

# Characterization of a Solid State DNA Nanopore Sequencer Using Multi-scale (Nano-to-Device) Modeling

Jerry Jenkins, Debasis Sengupta, and Shankar Sundaram

CFD Research Corporation 215 Wynn Drive,  
Huntsville, AL 35805  
{jwj, dxs, sxs}@cfdrcc.com

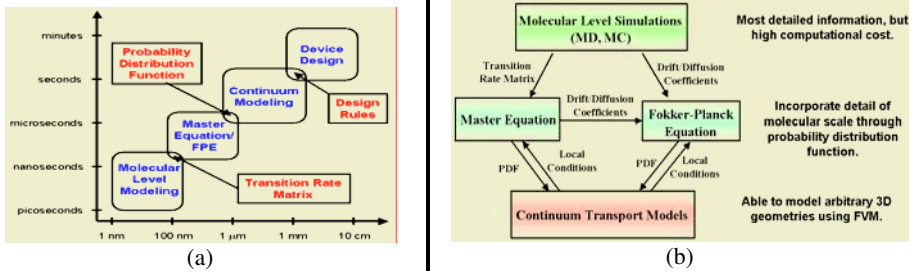
**Abstract.** Nanobiotechnology is a rapidly advancing frontier of science with great potential for beneficial impact on society. Unfortunately, design of integrated nano-bio systems is a complex, laborious task with large failure rates. Current models describing molecular level behavior are expensive, while device design codes lack the necessary nanophysics. The objective of this work is to demonstrate multiscale, multiphysics modeling of an integrated nanobio device, where nanoscale effects are efficiently integrated with a continuum model. A three-level modeling paradigm was developed for this purpose. The feasibility of this approach is demonstrated by characterizing a nanopore-based DNA sequencing device. In the demonstration calculations, the dependence of the device performance on the nucleotide sequence, pore diameter, and applied voltage was determined. Extension of the approach for describing biomolecular processes in other commercial nanobiosystems is discussed. The main conclusions of the device level simulations are presented along with an overview of future work.

## 1 Introduction

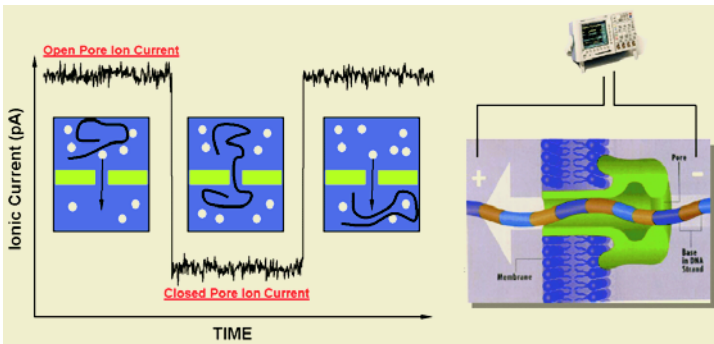
Research efforts directed at the problem of integrating nanotechnology and biology to form integrated nano-bio systems are becoming a priority. Integrated nano-bio systems have emerged as strong candidates for single molecule detection and genomic sequencing. Understanding the relationship between molecular behavior and system/device response, requires the development of modeling and simulation tools, which can simulate phenomena spanning a wide spectrum of length and time scales (Figure 1(a)). We have developed a hierarchical theoretical and computational framework (Figure 1(b)) that addresses the problem of coupling molecular level information into a continuum level device simulation. Each modeling and simulation level in the hierarchy is chosen with careful attention to a cost-benefit analysis.

The first level in the hierarchy is an atomistically detailed model of the nanoscale phenomena of interest. Molecular simulations enable a physical understanding of the fundamental mechanisms that underlie nanoscale phenomena. Results from the molecular level simulations are used to parameterize the second level stochastic

model. Stochastic models (Master Equation or Fokker-Planck Equation) are utilized to bridge molecular and continuum levels because they are able to correctly predict the average properties of a system as well as the fluctuations inherent in the nanoscale. The final level is a continuum model of the system of interest, which couples information from the preceding two levels. Stochastic models can be tightly integrated into the continuum model for high-fidelity design calculations. Insights obtained using this modeling hierarchy can guide the choice of candidate design strategies for specific integrated nanobiosystems.



