2022 IEEE International Electron Devices Meeting (IEDM), pp. 17.4.1-17.4.4.

DOI: 10.1109/IEDM45625.2022.10019451

The Nanopore-FET as a High-Throughput Barcode Molecule Reader for Single-Molecule Omics and Read-out of DNA Digital Data Storage

Koen Martens[1,2], David Barge[1], Lijun Liu[1,2], Sybren Santermans[1], Colin Stoquart[1,2], Jacobus Delport[1,2], Kherim Willems[1,2], Dino Ruic[1,2], Sanjin Marion[1], Juliette Gevers[1,2], Bert Du Bois[1], Alexandru Andrei[1], Liesbet Lagae[1,2], Mark Veugelers[4], Anne S. Verhulst[1,3], Simone Severi[1], Pol Van Dorpe[1,2]

[1]Imec, Leuven, Belgium, [2]KU Leuven Physics dept., [3]KU Leuven, Electr. Eng. Dept. [4]Enigma Life Sciences

***Abstract*—**Nanopore technology has augmented DNA sequencing, and nascent nanopore-based technology breakthroughs may lead to major steps forward in large-scale molecular analysis (omics) and DNA digital data storage. A major weakness of current nanopore technologies is the requisite recording of pico- to nanoampere ionic currents. Such low currents limit recording bandwidth in large integrated arrays and impose a limit on how quickly a detectable molecule can cross, or translocate the pore. The Nanopore Field Effect Transistor (NPFET) is a nanopore surrounded by a nanoscale field-effect transistor. The field-effect transistor senses single molecules translocating through the pore relying on much larger microampere currents. The NPFET potentially offers higher bandwidths in large integrated arrays and the breaking of the nanopore translocation speed limit. Moreover, because individual electrolyte contacts are not required for NPFETs they are more easily integrated into large arrays than nanopores. NPFETs may lead to major strides in high throughput reading of molecularly encoded information. We will introduce and elaborate on nanopore FET technology research for omics and DNA data storage and we will propose a scaling roadmap that aims to incrementally boost NPFET molecular read throughput.

I. Introduction – The Challenges of Omics and Reading DNA Data Storage

High-throughput next-generation DNA sequencing has fueled a revolution in genomics, with an ever-increasing impact on health care including cancer and covid19, and several other disciplines, such as plant biology and even archaeology[1]. While our understanding of the genome is increasing, we lack a similar understanding of the proteome, all proteins expressed by a cell, tissue, or organism at a certain time. High-accuracy and low-cost readout of our protein make-up is not yet feasible. Proteomics techniques are centralized in specialized institutes and are carried out with room-filling equipment run by highly trained personnel, with queuing and run-times of multiple days. Its application is limited to research purposes. Moreover, the limited throughput, dynamic range and limit of detection of the current mass spectrometry-based proteomics toolset does not allow to measure the most interesting low abundance proteins in clinical samples. Furthermore, comprehensive single - cell proteomics remains a challenge and is of importance to understand cancer. Moreover, multiomics, the combination of omics techniques (such as proteomics and RNA-based transcriptomics) is of growing interest. Meanwhile, silicon CMOS - based nanotechnology has made its way into DNA sequencing and has the untapped potential to realize proteomics-on-chip with high-throughput, high accuracy, and low-cost read-out.

Proteomics-on-chip techniques that are currently in development favor a single-molecule approach. Enormous concentration differences (= large dynamic range) occur in proteomes. Within a single cell this can be 7 orders of magnitude [2] (Fig. 1) and in blood plasma [3] (Fig. 2) this can exceed 10^{12}. To exceed the mass spectroscopy dynamic range (10^3-10^5) by multiple orders of magnitude with a single-molecule technique one needs to analyze billions of single molecules. To analyze a sample with 10^{12} dynamic range in an hour, billions of molecules per second need to be analyzed.

