Electronic Supplementary Material (ESI) for Nanoscale.


# Electronic Supplementary Information

## Direct observation of breathing dynamics at the mismatch induced DNA bubble at nanometre accuracy: A smFRET study

Tapas Paul, Subhas C. Bera, Padmaja P. Mishra*

Chemical Sciences Division, Saha Institute of Nuclear Physics, 1/AF Bidhannagar, Kolkata 700064, India

* To whom correspondence should be addressed. Tel: +91-33-2337-5345-49; Fax: +91-33-2337-4637; Email: Padmaja.mishra@saha.ac.in

## MATERIALS AND METHODS

### Chemicals, DNA oligonucleotides

All chemicals were of analytical grade, purchased from Sigma-Aldrich and used without further purification. HPLC purified custom made labelled/unllabeled oligonucleotides (Table S1) were purchased from Integrated DNA Technologies, Inc, Coralville, USA. All oligos are designed in such a way that the dsDNA formed by annealing of B1 and B2 would result a perfectly intact oligo (BL0). However, the result of the combination of B1 with either of B3, B4 or B5 would form dsDNA with three (BL3), seven (BL7) or eleven (BL11) number of mismatched base pair (bp) at the centre, keeping both of the end portion intact. The short oligo BBIO, having biotin tagged at the 5' end is designed to help in surface immobilization of the above dsDNAs to carry out the smFRET experiments, when annealed at the 3' overhang of B1. Similarly, combination of BR1, BR2 and BBIO would results eleven mismatch bps at the centre and twelve nucleotides clamped at the two ends, named as LnBL11 The appropriate pairs of oligonucleotides were annealed by mixing equimolar amounts of ssDNA in a buffer containing 10 mM Tris-HCl, (pH-7.5) 100 mM NaCl and 10 mM $Mg^{2+}$ (buffer-A). The mixture of oligonucleotides were thermally heated at 92°C for 4 minutes, followed by slowly cooling to room temperature and finally stored at -$20^0$ C till to be used for experimental purpose. All experiments were carried out using the same composition of buffer-A.

**Table S1. The sequence of the oligo nucleotides with modification**

| Name of the oligos | Sequence and Modification (5' to 3') |
|---|---|
| B1 | CGGGGCAAGTAT(**Cy3**)AAGTAGGCGGGACACACACAC |
| B2 | CCCGCCTACTTA(**Cy5**)TACTTGCCCCG |
| B3 | CCCGCCTACTAT(**Cy5**)AACTTGCCCCG |
| B4 | CCCGCCTAGAAT(**Cy5**)ATGTTGCCCCG |
| B5 | CCCGCCATGAAT(**Cy5**)ATGAAGCCCCG |
| BBIO | **Biotin**-AAAAAGTGTGTGTGT |
| BR1 | CGAAGGCGGGGCAAGTAT(**Cy3**)AAGTAGGCGGGAAGAGGTACACACAC |
| BR2 | CCTCTTCCCGCCATGAAT(**Cy5**)ATGAAGCCCCGCCTTCG |

**Fluorescence measurement with temperature variation**

Steady state fluorescence measurements were carried out taking individual dsDNA (either of BL0, BL3, BL7 or BL11) of 50 nM in buffer-A by a Varian Cary Eclipse fluorescence spectrophotometer equipped with a circulating water bath to control the experimental temperature. Fluorescence was recorded simultaneously by individual exciting Cy3 (donor) & Cy5 (acceptor) at 530 nm and 635 nm respectively keeping the slit width of 5 nm.

**Circular dichroism (CD) measurement**

Electronic circular dichroism (CD) spectra of individual sets of oligos (15 μM) with bubbles were recorded in a Chirascan spectropolarimeter (Applied Photophysics, U.K.) using standard quartz cell of 1 mm path length. Scans were taken from 220 to 320 nm for the intrinsic region and spectra were expressed in terms of molar ellipticity. For each spectrum, three consecutive readings were averaged at a constant temperature of 298 K with a 5 minutes equilibration before each sample.