**Fig. 1.** (a) Applicable range of scales for each modeling level. (b) Illustration of the different levels of modeling. The text next to the arrows is the information transferred at each level



**Fig. 2.** As negatively charged nucleic acids pass through the nanopore in single file, they block the flow of current in a manner characteristic of the polymer’s length and sequence [2]

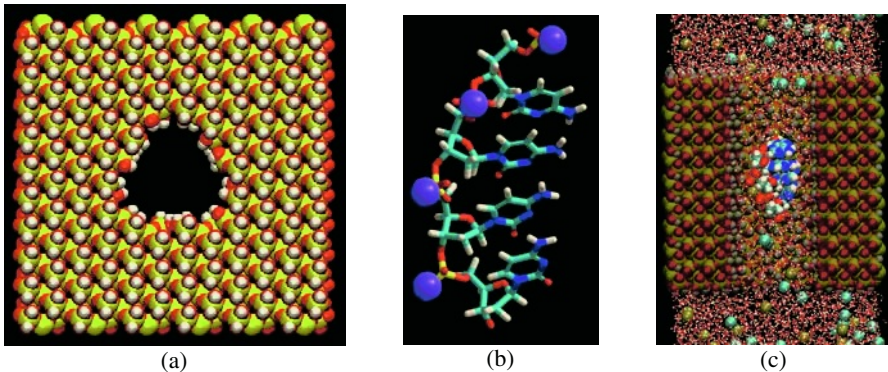
The demonstration case we have chosen (Figure 2) is the problem of DNA translocation through nanopores. This area has come under a great deal of investigation recently due to its promise for inexpensive ultra high-throughput sequencing of DNA. A voltage bias applied to a nanopore has been shown to induce charged single stranded DNA and RNA molecules to translocate through the pore. Each translocating molecule blocks the open pore ionic current, providing an electrical signal that depends on several characteristics of the molecule [1]. The natural system has limitations due to the fixed size of the pore, stability, and noise characteristics. These difficulties can be overcome using a solid state nanopore [2].

The paper begins with a description of the molecular modeling in §2, followed by the development of the stochastic model in §3. This information is then coupled to the device level calculations to evaluate a proposed ultrahigh throughput DNA sequencing device (§4). The main conclusions of the device level simulations are presented along with an overview of future work.

## 2 Molecular Level Modeling

Translocation times for DNA to cross a nanopore are strongly dependent on the pore diameter [1]. Nanopores with diameters of 1.5nm and 2.0nm, were constructed based on experiments presented in the literature [3]. The pores were created using an in-house software that constructs a cube of silicone dioxide using the method of Brodka [4]. The cube is truncated to form a solid block of with dimensions of 5nm x 5nm x 5nm [5]. An overhead view of the 2nm nanopore is shown in Figure 3(a).

Experimental work has also revealed that homopolymers of adenine and cytosine have significant differences in their translocation times [2]. For this reason, poly(dC)<sub>4</sub> and poly(dA)<sub>4</sub> were chosen for this computation with poly(dC)<sub>4</sub> shown in Figure 3(b). Once generated, the nucleotides were placed in the nanopore and aligned with the axis of the pore with the 5' end pointing in the positive z-direction. Water and KCl were added to the system using tcl scripting language built into the VMD molecular graphics software [6]. The final system consists of a 6.1nm layer of water on top of the 5nm crystal structure and a 2nm layer of water on the bottom of the crystal for a total height of 13.1nm (Figure 3(c)). The total number of atoms for each system is approximately 29,000.