Some groups focus on measuring the smallest possible differences on biopolymers and on discriminating single amino acids or bases. However, to obtain the highest possible single molecule throughput, requiring such resolution directly for the electronic readout system may not be beneficial.

The development of DNA data storage is propelled by the burgeoning drive to store cold data. Cold data, which is accessed infrequently, and is typically stored on tape, is growing the strongest of all data tiers [4]. DNA data is a strong contender for such long-term archival data storage due to its storage density, durability, low power consumption, low maintenance and hence low cost of ownership [4]. A key challenge for DNA memory is reading. Current DNA sequencing technology is expected to reach $130/gigabyte read cost within a few years. The IARPA targeted cost is $1/terabyte/day by 2030 with read and write costs assumed equal [4]. The human genome (0.75 gigabytes) can currently be sequenced in ~5 hours with a tabletop machine (Oxford Nanopore [5]) resulting in 0.3Mbit/s read speeds. The involved biological nanopore device reads DNA at 400-450 bases per second. Data read throughput is clearly a major bottle neck for developing DNA memory.

II. The Nanopore FET

The Nanopore FET can be seen as a combination of a nanopore and an electrolytically gated FET or bioFET. In a regular solid-state pore, a translocating molecule modulates the recorded ionic current passing through the pore (Fig. 3). In a bioFET a molecule docks close (1-10nm) to the FET surface (Fig. 4) and the molecule's charge modulates the FET channel charge and FET current. This Coulomb effect is present if the molecule is within a Debye length (typically a few nanometers)

of the FET surface, and the effect can be significant at low salinity. The nanopore FET, a nanoscale FET with a pore running through the channel (Fig. 5), senses single molecules translocating through the pore due to the so-called resistive divider effect [6]. The FET senses the potential between two electrolyte resistances of which at least one is modulated by the passing molecule. The sensing mechanism is hence strongly related to ionic current modulation in regular solid-state pores, and sensing remains strong at high salinity. At sufficiently low salinity also a Coulomb sensing effect as in bioFETs can occur.

The nanopore FET has been experimentally demonstrated in 2012 [6] and has taken many forms [7-13]. The nanopore FET's primary device-level advantage is its potential high > 100 MHz bandwidth with sufficient signal-to-noise ratio (SNR). This has been studied theoretically [8] but is yet to be proven experimentally. At low frequencies, the SNR in regular solid-state nanopores will be superior to the SNR of the nanopore FET [8]. However, at high frequencies where capacitive noise currents dominate the regular pore signal, FETs can improve the maximum bandwidth with SNR > 5, up to 100 MHz [8]. The second key advantage is the superior integrability of nanopore-FETs into large-scale arrays.

For the Nanopore-FET device as shown in Fig. 5 and operated at low salinity both a resistive divider effect and a Coulomb effect occur. A simulated translocation signal is shown in Fig. 6 [14]. The resistive divider effect by itself results in a typical sigmoid curve shape as in the case of Fig. 6 (bottom) for a neutral particle for which no Coulomb effect is present.

III. Barcode Molecules

DNA nanostructures containing a barcode [15] have been proposed to enable nanopore-based read-out of a highly multiplexed assay as for omics and of DNA data. Instead of reading out the individual bases of DNA as commonly done for assays making use of DNA barcodes at present, these structural DNA barcodes have larger molecular structures introduced along a double-stranded DNA backbone to encode a barcode. These larger structures enable read-out with solid-state pores, which are typically less sensitive than their biopore counterparts used for sequencing DNA, but which can offer substantially higher translocation speeds. In the work by the Keyser group [15][16] the DNA origami method is used to create the barcode molecules. See molecule in Fig. 8.

