August 31st, 2012

# IN-SILICO CHARACTERIZATION AND DESIGN OF NANOPORE-BASED DNA SEMICONDUCTOR SEQUENCING DEVICES

**California Institute of Technology**

**Andres Jaramillo-Botero, Qi An, Hai Xiao and William A. Goddard**

**Final report, (2nd semester of FY2)**



## FIRST-PRINCIPLES-BASED CALCULATIONS FOR CHARACTERIZING AND OPTIMIZING THE PERFORMANCE OF SOLID-STATE NANOPORE DEVICES FOR DNA SEQUENCING

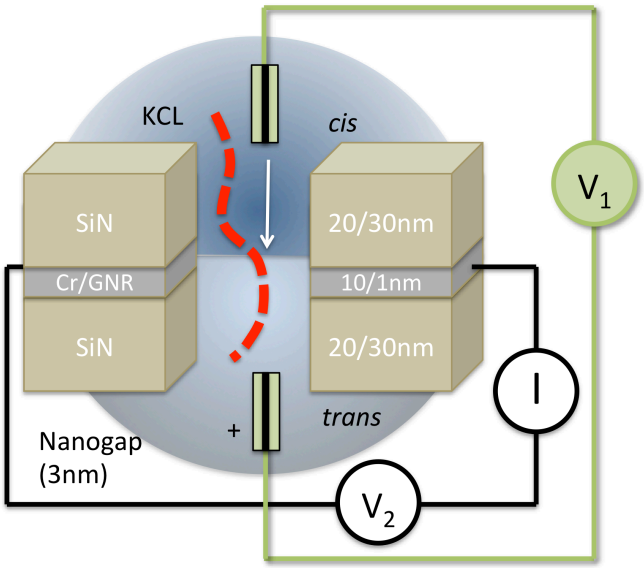



FIGURE 1 MODEL NANOSEQUENCING DEVICE DESIGN

On this report we describe our work on:

1. DNA translocation control via asymmetric versus symmetric ionic ($K^+$/$Cl^-$) concentrations under a bias voltage $V_1$ between the *cis* and *trans* reservoirs of a nanopore device setup (shown in Figure 1), and
2. Conductivity characterization of dsDNA in a two-probe setup with Cr versus Au electrodes with a bias voltage $V_2$ (shown in Figure 1).

Molecular dynamics simulations were used for 1 and quantum mechanics nonequilibrium Green's function (NEGF) calculations for 2. Details are provided in the corresponding sections below.

August 31st, 2012

## MOLECULAR DYNAMICS OF DNA TRANSLOCATION THROUGH SOLID-STATE MEMBRANE

***Model.*** Two crystalline Si3N4 membranes (~10 nm in thickness each) were constructed, each from a unit cell of β-Si3N4 crystal, and placed on top of each other. Each membrane has a length along the z-axis of 36 Si3N4 unit cells (10.45 nm) and a hexagonal cross section in the *xy* place of 12 Si3N4 unit cells. A cylindrical pore of 3nm in diameter was produced on the Si3N4 substrate slab. The middle layers of the membrane (z=0) that belong to one unit cell were removed and electrode segments were placed, in sandwiched configuration, at z =0 and separated by 3nm, which is equal to the chosen pore diameter. A 40 basepairs dsDNA helix was built from individual basepairs. The helix was oriented normal to the membrane and placed 1 nm outside of the pore. The Si3N4/DNA complex was then solvated in a volume of pre-equilibrated TIP3P water molecules at a bulk water density within the solvent accessible volume. K+ and Cl- ions were added in proportion to the desired molar concentration, for both symmetric (which ranged from 0.01M, 0.1M, 1.0M and 3M) and asymmetric concentrations, including the *cis/trans* concentration ratios of 1M/4M, 1M/0.2M, 2M/0.2M and 3M/0.2M). Figure 2 depicts the general setup without explicitly showing the water molecules.

