translocating DNA strand as a well-defined nano-disk of infinite resistance. This simplification allows for a systematic scrutiny of the variations in ionic current during the DNA translocation in the nanopore. This model is implemented in MATLAB and the simulation calculates the blockage ionic current under various input conditions. It allows for a quantitative evaluation of how most of the aforementioned parameters could affect the waveform and thereby sequencing outcome. It further assists to outline the design criteria for future nanopore solutions.

2. Model

Nanopores are typically below 10 nm in diameter constructed by proteins[34] or fabricated in solid-state membranes.[29,53] Although the model does not make a distinction between the two families of nanopores, it mainly refers to solid-state ones with the schematic representation in figure 1(a) defining some critical parameters for the model building. In operation, a voltage bias is applied between the two electrolyte reservoirs separated by the membrane and linked via the nanopore. The ionic current through the pore is measured. The nanopore region has a large pore resistance, R_p. In the vicinity of the entrance and exit of the nanopore, there are two hemispherical regions each with a relatively high resistance compared to that of the bulk electrolyte. This is caused by a reduction of ion diffusion near the pore mouth and is termed access resistance, R_a.[54] The total resistance of nanopore

system, $R_{t,o}$, becomes:[55]

$$R_{t,o} = R_{p,o} + 2R_{a,o} \tag{1}$$

$$R_{p,o} = \frac{4h}{\sigma\pi d_p^2} \tag{2}$$

$$R_{a,o} = \frac{1}{2\sigma d_p} \tag{3}$$

$$\sigma = q(c_+\mu_+ + c_-\mu_-) \tag{4}$$

where, d_p and h are, respectively, the diameter and thickness of the pore, σ the conductivity of bulk electrolyte, q the elementary charge, c_+ and c_- the ion concentration of cation and anion, and μ_+ and μ_- the mobility of cation and anion. The second subscript "o" is included to clarify that the resistances are the "open-pore" ones. Equation (3) is valid for cases where the diameter of the access region (d=2r, with the definition of r in figure 1(a)) is equal to that of the pore, *i.e.*, d=d_p, which is also a simplification adopted in the remainder of this work.

When an ssDNA strand translocates the nanopore, the pore is partially blocked and the ionic current is reduced. The degree of current blockade depends on a number of parameters and factors including the size of the nts on the translocating DNA strand. In terms of resistance changes in the pore and access regions, only one nt is considered first. No change in nt morphology is assumed during the translocation and no interaction between adjacent nts is considered. The resistance changes ΔR_p and ΔR_a caused by one nt are determined by the steric blocking of ion movement[55] and can be expressed as the following:

$$\Delta R_p = R_{p,b} - R_{p,o} = \frac{4d_s d_D^2}{\pi\sigma(d_p^2 - d_D^2)d_p^2} \tag{5}$$

$$\Delta R_a = R_{a,b} - R_{a,o} = \frac{\pi d_s d_D^2}{\sigma d_p^2(4d_p^2 - \pi d_D^2)} \tag{6}$$

With the second subscript "b", $R_{p,b}$ and $R_{a,b}$ represent the pore and access resistance when one nt is inside the respective regions. The other parameters are d_D and d_s as the diameter of the nt and distance between two adjacent nts, respectively. As no space is assumed between the nts, d_s is also the thickness of the nts. Finally, the formalism comprising equations (1)-(6) leads to a very simple representation of the total resistance of the nanopore system, $R_{t,b}$, by summing up all the resistance contributions of the nts of a translocating ssDNA:

$$R_{t,b} = R_{t,o} + \sum \Delta R_p + \sum \Delta R_a \tag{7}$$

For a fixed bias voltage U, the open-pore current $I_o = \frac{U}{R_{t,o}}$ and the blockage (residual) current $I_b = \frac{U}{R_{t,b}}$. In the remainder of this work, we will focus on how the normalized

blockage current:

$$\frac{I_b}{I_o} = \frac{R_{t,o}}{R_{t,b}} \tag{8}$$

will evolve with several aforementioned parameters. The computation is realized by implementing the model in the MATLAB environment and it does not require extensive resources.

3. Results and discussion

In the MATLAB simulation, the parameters are first assigned. The open-pore resistance $R_{t,o}$ is calculated. An ssDNA with a certain sequence is then set at one side of the nanopore and it is aligned to point to the pore. The front end of the strand is 6 nm away from the nanopore mouth. For a given translocation speed, the position of every nt on the ssDNA strand at any time can be identified. Then, the nts inside the pore and the access region are picked up and marked. Summing up the resistance increase contributed by these nts, we get the total resistance change at a certain time point. The normalized blockage current at each time point is calculated and the waveform of the translocation event can be plotted.

In order to obtain the physical dimensions of the nts, the following is performed. Since the bases in the nts can be regarded as a plane molecule, the cross section of each base is determined by the summation of its van der Waals area of all atoms. The cross section of the pentose and phosphate is estimated by the projected area in the direction perpendicular to the base plane. From the cross section area, we can calculate the effective diameter of the four nts as what has been done in the literature.[19,56] The parameters used in our simulation are presented in Table 1.

3.1. Waveform of I_b/I_o

The waveform of I_b/I_o is expected to be sensitive to many of the aforementioned parameters, *e.g.*, pore diameter and thickness, translocation speed, sampling rate, etc. Here, we analyze I_b/I_o for the translocation of an arbitrary 48-nt ssDNA with the sequence AAATTTGGGCCCATGCATGCATGCATGCAAATTTGGGCCCATGCATGCATGCATGC. By convention, A, T, G, and C represent the nts adenine, thymine, guanine, and cytosine, respectively. In the calculation, the noise level is set at 0.1% of I_o, which is a reasonable value as will be discussed in section 3.3 and section 3.5. We assume a colored Gaussian type of noise with a certain power spectral density (PSD). It is generated from the white Gaussian noise (Supplementary Information figure S1) and will be discussed in detail in section 3.3. Unless specified otherwise, all current and noise signals are normalized to I_o in this work.

The calculation I_b/I_o waveform is summarized in figure 1(b). The translocation is assumed to proceed stepwise, one nt at a time, instead of continuously at a constant speed. In the upper row, I_b/I_o is depicted for d_p=2, 3, 4, and 6 nm, at h=5 nm and translocation speed Sp=1 nt/ms, and with a sampling rate Sr=10 kHz. In the first 6 ms, I_b/I_o remains constant as the DNA strand has not reached the pore. Increasing d_p drastically increases I_b, as expected. For large-diameter nanopores, the very small signals, represented by the fluctuation

superimposed on the I_b/I_o section in the time interval 6-63 ms, become difficult to distinguish from the noise (right figure). In the middle row, I_b/I_o is shown for h=1, 2, 5, and 10 nm, at d_p=2 nm, and with the same Sp=1 nt/ms and Sr=10 kHz. Only a slight decrease in I_b/I_o with increasing h is observed in the time interval 6-63 ms. It is remarkable that the small signals superimposed on the I_b/I_o section in the time interval 6-63 ms are clearly discernable with the most aggressive pore dimensions h=1 nm and d_p=2 nm (left figure). Increasing h blurs the signals that eventually become featureless and undistinguishable from the noise (right figure). Finally in the lower row, I_b/I_o is examined in a close-in format for Sp=1, 5, 10, and 100 nt/ms at d_p=2 nm, h=1 nm, and Sr=10 kHz. As expected, increasing Sp leads to a much squeezed time window for I_b, see the respective insets. At Sp>10 nt/ms, which is inadequate for the 10 kHz sampling rate, the I_b/I_o waveform becomes severely distorted.

The simulation of I_b/I_o waveform provides a general picture of ion blockade caused by the ssDNA translocation. The reported trends in figure 1(b) are all within expectation. It also indicates that the model, though extremely simple in physics and mathematics, can indeed capture the essential features of nanopore sequencing by analyzing blockage ionic current.

