Chapter 6

Probing Single-Molecule Intermittent Coherence Conformation Dynamics and Multistate Transition Kinetic of DNA Holliday Junction by smFRET†

†This work is currently under review. "Probing Single-Molecule Intermittent Coherence Conformation Dynamics and Multistate Transition Kinetic of DNA Holliday Junction by smFRET. Pratibha Agarwala, Saikat Sadhukhan, Soumen Mondal, and Dibyendu K. Sasmal*

In this chapter, we have discussed our work on probing single-molecule intermittent coherence conformation dynamics and multistate transition kinetic of DNA HJ by smFRET experiment using pTIRF setup as discussed in Chapter 3.

6.1 Introduction

Conformational dynamics of HJ is involved with complex protein-DNA binding, branch migration, rotation of branch, and re-joining; typically, this structural dynamics has a high spatial and temporal resolution that is difficult to capture by steady-state fluorescence measurements.[94, 97, 104, 247-249] However smFRET is a powerful and state-of-the-art optical technology that can characterize such complex dynamics patterns one molecule at a time in real-time with a high spatial-temporal resolution of 1 nm and 1 ms.[4, 250] The conformation dynamics of HJ with connection to molecular reaction and pattern are a fundamental focus of investigation because of its potential to understand the key molecular mechanism of genetic diversity, and the topic is under intense investigation in the research area of single-molecule biophysics.[103, 111, 118, 123, 251] Conformational change of a biological system is often a critical question for understanding function, selectivity, activity, and many other biomolecular reactions because they can play different roles in changing the inherent conformation.[252-256] The answer to this question needs a critical understanding of the interaction pattern and an idea about the transition or intermediate state formation. Although crystal structure and Markovian dynamics of HJ have been reported previously, the nature of dynamics related to transition frequency to thermodynamically stable multistate and intermediate states remain unexplored. Here, we report an observation of intermittently coherent 3-state conformational dynamics of DNA HJ in the presence of Mg^{2+} ion with distinctly different conformation patterns: resonance, coherence, and decoherence. This observation shows that biological conformational dynamics has a dynamics pattern similar to nature with time-sequence when (1) amplitude increases first (resonance), (2) the amplitude remains constant for some time (coherence), and (3) then the amplitude fades out (decoherence).

Depending on the ion concentration and charge density, HJ may stay in the dynamic thermodynamic equilibrium position between open and closed state conformers through the electrostatic screening and the repulsion of the negatively charged backbone.[113, 123] Hence, conformational dynamics of HJ serves a critical role in homologous genetic recombination, as this structure can influence the binding orientation by determining the cleaved position of the junction-resolving enzyme.[257-260]

Recently, Lipfert and his group[261] reported the faster-unresolved dynamics of HJ under low or intermediate salt concentration by the X-ray scattering technique, which proves that in the low salt concentration, the HJ adopt open state without-the-plane geometry, and at intermediate ionic strength, it undergoes conformation changes by increasing the interduplex angle. However, the pattern of the dynamics of single-molecule HJ molecule is not reported yet.

In this chapter, we report 3-state intermittently coherent conformational fluctuation in addition to other oscillation pattern of the individual DNA molecule, which govern periodic

changes in the biological system. Such oscillation of dynamics is previously reported in enzymatic reactions.[262-269] We also used TCF, which provides more detailed information on kinetic network than the conventional Hidden Markov Model (HMM) analysis. From this analysis, we obtain information about the conformational transition pathways connecting the states, kinetic data for the transition and free energy minima of each state from the Arrhenius equation.

6.2 Methods

6.2.1 Auto-Correlation and Cross-Correlations Calculation

To calculate the cross-correlation of single molecule trajectory we used the following equation[270]

$$C_{Cross}(\tau) = \frac{\langle \delta I_D(t) \delta I_A(t+\tau) \rangle}{\langle I_A(t) \rangle \langle I_D(t) \rangle} = \frac{\langle I_D(t) I_A(t+\tau) \rangle}{\langle I_D(t) \rangle \langle I_A(t) \rangle} - 1 \quad (6.1)$$

In the equation the intensity of donor $I_D(t)$and acceptor I_A (t), and $\langle I(t) \rangle$ is the time average of I(t) and $\delta I(t)$ is the difference between, I(t) from $\langle I(t) \rangle$.
To calculate autocorrelation function, when I_A (t) = I_D (t)

$$C_{auto}(t) \text{of Acceptor} = \frac{\langle \delta I_A(t) \delta I_A(t+\tau) \rangle}{\langle I_A(t) \rangle \langle I_A(t) \rangle} = \frac{\langle I_A(t) I_A(t+\tau) \rangle}{\langle I_A(t) \rangle \langle I_A(t) \rangle} - 1 \quad (6.2)$$