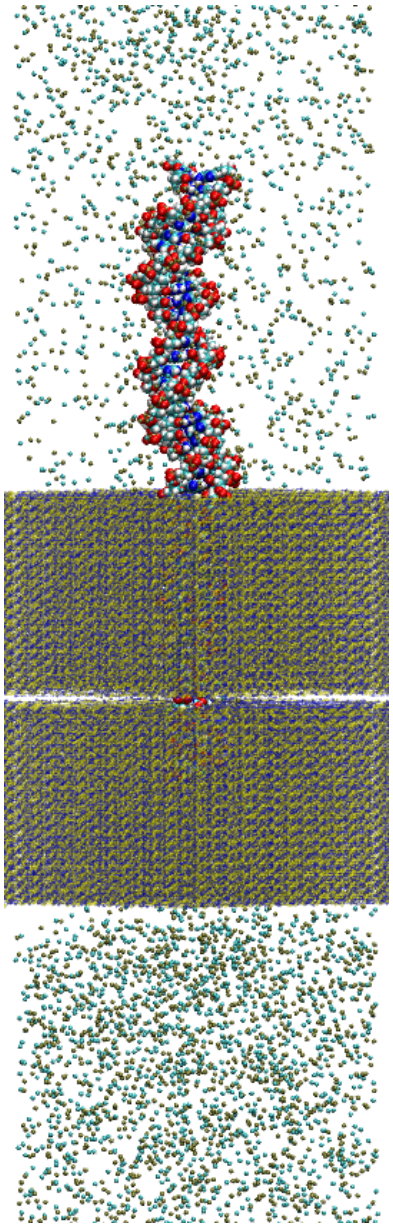


FIGURE 2 (LEFT) SI3N4 PORED-MEMBRANE AND THE DSDNA IN ASYMMETRIC CONCENTRATION SOLVENT SETUP. TWO SI3N4 SLABS SANDWICH GNR/CR ELECTRODES (NOT SHOWN). WATER NOT SHOWN. CL- ATOMS SHOWN IN CYAN AND K+ ATOMS SHOWN IN BROWN.

MD simulations were carried out using the program NAMD2 ([1]), with the CHARMM[2] force field describing the Si3N4 nucleic acid, water, graphene and ions, and the van der Waals parameters and charges of Si3N4 atoms were taken from our group's force field[3]. Every Si atom in the Si3N4 membrane was restrained by a harmonic force to its initial location in the crystalline membrane with the force constant of 0.01 or 0.1 kcal/(nm$^2$ mol) for bulk and surface atoms, respectively. The electrode layers were also restrained to the Si3N4 surface by a harmonic force using a force constant of 0.05 kcal/(nm$^2$ mol). The spring constant of the harmonic bonds connecting silicon and nitrogen atoms in Si3N4 was adjusted to yield the relative permittivity of bulk Si3N4 of 7.5[4]. The DNA-specific force generating from grid-steered molecular dynamics method[5] is used to reduce the interaction between the DNA and the pore surface and thus prevent irreversible binding of DNA to the pore walls.

We performed 2,000 steps of minimization, or until convergence to either 1e-4 in energy difference or 1e-5 in RMS force occurred, followed by a 500 ps NPT simulation to equilibrate the system. We then applied a uniform electric field with 6.5 V along *z* directions and ran the simulations with the ssDNA translocating through the membrane. In all simulations, the temperature was kept at 295 K by applying Langevin thermostat with the damping constant of 1.0/ps.

The total size of the modeled solvated system is >220,000 atoms, which varies depending on the particular concentration of ions chosen. This heavily weights on the computational cost of these simulations, considering the integration time step for the fully periodic system is 1fs. Simulations ran for an average of 3 weeks (per

model system) using an average of 196 processors per run in the SHC cluster at Caltech (http://www.cacr.caltech.edu/main/?page_id=101).

## TRANSLOCATION UNDER SYMMETRIC IONIC CONCENTRATIONS

The concept here involved testing the effect of different ionic concentrations on DNA (ss and ds) translocation speed through the solid-state nanopore. From Figure 3(right) we observe that: 1) there is a sharp change in translocation speed for small increases in concentration, above and below a sweetspot, 2) the concentration sweet-spot that maximizes the translocation speed is between 0.1 and 1.0M, 3) the ssDNA translocates even at low ionic concentrations due to the net negative charge of its backbone phosphate groups, 4) there is a linear decrease in translocation velocity after the sweet-spot, and the relationship between concentration and speed becomes inversely proportional, and pore DNA capture is maximum at the sweetspot concentration.

