DNA Sequencing by Synthesis Using Fluorescent, Raman, and Nanopore Detectable Tagged Nucleotides

Jingyue Ju Laboratory Members and Collaborators

Columbia Genome Center, Columbia University College of Physicians and Surgeons, New York, NY 10032; and Departments of †Chemical Engineering, ‡Mechanical Engineering, and §Chemistry, Columbia University, New York, NY 10027

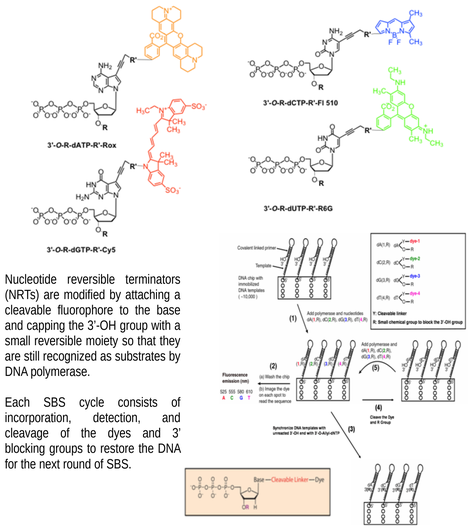
Abstract

DNA sequencing is a fundamental tool in biological and medical research. High throughput DNA sequencing is especially important for the paradigm of precision medicine. The approach that has dominated the second generation of sequencing technologies is sequencing by synthesis (SBS), whereby the DNA sequence is determined as each nucleotide adds to the growing strand of DNA in the polymerase reaction. This is made possible by the presence of detectable tags that distinguish the 4 bases of DNA (A, C, G and T). In the Ju laboratory, we have pioneered the development of the SBS approach using a wide variety of tagged nucleotides.

- Sets of nucleotides with molecular tags detectable by mass spectroscopy
- Sets of nucleotides with fluorescent tags detectable by optical methods
- Sets of nucleotides with molecular tags detectable by surface enhanced Raman spectroscopy (SERS)
- Sets of nucleotides with polymeric tags detectable at single molecule level by their effect on ion transport through nanopores

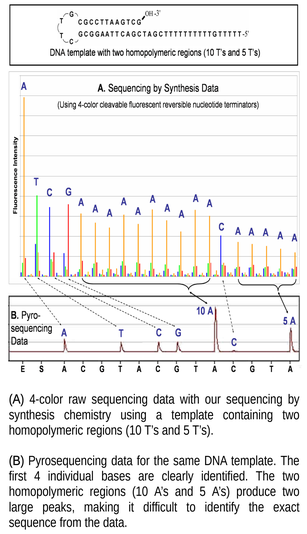
In this poster, we present some of our work on the latter 3 methods. The fluorescent SBS approach (upper center) underpins the most popular second generation sequencing platform. SERS-SBS (upper right) may offer very high sensitivity. The nanopore-SBS real-time single molecule approach (lower left) offers the opportunity to achieve very long sequence reads at low cost and high speed.

Fluorescence SBS Using Nucleotide Reversible Terminators (NRTs)



Nucleotide reversible terminators (NRTs) are modified by attaching a cleavable fluorophore to the base and capping the 3'-OH group with a small reversible moiety so that they are still recognized as substrates by DNA polymerase.

Each SBS cycle consists of incorporation, detection, and cleavage of the dyes and 3' blocking groups to restore the DNA for the next round of SBS.



(A) 4-color raw sequencing data with our sequencing by synthesis chemistry using a template containing two homopolymeric regions (10 T's and 5 T's).

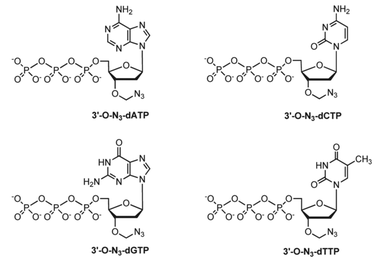
(B) Pyrosequencing data for the same DNA template. The first 4 individual bases are clearly identified. The two homopolymeric regions (10 A's and 5 A's) produce two large peaks, making it difficult to identify the exact sequence from the data.