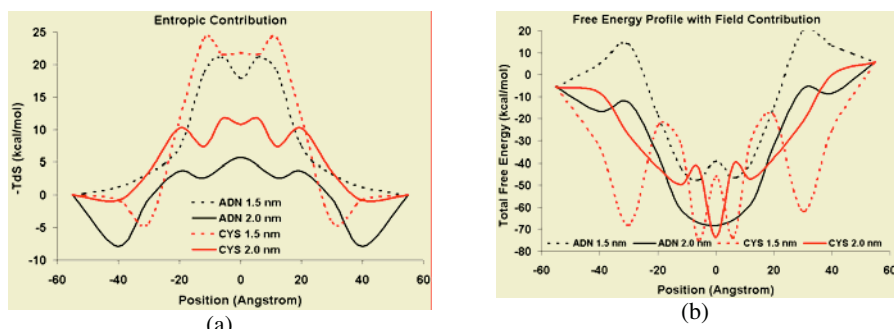


**Fig. 3.** Generation of (a) the nanopore in a Block of SiO<sub>2</sub> with dimensions of 5x5x5 nm. (b) 4-mer of Cytosine with potassium counter ions. (c) Final system (cutaway view) with nucleotide in the nanopore including counter-ions and solvent water

Molecular Dynamics (MD) simulations were performed in three phases following the work of Auffinger [7]. The phases consist of: (a) the initial heating of the water and ions, (b) equilibration of the entire system, and (c) production. The NAMD

software was used for all molecular simulations. (NAMD was developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign) [8]. The Verlet-I or r-RESPA Multiple Time Stepping (MTS) algorithm was used in order to increase the time step to 2fs [9]. During the production simulations a Langevin based temperature controller was used to maintain the temperature [10], but was not coupled to hydrogen atoms. The Particle Mesh Ewald (PME) algorithm was used to model the periodic boundary conditions, with grid spacing of 64, 64, and 128 in the x, y, and z directions, respectively. All forcefield parameter information was taken from the CHARMM27 all hydrogen nucleic acid forcefield, derived from the work of [11] and [12]. Parameters for the potassium and chlorine ions were taken from [13]. The charge and Lennard-Jones parameters for the  $\text{SiO}_2$  were taken from [14].

Entropic affects can contribute significantly to the free energy, especially if a molecule is moving from solution into a confined space. The method of Schlitter [15] was used to compute the state entropy from the production phase of the molecular dynamics trajectories. The method is uses an approximate Gaussian form to approximate the canonical configurational distribution function. This method has been widely used to determine the entropy change for protein stability [16], and the stability of double stranded DNA [17]. A striking feature (Figure 4(a)) is that the 1.5nm pores have larger entropy changes relative to the surroundings when compared to the 2.0 nm pores. The largest entropic penalty was felt by poly(Cytosine)<sub>4</sub> in the 1.5nm pore. This contribution to the free energy is approximately 25 kcal/mol, while the 2nm pore case for cytosine only has a 11 kcal/mol entropic barrier. The adenine shows a similar pattern of higher entropy outside of the pore and higher entropy change.



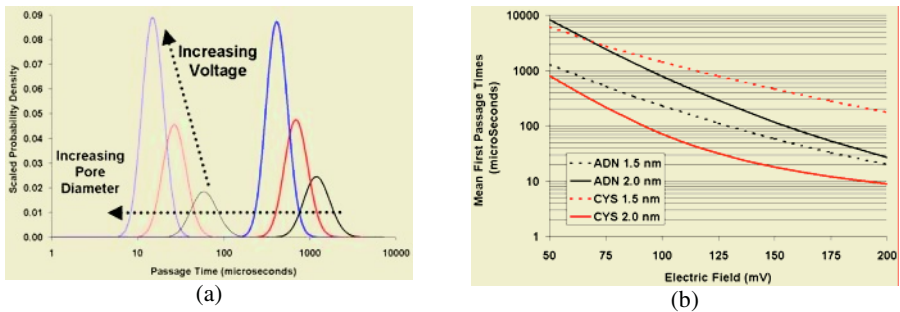
**Fig. 4.** (a) Prediction of the entropic component of the overall free energy of the system. (b) Free Energy Profiles ( $\Delta U - T\Delta S$ ) of translocation for poly(dC)<sub>4</sub> and poly(dA)<sub>4</sub> translocating through the 1.5 and 2.0nm pores