At 3M the translocation rate is significantly reduced, when the DNA is in the pore (slope of curves in Figure 3(left), after DNA capture by the pore). This provides evidence of an electrokinetic effect that can be used to control the nucleotide translocation rate, i.e. under the influence of an external field applied between the two ionic reservoirs above and below the nanopore membrane (i.e. positive voltage to the *trans* chamber), diffusion of Cl- ions from the *trans* to *cis* chamber is hindered since these must overcome the applied electric potential energy. In contrast, K+ ions move along the direction of both the chemical and electrical potential gradients, from *trans* to *cis*.

On the other hand, the pore capture efficiency is improved for the higher concentrations as seen from the time of capture for each concentration case in Figure 3.

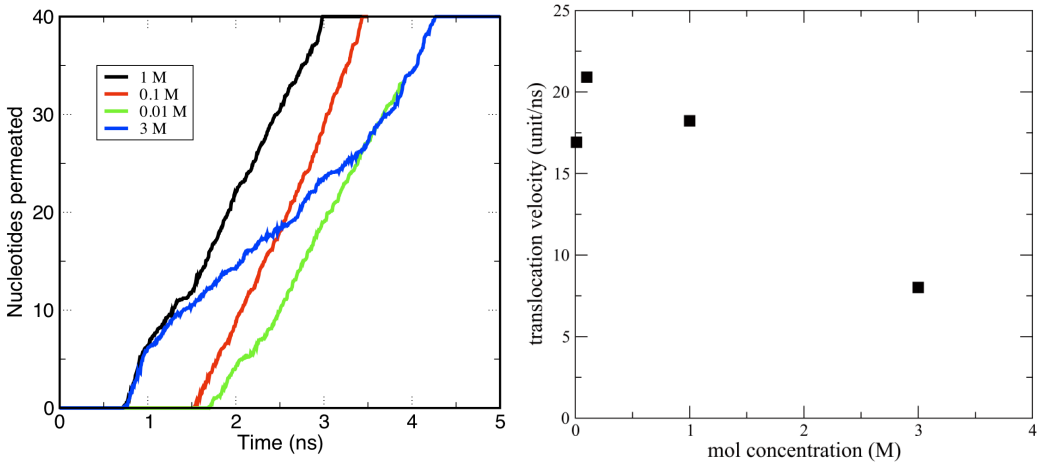



FIGURE 3 (LEFT) NUMBER OF NUCLEOTIDES PASSING THE PORE CENTER AS A FUNCTION OF TIME FOR DIFFERENT IONIC SOLUTION CONCENTRATIONS (3M HAS NOT COMPLETED), AND (RIGHT) ESTIMATED TRANSLOCATION VELOCITY AS A FUNCTION OF MOLAR CONCENTRATION FOR THE IONIC SOLUTION.

## TRANSLOCATION UNDER ASYMMETRIC IONIC CONCENTRATIONS

The effect of the bias voltage, $V_{bias}$, between the reservoirs separated by the nanopore (see Figure 1) depends on the ionic environment in the vicinity of the pore. Under symmetric ionic strengths, the electric field across the pore equally drives K+ ions from *trans* to *cis* and Cl- ions from *cis* to *trans*, creating a symmetric steady-state ion flux (neglecting electro-osmosis). In contrast, under asymmetric salt conditions, the applied voltage may result in cation selectivity. For example, when the salt concentration in the *cis* chamber is much lower than in the *trans*; under a typical field magnitude and polarity applied to the nanopore (i.e. positive voltage to the *trans* chamber), diffusion of Cl- ions from the *trans* to *cis* chamber down the concentration gradient is hindered by the applied potential (a transported Cl- ion must overcome an electric potential energy of ~0.3 eV). In contrast, K+ ions are pumped from *trans* to *cis*,

moving in the direction of both the chemical and electrical potential gradients. As positive ions are continuously pumped into the *cis* chamber, the pore vicinity is effectively polarized and the effective magnitude of $V_{bias}$ within a cutoff distance increases (V(r)). This effect enhances the capture rate in both the diffusion-limited and barrier-limited regimes. This can be better seen in the movie provided at http://www.wag.caltech.edu/multiscale/movies/dna_1_4.mpg