Raman Measurements of NRTs on SERS Substrate

Azidomethyl Modified NRTs

In our SERS-SBS approach, nucleotides were modified with 3'-azidomethyl moiety to:

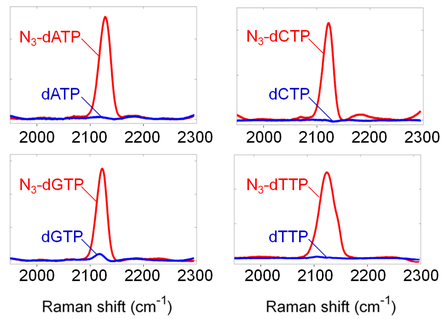
- **temporarily terminate** polymerase reaction after incorporation
- serve as the **reporter group** with distinct Raman shift at 2080-2170 cm^{-1}, in a spectral region where DNA and proteins do not elicit signals



Structures of the nucleotide reversible terminators.

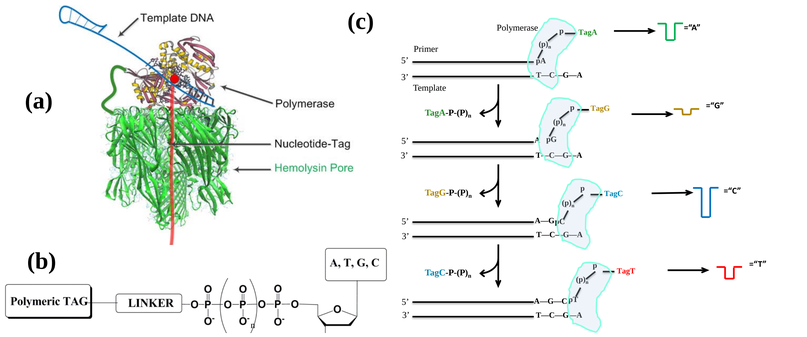
SERS Measrurements

- Four 3'-*O*-N_3-dNTPs display enhanced Raman scattering at 2125 cm^{-1} on Klarite SERS substrates
- Natural dNTPs produce only a background signal
- 10^2-fold signal increase at the expected Raman shift of 2125 cm^{-1} due to the N_3 group

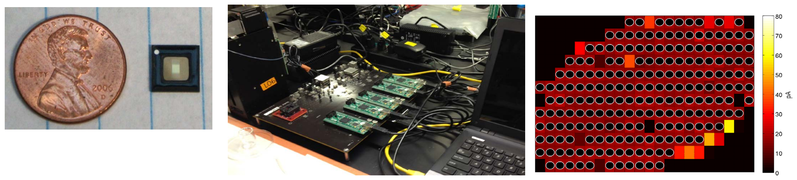


Raman signal of 3'-*O*-N_3-dNTPs (red) and natural dNTPs (blue).

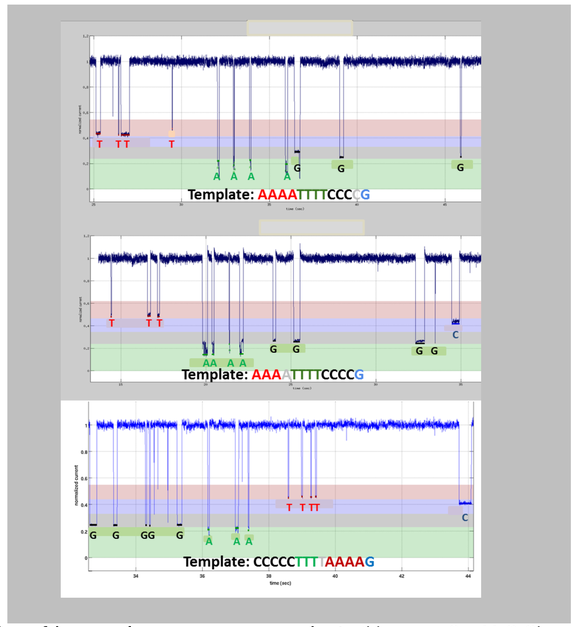
Single Molecule Electronic Sequencing by Synthesis (Nanopore-SBS)