IV. Scaling Roadmap

Here, we propose a scaling roadmap (Fig. 8) for molecularly encoded data coupled to a Nanopore-FET high-throughput read-out. As in hard drives or tape, the roadmap is characterized by the bit density on the medium and we derive the scaling of the NPFET reader. Importantly, despite the larger molecular structures on the barcode molecules, the NPFET reader with its high bandwidth allows much higher read throughput than regular solid-state pores or biopores (see Fig. 9) by reading freely translocating barcodes at high translocation speed. The bit length along the DNA backbone is defined in base pairs or monomers. The bit length is scaled down progressively starting here at ~200 base pairs or ~70nm. As in CMOS scaling a "red brick wall" is discerned for future nodes where issues such as Brownian motion may hinder further scaling. To allow reading downscaled barcode molecules, the NPFET reader will be scaled along (Fig. 10). The increasing NPFET array size will boost throughput further. Barcode molecule concentration will co-determine how frequently molecules are captured and translocate. For optimal throughput, the translocation duration preferably dominates the wait time between successive translocating barcode molecules. To realize high-throughput omics the proteome is converted to structural barcodes which are isolated in a purified mixture for quick read-out by an NPFET array (Fig. 7). Proteomics assays that translate the proteome to DNA sequences already exist or are in development such as e.g., Somalogic [17] and Encodia [18] assays. Progress in DNA data read throughput promises to boost the dynamic range of omics techniques and contributes to materializing DNA memory.

V. Recent Work and Outlook

In our recent simulation work making use of a dedicated nanofluidic nanodevice simulator [19] we have investigated the NPFET signal due to molecular structures as a function of separation to determine resolution (Fig. 11). Moreover, based on this simulator we also reveal that sensitivity can be boosted exponentially by increasing cis-trans bias [20] (Fig. 12).

We have fabricated and characterized nanopores (Fig. 13 [21]), bioFETs [22] and a Nanopore-FET milestone device, which has an opened nanopore but does not have a bottom reservoir yet. See Fig. 15-16 for process flow and TEM images. On this milestone device we obtain excellent electrolytically gated FET transfer and output characteristics (Fig. 17, 18) demonstrating a low subthreshold swing of 67mV/decade, indicating excellent device quality. Simulations (Fig. 19, 20) show that the obtained low subthreshold swing can only be obtained if the nanopore is wetted. Next step is the electrical demonstration of the full NPFET.

In sum, the nanopore-FET promises to enable progressive increases in molecularly encoded data read throughput far beyond what is possible at present.

References

[1] L. Orlando et al., Ancient DNA analysis. *Nature Reviews Methods Primers*, 1, 1, pp.1-26, 2021
[2] B. Schwanhäusser et al., *Nature* 473, pp. 337-342., 2011
[3] P.M. Kosaka et al. *Seminars in Cancer Biology*, 52, Part 1, pp.26-38, 2018
[4] The DNA Storage Alliance, An introduction to DNA Data Storage, www.dnadataalliance.com, June 2021
[5]https://med.stanford.edu/news/all-news/2022/01/dna-sequencing-technique.html#:~:text=A%20research%20effort%20led%20by,5%20hours%20and%202%20minutes, 2022
[6] P. Xie et al., *Nature nanotechnology,* 7, 2, pp. 119-25, 2012
[7] F. Traversi et al., *Nature nanotechnology,* 8, 12, p. 939, 2013
[8] W. Parkin and M. Drndić, *ACS sensors,* 3, 2, p. 313, 2018
[9] M. Graf et al., *Nano letters*, 19, 12, pp. 9075-9083, 2019
[10] S. Heerema et al., *ACS nano,* 12, 3, pp. 2623-33, 2018
[11] B. Takulapalli, Patent US 2015/0060952 A1, 2013
[12] I. Yanagi et al., *IEEE IEDM*, pp. 14-3, Dec. 2013
[13] I. Yanagi et al., *Nanotechnology*, 27, 11, p.115501, 2016
[14] D. Ruic et al., *S3IC*, Barcelona, Nov., 2020
[15] N.A. Bell et al., *Nature nanotechnology*, 11, 7, p. 645, 2016
[16] K. Chen et al., *Nano letters*, 19, 2, pp.1210-1215, 2018
[17] L. Gold et al., *Nature precedings*, pp. 1-1, 2010
[18] M. Chee, K. Gunderson, WO 2017/192633, 2017