**Single-molecule FRET measurements and data analysis**

Single-molecule FRET experiments were performed using plasma cleaned slide and cover slip based micro-fluidic sample chamber coated with PEG/biotin-PEG (100:1) (MW 5,000, Laysan Bio, USA). Sample chambers were treated with 0.2 mg/ml of streptavidin (Invitrogen) and waited for 6-7 minutes, then washed with buffer-A. The annealed dsDNA labeled with biotin, Cy3 and Cy5 were diluted to ~ 50 pM and then ~ 50 μl was injected into the sample chamber for immobilization. After 5-7 minutes incubation, buffer-A was flushed into the sample chamber to remove unbound DNAs. 1 μL of oxygen scavenging system composed of 2 mM trolox with protocatechuic acid/protocatechuate-3,4-dioxygenase in buffer-A has been used to avoid blinking and improved dyes stability. Detailed standard procedure of these experimental precautions was published elsewhere.[1, 2]

We have used a home built prism-type total internal reflection (PTIR) inverted fluorescence microscope (Olympus IX 71) as described earlier.[3] A solid state 532 nm diode laser (Laser Quantum, U.K.) was used to excite Cy3 and the fluorescence from Cy3 and Cy5 were simultaneously collected using water immersion objective (60X appo, 1.2 NA, Olympus). The donor and acceptor signals were passes through a long-pass filter (550 nm, Chroma), then it is divided using a 550−630 nm dichroic mirror (640 DCXR, Chroma) and finally

projected onto the electron-multiplied charge-coupled-device camera (EMCCD, Ixon3+ 897, Andor Technologies, South Windsor, CT) with 30 ms frame integration time. Visual C++ based data acquisition software was used to record the raw movie file, a generous gift from either Tae-Hee Lee (Pensylvania State University, USA) or T. J. Ha (University of Illinois, USA). Initially, each image frame contained approximately 300 Cy3/Cy5 single-molecule features. During data analysis, the bleeds through of Cy3 emission to the Cy5 channel (typically <5%) is rectified. Direct excitation of Cy5 by using 532 nm green laser almost negligible. All experimental data were recorded at room temperature (~ 25 $^0$C).

Raw movie files were analysed using codes written in the Interactive Data Language (IDL) based programming language. The FRET efficiency ($E_{\text{fret}}$) was calculated after correcting the cross talk between donor and acceptor channels using the relationship:

$$E_{\text{fret}} = \frac{I_A}{I_D + \gamma I_A} \dots\dots\dots\dots\dots 1$$

Where $E_{\text{fret}}$ is the FRET efficiency and $I_D$ and $I_A$ is intensity of Cy3 and Cy5 respectively and $\gamma$ represents the difference in quantum yield and instrumental detection efficiency of donor and acceptor. $\gamma$ is calculated by the ratio of change in the acceptor intensity ($\Delta I_A$) to change in the donor intensity ($\Delta I_D$) upon acceptor photobleaching and found to be 0.95 ± 0.12.[4] Five adjacent data points in single-molecule time traces were averaged out to suppress the noise. Individual trajectories were chosen with clear anti-correlated donor–acceptor dynamics and single-step donor and/or acceptor photobleaching, which indicates one dye per DNA and stored for further analysis. Approximately 100 single-molecule time traces containing both active donor and acceptor fluorophores were compiled to construct the FRET histogram. FRET trajectories showed conformational dynamics were analysed by using hidden Markov model (HMM) algorithm and extract all breathing kinetics information of the individual states. The details of HMM can be found elsewhere.[5] The probabilities of transition between different states are reported. HaMMy software was used to calculate the average lifetime of

each breathing states. The distance between the two breathing strand was calculated using the following equation

$$E_{\text{fret}} = \frac{R_0^6}{R^6 + R_0^6} \ldots\ldots\ldots\ldots\ldots 2$$

Where, $R_0$ and R were the Förster distance and distance between the donor-acceptor respectively.