3.2. Range of signal

It is worth recalling that the small signals superimposed on the I_b/I_o section in the time interval 6-63 ms are generated by the differences in the nts. It is, therefore, such signals that are useful for identification of the nts on a translocating ssDNA strand. The shape of the I_b/I_o waveform in the current blockade region, *e.g.*, from 6 to 63 ms in figure 1(b), should be carefully analyzed. As an example, the zoomed-in I_b/I_o at the beginning of ssDNA translocation, in a nanopore of d_p=2 nm and h=2.4 nm, in figure 2(a) shows how stepping different nts into the pore generates staircase-like current levels, not peaks or valleys. The choice of h=2.4 nm is to simultaneously accommodate 3 nts in the pore. The test ssDNA has a sequence of AAATTTGGGCCCATGC. To help illustrate the sequencing process, cartoon pictures are inserted at the various staircase levels. In order to simplify the analysis, R_a is excluded in the first calculations. Each step between two successive current levels is, in fact, primarily determined by the difference between two nts, one entering and one exiting the pore. As anticipated, the size of the step is proportional to the difference between these two nts. The size of the step is also sensitively dependent on the number of nts in the pore region. It is apparent that a successful sequencing leading to decoding of a DNA needs to resolve such small steps.

Since the model is simply geometry-based, the largest current step should involve the coupled events with C-in (or G-in) and G-out (or C-out) with respect to the nanopore, because the size difference between C and G is the largest, see Table 1. To achieve the largest step, the nts in the pore region should all be C. Under such circumstances, the current step can be calculated by:

$$\Delta I_{\max} = | I_{\underbrace{CCCC\ldots C}_{n-1}} - I_{\underbrace{CCC\ldots CG}_{n-1}} | \tag{9}$$

where, n is the number of nts in the pore region. $I_{N\ldots N}$ (N=A, T, G, C) represents the blockage ionic current when the section of an ssDNA with sequence N…N is found in the pore. After certain mathematical rearrangements, equation (9) becomes,

$$\Delta I_{\max} = \frac{nI_C^2}{[I_C + (n-1)I_C][I_C + (n-1)I_G]} | I_C - I_G | \quad (10)$$

where, I_N (N=A, T, G, C) is the blockage ionic current when an ssDNA with an N homopolymer is in the pore. Similarly, the smallest current step is found for the coupled events with A-in (or G-in) and G-out (or A-out) along with all nts in the pore region being G. Mathematically, the following is obtained:

$$\Delta I_{\min} = | I_{A\underbrace{GGG...G}_{n-1}} - I_{\underbrace{GGG...GG}_{n-1}} | \quad (11)$$

$$\Delta I_{\min} = \frac{nI_G^2}{[I_G + (n-1)I_G][I_G + (n-1)I_A]} | I_A - I_G | \quad (12)$$

After normalization of $\Delta I_{\max}$ and $\Delta I_{\min}$ to I_o, the distribution of the normalized signal $\Delta I/I_o$ (ΔI to represent both $\Delta I_{\max}$ and $\Delta I_{\min}$) with d_p and h is shown in figure 2(b-d). The two curved surfaces in figure 2(b) define the upper and lower bounds of the normalized signal, *i.e.*, the relative change in blockage ionic current, when the nts are stepwise fed into the nanopore region. The two-dimensional projection of these two surfaces is shown in figure S3 of Supplementary Information. As expected, $\Delta I/I_o$ is larger with smaller d_p and/or h; the best being that with both d_p and h very small. The converse is true that the decrease in $\Delta I/I_o$ is sharp with increasing both d_p and h, thus rendering the identification of the steps (*i.e.*, normalized signals) a difficult task. To appreciate the simulation results, it is important to bear in mind that the ionic current of an unblocked nanopore of sub-5 nm diameter is usually below 10 nA.[54,57,58] Therefore, a successful sequencing should avoid large pore sizes (d_p and h) so as to focus on signals (ΔI) in the 10-100 pA range.

To confirm that equations (10) and (12) define, respectively, the maximum and minimum normalized signals (steps), translocation results of a randomly generated 10000-nt sequence are simulated. The black squares with error bars in figure 2(c-d) represent the statistical results of the 10000 samples from the simulation. They all lie in-between the two curves that are two-dimensional projections of figure 2(b) at d_p=4 nm (2c) and h=2.4 nm (2d). The observation of a constant $\Delta I/I_o$ for h<0.8 nm in figure 2(c) is simply due to d_s=0.8 nm (Table 1), *i.e.*, no change in blockage current is expected by this model if the pore thickness is smaller than the nt thickness 0.8 nm since there is always only one nt found in the pore.

The simulation results in figures 1 and 2 do not include the influence of R_a (access resistance). In adverse cases, R_a can dominate the total resistance, especially for very thin nanopores. When R_a is considered for the case with $2r=d_p$ (cf. figure 1(a)), an overall decrease in $\Delta I/I_o$ is evident in figure 3(a) in comparison with figure 2(b). The decrease is the worst for pores of both large d_p and large h and the amplitude of decrease can be substantially more than 10 times. Apart from speed, the manner of translocation also plays a role in affecting the step size, $\Delta I/I_o$. Thus far, a step-like nt feeding scheme (to the pore region) is assumed, as illustrated by the inset of figure 3(a). In reality, such an ideal stepwise feeding is hardly encountered without special arrangements. Such an arrangement will be discussed in section 3.5. For a uniform and continuous translocation at a constant speed, which is rather common in solid-state nanopores, the waveform of one nt translocating the pore is illustrated in the insert of figure 3(b). This continuous translocation leads to a further overall

decrease in $\Delta I/I_o$ in figure 3(b), in comparison to figure 3(a). The simulation in figure 3(b) has also considered R_a. The continuous translocation can be regarded as being equivalent to increasing R_a. Sequencing with the consideration of R_a and non-stepwise nt feeding would demand a current measurement system that has a wider dynamic range, higher resolution and smaller noise level.

3.3. Bandwidth and noise level

A key performance factor of an ionic current measurement system is noise character, which can be characterized by its signal to noise ratio (SNR). By adopting a widely used signal extraction algorithm CUSUM,[59] we have found that the signal can be extracted without distortion when its SNR is larger than 4 (see figure S2 and related discussion in Supplementary Information). By accumulating the difference between a signal and its estimated level, CUSUM can effectively discern abrupt changes in signal level. When the nanopore setup and the measurement system are considered as a whole, noise mainly comes from the nanopore-electrolyte system.[60-62] The noise level of the measurement system can be very low when proper measures are taken.[61] In general, use of a wider bandwidth of a system invites a higher noise level.[61,63] However, the choice of bandwidth is not without constraints; it should be larger than the frequency of the nt translocation (*i.e.*, reciprocal of the translocation speed). It is difficult to make bandwidth very small by utilizing filters in the measurement circuit, since, otherwise, details in ΔI or $\Delta I/I_o$ risk to be filtered out as well. The noise PSD in pA^2/Hz of a nanopore can be described as:[64-66]

$$S_I = a_1 + a_2 f + a_3 f^2 + a_4 \frac{1}{f} \tag{13}$$

where, f is frequency and a_i (i=1, 2, 3, 4) are four constants specific of the nanopore-electrolyte system. The first term a_1 describes the intensity of thermal white noise, which comes from the resistance of bulk electrolyte especially in the access and pore region that have a relatively high resistance.[64] The second and third terms account for high-frequency noise caused by parasitic capacitances and their distribution on the membrane.[64,65] The origin of the $1/f$ noise, the last term, is still under debate. Non-equilibrium dynamics of charged carriers in the electrolyte in the pore,[67] surface charge fluctuation,[45] and salt concentration variations in the pore region[68] may all contribute to the $1/f$ noise.