[19] Source code of npfetFOAM, https://gitlab.com/dinoruic/npfet_s3ic2020
[20] A.S. Verhulst et al., *IEEE Sensors Journal*, 22, 6, pp.5732-5742, 2022
[21] J. Delport et al., *MNE*, Leuven, Sep., 2022
[22] S. Santermans et al., 2020 *IEEE IEDM*, pp. 35.4.1-35.4.4, Dec., 2020
[23] L Liu et al., *MNE*, Leuven, Sep., 2022

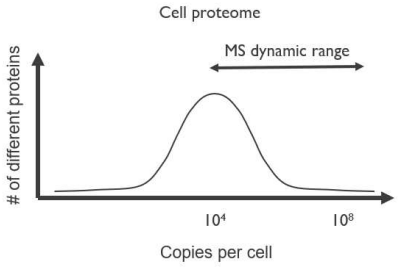

Fig. 1. Schematic of the number of different protein species present at different concentrations in a typical cell illustrating the concentration dynamic range of a cell. Typical Mass Spectroscopy (MS) measurable concentration dynamic range indicated by the arrow.

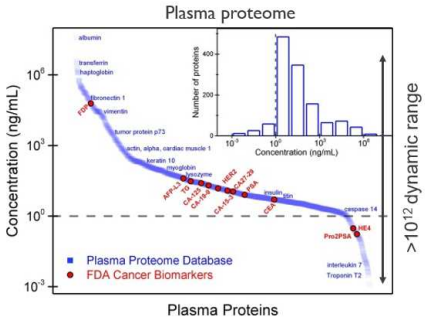

Fig. 2. Concentration vs. cumulative amount of different protein species (x-axis) illustrating the concentration dynamic range of proteins in blood plasma. Plasma can have dynamic ranges exceeding 10^{12}. From [3] (CC BY NC ND https://doi.org/10.1016/j.semcancer.2017.08.011).

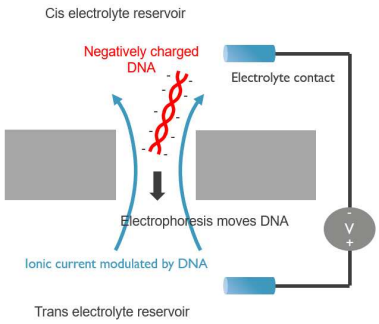

Fig. 3. The solid-state nanopore. A voltage is applied between a cis- and trans electrolyte (top and bottom fluidic chamber resp. separated by the nanopore) Electrophoresis drives a DNA molecule from cis to trans. The passing molecule modulates the ionic current running through the pore.

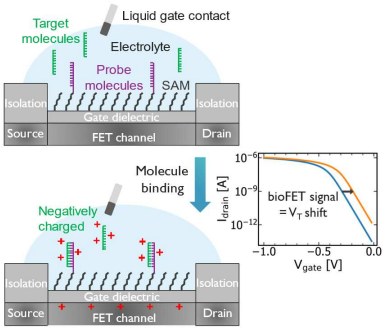

Fig. 4. Illustration of bioFET functionality. Target molecules bind specifically to probes grafted on the bioFET surface. The target charge modulates FET channel charge.

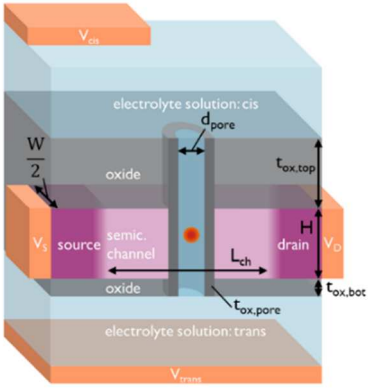

Fig. 5. Illustration of the nanopore FET, an electrolytically gated FET with a nanopore running through its channel.

