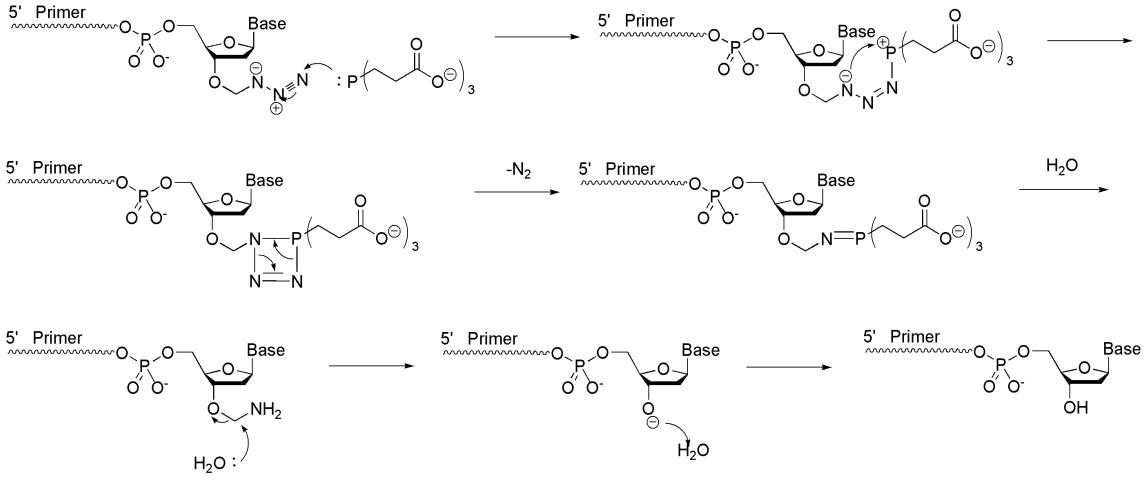
Electronic Supplementary Information

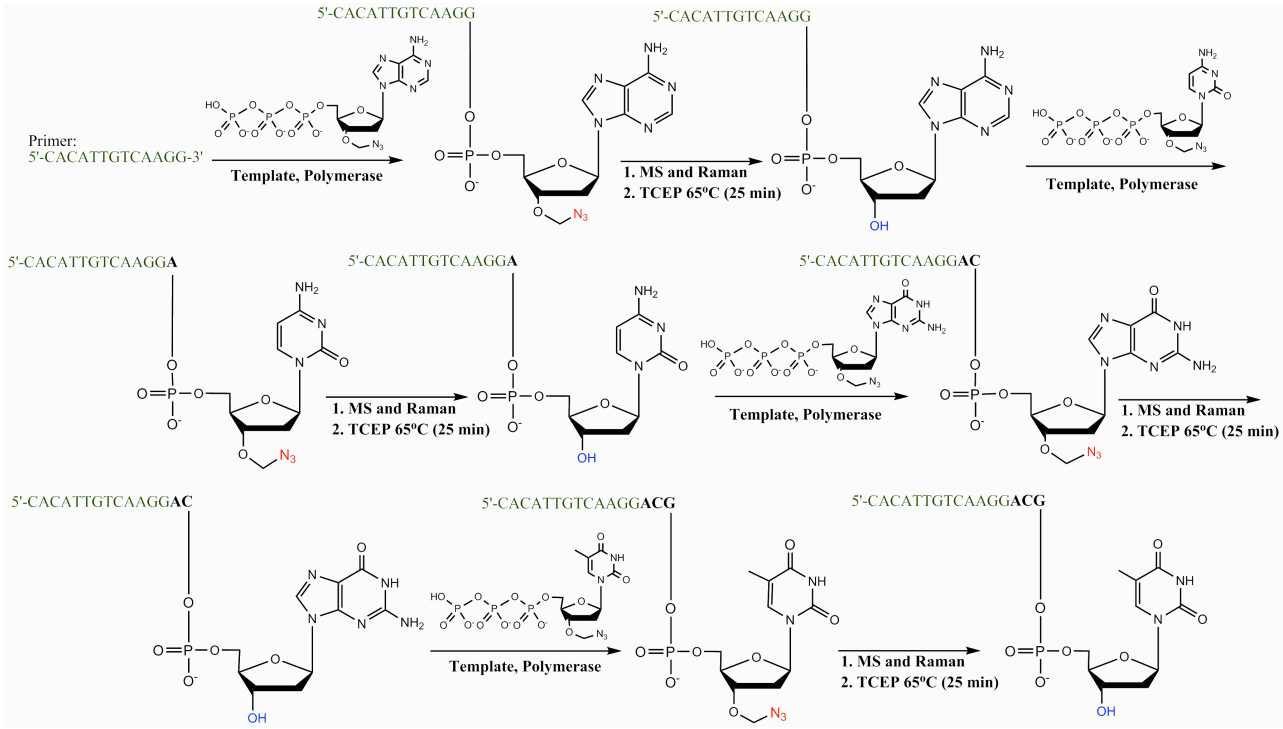
DNA sequencing by synthesis using 3′-*O*-azidomethyl nucleotide reversible terminators and surface-enhanced Raman spectroscopic detection