To facilitate the discussion leading to guidelines for nanopore design, we choose a set of typical values for these parameters:[64] $a_1=2.2\cdot10^{-4}$ pA^2/Hz, $a_2=8\cdot10^{-8}$ pA^2/Hz^2, $a_3=5.7\cdot10^{-12}$ pA^2/Hz^3, and $a_4=1$ pA^2. With the PSD-f relation shown in figure 4(a), the bandwidth dependency of the noise level, in terms of root mean square (RMS) of the noise, is obtained in figure 4(b). As expected, the noise level increases monotonously with frequency. In order to achieve a successful sequencing, the noise level shown in figure 4(c) should be at least four times smaller than the lower bound of the normalized signal $\Delta I/I_o$ in figure 2(b). In other words, a successful sequencing without being influenced by noise can be achieved if the system has a noise level below this curved surface in figure 4(c).

Moreover, the requirements for the operational bandwidth can be obtained by combining the information in figure 4(b) with 4(c). The results are given in figure 4(d).

The bandwidth of a system should be smaller than the specified value for a given pore size in order to ensure that the noise is below the surface in figure 4(c). Furthermore, the frequency of nt translocation should be kept below the bandwidth. For good resolution and acquisition of enough details, the translocation frequency is suggested to be 3-5 times smaller than the bandwidth.[61]

3.4. Nanopores in MoS_2 monolayer

The nano-disk model is validated by analyzing the recently reported very extensive experimental data on ssDNA sequencing with nanopores in a MoS_2-monolayer membrane.[49] A MoS_2 monolayer is 0.7 nm thick.[69] Blockage ionic currents are available for the four nts translocating a series of MoS_2 nanopores of diameter ranging from 2 nm to 6.1 nm.[49] From the given open-pore current of 9.4 nA at 200 mV for the nanopores of 4 nm diameter, the open-pore currents for other pore diameters can be calculated according to equations (1-4) and (8). A noise RMS value of 89 pA at 200 mV is also given.[49] From the blockage current data, the diameter of the four nts can be found and the average values along with the standard deviations are displayed in Table 2. Small variations, less than 10%, are obtained, which supports the model as well as the applicability of equations (5-7). These parameters facilitate a direct comparison in figure 5(a) between the experimentally measured blockage currents for the four deoxynucleotide monophosphates (dNMP, with N=A, T, G, C) and the calculated ones according to the model. For each dNMP case, variations (shown as gray bands) are calculated by considering measurement errors and uncertainties provided in the reported work.[49] An excellent agreement is attained except for the 6.2 nm case, which provides values closer to a 4.6 nm pore of the model.

The nt diameters extracted from the experimental data for MoS_2 in Table 2 significantly differ from their steric diameters determined from the van der Waals areas in Table 1. The current blockade is not just caused by the steric size of the nt. It is also affected by a list of other factors such as charge distribution on ssDNA,[51,70,71] induced surface charge on the pore sidewall,[72,73] interaction of the translocating ssDNA with the pore,[19,74] influence of the outer electrical filed on the electric double layers on the nanopore sidewall and around the DNA strand,[75,76] mechanical flexibility of ssDNA strands,[77] etc. Furthermore, these non-steric effects can have a dominant influence on the total current blockade.[19] The diameters extracted from the experiments[49] have inherently already considered all such effects and they can be termed effective diameters. To evaluate the overall contribution of the non-steric effects to the total current blockade, a dimensionless parameter is introduced:

$$m = \frac{d_{eff}^2 - d_{st}^2}{d_{eff}^2} \tag{14}$$

where, d_{eff} and d_{st} are the effective and steric diameter, respectively. Squared quantities are used since it is conductance or blockage current that is relevant, and m=-1.3, -1.7, -0.8, and -2.1 for A, T, G, and C, respectively. A couple of observations can be made. First, the minus sign indicates that the non-steric effects oppose the effect of the steric blocking, *i.e.*, they tend to weaken the blocking of the ionic current. Second, the absolute size of m inversely follows the size of the nts, and it is the largest for the smallest C and the smallest for the largest G. Hence, the non-steric effects are the strongest for C.

The normalized signal $\Delta I/I_o$ versus d_p shown in figure 5(b) clearly indicates the importance with a small h (only 0.7 nm for MoS_2), since the $\Delta I/I_o$ here are about 10 times those in figure 2(d) with h=2.4 nm. Similarly, both ΔI_{max} and ΔI_{min} are included and each with its variations calculated again by considering the experimental errors and uncertainties.[49] By applying the same criterion of SNR=4 above which the signal can be extracted without distortion, the upper and lower bounds for noise requirement are depicted in figure 5(c). The experimental noise data as black squares, after normalization, are also included in figure 5(c) for comparison. For the small diameters d_p=2 and 2.3 nm, the experimental noise lies between the upper and lower bounds. Thus, the signal whose noise is above the black squares but below the upper bound can be distinguished from the background noise. Otherwise, the signal whose noise is below the black squares and although above the lower bound becomes buried in the noise. For the large diameters d_p>2.8 nm, the black squares are all above the required range and the signal will be completely distorted by the noise, thus rendering no chance to attain any useful signal.

The model can also predict the possibility of DNA sequencing using nanopores in MoS_2-multilayer membranes that are above 0.7 nm in thickness. The upper and lower bounds of $\Delta I/I_o$ are shown in figure 5(d) for different diameter and thickness combinations. The trend resembles that in figure 2(b), as expected, although absolute values can be slightly different.

3.5. Way forward

The simulation above is based on an over-simplified rigid nano-disk model and has neglected a number of subtle yet potentially critical details such as the complex morphology of the translocating DNA strand and the intricate ionic environment around the nanopore. The simulation results are, hence, considered to represent the best-case scenarios. In spite of the over-simplification nature of the model, the simulation results agree, qualitatively at least, well with what has been established in the literature with respect to nanopore DNA sequencing based on analysis of ionic current.[19,29,34,53,78] Signal and noise are the two key considerations to strive for, as expected, and the design of pore size, translocation speed, membrane design, etc., needs to conform with them. With its analytical and quantitative nature, the simple nano-disk model reveals some technical details that can be helpful for the design and operation of nanopore technologies. To begin with, it is a very small number of nts on a DNA strand that contribute to and determine the total current blockade. It includes nts in both pore and access regions. The fewer nts involved, the clearer and more distinguishable signal is achievable. Ultrathin membranes are, then, desirable as the starting material for nanopore formation, as clearly seen in the middle row of figure 1(b), figure 2(b-d), figure 3, and figure 5(b) and 5(d). This is reflected by the migration from SiN_x and SiO_2 of tens of nanometers to graphene and MoS_2 monolayers of sub-nm thickness. However, high noise with graphene nanopores has kept this approach from reaching the single-nt resolution.[46,57] Graphene is flexible and tends to strongly interact with DNA,[74,77] both of which can be the source of high noise.[46,79] The noise consideration has called for the study of MoS_2 monolayers as a low-noise material in very resent years and significant progress has already been made.[48,49,80] For instance, MoS_2 nanopores have demonstrated the ability to distinguish the four nts in single monomers or homopolymer

ssDNA strands[49] although sequencing a natural DNA still awaits further progress.

In general, research in solid-state nanopores still drags behind that in bio-nanopores; the latter has already achieved sequencing of natural ssDNA strands although error rates remain high.[30,32,33] Much can be learnt from the success of the bio-nanopore technology and a few of them are discuss as follows.

(a) Minimizing the contribution of R_a via biochemical approaches leading to the dominance of R_p is central to the success with the bio-nanopores.[31,32,81,82] In contrast, high R_a persists for solid-state nanopores especially for those of 0.5-0.7 nm thickness.[55,79] The detrimental effect of R_a on suppressing the sequencing signal is clearly shown in figure 3(a) in comparison with figure 2(b).