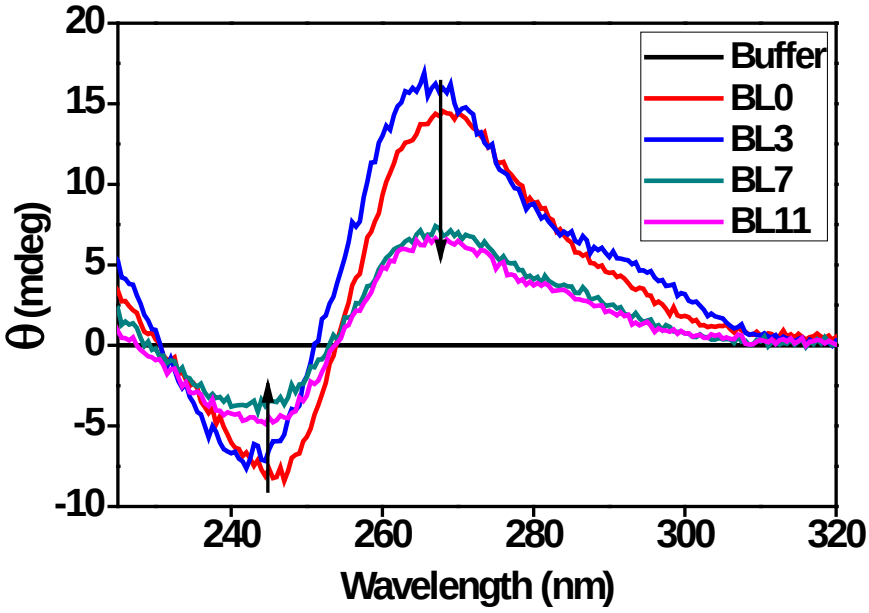
The cross-correlation (CC) and autocorrelation (AC) analysis of donor and acceptor intensity time traces was performed as previously described.[6] The cross-correlation and autocorrelation is defined by the following equation

$$CC(\tau) = \frac{\langle \delta I_d(t) \delta I_a(t+\tau) \rangle}{\langle I_d(t) \rangle \langle I_a(t) \rangle} = \frac{\langle I_d(t) I_a(t+\tau) \rangle}{\langle I_d(t) \rangle \langle I_a(t) \rangle} - 1 \ldots\ldots\ldots\ldots\ldots 3$$

$$AC(\tau) = \frac{\langle \delta I(t) \delta I(t+\tau) \rangle}{\langle I(t) \rangle \langle I(t) \rangle} = \frac{\langle I(t) I(t+\tau) \rangle}{\langle I(t) \rangle \langle I(t) \rangle} - 1 \ldots\ldots\ldots\ldots\ldots 4$$

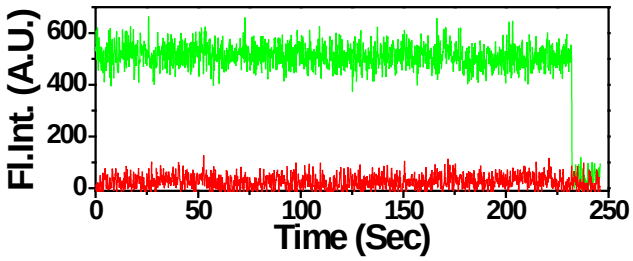
Where $\langle I(t) \rangle$ is the time average of $I$ (t), $\delta I$ (t) is the difference of $I$ (t) from $\langle I(t) \rangle$ and $\tau$ is the lag time. The cross-correlation analysis was performed using the same traces that were used for the histogram and the data was fitted in single exponential function. From the fitted curve determined the characteristic time of the exponential (t) and the amplitude.