For this calculation of two dimensional cross corelation function, we have chosen a start time (t_{start}) and stop time (t_{stop}) to calculate the cross-correlation function from the time trajectory of $I_D(t)$ donor and $I_A(t)$ acceptor.[268]

$$C_{Cross}(\tau, t_{start} : t_{stop}) = \int_{t_{start}}^{t_{stop}} I_A(t) I_D(t-\tau)\, dt = \sum_{t_{start}}^{t_{stop}} I_A(t) I_D(t-\tau) \quad (6.3)$$

6.2.2 Calculation of 2nd Order TCF

The 2nd order TCF can be fitted in the form

$$C^2(\tau) = A_2 e^{-\tau/\tau_{fast}} + A_3 e^{-\tau/\tau_{slow}} \quad (6.4)$$

In equation 6.4 τ_{fast} and τ_{slow} are the characteristic decay constant of the amplitudes A_2 and A_3 respectively.

6.2.3 Macrostate and Free Energy Minima Calculation

The DNA HJ can adopt different conformation states; if we assume there are N number microstates, then according to the Boltzmann distribution function, for an i-th microstate, the probability is given by[271]

$$P_i = \frac{1}{Z} e^{-G_i/k_B T} \quad (6.5)$$

The Z term is the partition function which is given by $Z = \sum_{i=1}^{N} e^{-G_i/k_B T}$ Gibbs free energy of i-th Macrostate and is represented by G_i, k_B is the Boltzmann constant at a given temperature T. In this case, it is expected that the normalized probability is equal to unity.

$$\sum_{i=1}^{N} P_i = 1 \quad (6.6)$$

Here the PDF is a bi-gaussian model by considering that the E_{FRET} value as a sum of N macrostates as $P(E_{FRET}) = \sum_{j=1}^{N} P_j(E_{FRET})$,

$$P_j(E_{FRET}) = A_j \, e^{[(-(E_{FRET}-\langle E_{FRET}\rangle_j)^2/2\sigma_j^2]} \quad (6.7)$$

Where $\langle E_{FRET}\rangle_j$ is the mean FRET value, σ_j is the standard deviation with an amplitude A_j. Here the relative probability of the j-th state at equilibrium is given by the area under the Gaussian component, which is represented by,

$$P_j^{equ} = A_j \times \sigma_j \times \sqrt{2\pi} \quad (6.8)$$

the sum of the normalized probability of equilibrium macrostates is given by unity.
The Gibbs free energy of the j-th macrostate in terms of FRET efficiency is given by (in the unit $k_B T$),

$$\frac{G_j(E_{FRET})}{k_B T} = -\ln\left[P_j(E_{FRET})\right] = -\ln\left(\frac{P_j^{equ}}{\sigma_j\sqrt{2\pi}}\right) + \frac{1}{2\sigma_j^2}\left[E_{FRET} - \langle E_{FRET}\rangle_j\right]^2 \quad (6.9)$$

Here the total Gibbs free energy of all possible macrostates is given from the individual contribution of each macrostate,[272]

$$G(E_{FRET}) = -k_B T \ln\left[\sum_{j=1}^{N} P_j\,(E_{FRET})\right] \quad (6.10)$$

We calculated the relative PDF of Energy efficiency (E_{FRET}) by fitting this in the bi-gaussian model as per equation 6.7. We have found three Gaussian peaks representing each macrostate: state 1 (S1), state 2 (S2) and state 3 (S3), and the normalized area under each Gaussian component is equal to the equilibrium probability P_j^{equ}, where the three P^{equ} where the sum of three P_j^{equ} is unity as per Table 6.1.

6.2.4 The Analytical Expression of The Kinetic Rate Calculation for N= 3 System

The Master equation for the N=3 system was

$$\dot{p}_t = Kp(t) = \begin{bmatrix}\dot{p_1}\\ \dot{p_2}\\ \dot{p_3}\end{bmatrix} \equiv \begin{bmatrix}-k_{12}-k_{13} & k_{21} & k_{31}\\ k_{12} & -k_{21}-k_{23} & k_{32}\\ k_{13} & k_{23} & -k_{31}-k_{32}\end{bmatrix}\begin{bmatrix}p_1\\ p_2\\ p_3\end{bmatrix} \quad (6.11)$$

where $p(t)$ was the time-dependent macrostate population vector composed of $[p_1, p_2, p_3]$ and the time derivative was represented as $\dot{p}_t$. The 3×3 matrix represented by K, which contains the rate constant k_{ij}.
Then, the general solution of the Equation was[273]

$$p(t) = p^{eq} + c_2 v_2 e^{\lambda_2 t} + c_2 v_3 e^{\lambda_3 t} \quad (6.12)$$

