GA

TN

EPIGENETIC IDENTIFICATION AND MAPPING USING SOLID-STATE 2D NANOPORES

Allocation: Illinois/418 Knh
PI: Jean-Pierre Leburton[1]
Co-PI: Lav Varshney[1]
Collaborators: Nagendra Athreya[1], Aditya Sarathy[1]

[1]University of Illinois at Urbana-Champaign

EXECUTIVE SUMMARY

Epigenetic modifications, such as heritable alterations of the human genome, are believed to play a critical role in gene regulation, causing diseases such as cancer and various autoimmune and neurological disorders. In the present work, we develop computational techniques to understand and assess the efficiency of epigenetic detection of methylation sites and their mapping on the DNA strands with the use of 2D atomically thin nanopore membranes in electrolytic cells. We consider various detection scenarios involving ionic current blockade of the pore as well as monitoring the transverse electronic current variations across the membrane. The simulated current signatures are obtained by coupling all-atom molecular dynamics (MD) simulations to a combination of self-consistent Poisson–Boltzmann electrostatics and electronic transport calculations. Additionally, to overcome the inherently low signal-to-noise ratio (SNR) during these detections, we have developed statistical signal processing algorithms recognizing and distinguishing DNA nucleotides and various methylated sites on the DNA strands.

RESEARCH CHALLENGE

Aside from sequencing DNA molecules, the identification of traits of the human genome, such as methylation, is crucial for diagnosis of epigenetic diseases. Recent experimental evidence of DNA methylation alterations linked to tumorigenesis suggests that DNA methylation plays a major role in causing cancer by silencing key cancer-related genes [1–3]. Until now, detection and mapping of such DNA methylation patterns using solid-state nanopores have been unsuccessful due to rapid conformational variations generated by thermal fluctuations that result in low SNR. To overcome these drawbacks, a versatile, general sensor technology for detecting methylation patterns is desirable. For this reason, we propose an integrated approach that combines MD with device physics-based electronic modeling and statistical signal processing techniques to assess the resolution limit of solid-state nanopore sensing. In addition, we further develop algorithms for epigenetic marker classification at the fundamental limits of SNR improvement for biodetecting membranes.

METHODS & CODES

Our research consists of a two-step process that first uses MD simulations with the latest NAMD version, and then exploits the MD data to calculate the current variations due to DNA translocation through the nanopore via electronic transport modeling. The system is built, visualized, and analyzed using VMD [4]. The protein and DNA are described by the CHARMM22 force field with CMAP corrections [5] and the CHARMM27 force field [6], respectively. An external electric field is applied to the system to drive the DNA–protein complex through nanopores.

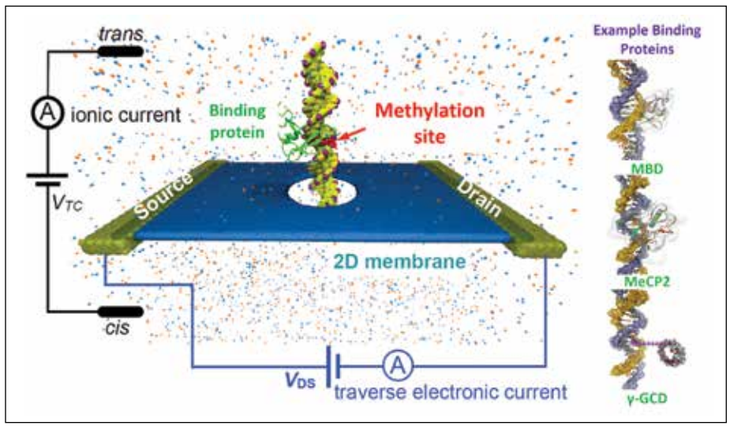

Figure 1: Schematic of the experimental setup for epigenetic detection and mapping with a 2D nanopore membrane. Frozen (ideal) and nonfrozen (noisy) signatures of three different biomarker proteins (right) binding to the methylation sites are considered for use in the matched-filter algorithm.