Figure 4(b) illustrates that the free energy profiles are similar for all four cases. Focusing in on the portion of the graph that contains the pore (-25 to 25 on the x-axis), we can draw some conclusions pertaining to the contribution to translocation times from the free energy. The relative free energy profiles show that Cytosine has a

lower free energy relative to Adenine within the pore, for both pore diameters, having the lowest free energy in the 1.5nm pore. Adenine shows a lower free energy within the 2.0nm pore relative to the 1.5nm pore, while Cytosine shows a similar free energy for both pores. The barriers extracted from the nanopores are very similar to the shape of the energy profiles extracted from ion channels [18]. The contribution to the electric field tilts the free energy profiles 11.5 kcal/mol assuming a 120 mV bias across the 5nm thickness of the nanopore.

### 3 Stochastic Model of DNA Translocation

The stochastic model forms the second modeling scale (complementary to the molecular modeling described above). The inputs to this level are the translocation free energy profiles and the dependence of the profile on applied voltage. The goal is to compute the distribution of translocation times and the mean translocation time. The distribution of translocation times was computed using the method of Lubensky [19], and are shown in Figure 5(a). The mean translocation times were then computed from the distribution and checked against the method of Schulten [20].

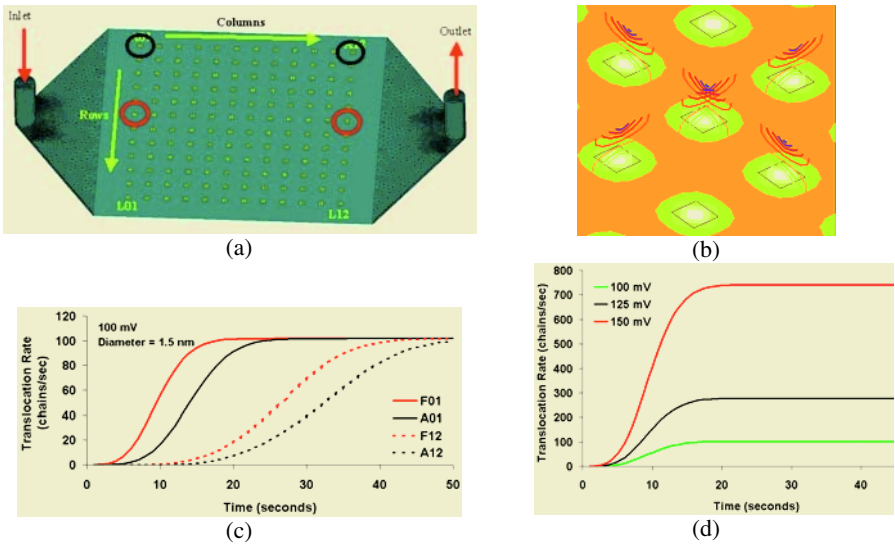


**Fig. 5.** (a) Variation in the distribution of passage times with applied field and pore diameter for cytosine. (b) Mean passage time dependence on applied voltage for all four cases

For all four cases, Figure 5(b) shows an exponential drop in the mean passage time with applied voltage, in agreement with several other researchers [21, 22]. Looking at the individual nucleotides, the cytosine shows a longer passage time with the 1.5 nm pore as opposed to the 2.0 nm pore. However, the Adenine chain shows a longer passage time with the 2.0 nm pore as opposed to the 1.5 nm pore. Also to note is the fact that there is a much larger difference in passage time comparing pore diameters with cytosine than with adenine. The reason for the differences is not immediately clear until we look at the free energy profiles (Figure 4(b)). The Adenine profiles show a large difference in the depth of the minimum on the order of 20% of the total depth. While the cytosine nucleotides do not show the same dependence. The information computed in this section will be coupled in directly with the continuum level model presented in the next section.