Where the eigenvectors were v_2 and v_3, and eigenvalues λ_2 and λ_3 were given by $\lambda_2 = a + b$, $\lambda_3 = a - b$. Definitions of a and b were:

$$a = \frac{1}{2}(k_{12} + k_{21} + k_{23} + k_{32} + k_{13} + k_{31}) \quad (6.13)$$

$$b = \frac{1}{2}[(k_{12}-k_{23})^2 + (k_{13}-k_{32})^2 + (k_{21}-k_{31})^2 + 2k_{12}(k_{13}+k_{21}-k_{31}-k_{32}) + 2k_{13}(k_{31}-k_{21}-k_{23}) + 2k_{23}(k_{21}-k_{31}+k_{32}) - 2k_{21}k_{32} + -2k_{31}k_{32}]^{1/2} \quad (6.14)$$

To satisfy the detailed, balanced condition, one rate constant must depend on others such that

$k_{31} = k_{13}k_{32}k_{21}/k_{23}k_{12}$

The equilibrium populations $p^{eq} = [p_1^{eq}, p_2^{eq}, p_3^{eq}]$ were obtained by solving Equation 6.11 with the boundary condition $\dot{p}_t = 0$. This gives the results

$$p_1^{eq} = \{1 + [(k_{12} + k_{13})/k_{21}] + [(1 - k_{31})/k_{21}].(k_{21} + k_{23})(k_{12} + k_{13}) - k_{12}.k_{21}/[(k_{21} + k_{23})k_{31} + k_{32}k_{21}]\}^{-1} \quad (6.15)$$

$$p_2^{eq} = k_{21}^{-1}.[p_1^{eq}(k_{12} + k_{13}) - p_3^{eq}k_{31}] \quad (6.16)$$

$$p_3^{eq} = p_1^{eq}.[(k_{21} + k_{23})(k_{12} + k_{13}) - k_{12}k_{21}].[(k_{21} + k_{23})k_{31} + k_{32}k_{21}]^{-1} \quad (6.17)$$

We have solved the equation 6.13-6.17 to get the rate constant values tabulated in Table 6.2. From the rate constants, we have also calculated the free energy of activation using the Arrhenius equation, [271]From the rate constants, we have also calculated the free energy of activation using the Arrhenius equation, [271]

$$G_{ij}^{0\#} = -k_B T \ln(\frac{k_{ij}}{k_{max}}) \quad (6.18)$$

According to the Arrhenius equation $k_{ij} = k_{max}e^{-G_i/k_BT}$. Here, the pre-exponential factor in the equation is considered to k_{max}.[271] The fastest rate conversion in the HJ conformational transition between macrostates is corresponds to $k_{13} = 0.66\ ms^{-1}$ s per Table 6.2. For the purine-pyrimidine (G-C) nucleotides the activation barrier is ~2 kcal/mol which is equivalent to 3.4 k_BT.[274] So, the pre-exponential factor becomes $k_{max} = k_{13} \times e^{3.4} = 0.66 \times 29.9 = 19.77\ ms^{-1}$.
We have considered the fastest rate constant as a pre-exponential factor so that the remaining activation barriers are calculated relative to this value to make a comparison in the three-state kinetic model.
The activation barrier G_i, here considered as ~2 kcal/mol correspond activation barrier of purine-pyrimidine base pair because in our sequence of HJ the junction points consist of G-C nucleotide sequence.

6.3 Results and Discussion

6.3.1 . smFRET to Study the Conformation Heterogeneity

To explore the conformation heterogeneity, kinetics of the HJ and how the kinetics of dynamics plays a pivotal role in the dynamic changes, we have performed the smFRET experiment,[124] with HJ labelled with Cy3 and Cy5 (Figure 6.1.A). The biotin was incorporated into another arm of HJ for immobilization using the biotin-streptavidin interaction described in Chapter 3 section 3.4.11 (Figure 6.1.A). Figure 6.1.B represents fluorescence images of donor channel and acceptor channel. The individual peak represents the single molecule in both donor and acceptor channels. Here, donor and acceptor channels were colocalized with each other, which reflects that the HJ was completely immobilized on the glass surface with the biotin-streptavidin interaction. Figure 6.1.C representative fluorescence intensity time trajectories of single donor (green) and acceptor (red) molecule in the presence of 50 mM Mg^{2+}, showing high dynamic fluctuation of DNA HJ. When the donor and acceptor come closer to each other, the donor can transfer more energy to the acceptor, and the acceptor signal goes higher decreasing donor signals at the same time. The anticorrelated pattern in the donor and the acceptor signal was a clear signature of smFRET.[146] Figure 6.1.D shows the FRET efficiency time trajectory (black line) calculated from equation 3.8 from the corresponding donor and acceptor signal shown in Figure 6.1.C.