Under asymmetric salt conditions, Wanunu et al [6] predicted that, to a first approximation, V(r) is proportional to the ratio of bulk ionic concentrations in the *cis* and *trans* chambers. Accordingly, the value of r is modified proportionally to the salt concentration ratio. The rationale behind this is that the voltage drop in the solution of lower conductance is higher and, therefore, a high/low *trans*/*cis* salt gradient will yield an asymmetric potential profile, with more of the voltage dropping in the *cis* chamber than the *trans* chamber.

To corroborate the predictions from [6] we prepared and ran molecular dynamics simulation of dsDNA capture and translocation through a Si3N4 3nm nanopore membrane in K+ and Cl- solution at four different $C_{cis}/C_{trans}$ concentration ratios: 1M/4M, 1M/0.2M, 2M/0.2M and 3M/0.2M. The results and corresponding analysis are provided below.

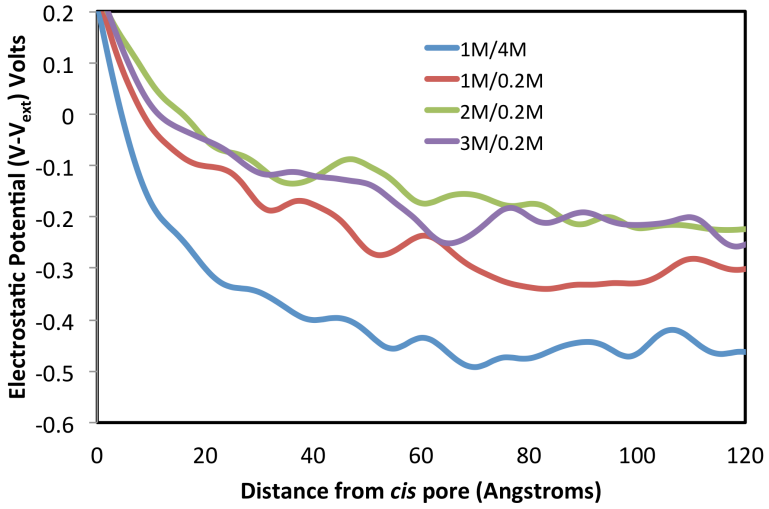



FIGURE 4 DEPICTS THE INCREASE IN THE EFFECTIVE ELECTROSTATIC POTENTIAL FIELD, I.E. ($V-V_{EXT}$), FELT NEAR PORE IN THE CIS SIDE AS THE $C_{CIS}/C_{TRANS}$ RATIO DECREASES (E.G. 1M/4M).

Our findings coincide with those from [6]. That is, the electric potential distribution in the *cis* side of the pore affects the capture rate of DNA (ss and ds), for both symmetric and asymmetric KCl concentrations cases. This effect is more evident for the asymmetric salt concentration results, which shows a larger effective electrostatic potential magnitude (see Figure 4) and range outside the *cis* pore for smaller $C_{cis}/C_{trans}$ ratios. These ratios were calculated as added electrostatic charge (with respect to $V_{ext}$) as a function of distance from the pore during the translocation simulations. The axial dependence of the potential as a function of distance from the pore mouth, shown for four different salt gradients, is indicated in Figure 4 (albeit it is limited by the size of our modeled unit cell).

## TRANSLOCATION RATE AND IONIC CURRENT UNDER ASYMMETRIC CONCENTRATIONS

We calculated the translocation rate (Figure 5) and the ionic current blockade (Figure 6) during the translocation events for the different $C_{cis}/C_{trans}$ concentration ratios.