(b) A low translocation speed translates to a small bandwidth of the system and thereby a low noise level. The ultrahigh-speed translocation in solid-state nanopores[83] can be substantially slowed down either by viscosity control[49] or by charge control via counterion binding to the DNA strand[84] or photo-induced charge on the pore surface.[85] However, also of importance is the stepwise movement that generates the desired step-like change in blockage ionic current and therefrom a significantly enhanced signal, as demonstrated in figure 3. Thus, an enzyme is incorporated in a bio-nanopore to both control the translocation speed and to make the DNA movement stepwise.

(c) The natural anti-adsorption property of a protein surface could be the reason for a much low noise level achieved with bio-nanopores.[29,64] On the contrary, the surface of solid-state membranes can easily adsorb ions, small molecules, etc. The adsorption-desorption process on the surface can induce large noise,[60] which is also regarded as parasitic capacitive effects.[65,68] Moreover, the sidewall of solid-state nanopores can induce charge in an electrolyte. It can disturb the distribution of ions in the electrolyte and cause the electrolyte conductivity to fluctuate, resulting in a $1/f$ type of noise.[66] One approach to decreasing such noise is to coat the solid-state surface with an organic polymer layer. This approach has indeed been shown effective.[64]

Overall, sequencing using solid-state nanopores is characterized by a lower signal and more blurry signal changes than using bio-nanopores. Solid-state nanopores also suffer from a more noisy background than bio-nanopores. The simulation results in figure 3(b) are referred to for an idea where the ballpark could be. For I_o=10 nA, which is a normal open-pore current, the minimum signal change ΔI_{min} is only 0.1 pA for a nanopore of d_p=5 nm and h=5 nm. This pore size is achievable with MoS_2 monolayers but difficult to realize with SiN_x membranes.[53,86,87] However, the noise level for solid-state nanopores is seldom smaller than 1 pA,[61,64,65,88] which is also seen in figure 4(b). Even we consider a more ideal case by referring to figure 3(a), ΔI_{min}=1 pA is found for the same nanopore of d_p=5 nm and h=5 nm. In order to achieve a reliable sequencing using solid-state nanopores, suppressing noise is highly necessary. Concurrently, the signal in form of ΔI should be substantially enhanced.

A conceptual approach to boosting the signal is considered in figure 6. It involves DNA strand synthesis from single-stranded (ss) to double-stranded (ds) facilitated by a DNA polymerase such as enzyme phi29 DNA polymerase[33,89,90] or polynucleotide phosphorylase.[91] The translocation becomes stepwise and the DNA strand move one base at a time in accordance to the completion of the strand synthesis from ss to ds. This strategy

forms the base for several commercial sequencing technologies such as semiconductor genome sequencing[33] and zero-mode waveguides.[36] The strand synthesis further engages a sequential supply of the four deoxynucleotide triphosphates (dNTPs), *i.e.*, each time only one specific type of the dNTPs is provided as a probe to the translocating ssDNA strand. This arrangement will allow us to concentrate on signal detection since identification of the nts on the translocating DNA strand will be realized by correlating the known probing dNTP to the signal detected, precisely as in the semiconductor genome sequencing.[33] The normalized signal $\Delta I/I_o$ is studied in three representative cases in figure 6, with synthesis occurring: (a) at the entrance on the *cis* side, (b) in the middle, and (c) at the exit on the *trans* side, of the nanopore. For simplicity, the simulation uses the following parameters: diameter for ssDNA 1 nm, diameter for dsDNA 2 nm,[92] thickness for both ssDNA and dsDNA 0.8 nm, Sp=1 nt/ms, and Sr=10 kHz. Overall, this strategy leads to a 10-fold signal boost in comparison with the ordinary nanopore sequencing with a plain pore translocation of ssDNA or dsDNA strands. In detail, the precise location at which the synthesis occurs plays a crucial role. The case in figure 6(c) is obviously the most effective in amplifying the signal, as expected. While d_p is a rather sensitive parameter differentiating the three cases with respect to $\Delta I/I_o$, h plays a minor role. An experimental realization of this amplification strategy requires immobilization of a DNA polymerase onto or inside the nanopore, which poses significant technological challenges. Nonetheless, the simulation results are encouraging and may motivate further proposals to boost the signal.

4. Conclusions

We have established a simple nano-disk model to simulate ionic current changes in a nanopore when a DNA strand translocates it. The model is implemented in the MATLAB environment. In the model, each nucleotide on a single-stranded DNA strand is assumed as a nanoscale disk of a well-defined size (diameter and thickness) available in the literature. An ssDNA strand is, therefore, composed of a series of disks of varying diameters. The simplicity nature of the model allows for a quick, quantitative assessment of DNA sequencing under various combinations of a large number of parameters. The parameters include nanopore diameter and thickness, nucleotide diameter and thickness, ionic strength and ion types in electrolytes, translocation speed and manner, sampling frequency, background noise, and bandwidth. The model is used to analyze the literature experimental results on nucleotide sequencing using nanopores in monolayer-MoS_2 membranes and an excellent agreement in terms of signal and noise performance is found. However, a best-case scenario analysis with an ideal combination of the aforementioned parameters unveils severe limitations, with respect to achieving the needed single-nucleotide resolution, of the plain nanopore sequencing practice without recourse to any signal amplification scheme. The model is subsequently used to examine a conceptual approach aiming at signal amplification. Despite its rudimentary character involving no advanced physics and with extremely simple mathematics, the model is robust as it captures the basic features of nanopore sequencing and is found to be useful in outlining strategies towards improving the DNA sequencing technology.

Acknowledgements

This work was supported by the Swedish Research Council (621-2014-6300). C.W. was a scholarship recipient of the China Scholarship Council.

References:

[1] Collins F S and Varmus H 2015 *The New England Journal of Medicine* **372** 793

[2] Mardis E R 2011 *Nature* **470** 198

[3] Green E D, Watson J D and Collins F S 2015 *Nature* **526** 29

[4] Green E D, Guyer M S and Institute National Human Genome Research 2011 *Nature* **470** 204.

[5] O'Rourke D H, M. Hayes G and Carlyle S W 2000 *Annual Review of Anthropology* **29** 217

[6] Liu G E 2009 *Recent Patents on Food, Nutrition & Agriculture* **1** 75

[7] Kuder M, Ling S T W and Li S K L 2011 *Singapore Law Review* **29** 35

[8] Voelkerding K V, Dames S A and Durtschi J D 2009 *Clinical Chemistry* **55** 641

[9] Shokralla S, Spall J L, Gibson J F and Hajibabaei M 2012 *Mol. Ecol.* **21** 1794

[10] Morozova O and Marra M A 2008 *Genomics* **92** 255

[11] Mardis E R 2008 *Annu Rev. Genomics Hum. Genet.* **9** 387

[12] www.genome.gov/sequencingcosts

[13] Rothberg J M *et al* 2011 *Nature* **475** 348

[14] Levene M J, Korlach J, Turner S W, Foquet M, Craighead H G and Webb W W 2003 *Science* **299** 682

[15] Ivanov A P, Instuli E, McGilvery C M, Baldwin G, McComb D W, Albrecht T and Edel J B 2011 *Nano. Lett*. **11** 279.