### 4 Continuum Level Model

This forms the final element in the three-tiered paradigm A pore model is constructed that takes into account the arrival rate of DNA in the proper conformation, and the translocation of DNA across the pore. A boundary condition was implemented in the multiphysics modeling software CFD-ACE+ (marketed by ESI Group), which accounts for the rate of capture as well as the rate of translocation (Figure 2) The translocation model contains two steps, with the first step being the rate of capture of a DNA (**R**) chain into the nanopore. This rate was not computed, but was taken from the literature [2]. The second step is the translocation of the captured DNA across the pore to free up the pore. This is assumed to occur at a rate of  $1/\tau$  where,  $\tau$  is the mean translocation time computed using the free energy profiles (Figure 5(b)). Both  $\tau$  and **R** are dependent on the applied voltage.



**Fig. 6.** (a) Three dimensional geometry of the notional device showing the labeling of the nanopore containing patches. The patches are shown in yellow. (b) Zoom in view of the electric field on each patch optimized to minimize cross talk. (c) Dependence on the Translocation Rates on the Placement. For the case of *CYS* 1.5 nm at 100, 125, and 150 mV. (d) Dependence on the Translocation Rates on the Voltage for the Case of F01 Spot

The device chosen for demonstration (Figure 6(a)) is a 21 mm by 7mm by 0.25 mm thick microfluidic chip containing 144 nanopore bearing patches. Each patch is uniquely identified (similar to a standard spreadsheet) with a letter and a number. Within the computational domain of the chip, the Navier-Stokes equation for an incompressible fluid is solved along with the equation of mass transport for DNA. A sink term is added at each patch based on the rate of translocation. Initially there is assumed to be no DNA in the chip. At time zero, DNA begins to flow into the chip at a concentration of  $10^{-6}$  Molar being carried by the fluid that is being pumped at a rate

of 100 microL/min. The total simulated time is 50 seconds, long enough for the pore translocation to reach a steady state. Also, a constant voltage was held across each patch, and the solution for the electric field was computed during the simulation. This calculation allowed for the determination of the optimal spacing of the pore containing patches to minimize electronic cross talk (Figure 6(b)).

To understand the dependence of the output signals upon pore placement, four pores were monitored. A01, A12, F01, and F12 (see Figure 6(a) for locations). These represent extremes in terms of placement within the flow field and placement along the line of flow. Figure 6(c) shows the evolution of the translocation rate for three voltages and the four candidate spots. All three voltages show a similar trend of F01 and A01 being the first to reach steady state, while F12 and A12 show a nearly 20 second lag. This lag time will be strongly dependent upon the inlet flowrate of the device, with an increasing flowrate decreasing the lag time. An increase in the applied voltage increases the point at which the patch translocation rate reaches a steady state. Figure 6(d) shows the dependence of the DNA translocation rates on the applied pore voltage. The translocation rates are shown to increase exponentially in agreement with [21].

## 5 Summary and Conclusions

The design of devices that incorporate both nanoscale and continuum scale physics is challenging due to the nature of simulation techniques for each scale. This paper presents a general three-tiered modeling paradigm for use in simulating devices containing a nanoscale component. The models presented will be incorporated into a continuum code and ultimately used to simulate integrated nanobio devices. Future objectives for this research include the development of a generic framework for simulation of systems of Master Equations and Fokker-Planck Equations. The ME and FPE solver will be designed as a separate module to be integrated with a continuum code for simulations of integrated nano-bio systems. Future work will focus on the demonstration of the feasibility and efficiency of the proposed approach on alternative test systems.