**Fig. S1**



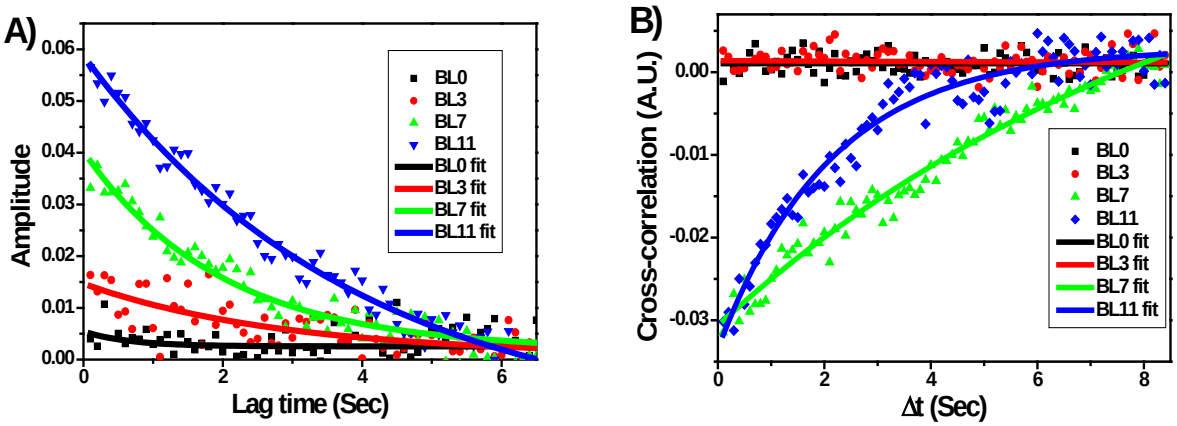

The CD spectrum of all the four sets of DNA with different bubble size show positive peaks at approximately 278 nm and negative peaks at 247 nm, as expected for a B-DNA. However, the CD spectra show that the ellipticity (θ) of positive peak decreases (right arrow) and negative peak increases (left arrow) with increasing bubble size. This indicates that the hydrogen bonding and helicity would decreases and it clearly identifies the different size of bubble present.

**Fig. S2**



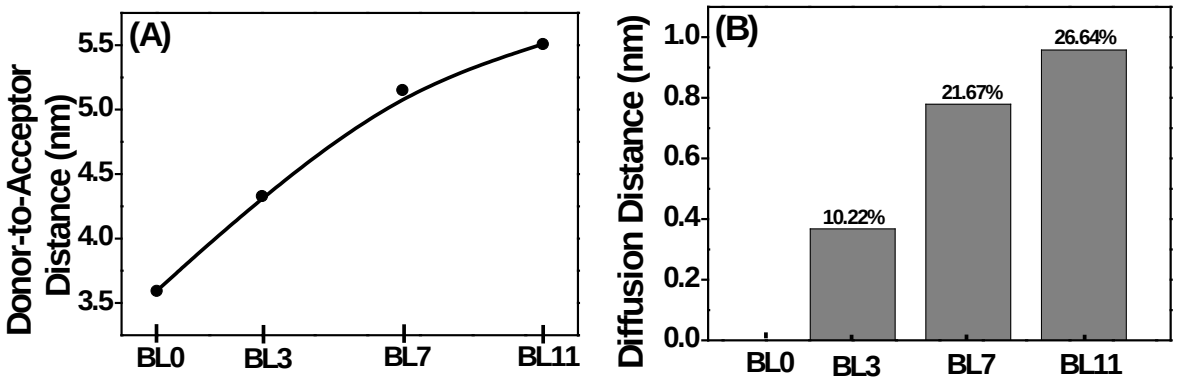

The single molecule time traces of the same construct of BL0 DNA without Cy5 dye. The green line represents the fluorescence intensity of Cy3. For a certain laser power, the fluctuation emission of Cy3 remains constant at a fixed intensity until photobleach.