Mirkó Palla,[a,b] Wenjing Guo,[a,c] Shundi Shi,[a,c] Zengmin Li,[a,c] Jian Wu,[a] Steffen Jockusch,[e] Cheng Guo,[a,c] James J. Russo,[a,c] Nicholas J. Turro,[e] Jingyue Ju[a,c,d,*]

[a]Center for Genome Technology and Biomolecular Engineering, Departments of [b]Mechanical Engineering, [c]Chemical Engineering, [d]Pharmacology and [e]Chemistry, Columbia University, 3000 Broadway, New York, NY 10027, USA.



Scheme S1. Mechanisms to cleavethe 3'-*O*-azidomethyl group from the DNA extension products with TCEP to regenerate the 3'-OH group.



Scheme S2. Polymerase DNA extension reaction using 3'-*O*-N_3-dNTPs as reversible terminators.

Measurement and Calculation of the SERS Enhancement Factor

The test compound 2-azido-3-(benzyloxy)propanoic acid containing an azido (N_3) moiety was dissolved in methanol to form a 100 mM solution, which was then stepwise diluted to lower concentrations for SERS measurement. An aliquot (3 μL) was deposited onto SERS substrates (Reinshaw, UK) and then dried in ambient air to obtain uniform molecular deposition. To calculate the enhancement on our SERS samples, a piece of control substrate without Raman enhancement, which has 50 nm aluminum deposited on a flat silicon wafer by thermal evaporation at 0.1 nm/s, was also processed in the same way described above for the SERS samples as a reference. Once the suspension had completely dried, a set of SERS and reference measurements (typically N = 10) were collected. Fig. S1 shows the averaged and background removed Raman peaks for the azido model compound for both Klarite (red) and aluminum reference (blue) substrates, respectively.

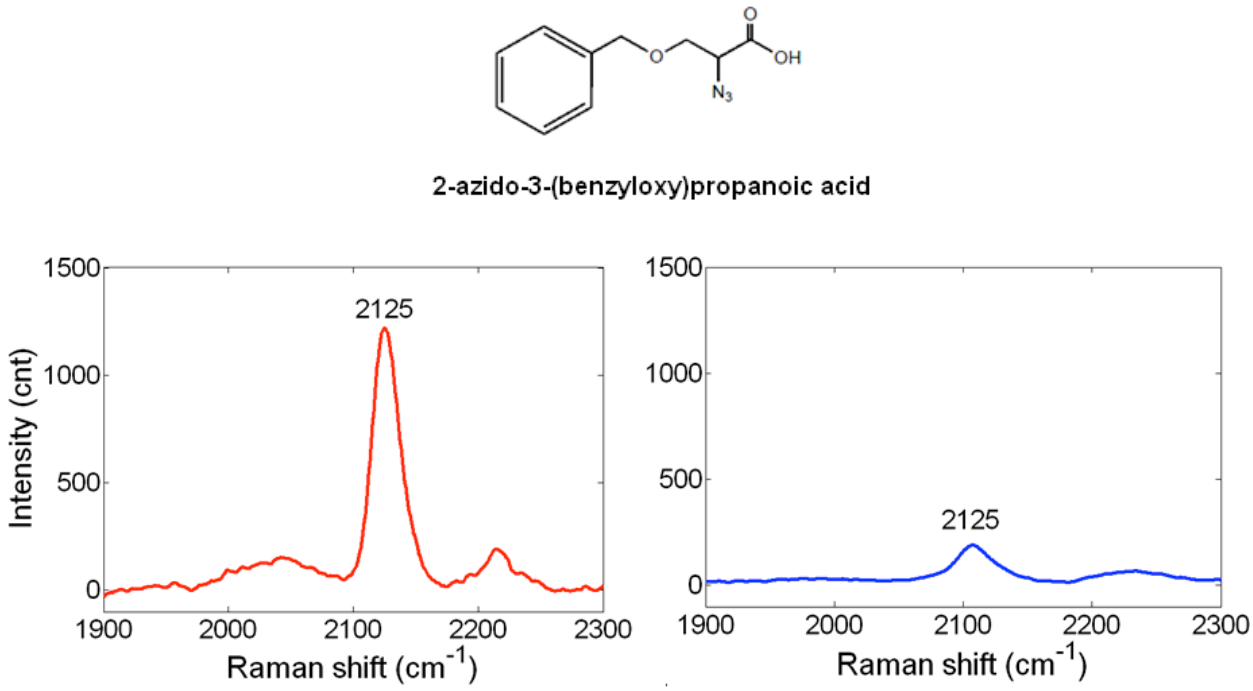


Figure S1: Raman scattering of the azido group from 2-azido-3-(benzyloxy)propanoic acid. (left) Typical experimental SERS of 1 μM (3 pmol) solution deposited on Klarite substrate and (right) the corresponding reference Raman spectrum of 100 mM (300 pmol) solution deposited on an aluminum coated slide.