[16] Tsutsui M, Taniguchi M, Yokota K and Kawai T 2010 *Nat. Nanotech.* **5** 286

[17] Prasongkit J, Feliciano G T, Rocha A R, He Y, Osotchan T, Ahuja R and Scheicher R H 2015 *Sci. Rep.* **5** 17560

[18] Sarma S D, Adam S, Hwang E H, Rossi E 2011 *Rev. Mod. Phys.* **83** 407

[19] Zwolak M and Ventra M D 2008 *Reviews of Modern Physics* **80** 141

[20] Lindsay S 2016 *Nat. Nanotech.* **11** 109

[21] Taniguchi M and Ventra M D 2016 *Nat. Nanotech.* **11** 117

[22] Min S K, Kim W Y, Cho Y and Kim K S 2011 *Nat. Nanotech.* **6** 162

[23] Rajan A C, Rezapour M R, Yun J, Cho Y, Cho W J, Min S K, Lee G and Kim K S 2014 *ACS Nano* **8** 1827

[24] Thomas S, Rajan A C, Rezapour M R, Kim K S 2014 *J. Phys. Chem. C* **118** 10855

[25] Heerema S J and Dekker C 2016 *Nat. Nanotech.* **11** 127

[26] Arjmandi-Tash H, Belyaeva L A and Schneider G F 2016 *Chem. Soc. Rev.* **45** 476

[27] Kasianowicz J J., Brandin E, Branton D and Deamer D W 1996 *Proc. Natl. Acad. Sci. USA,* **93** 13770

[28] Akeson M, Branton D, Kasianowicz J J, Brandin E and Deamer D W 1999 *Biophysical J.* **77** 3227

[29] Muthukumar M, Plesa C and Dekker C 2015 *Physics Today* **68** 40

[30] Bayley H 2015 *Clin. Chem.* **61** 25

[31] Clarke J, Wu H-C, Jayasinghe L, Patel A, Reid S and Bayley H 2009 *Nat. Nanotechnol.* **4** 265

[32] Jain M, Fiddes I T, Miga K H, Olsen H E, Paten B and Akeson M 2015 *Nat. Methods* **12** 351

[33] Laszlo A H *et al* 2014 *Nat. Biotechnol.* **32** 829

[34] Deamer D W and Branton D 2002 *Acc. Chem. Res.* **35** 817

[35] Howorka S and Siwy Z 2009 *Chemical Society Reviews* **38** 2360

[36] Li J, Gershow M, Stein D, Brandin E and Golovchenko J A 2003 *Nat. Mater.* **2** 611

[37] Chang H, Kosari F, Andreadakis G, Alam M A, Vasmatzis G and Bashir R 2004 *Nano Lett.* **4** 1551

[38] Fologea D, Gershow M, Ledden B, McNabb D S, Golovchenko J A and Li J 2005 *Nano Lett.* **5** 1905

[39] Sigalov G, Comer J, Timp G and Aksimentiev A 2008 *Nano Lett.* **8** 56

[40] Gershow M and Golovchenko J A 2007 *Nat. Nanotechnol.* **2** 775

[41] Iqbal S M, Akin D and Bashir R 2007 *Nat. Nanotechnol.* **2** 243
[42] Kim M J, Wanunu M, Bell D C and Meller A 2006 *Adv. Mater.* **18** 3149
[43] Li J, Stein D, McMullan C, Branton D, Aziz M J and Golovchenko J A 2001 *Nature* **412** 166
[44] Storm A J, Chen J H, Ling X S, Zandbergen H W and Dekker C 2003 *Nat. Mater.* **2** 537
[45] Hamai C, Tanaka H and Kawai T 2000 *J. Phys. Chem. B* **104** 9894
[46] Merchant C A *et al* 2010 *Nano Lett.* **10** 2915
[47] Nelson T, Zhang B and Prezhdo Ol V 2010 *Nano Lett.* **10** 3237
[48] Liu K, Feng J, Kis A and Radenovic A 2014 *ACS Nano* **8** 2504
[49] Feng J, Liu K, Bulushev R D, Khlybov S, Dumcenco D, Kis A and Radenovic A 2015 *Nat. Nanotechnol*. **10** 1070
[50] Comer J and Aksimentiev A 2012 *J. Phys. Chem. C: Nanomater Interfaces* **116** 3376
[51] Bond P J, Guy A T, Heron A J, Bayley H and Khalid S 2011 *Biochemistry* **50** 3777
[52] Aksimentiev A, Heng J B, Timp G and Schulten K 2004 *Biophys. J.* **87** 2086
[53] Healy K, Schiedt B and Morrison A P 2007 *Nanomedicine* **2** 875
[54] Kowalczyk S W, Grosberg A Y, Rabin Y and Dekker C 2011 *Nanotechnology* **22** 315101
[55] Carlsen A T, Zahid O K, Ruzicka J, Taylor E W and Hall A R 2014 *ACS Nano* **8** 4754
[56] Zwolak M and Ventra M D 2004 *Nano Lett.* **5** 421
[57] Garaj S, Hubbard W, Reina A, Kong J, Branton D and Golovchenko J A 2010 *Nature* **467** 190
[58] Ivanov A P, Instuli E, McGilvery C M, Baldwin G, McComb D W, Albrecht T and Edel J B 2011 *Nano Lett.* **11** 279
[59] Raillon C, Granjon P, Graf M, Steinbock L J and Radenovic A 2012 *Nanoscale* **4** 4916
[60] Hoogerheide D P, Garaj S and Golovchenko J A 2009 *Phys. Rev. Lett.* **102** 256804.
[61] Uram J D, Ke K and Mayer M 2008 *ACS Nano* **2** 857
[62] Smeets R M, Keyser U F, Dekker N H and Dekker C 2008 *Proc. Natl. Acad. Sci. USA* **105** 417
[63] Wanunu M 2012 *Phys. Life Rev.* **9** 125
[64] Tabard-Cossa V, Trivedi D, Wiggin M, Jetha N N and Marziali Andre 2007 *Nanotechnology* **18** 3505
[65] Balan A, Machielse B, Niedzwiecki D, Lin J, Ong P, Engelke R, Shepard K L and Drndić M 2014 *Nano Lett.* **14** 7215
[66] Tasserit C, Koutsioubas A, Lairez D, Zalczer G and Clochard M C 2010 *Phys. Rev. Lett.* **105** 260602
[67] Powell M R, Vlassiouk I, Martens C and Siwy Z S 2009 *Phys. Rev. Lett.* **103** 248104
[68] Smeets R M, Dekker N H and Dekker C 2009 *Nanotechnology* **20** 095501
[69] Zhang Y, Li H, Wang H, Liu R, Zhang S-L and Qiu Z-J 2015 *ACS Nano* **9** 8514
[70] Chaudhry J H, Comer J, Aksimentiev A and Olson L N 2014 *Commun. Comput. Phys.* **15** 1
[71] He Y, Tsutsui M, Fan C, Taniguchi M and Kawai T 2011 *ACS Nano* **5** 5509
[72] He Y, Tsutsui M, Fan C, Taniguchi M and Kawai T 2011 *ACS Nano* **5** 8391
[73] Luan B and Stolovitzky G 2013 *Nanotechnology* **24** 195702.
[74] Kong Z, Zheng W, Wang Q, Wang H, Xi F, Liang L and Shen J-W 2015 *J. Mater. Chem. B* **3** 4814
[75] Wong C T and Muthukumar M 2007 *J. Chem. Phys.* **126** 164903
[76] Ding K, Sun W, Zhang H, Peng X and Hu H 2009 *Appl. Phys. Lett.* **94** 014101
[77] Wells D B, Belkin M, Comer J and Aksimentiev A 2012 *Nano Lett.* **12** 4117
[78] Bayley H 2006 *Curr. Opin. Chem. Biol.* **10** 628
[79] Heerema S J, Schneider G F, Rozemuller M, Vicarelli L, Zandbergen H W and Dekker C 2015 *Nanotechnology* **26** 074001
[80] Farimani A B, Min M and Aluru N R 2014 *ACS Nano* **8** 7914