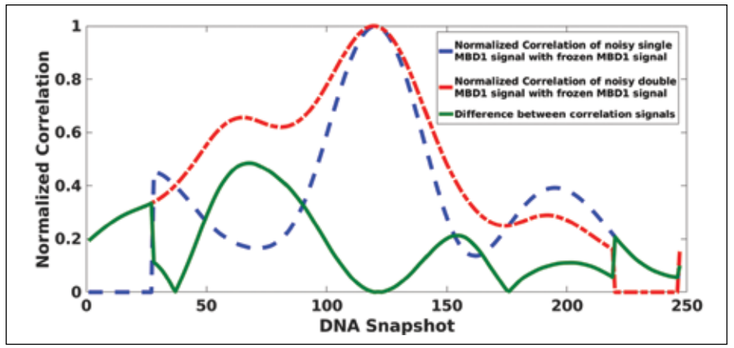

Figure 2: Normalized correlations of noisy electronic sheet current signals obtained from the matched-filter algorithm. Two signals, one corresponding to single-MBD1 (blue) and the other corresponding to multiple-MBD1 (red), are aligned and the difference between them (green) is used to determine and map the presence of a second protein on the DNA.

For each frame of the trajectory files obtained from the MD simulations, ionic current blockade is calculated instantaneously [7]. Further, electrostatic potential induced by the biomolecule around the pore is obtained by using the self-consistent Poisson–Boltzmann equation (PBE) formalism. The PBE is solved numerically using the multigrid method until the convergence criterion is met. Once the electrostatic potential co-planar to the membrane is obtained, the transverse conductance across the 2D nanopore, graphene, or MoS_2 membrane is computed by using the nonequilibrium Green's function formalism [8] or a semiclassical thermionic emission technique [9], respectively. The PBE and the electronic transport code are written and maintained by the Leburton Group at the University of Illinois at Urbana-Champaign.

RESULTS & IMPACT

Previously, we showed the ability of 2D material nanopores to detect DNA methylation sites labeled by MBD1 proteins by two techniques: through the ionic current blockade and transverse electronic sheet conductance. Further, our combined MD device-modeling approach showed that multiple methylation sites could be distinguished in a single ionic current measurement, provided that they are separated by at least 15 base-pairs (bps), whereas single transverse sheet current measurements resulted in a better identification resolution of 10 bps. In the latter case, the superior performance of electronic detection is due to the ability of the single transverse sheet current method to capture local protein charge variation within the membrane nanopore [10,11].

In our scenario, we built systems where the methylated cytosines are complexed by attaching either a methyl-CpG binding domain (MBD-1) protein or a methyl CpG binding protein 2 (MeCP2), whereas methylated adenines are attached to an oligosaccharide, γ-cyclodextrin (γ-GCD) [12]. "Noise-free" electronic currents (ideally obtained by frozen biomolecules artificially translocated through the nanopore) were calculated for all the biomarkers. These signatures were compiled into a set of dictionary signals for each of the marker proteins. When a target noisy signal includes the stochastic conformational fluctuations obtained from MD simulations and electronic transport calculations whose marker protein is unknown is fed into the matched filter, it is correlated with the different dictionary signals to identify the marker-protein type. We anticipate that this algorithm can be extended to the detection of multiple markers attached to the same DNA molecule.

These biomarkers are important for the recognition of different cancer segments. MBD1 and MeCP2 are proteins in humans that are capable of binding to hypermethylated sites along the DNA strand and that also repress transcription from methylated gene promoters [13]. MeCP2 mutations are thought to be responsible for Rett syndrome, and polymorphisms of MBD1 are associated with increased lung cancer risk. Alternatively, hypomethylation, identified using γ-GCD (a synthetic biomarker for the N6-methyladenine), has been linked to cancers of the stomach, prostate, breast, pancreas, and kidney [14]. Therefore, identification and differentiation of these different proteins are critical, as their interactions with DNA play important roles in breast, lung, and other kinds of cancers.

WHY BLUE WATERS

Investigation of the interactions of biomolecules with solid-state materials, characterization of the stochastic structural fluctuations of the epigenetic biomarker complexed with DNA translocating through solid-state nanopores, and acquisition of the electronic response using all-atom MD simulations coupled with electronic transport calculations are only possible with petascale computing resources such as Blue Waters. Our systems are about 500,000 atoms in size, each requiring multiple MD simulation (NAMD) runs. With NAMD code efficiently deployed on XE/XK nodes to run highly parallel simulations of large biomolecular systems, Blue Waters is well suited for our research needs.

PUBLICATIONS & DATA SETS

Sarathy, A., N.B. Athreya, L.R. Varshney, and J.-P. Leburton, Classification of Epigenetic Biomarkers with Atomically Thin Nanopores. *Journal of Physical Chemistry Letters*, 9 (2018), pp. 5718–5725.