There was a fluctuation in FRET efficiency from an average 0.2 FRET efficiency to an average 0.8 FRET efficiency. The higher FRET efficiency reflects that two arms of DNA HJ tagged with the donor and acceptor were close to each other as the FRET efficiency was inversely proportional to the 6th power of the distance between donor and acceptor.[146]The low FRET reflects the two arms of the HJ tagged with donor and acceptor were far away from each other.

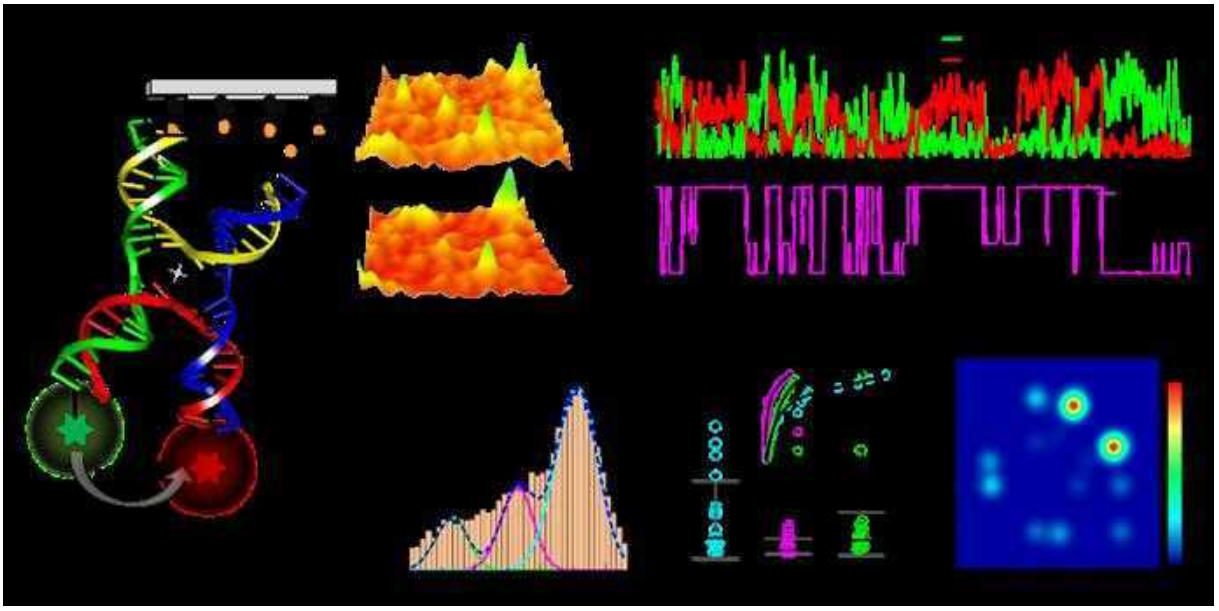


Figure 6.1 (A) Biotin-labelled BSA is attached to the quartz surface, which is labeled with Streptavidin. Then, biotinylated DNA HJ is coupled to on Streptavidin for smFRET experiments. The sequence of four DNA oligos is given in chapter 3 section 3.4.8. Two arms of DNA HJ are labeled with FRET pair dyes - Cy3 (green) and Cy5 (red). (B) Single-molecule fluorescence image (13 μm × 13 μm) of DNA HJ labeled with Cy3 (top) and Cy5 (bottom) obtained in the presence of [Mg^{2+}] = 50 mM. Co-localization of the donor and acceptor was evident in both images. (C) Single-molecule fluorescence intensity time-trajectories of Cy3 (Donor, Green) and Cy5 (Acceptor, Red) attached to DNA HJ. The anticorrelated pattern of the two signals was a clear signature of smFRET. (D) FRET efficiency fluctuation (black line) was calculated from the data shown in (C) with the fitting line (pink line) by the HMM. (E) The Histogram of FRET efficiency shows three prominent states corresponding to, E_{FRET} = 0.2 (S3), 0.56 (S2) and 0.77 (S1). (F) Median dwell-time ± standard deviation (SD) of three states corresponding to data in (D) calculated from $\mathbf{E_{FRET}}$ in (D). Inset shows the cumulative distribution function (CDF) of dwell times of three conformational states of HJ. (G) Transition density plot representing the dynamic changes between three states, majority of transition had taken place from EFRET = 0.3 to 0.5, and 0.5 to 0.8.