**Fig. S3**




A) The smFRET autocorrelation decay with increasing amplitude indicating a lowering rigidity or higher flexibility in the system that provides clear evidence of the dynamic nature of bubble and it increases with increasing the bubble size. B) Cross-correlation curve of four different DNA bubble systems indicates that anti-correlated fluorescence intensity fluctuation between donor and acceptor was observed only for BL7 and BL11. Solid line is the fitted single exponential function.

**Fig. S4**




(A) The variation of the distance between the two strands of the bubble region with increase in the bubble size, calculated from the FRET values. The bubble size has been estimated from the corresponding dye-to-dye distance. The strands separation distance increase linearly with the bubble size (where BL0, BL3, BL7 and BL11 signify that the 0, 3, 7 and 11 mismatches respectively).

(B) Change in distance between the two strands of a bubble and % of diffusion from the intact DNA (BL0) for all the DNAs with different size of bubble. The % of diffusion calculated as follows:

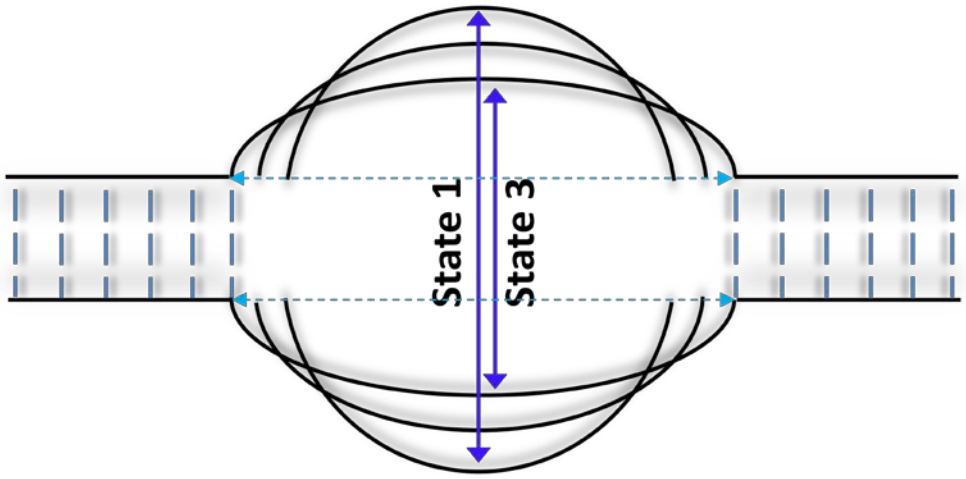
Diffused distance generated due to bubble formation ($D_i$) = Distance between the two strands with Bubble (D) = Distance between the two strands of the Intact DNA ($D_0$)

$$\% \text{ of diffusion} = \frac{Extra\ Distance\ travelled\ by\ a\ strand\ due\ to\ diffusion}{Distance\ between\ the\ strands\ with\ out\ dissusion} \times 100$$

$$= D\frac{D_1/_2}{D} \times 100$$

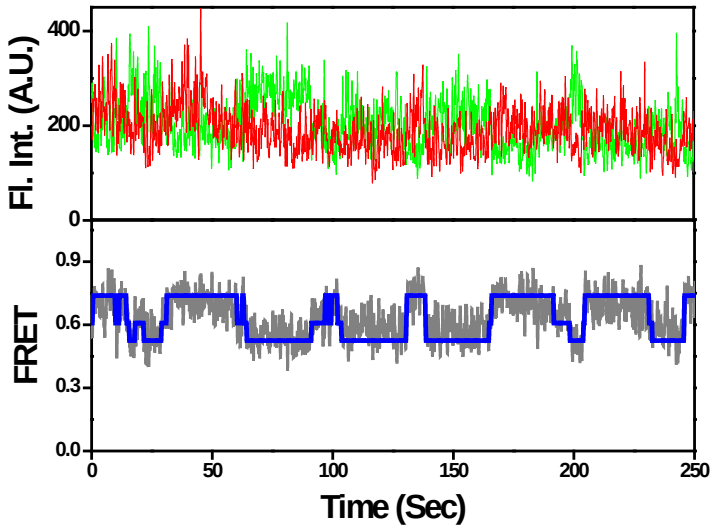
Calculating in this way, indicate 10%, 22% and 27% diffusion for BL3, BL7 and BL11 system respectively with respect to control.