[81] Maglia G, Restrepo M R, Mikhailova E and H Bayley 2008 *Proc. Natl. Acad. Sci. USA* **105** 19720
[82] Derrington I M, Butler T Z, Niederweis M, Manrao E, Pavlenok M, Collins M D and Gundlach J H 2010 *Proc. Natl. Acad. Sci. USA* **107** 16060
[83] Venkatesan B M and Bashir R 2011 *Nat. Nanotechnol.* **6** 615
[84] Kowalczyk S W, Wells D B, Aksimentiev A and Dekker C 2012 *Nano Lett.* **12** 1038
[85] Di Fiori N, Squires A, Bar D, GIlboa T, Moustakas T D and Meller A 2013 *Nat. Nanotechnol.* **8** 946
[86] Deng T, Li M, Wang Y and Liu Z 2015 *Science Bulletin* **60** 304
[87] Rhee M and Burns M A 2007 *Trends in Biotechnology* **25** 174
[88] Powell M R, Martens C and Siwy Z S 2010 *Chemical Physics* **375** 529
[89] Manrao E A, Derrington I M, Laszlo A H, Langford K W, Hopper M K, Gillgren N, Pavlenok M, Niederweis M and Gundlach J H 2012 *Nat. Biotechnol.* **30** 349
[90] Cherf G M, Lieberman K R, Rashid H, Lam C E, Karplus K and Akeson 2012 *Nat. Biotechnol.* **30** 344
[91] Ayub M, Hardwick S W, Luisi B F and Bayley H 2013 *Nano Lett.* **13** 6144
[92] Watson J D and Crick F H C 1953 *Nature* **171** 737
[93] Wanunu M, Dadosh T, Ray V, Jin J, McReynolds L and Drndic M 2010 *Nat. Nanotechnol.* **5** 807

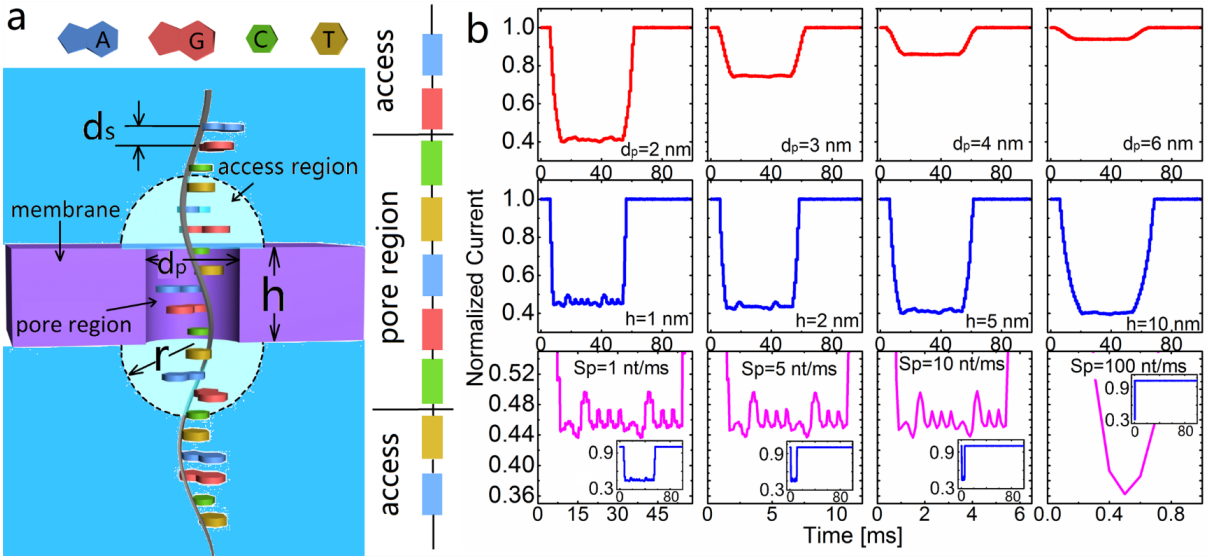

Figure 1. (a) Schematic translocation of an ssDNA strand in a nanopore, with the critical dimensions defined and the illustrations of the four nucleotides marked in different colors. (b) Waveform of the normalized ionic current (to I_o) with different setups: variation with pore diameter (upper row), pore thickness (middle row), and translocation speed (lower row), all without consideration of the access resistance.

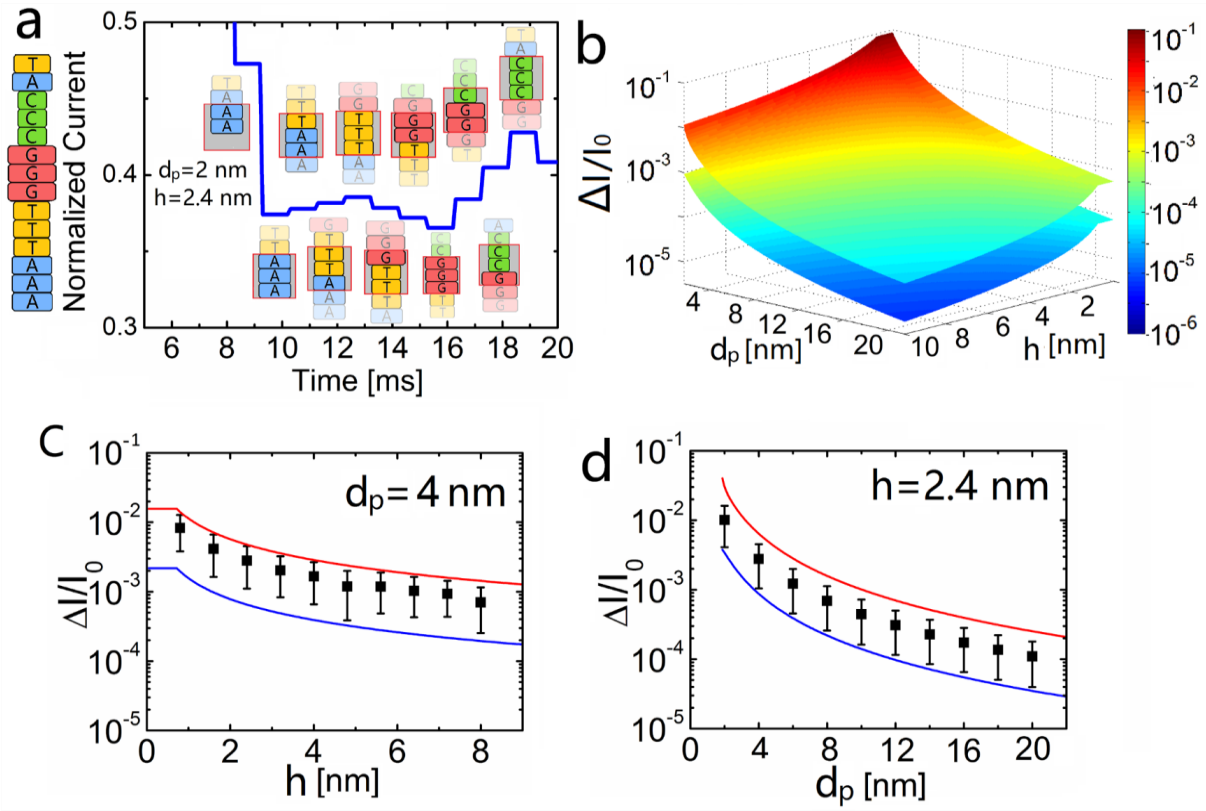

Figure 2. (a) Details of the normalized blockage ionic current (to I_o) assisted by a series of cartoons to correlate the position of the nucleotides on a translocating ssNDA to the current level. (b) Normalized signal $\Delta I/I_o$ defining the upper bound ($\Delta I_{max}/I_o$) and lower bound ($\Delta I_{min}/I_o$) shown as two curved surfaces. Two-dimensional projection of (b) at d_p=4 nm (c) and h=2.4 nm (d), with the black squares and error bars depicting statistical results of an ssDNA strand of 10000 randomly chosen nucleotides. No consideration of the access resistance.