We observe from Figure 5 that the average time for the dsDNA to reach the middle of the nanopore, starting from the pore mouth, is ~1.5ns, for all $C_{cis}/C_{trans}$ concentration ratios. Past this point the translocation rate is approximately constant for all cases, i.e. by ~3ns the pore is completely occupied by

the dsDNA. After 3ns, there is a clear difference in translocation speed observed for the 1M/4M $C_{cis}/C_{trans}$ concentration case that we attribute to the electrokinetic effects described previously. This reasoning seems to be consistent with ionic current blockade profiles depicted in Figure 6 and explained below.

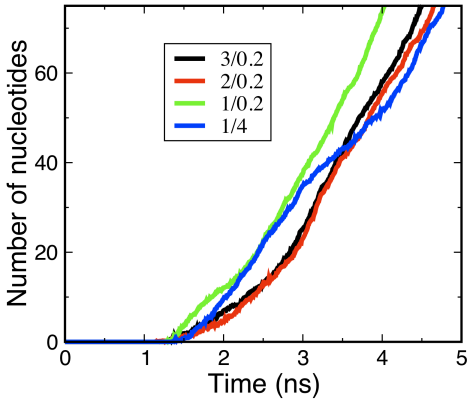



FIGURE 5 TRANSLOCATION RATE AS A FUNCTION OF $C_{CIS}/C_{TRANS}$ RATIO CONCENTRATION. COUNT OF TRANSLOCATED NUCLEOBASES IS DONE AFTER THE DSDNA REACHES THE ELECTRODES, I.E. MIDWAY THROUGH THE NANOPORE.

Figure 6 shows the ionic current measured within the membrane pore for the different asymmetric concentrations. Because the membrane has a thickness of 10 *nm*, we compute the current over -4.5 nm < z < 4.5 nm. The instantaneous ion current has large fluctuations due to thermal noise. To obtain reliable current values that can be compared with experiment, we average the instantaneous current over long trajectories. Here we use a time window average of 0.1 ns. Full current blockage is not observed, i.e. there is a finite leakage current.

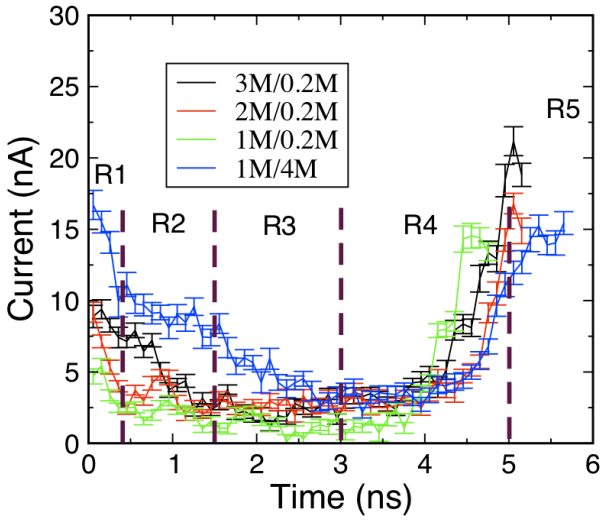



FIGURE 6 AVERAGE IONIC CURRENT IN THE NANOPORE AS A FUNCTION OF TIME FOR THE DIFFERENT ASYMMETRIC $C_{CIS}/C_{TRANS}$ RATIO OF SALT CONCENTRATIONS.

From Figure 6 we observe a faster and larger drop in the ionic current for the $C_{cis}/C_{trans}$ concentration ratio of 1M/4M, during the first 0.5ns. This may indicate a faster dsDNA capture into the pore, but we would have to do additional calculations with the dsDNA placed initially at a more distant location from the *cis* pore entry to decisively conclude this. Once the dsDNA has reached the middle of the pore (by ~1.5ns) the translocation rate is relatively the same for all cases, except that there is more ionic leakage current for the 1M/4M case. By ~5ns all the cases have emerged out of the pore.

We can therefore distinguish 5 different regimes from our simulations marked in Figure 6 as,

R1) initial capture,

R2) pore halfway occupied,

R3) pore fully occupied and

R4) pore halfway unoccupied, and

R5) pore open.