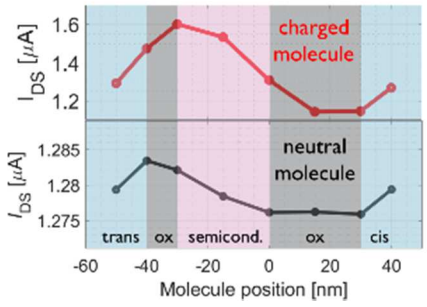

Fig. 6. Drain current signal profile for V_{DS} = 0.1V, $V_{cis;S}$ = 0.58V,$V_{trans,S}$ = 0.5V, W = 25nm, L ≈ 30nm at pH = 4 and 1mM NaCl for a molecule with a diameter of 3nm with +10 elementary charges (top) and no charge (bottom) [14].

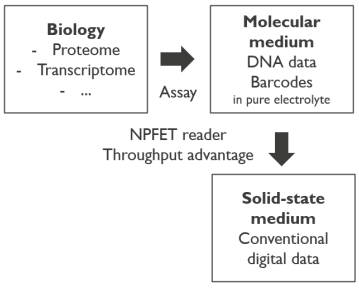

Fig. 7 Biological information as e.g. contained in a proteome can be converted with an assay to molecularly encoded data in a purified NPFET-friendly electrolyte for fast NPFET read-out.

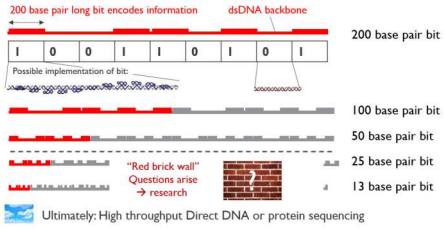

Fig. 8 Schematic illustration of the roadmap which scales down the size of the bits on the information medium, a DNA barcode molecule. A possible molecule implementation is shown as in [15].

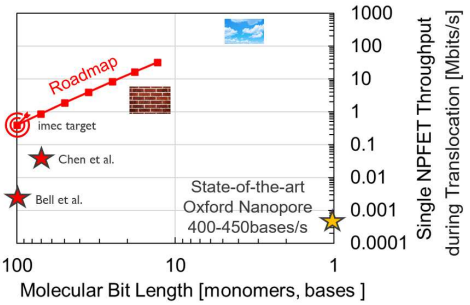

Fig. 9 Data throughput per device along the proposed roadmap compared with the research work of the Keyser group (Red Star) and State-of-the-Art Oxford Nanopore biopore-based DNA sequencing.

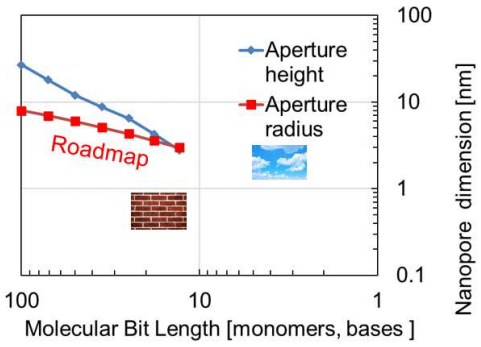

Fig. 10 Indicative NPFET feature size downscaling along the proposed roadmap.

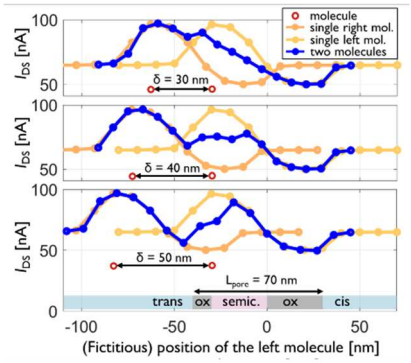

Fig. 11 Simulated NPFET signal for two positively charged labels (molecules) separated by varying δ for the NPFET structure shown in Fig. 5. The two labels can be resolved if $\delta > \min(L/2, H_{Si})$ with L the pore height and H_{Si} the semiconductor height.