The SERS enhancements of the adsorbed samples on Klarite substrate were compared with the aluminum reference substrate, which is assumed to have no enhancement to Raman signal. Here, the analytical enhancement factor is defined as:

$$\mathrm{AEF} = \frac{I_{\mathrm{SERS}}/c_{\mathrm{SERS}}}{I_{\mathrm{RS}}/c_{\mathrm{RS}}} \quad (1)$$

where I_{SERS} is the SERS signal intensity; I_{RS} is the Raman signal intensity of the aluminum reference; c_{SERS} is the analyte solution concentration on the SERS substrate and c_{RS} is the analyte concentration under non-SERS conditions. In the calculation of enhancement factors, I_{SERS} and I_{RS} are measured as the maximum value of the azido group's ~2125 cm^{-1} Raman peak fitted from the baseline-removed Raman spectra using a fourth degree polynomial fit. In order to achieve comparable Raman signal intensities on various samples, we used the same excitation powers and exposure times on different SERS substrates and the reference measurements. Moreover, analytes applied onto the SERS and the reference substrates had different concentrations, *i.e.*, 1 μM for all SERS substrates compared with 100 mM for reference substrates. Through a systematic experimental study using various analyte dilutions and parameter optimization of SERS measurements, we achieved an area-averaged SERS enhancement of ~8 $\times 10^5$ over the entire SERS-active area for the azido compound as defined by Equation (1). Our experimentally obtained enhancement factor is close to the manufacturer defined theoretical maximum of ~10^6 for surface enhancement on Klarite substrates.

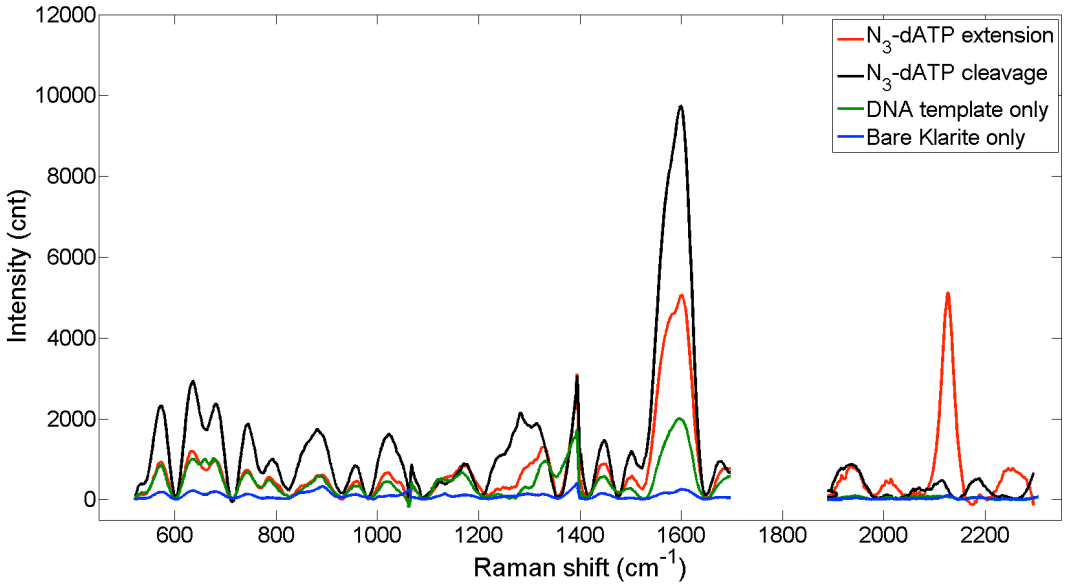


Fig. S2: Extended Raman spectra of a typical SBS cycle and the substrate background. In the spectral window from 2000-2300 cm^{-1} only the DNA product that has the azido group shows a strong Raman signal (red) while the others only show background signal. However, in the spectral window of 600-1700 cm^{-1} there is an overlay of many vibrational modes which can not be used to distinguish the DNA product containing the azido group from the one without. Note that each spectrum was generated from 4 separate acquisition windows and no data were collected in the range 1700-1900 cm^{-1}.