Not only the high FRET and low FRET state there was a possibility of some intermediate conformation as the FRET efficiency shows intermediate FRET at 0.4-0.5 range.[104, 113, 275] In Figure 6.1.D, we have seen that between 5-15 s, the molecule shows high dynamic fluctuation, which was going from a low FRET state to a high FRET state via some possible intermediate states; in the 15-25 s, a slower dynamic behavior may correspond to resting state, and after 25 s it shows that the same molecule acceptor undergoes single step photobleaching, which was also another clear signature of smFRET. After that, we used the ebFRET interface to idealize the raw FRET efficiency vs time data by the Viterbi Algorithm of the HMM.[276-277] After the idealization of the FRET trajectory, information regarding the variable state is generated, which was further used to calculate a particular state's dwell time, FRET efficiency histogram and transition density plot.

The HMM fit model clearly shows that there is a possibility of three state conformation changes, which was further explored by the TCF analysis.

The distribution of FRET efficiency was represented in Figure 6.1.E, where we have considered that the conformational dynamics can happen with 3 possible states i) S1 state, which shows high FRET efficiency, ii) State S3, which shows low FRET efficiency, and iii) State S3 which shows intermediate FRET efficiency. The three-state FRET histogram reflects that there was a possibility of changing the conformation from low FRET state (S3) to High FRET state (S1) via an open or intermediates state (S2) or vice versa, which we further explored by the statistical analysis in the following section.

The average dwell time of the same trajectory (Figure 6.1.F) in state S1 was 0.97 ± 0.06 s, whereas for state S2, it was 0.20 ± 0.10 s, and for state S3, the average waiting time was 0.56 ± 0.11 s, which shows the dwell time of High FRET state was much higher that the intermediate and low FRET state. From the cumulative density function also, it was observed that CDF shifted towards a longer time scale for the S1 state, which signified the increasing dwell time in the high FRET state.

We have also calculated the transition density plot (TDP) (Figure 6.1.G) to characterize the switching between states quantitatively. The transition density plot shows the distribution of FRET transition between all observed states or the most likely state probabilities. From the TPD plot (for all trajectories), it is indicated that most of the transition had taken place from FRET = 0.3 to 0.5, and 0.5 to 0.8. The symmetry in the transition density plot indicates that the transition from the observed state and vice-versa is possible. The cumulative density plot of the dwell time is also plotted in the inset of Figure 6.1.F which reflects that the dwell time shifted towards a longer time scale for the S1 state.

6.3.2 Intermittently Coherent Conformational Fluctuation in DNA

The dynamics of DNA HJ depends heavily on minute modifications in the structural mobility. We employed cross correlation analysis to retrieve the conformation dynamics information from the donor and acceptor intensity fluctuation time trajectory caused by the FRET process. We have constructed a mathematical model to describe three patterns: resonance, coherence, and decoherence, which discuss the structural heterogeneity of a biological system and allow for the investigation of the intermittent conformational change and the motion pattern of dynamics along with time.