The example system is one of the simulation of a nanopore-based DNA sequencing device. The results from the device level calculation show that the overall translocation rate is exponentially dependent on the applied voltage. The placement of the patches of nanopores affects the rate at which they reach a steady translocation rate. The final result is that the nanopore diameter had only a marginal affect on the overall translocation rates. The main conclusion from the device level modeling is that in order to improve the performance of this type of a device it is desirable to increase the rate at which DNA is captured. This can be accomplished via the addition of a surface treatment around the pore via functionalization.

## Acknowledgement

The authors gratefully acknowledge NASA-AMES (NNA04CB26C, Technical Monitor: Dr. Harry Partridge) for funding, and thank Dr. Harry Partridge and Dr. Jun Li for their many helpful discussions.

## References

1. Stolc, V., Cozmuta I., Brock, M., O’Keeffe, J.: Presentation given at the DARPA Joint BIOFLIPS/SIMBIOSYS Principal Investigators Meeting, Monterey, CA (2003).
2. (a)Meller, A., Branton D.: Electrophoresis, 23 (2002) 2583-2591 and (b) Meller, A., Nivon L., Branton D.: Physical Review Letters, 86 (2001) 3435-38.
3. Li, J., Stein, D., McMullan, C., Branton, D., Aziz, M., and Golovchenko, J.: Nature 412 (2001) 166-8.
4. Brodka, A. and Zerda, T.: Journal of Chemical Physics 104 (1996) 6319-6326.
5. Li, J., Gershow, M., Stein, D., Brandin, E., Golovchenko, J.: Nature Materials, 2 (2003) 611-15.
6. Humphrey, W., Dalke, A. and Schulten, K.: "VMD - Visual Molecular Dynamics", J. Molec. Graphics, 14 (1996) 33-38.
7. Auffinger, P. and Westhof, E.: Journal of Molecular Biology, 269 (1997) 326-341.
8. Kalé, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K., and Schulten, K.: NAMD2: Greater scalability for parallel molecular dynamics. Journal of Computational Physics, 151 (1999) 283-312.
9. Izaguirre, J., Reich, S. Skeel, R.: Journal of Chemical Physics, 110 (1999) 9853-64.
10. Skeel, R., Izaguirre, J.: Molecular Physics 100 (2002) 3885-91.
11. Foloppe, N. and MacKerell, Jr., A.D.: Journal of Computational Chemistry, 21 (2000) 86-104.
12. MacKerell, Jr., A.D. and Banavali, N.: Journal of Computational Chemistry, 21 (2000) 105-120.
13. Beglov, D. and Roux, B.: Journal of Chemical Physics, 100 (1994) 9050-9063.
14. Rovere, M., Ricci, M., Vellati, D., and Bruni, F.: Journal of Chemical Physics, 108 (1998) 9859-9867.
15. Schlitter, J.: Chemical Physics Letters, 215 (1993) 617-21.
16. Malliavin, T., Jocelyn, G. Snoussi, K. and Leroy, J.: Biophysical Journal, 84 (2003) 1-10.
17. Rueda, M., Kalko, S., Luque, J., and Orozco, M.: JACS, 125 (2003) 8007-14.
18. Edwards, S., Corry, B., Kuyucak, S., Chung, S.: Biophysical Journal, 83 (2002) 1348-60.
19. Lubensky, D. and Nelson, D.: "Driven Polymer Translocation Through a Narrow Pore" Biophys. J. 77 (1999) 1824-38.
20. Park, S., Sener, M., Lu, D., Schulten, K.: J. Chem. Phys. 119 (2003) 1313:19.
21. Ambjornsson, T. Apell, S. Konkoli, Z. Marzio, E. Kasianowica, J.: Journal of Chemical Physics. 117 (2002) 4063-73.
22. Chuang, J., Kantor, Y., Kardar, M.: Phys. Rev. E. E65 (2002) 11802(1-8).