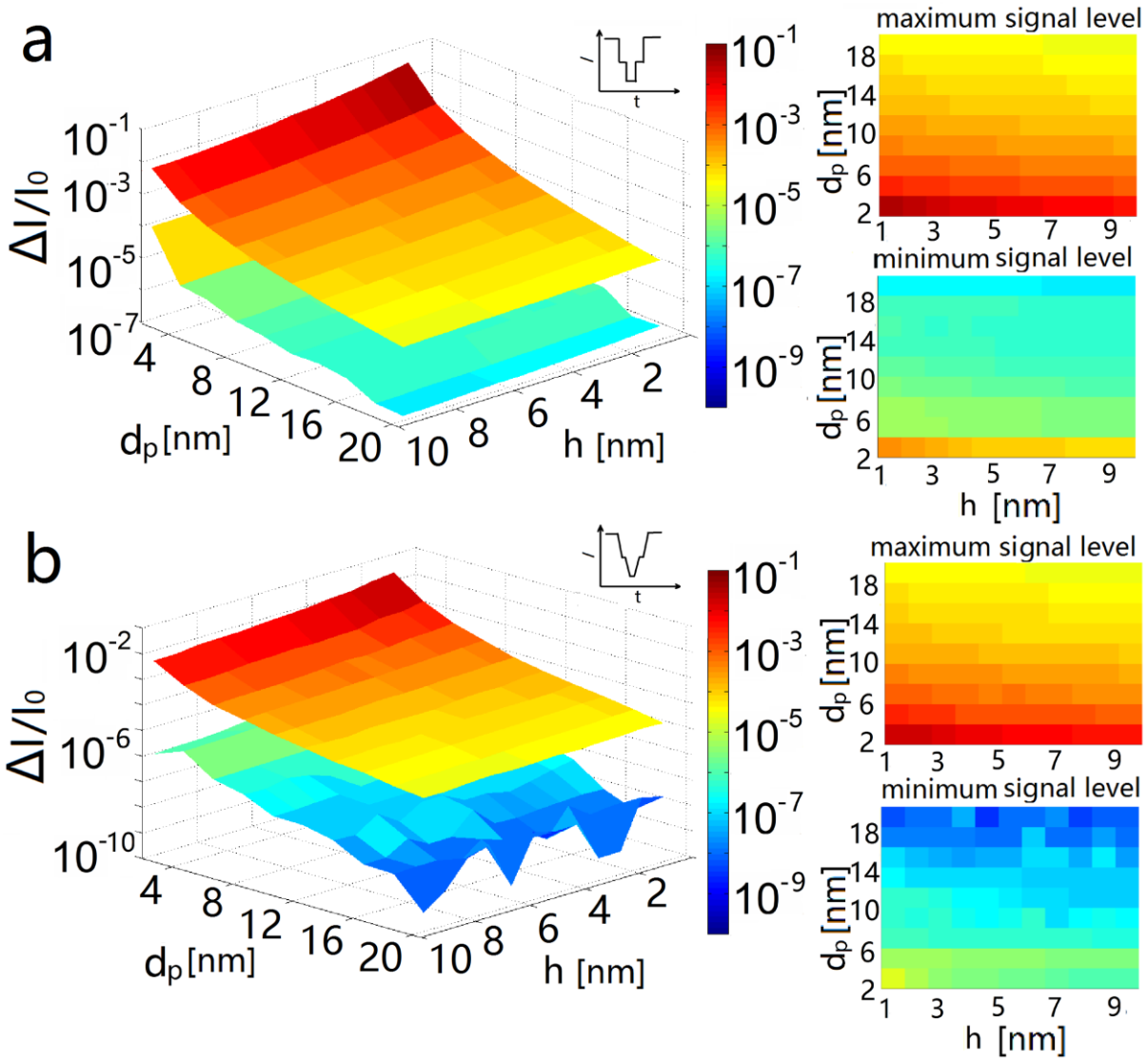

Figure 3. Normalized signal $\Delta I/I_o$ defining the upper bound ($\Delta I_{max}/I_o$) and lower bound ($\Delta I_{min}/I_o$) shown as two curved surfaces with the two access resistances included, but assuming (a) a step-like nucleotide translocation and (b) a uniform and continuous translocation at a constant speed.

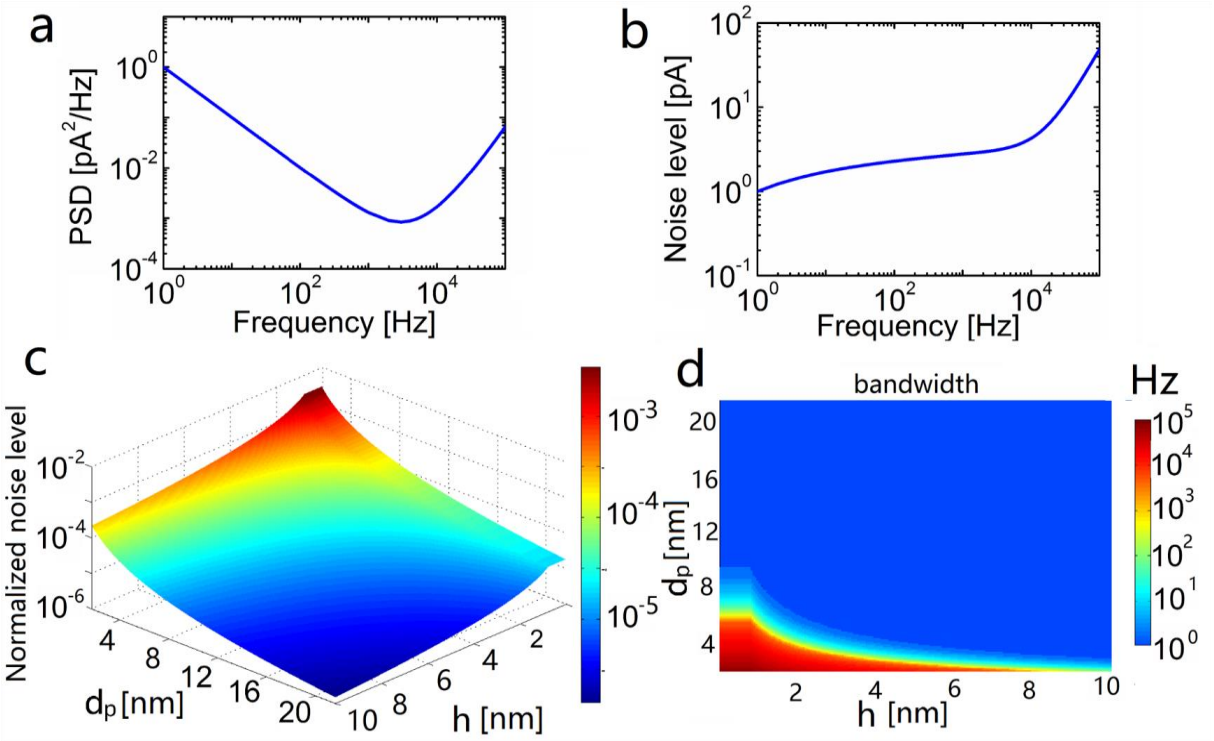

Figure 4. (a) Typical power spectrum density of noise in a nanopore system adopted from the literature.[64] (b) Noise level at different bandwidth frequency obtained from (a). (c) Normalized noise level, to I_o, for nanopores of different sizes, with reference to (b). (d) Maximum bandwidth of the measurement system for nanopores of different sizes.