During regime 3, we observe that the 1M/4M case translocates the slowest through the pore, which is consistent with the explanation offered for the ssDNA translocation electrokinetic effects (Figure 3). As we mention, this may provide a mean to control translocation rates once the dsDNA has been captured in the pore. On the other hand, this same concentration offers improved capture, i.e. longer range and magnitude of the electrostatic driving potential with respect to the pore mouth in the *cis* end.

Considering these 5 different regimes indicates that dsDNA throughput would depend on several factors, including: bias voltage ($V_b$), capture time ($t_c$), ionic current leakage ($I_l$), cost of filling the pore ($t_f$), rate of translocation ($r_t$), and average DNA length given in terms of total pore length ($d_l$). If we assume for simplicity that $T=t_f=t_e=t_c$, that $t_r$ is constant (in practice it will be a function of $V_b$ and $I_l$), and that there is a capture probability ($c_p$) then a general throughput expression is determined by,

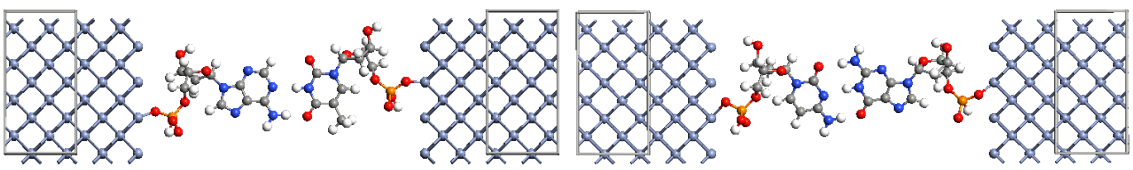
$$throughput = \frac{n}{\left[(s-1)/c_p + d_l n\right] r_t T}$$

In the limit of 100% capture probability (i.e. one chain can be capture immediately after the tail of another one enters the pore) and an infinite number of chains in the *cis* reservoir, the throughput would be equal to $1/d_l r_r T$. We may also deduce that capture efficiency, rate of translocation and DNA length all play a significant role in the throughput. We have focused here on mechanisms to improve capture efficiency and to control the translocation rate, both of which can be managed via regulation of the salt concentrations in the *cis* and *trans* reservoirs (the asymmetric case would require a salt pump for the device to work properly). We have not yet discussed the effect of DNA length on throughput and this may be addressed later. Parallel pore arrays have been also used to improve the net capture efficiency.

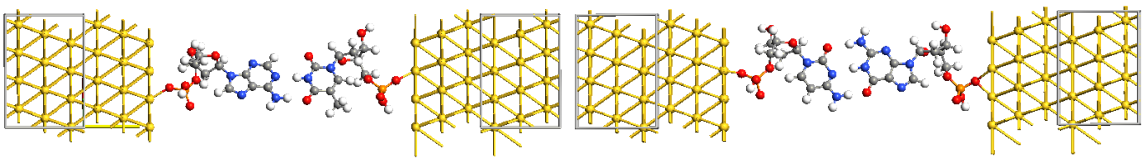
We are currently performing additional calculations to understand the field effect enhancement range from asymmetric salt concentrations, by calculating the capture efficiency for dsDNA at varying distances from the *cis* pore, and will start calculations to test the DNA-length effects on capture, in order to completely characterize the asymmetric concentration ratio on the capture efficiency. These will be updated as we move forward.

## CROMIUM ELECTRODES IN TUNNELING CONDUCTIVITY OF $_{DS}$DNA

This study explores the use of Cr electrodes versus Au electrodes for the two-probe setup used in extracting tunneling signatures from dsDNA (we had partially reported on sdDNA base IVs in the midterm report, but we shifted to dsDNA by request of our SAIT contact). Four different configurations where prepared to determine the transverse electrical current through the dsDNA as a function of bias voltage between the metal electrodes (see Figure 7).



a) b)



c) d)

FIGURE 7 SHOWS THE TWO-PROBE SETUP USED FOR TUNNELING CALCULATIONS THROUGH THE DSDNA BETWEEN TWO METAL ELECTRODES, SHOWN AS GRAY BOXES. A-B) CR AT AND GC, RESPECTIVELY, AND C-D) AU AT AND GC, RESPECTIVELY.