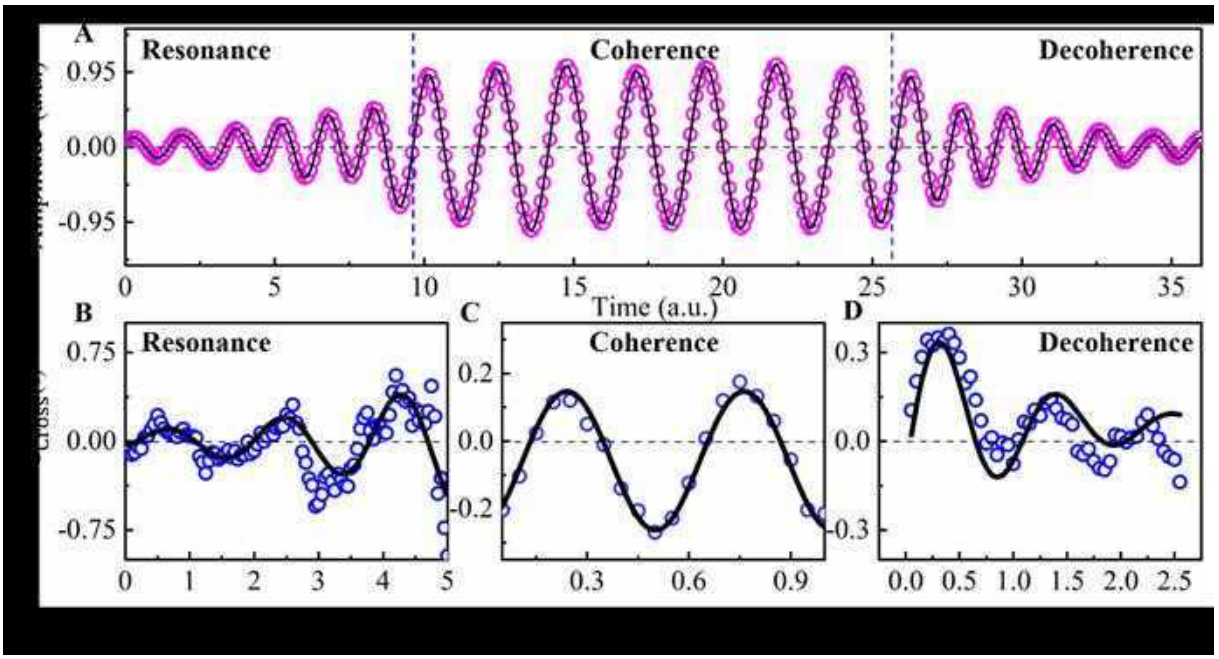


Figure 6.2 (A) A representative mathematical modelling of three different phase patterns of an intermittent oscillation: resonance, coherence and decoherence, calculated using equation, $y = \mathbf{Ae^{kt}\sin(\omega t + \phi)} + \mathbf{y_0}$ where A represents amplitude, $\boldsymbol{\omega}$ = angular frequency, $\boldsymbol{\phi}$ = phase shift and yo was the offset. k was the modulating coefficient, which was responsible for modulating the resonance, coherence or decoherence pattern. when k = positive represents the resonance pattern, k = negative corelates with the decoherence pattern and k = zero reflects the coherence phenomenon. The black line is the connecting line of the cross corelation represents here as fitting line. (B-D) Cross-correlation function analysis of donor and acceptor, which shows resonance, coherence, and decoherence patterns, respectively, of conformational fluctuation. This magenta circle is the cross-correlation data calculated from donor and acceptor intensity trajectories and the black line is the fitting line as per equation 6.19.

Resonance is a phenomenon that occurs when an external stimulus with a comparable frequency causes an increase in the amplitude of the system being studied. The coherence frequency is determined by the similarity in amplitude frequencies and the continuous phase difference between them. Decoherence occurs when the inherent coherence of a system is gradually lost over time. It can be described as the leakage or induction of coherent phenomena by external factors.

First, we used following mathematical model to describe these three pattern of dynamics.

$$y = Ae^{kt}\,\mathrm{Sin}(\omega t + \phi) + y_0 \qquad (6.19)$$

Where ω was the frequency of an oscillation pattern with time, ϕ was the phase shift, and A was the amplitude of the frequency. The exponential term introduced here to describe the resonance and decoherence features, where k was called the modulating coefficient which changes the amplitude of a wave nature.

$$k = \begin{cases} = 1, \text{ resonance} \\ = 0, \text{ coherence} \\ = -1, \text{ decoherence} \end{cases}$$

When the value of k is equal to 1, the frequency exhibits a pattern of increasing amplitude, and vice versa. Conversely, when the modulating coefficient is negative, a decoherent pattern is observed. Additionally, when k is equal to 0, meaning there is no exponential term, the system behaves as an oscillation wave with the same frequency and amplitude.

Figure 6.2.A illustrates the mathematical model that explains the wave pattern of resonance, coherence, and decoherence frequency. The key finding of this study was the discovery that the resonance phase (Figure 6.2.B) has a frequency of 3.48 s^{-1}, a phase shift of 0.62, and an amplitude of 0.08. The modulating factor, which governs the occurrence of resonance, was determined to be 0.37. The coherence phenomena (Figure 6.2.C) have a lower frequency fluctuation compared to the resonance pattern of 2.07 s^{-1}. It has a phase shift of 0.9 and a constant amplitude of 0.21. The modulating factor, which is 0, alters equation 6.19 as $y = A\sin(\omega t+\phi) + y_0$. The decoherence phase in Figure 6.2.D exhibits a frequency of 5.8 s^{-1}, which is higher than the frequency of the coherence phase. The decoherence phase also has a phase shift of -0.37 and a negative modulation coefficient of -0.9. Additionally, it demonstrates a decreasing amplitude factor with an amplitude of 0.38. The varying frequency suggests the presence of several types of conformational changes, such as local conformational changes or partial global structural alterations.

The coherence pattern indicates the presence of conformational heterogeneity in the system, which repeats after a specific time. Conformation dynamics may not consistently exhibit the coherent feature due to numerous factors influenced by the biological environment. The data indicates that the correlation oscillation pattern undergoes occasional alterations with varying frequency. The precise cause for the observation remains unknown, however it is likely because to the variation in conformation generated by Mg^{2+}, which changes with time. Here, we would

like to emphasize that the observation of intermittent change in the conformational dynamics is purely because of concerted motion of two branches of HJ around the junction point. Three patterns may also intermittently appear in the correlation function which is calculated from the fluorescence intensity time-trajectories. The following section specifically describes the coherence motion of dynamics which is most abundant in the observed dynamics heterogeneities.