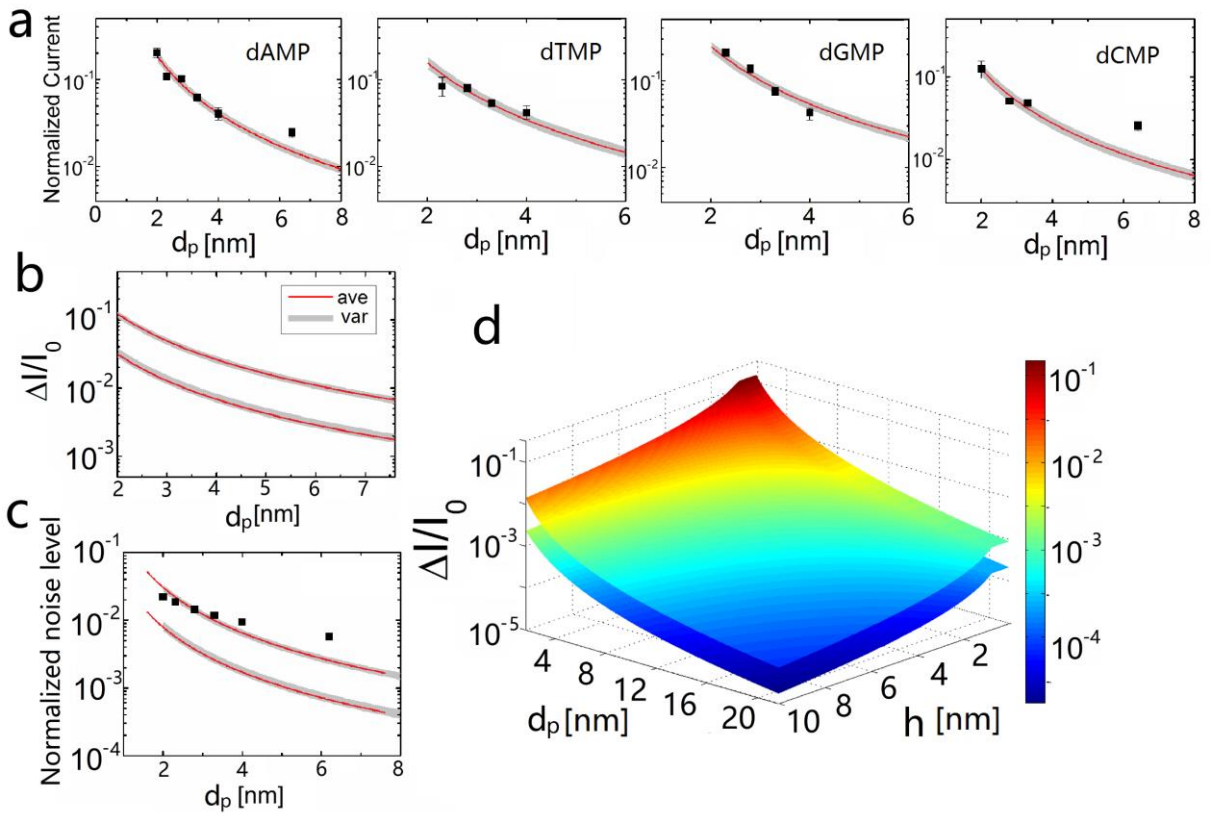

Figure 5. (a) Comparison between experimental (black squares) and calculated (continuous lines) blockage ionic current after normalization (to I_o) for the four different nucleotides dNMPs (N=A, T, G, C). (b) Normalized signal $\Delta I/I_o$ defining the upper bound ($\Delta I_{max}/I_o$) and lower bound ($\Delta I_{min}/I_o$) for h=0.7 nm (MoS_2 monolayer), with variations calculated by taking into account experimental errors and uncertainties under different settings. (c) Comparison between experimental (black squares) and calculated (continuous lines) noise level after normalization (to I_o) for nanopores of different sizes, again with variations for each set calculated by taking into account experimental errors and uncertainties under different settings. (d) Normalized signal $\Delta I/I_o$ defining the upper bound ($\Delta I_{max}/I_o$) and lower bound ($\Delta I_{min}/I_o$) shown as two curved surfaces for MoS_2 multilayers.

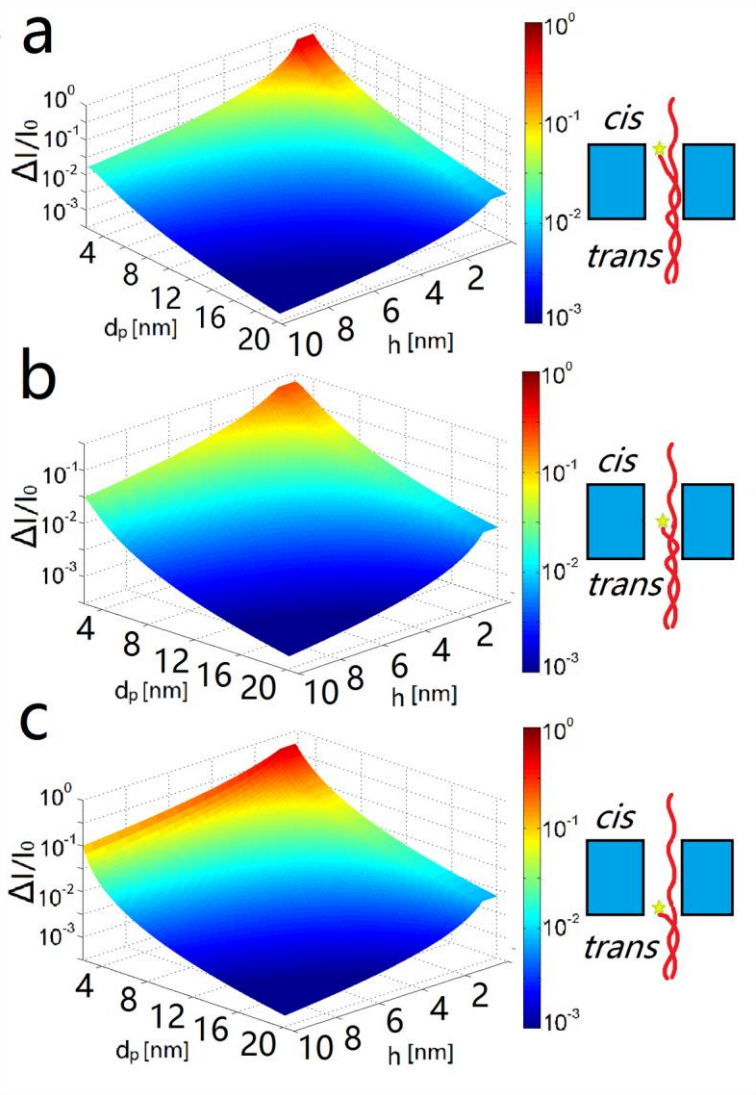

Figure 6. Normalized signal $\Delta I/I_o$ surface for a conceptual signal amplification scheme involving DNA strand synthesis from single-stranded to double-stranded occurring (a) at the entrance on the *cis* side, (b) in the middle, and (c) at the exit on the *trans* side of the nanopore, as schematically shown in the right panel of each case.

Table 1 Parameters used for the simulation. The subscripts in the nt diameters denote adenine (A), thymine (T), guanine (G), and cytosins (C). The mobility data are taken from the representative KCl electrolyte as an example.

Parameter [unit]	q [C]	μ_+ [m^2/Vs]	μ_- [m^2/Vs]	c_+, c_- [M]	d_s [nm]	d_A [nm]	d_T [nm]	d_G [nm]	d_C [nm]
Value	$1.6 \cdot 10^{-19}$	$6.95 \cdot 10^{-8}$	$7.23 \cdot 10^{-8}$	1	0.8	1.582	1.568	1.593	1.513
Reference	-	[93]	[93]	-	[45]	[19]	[19]	[19]	[19]

Table 2 Extracted nt diameters from the experimental results in Ref. [49].

nt	C	T	A	G
Average [nm]	0.859	0.958	1.039	1.188
STD [nm]	0.065	0.098	0.067	0.103
Variation[1] [%]	7.5	10.2	6.4	8.7

[1.] Variation=(STD/Average)100%