Since the tunneling probability would be dependent on the molecular orbitals of the dsDNA and the electrical resistivity of Cr-DNA interface and the bulk Cr electrodes, we focus on the binding affinity at the interface between the metal and the organic molecule. Considering the electrical resistivity of bulk Cr is almost 6x higher than that of gold in the bulk, we speculate that the molecule-metal binding ability of Cr could be the answer to experimental observations by the SAIT group.

***Computational details.*** Gold electrodes were prepared from a cleaved {111} gold surface in fcc stacking sequence, repeated to a 4x4x4 size and mirror-copied and displaced to make the two-probe device with enough vacuum space for the dsDNA pairs. Alignment points were created on each electrode surface and the DNA pairs were attached to these. Similarly, Chromium electrodes were made from 4x4x4 {100} cleaved bcc Cr and the corresponding two-probe systems built for each molecular nucleobase pair junction (i.e. AT and GC) as shown in Figure 7. Organic molecules were prepared separately, minimized with a force field and incorporated into the two-probe QM model.

The bonding anchor between the metal and the dsDNA phosphate backbone was chosen from breaking O-H bonds on Phosphorus (since these are easily broken or ionized in a solvated environment). This is different from other published work in which the C-H is used as anchor [7], in spite of its relatively high dissociation energy.

Calculations were performed using Density Functional Theory (DFT) with PBE-GGA exchange correlation and a PZ functional and a double zeta polarized LCAO basis set. Electron temperature was set to 300K. Geometries were optimized using a Quasi-Newton optimizer (Broyden–Fletcher–Goldfarb–Shanno [BFGS] method) with convergence criteria set to a threshold force of 0.01eV/Angstroms (~$4x10^{-4}$ au/A) and a maximum step size of 0.5 Angstroms. Transmission spectra and voltage drops were calculated self-consistently on the biased configurations over an energy range of -1:1 E/eV using a Monhorst Pack grid with 4x4x101 k-point sampling in x,y,z directions, respectively, were z is the transmission direction.

The Cr cases are generally much harder to converge than Au cases, because DFT has difficulty with first row transition metals and because Cr metal is anti-ferromagnetic in ground state, which significantly complicates the calculation; in particular, considering its Néel or magnetic ordering temperature (308 K) is close to room temperature. For this reason we decided to start with non-magnetic optimization on the unbiased system, and then use the resulting geometry for full optimization. This has proven to work for the Cr-AT-Cr case as shown in Figure 8. The first few attempts to optimize the Cr-GC-Cr case failed, and by using the procedure described previously we expect to have it fully optimized in a couple of weeks. We are exploring other options to improve convergence of Cr geometry optimizations in two-probe calculations using ATK (v12.2), beyond atomic constraints, but so far they involve coding low-level routines that would need to be compiled and wrapped in python to expose them to the atkpython library.

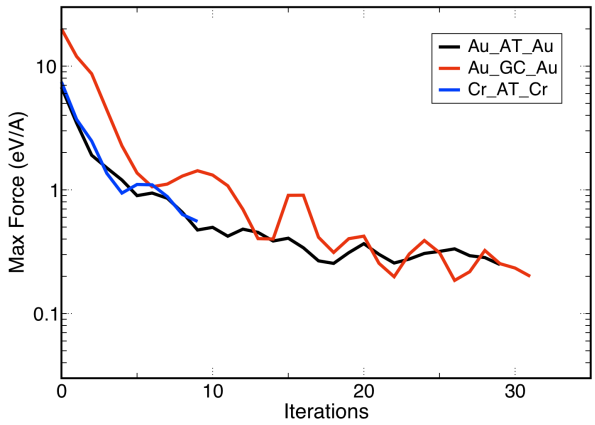



FIGURE 8 MAX FORCE DURING DFT GEOMETRY OPTIMIZATION OF OUR TWO-PROBE SYSTEMS SHOWN IN FIGURE 7. THE CR-AT CASE IS STILL RUNNING AND IS EXPECTED TO FULLY CONVERGE TO THE ~0.1EV/ANG MAX FORCE WITHIN DAYS OF THIS REPORT.