6.3.3 Coherence Frequency in DNA HJ Conformations

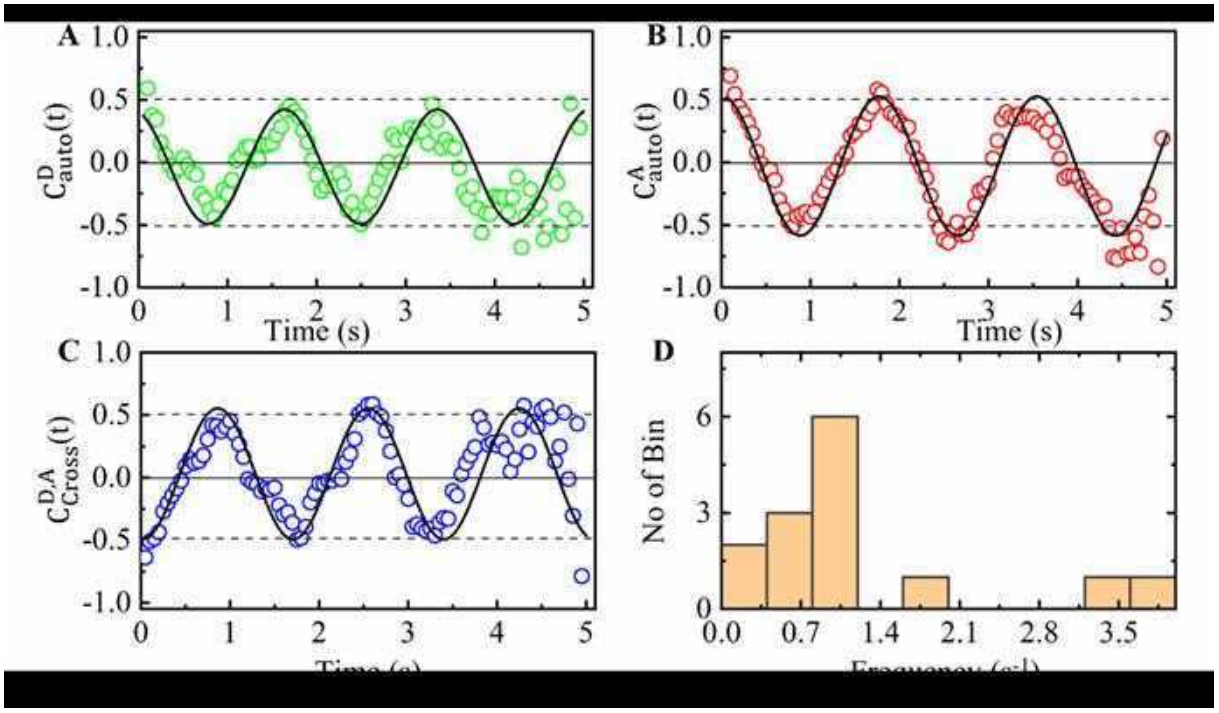


Figure 6.3 (A-B) Autocorrelation functions of donor (green) and acceptor (red) from single-molecule intensity time-trajectories represented in Figure 6.1.C. Data were fitted with the following equation $y = \mathbf{A} * \mathbf{Sin}(\boldsymbol{\omega t} + \boldsymbol{\phi}) + \mathbf{y_0}$, (in black) and obtained oscillatory frequency for both donor and acceptor were 3.8 s-1. (C) Cross-correlation function analysis clearly shows that two signals were anticorrelated, which was also evident from data in Figure 6.1.C with frequency 3.8 s-1. (D) The distribution of coherence frequency calculated from cross correlation function Figure 6.3.C.

The autocorrelation and cross-correlation functions, which are employed to analyze the intricate characteristics of coherence dynamics from individual molecule trajectories, are computed according to equations 6.1 and 6.2.

The donor and acceptor autocorrelation function were estimated in Figures 6.3.A and 6.3.B, revealing a distinct coherent characteristic at a one-second time scale. The coherence feature in conformational dynamics arises from the stochastic conformational behavior of DNA HJ, indicating a repetitive pattern of conformational changes in HJ. The plot in Figure 6.3.C displays the cross-correlation function, which exhibits an initial transition from negative to positive amplitude, followed by zero. This pattern reflects the characteristic anticorrelated behaviors observed in the single molecule trajectory. Figure 6.3.D displays a histogram representing the distribution of coherence frequencies within the range of 0.2-3.8 s^{-1}. The consistent coherence frequency pattern seen in the donor, acceptor autocorrelation, and cross-correlation provides confirmation that it arose from the same process of conformation fluctuation. The coherence pattern remains stable for a duration of 4 seconds, indicating conformational changes in the various sub-states.