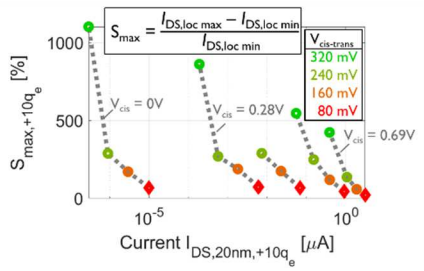

Fig. 12 An exponential sensitivity (S_{max}) boost can be obtained with the NPFET by increasing cis-trans bias [20] for the NPFET structure shown in Fig. 5.

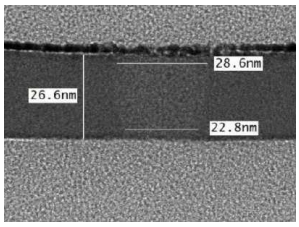

Fig. 13 Cross-section TEM of a regular solid-state nanopore device [22].

- SOI thinning & channel doping
- Nanowire formation
- Source and drain pattern, implant & activation
- Nanopore opening by dry etch
- Nanopore gate oxide formation
- Filling of pore with sacrificial material
- Metal interconnect
- Dielectric passivation
- Removal of sacrificial material

Fig. 14 NPFET process flow [23].

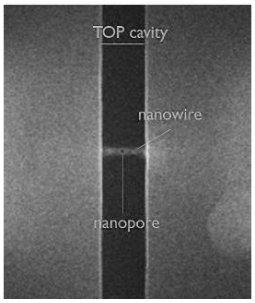

Fig. 15 a) b) top view SEM of semi-nanopore FET [23].

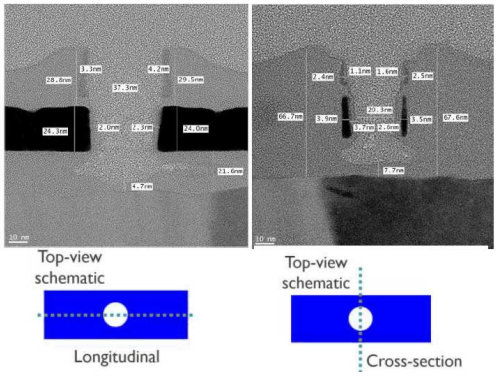

Fig. 16 Cross-section and longitudinal TEM of a semi-nanopore FET milestone device [23].

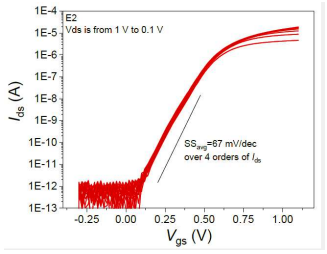

Fig. 17 Transfer characteristic curve of a typical NPFET with different bias Voltages V_{ds} [23].

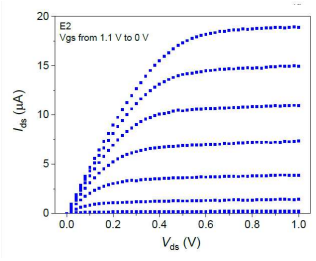

Fig. 18 Output characteristic curve of the same NPFET with different gate voltage V_{gs} [23].

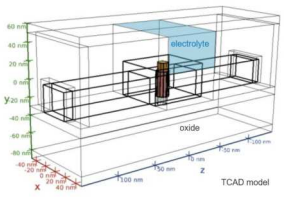

Fig. 19 Simulation structure for NPFET [23].

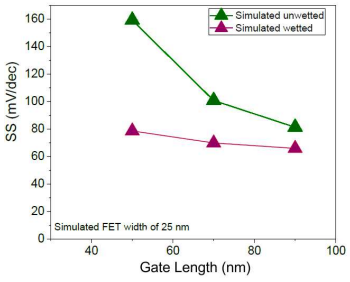

Fig. 20 Simulated results of subthreshold swing for different gate lengths [23].