**Fig. S5.**

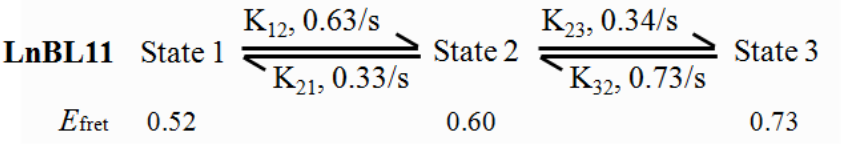



Different states present during the fluctuation of a particular DNA, considering a synchronised fluctuation of the strands two bubble strands were continuously fluctuated having different state of a particular system. We have assigned state 1 and state 3 as the farthest and the closest possible approach between the two bubble strands. State 2 appears as an intermediate for bubbles with 7 &11 mismatches. As the length of the bubble strand is fixed, to attain these states, the start-to-end distance would experience an increase or decrease accordingly.

**Fig. S6**

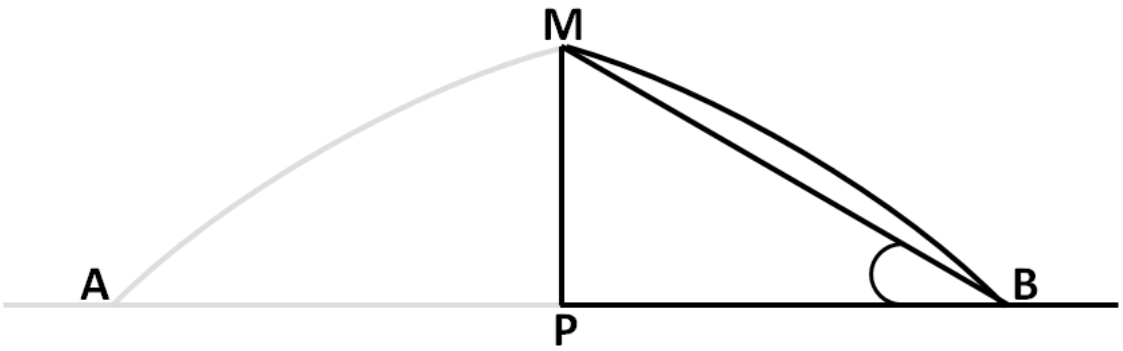



The single-molecule time traces trajectories of LnBL11, red and green represent the fluorophore intensity of acceptor and donor respectively. Below, the corresponding FRET efficiency is shown in grey and the blue lines represent the idealized FRET having $E_{fret}$ 0.73, 0.60, and 0.52 respectively.




Three distinct FRET state fluctuating of LnBL11 system with kinetic rates obtained from HMM analysis.

**Fig. S7**




Representative diagram of one of the bubble strands. AMB is the one strand of a diffused bubble and AB is the start-to-end diffused bubble distance. Considering M and P as the midpoint of arc AMB formed due to diffusion and straight AB is the corresponding distance without diffusion, and assuming BM as a straight line, then PBM forms a right angle triangle. From that triangle PBM, the bubble angle $\angle PBM$ and the straight distance AB between the zipper forks calculated which provide the overall bubble shape determination.

**Speculation of Bubble shape:**

Our DNA construct containing bubble due to presence of Non Watson-Crick base pairs at the middle are clamped with Watson-Crick base pairs towards both the ends. The experimental results support that the separation of two bubble strand increases with increase in bubble size, inducing a particular conformational shape. That could be assigned as either a 'ellipse' or a 'circle'. Comparing the calculated parameters to that of ellipse (Table S2) and circle (Table S3) indicates the shape to be more similar to an ellipse. However, bubble of smaller size resembles equally to a circle shape as well.