Principle of the Nanopore SBS Sequencing Method. (a) A single polymerase molecule covalently attached to the α-hemolysin nanopore heptamer. Template DNA and primer bind, along with a tagged nucleotide during incorporation. (b) A generalized diagram of the tagged nucleotides. (c) SBS schematic showing the sequential capture and reading of nucleotide tags as they are incorporated into the growing primer strand.



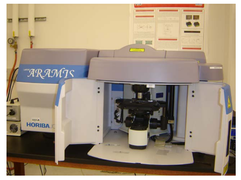
Automated Nanopore Array Setup: In 3 automated steps, the 264 sensor IC electrodes are first coated with lipid and salt solution to create bilayers, pore solution is added to insert single pores, and when the pores are ready, the remaining reagents are added to initiate sequencing. Shown are photographs of the chip, the device and a heat map of 200 single pores over an array of sensors. Black circles represent single pores over single sensors.



Examples of homopolymer sequence reads. In this case, 3 mM $SrCl_2$ and 0.1 mM $MnCl_2$ was present on the *trans* and 0.1 mM $MnCl_2$ on the *cis* side of the membrane. Deflections less than 10 ms were ignored. Letters shown in gray in the template indicate bases that were missed in these reads.

Nucleotide	Blockade Level	Tdwell [ms]
dT6P-dSp8-	0.5 to 0.6	17
dC6P-dSp3-	0.4 to 0.5	30
dG6P-dT30-	0.3 to 0.4	29
dA6P-FldT-	0.2 to 0.3	17

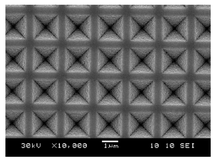
Instrumentation and SERS Substrate



Jobin-Yvon LabRam ARAMIS Raman microscope.

Data Acquisition Parameters:

- Laser wavelength: 785 nm
- Magnification: 50x (NA=0.5) dry objective
- Exposure time: 10 s, 5 accumulations
- Power:100 mW (before objective)



Scanning electron micrograph of Au-coated, lithographically patterned substrate – Klarite.

Conclusions and Other Projects

We carried out proof of principle studies demonstrating the ability to conduct SBS using a variety of novel tagged nucleotide analogues. Depending on the chemical structure of these tags, the specific nucleotide added to the growing strand can be detected optically (fluorescence based SBS), by Raman spectroscopy (SERS-SBS), or electronically at single molecule level in real time (Nanopore-SBS).

With our multidisciplinary team of organic chemists, molecular biologists, biochemists, chemical engineers and bioinformaticians, we continue to further develop these methods in our laboratory, with the aim of enhancing accuracy, lowering cost, and reducing time to obtain sequences for a wide variety of biomedical and biological applications (e.g., whole genome and whole exome sequencing, disease-specific diagnostic sequencing and SNP genotyping, single cell RNA expression studies, and metagenomic sequencing of the human microbiome. Other lab projects at various stages of development include new technologies for characterizing the epigenomic landscape (e.g., methylated and hydroxmethylated cytosines) and for detecting non-DNA biomarkers.

Ju Laboratory Publications for Material Discussed in This Poster

Fluorescent SBS: Seo et al (2005) PNAS 102: 5926; Wu et al (2007) PNAS 104:16462; Guo et al (2008) PNAS 105:9145.

SERS SBS: Palla et al (2014) Royal Soc Chem Adv 4:49342.

Nanopore SBS: Kumar et al (2012) Science Reports 2:684; Fuller et al (2016) PNAS 113:5233; Stranges et al (2016) PNAS 113:E6749.

This work was supported by National Institutes of Health Grants P50 HG002806, R01 HG003582, and R21HG004404, by the National Science Foundation Grant NSF-CHE-11-11398 and by the Packard Fellowship for Science and Engineering.