I-V and voltage-drop calculations are now being performed, for a range of -1.5-1.5V, using the resulting optimized two-probe structures, with SCF but without geometry optimization under bias since it has little effect on the transmission (and is very time consuming). We've have had to perform the geometry optimization calculations several times to secure convergence and no dsDNA geometry deformations. For the Cr-GC-Cr case we had to increase the electrode size to avoid periodic interactions in the dsDNA images that led to incorrect molecule geometries. These (optimization, transmission spectrum, and voltage drop) calculations take a significant amount of time (3-4 weeks) and are being executed each on 3-nodes with 8-processors/node, i.e. a total of 24 cores. We do not include the partial results here, since we can't compare the IV characteristics between the two electrode types with the results obtained at this stage. We expect to have these within the upcoming month, and we will submit these to our SAIT contact at that time.

To explain the nature of relatively high currents (300~600 pA transverse currents for dsDNA), with respect to Au, through the Cr electrodes reported by SAIT, we contend the following. Although being one of the best conductors, Au is the least reactive metal, in the sense that chemical adsorption on its surface normally is not favored[8]. Thus, when our dsDNA target molecule is introduced between two gold electrodes, electronic current transmission (tunneling or hole-based) is dependent on the molecule itself, i.e., the selectivity relies only on the difference between targets, hence we observe different IV characteristics between AT and GC dsDNA pairs. On the other hand, the Cr metal surface chemically binds very well with various species [9], including our organic molecules. Since chemical bonds are generally highly selective, we speculate that the difference in transmission spectrum between our targets is enhanced by chemisorption on Cr electrodes, thereby leading to a better electrode material for our dsDNA sequencing solid-state devices. To corroborate this we prepared independently optimized geometries for Au and Cr electrodes and an AT dsDNA pair molecule. We are now computing the binding affinity between the dsDNA and each of the electrode types. We will submit these results as we move forward, in conjunction with the IV comparisons.

## REFERENCES

1. Kale, L., R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten, *NAMD2: Greater scalability for parallel molecular dynamics.* Journal of Computational Physics, 1999. **151**(1): p. 283-312.
2. Brooks, B.R., R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus, *Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations.* Journal of Computational Chemistry, 1983. **4**(2): p. 187-217.

3. Wendel, J.A. and W.A. Goddard, *The Hessian Biased Force-Field for Silicon-Nitride Ceramics - Predictions of Thermodynamic and Mechanical-Properties for Alpha-Si3n4.* Journal of Chemical Physics, 1992. **97**(7): p. 5048-5062.
4. Xu, D., J.C. Phillips, and K. Schulten, *Protein response to external electric fields: Relaxation, hysteresis, and echo.* Journal of Physical Chemistry, 1996. **100**(29): p. 12108-12121.
5. Wells, D.B., V. Abramkina, and A. Aksimentiev, *Exploring transmembrane transport through alpha-hemolysin with grid-steered molecular dynamics.* Journal of Chemical Physics, 2007. **127**(12).
6. Wanunu, M., W. Morrison, Y. Rabin, A.Y. Grosberg, and A. Meller, *Electrostatic focusing of unlabelled DNA into nanoscale pores using a salt gradient.* Nature Nanotechnology, 2010. **5**(2): p. 160-165.
7. Jauregui, L.A., K. Salazar-Salinas, and J.M. Seminario, *Transverse Electronic Transport in Double-Stranded DNA Nucleotides.* Journal of Physical Chemistry B, 2009. **113**(18): p. 6230-6239.
8. Hammer, B. and J.K. Norskov, *Why Gold Is the Noblest of All the Metals.* Nature, 1995. **376**(6537): p. 238-240.
9. Frese, K.W., *Calculation of Surface Binding-Energy for Hydrogen, Oxygen, and Carbon-Atoms on Metallic Surfaces.* Surface Science, 1987. **182**(1-2): p. 85-97.