**Table S2.** Considering an "ellipse shape"

| **Taking shape as "Ellipse"** | | | | |
|---|---|---|---|---|
| Different systems (distance) | Distance (MN) (nm) | Half distance increased (PM) (nm) (minor axes) | Major axes (AB) (nm) | % of bending (end to end distance) |
| BL0 | 3.59 | | | |
| BL3 (3x0.7=2.1nm) | i) 4.09 | 0.25 | 1.82 | 13.1% |
| | ii) 4.33 | 0.37 | 1.74 | 17.0% |
| BL7 (7x0.7=4.9 nm) | i) 4.37 | 0.39 | 4.34 | 11.3% |
| | ii) 4.57 | 0.49 | 4.30 | 12.2% |
| | iii) 5.15 | 0.78 | 4.13 | 15.7% |
| BL11 (11x0.7=7.7 nm) | i) 4.68 | 0.55 | 6.85 | 11.0% |
| | ii) 5.01 | 0.79 | 6.75 | 12.3% |
| | iii) 5.50 | 0.96 | 6.67 | 13.4% |

**Table S3.** Considering a "circle shape"

| Taking shape as "Circle" | | | | |
|---|---|---|---|---|
| Different systems (distance) | Distance (MN) (nm) | Radius (r) (MN/2) (nm) | Chord distance (AB) (nm) AB=2 x $\sqrt{(r^2-t^2)}$ | % of bending (end to end distance) |
| BL0 | 3.59 | | | |
| BL3 (3x0.7=2.1nm) | i) 4.09 | 2.05 | 1.96 | 6.4% |
| | ii) 4.33 | 2.16 | 2.41 | 14.9% |
| BL7 (7x0.7=4.9 nm) | i) 4.37 | 2.18 | 2.49 | 49.1% |
| | ii) 4.57 | 2.29 | 2.83 | 42.2% |
| | iii) 5.15 | 2.57 | 3.69 | 24.6% |
| BL11 (11x0.7=7.7 nm) | i) 4.68 | 2.34 | 3.01 | 60. 9% |
| | ii) 5.01 | 2.59 | 3.74 | 51.4% |
| | iii) 5.50 | 2.75 | 4.17 | 45. 8% |

## REFERENCES


(1) Joo, C.; Ha, T. Preparing Sample Chambers for Single-Molecule FRET. *Cold Spring Harbor Protocols* **2012**, *2012*, pdb. prot071530.
(2) Aitken, C. E.; Marshall, R. A.; Puglisi, J. D. An Oxygen Scavenging System for Improvement of Dye Stability in Single-Molecule Fluorescence Experiments. *Biophys. J.* **2008**, *94*, 1826-1835.
(3) Bera, S. C.; Sanyal, K.; Senapati, D.; Mishra, P. P. Conformational Changes Followed by Complete Unzipping of DNA Double Helix by Charge-Tuned Gold Nanoparticles. *J. Phys. Chem. B* **2016**, *120*, 4213-4220.
(4) Roy, R.; Hohng, S.; Ha, T. A Practical Guide to Single-Molecule FRET. *Nat. Methods* **2008**, *5*, 507-516.
(5) Lee, T.-H. Extracting Kinetics Information from Single-Molecule Fluorescence Resonance Energy Transfer Data Using Hidden Markov Models. *J. Phys. Chem. B* **2009**, *113*, 11535-11542.
(6) Kim, H. D.; Nienhaus, G. U.; Ha, T.; Orr, J. W.; Williamson, J. R.; Chu, S. Mg2+-Dependent Conformational Change of RNA Studied by Fluorescence Correlation and FRET on Immobilized Single Molecules. *Proc. Natl Acad. Sci. USA* **2002**, *99*, 4284-4289.