6.3.4 Analysis of Stochastic Correlation Frequency and 2D Regional Correlation

In addition to three different patterns of dynamics, we also found stochastic pattern of correlation and its frequency. Here, we discussed mainly the conformational fluctuations of a single molecule trajectory, which contain critical information about the molecular level mechanism that does not follow a specific pattern, which is stochastic dynamics. Additionally, we have also applied the idea of two-dimensional regional function analysis to identify the map of the time segment when a molecule behaves in a correlated and anticorrelated way by cross correlation calculations.[278] In a conventional correlation calculation, it's considered the whole trajectory where the pattern was unable to reveal some complex behavior in case of a complex system with noisy background.[269] \A common approach in a single molecule spectroscopy was that when a fluctuation originates from a FRET pair, donor and acceptor will follow the anticorrelated pattern. But in a real experiment, there were thermal fluctuations also exist which affect the local environment, and, in that case, molecules behave in correlated patterns. Additionally, when fluctuations occur at timescales faster than the experimental frame rate or when noise dominates, the donor and acceptor signals may exhibit uncorrelated behavior.

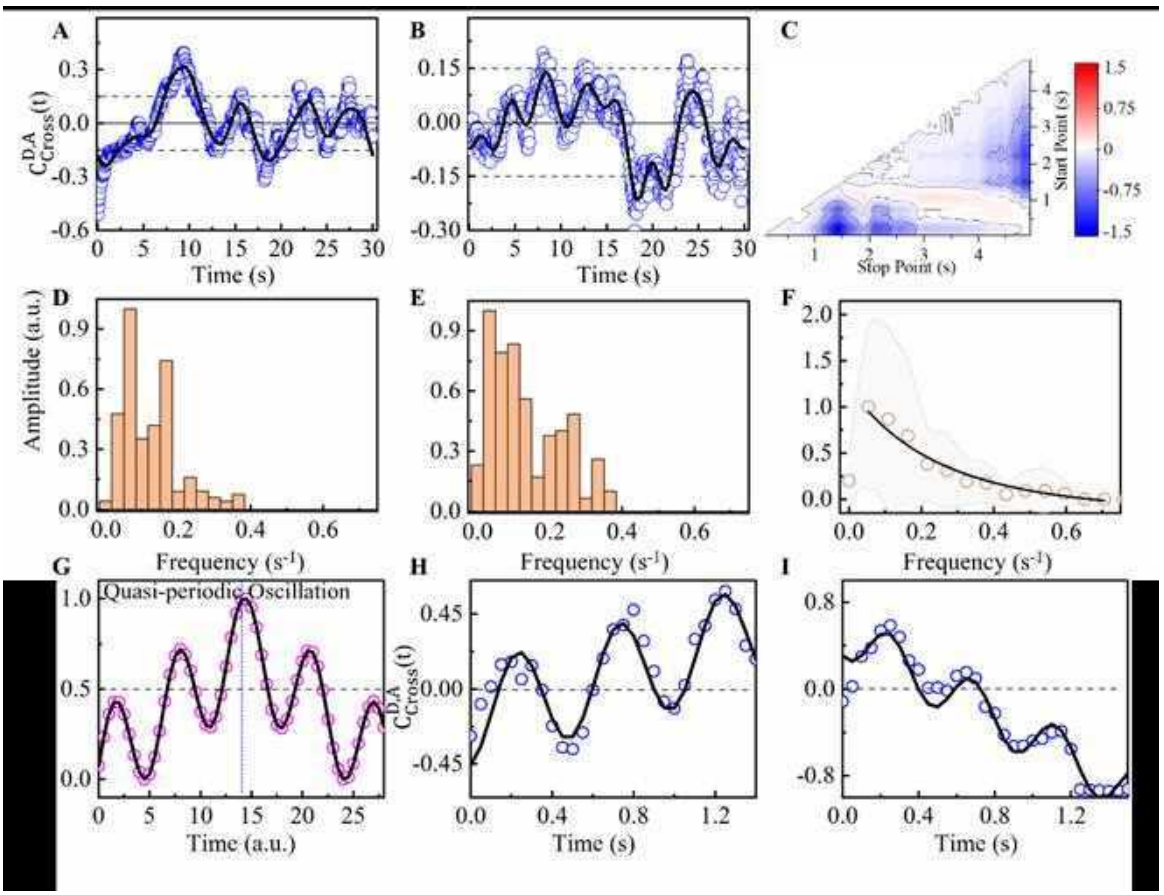


Figure 6.4 (A-B) Representative Cross-correlation function analysis, which shows stochastic frequency. (C) 2D regional correlation analysis by two-dimensional cross-correlation function amplitude distribution mapping. The blue colour represents negative amplitude, and the red colour represents positive amplitude. (D-E) The frequency (s-1) plot from the fast Fourier transform (FFT) analysis calculated from Figure 6.4.A and 6.4.B, respectively. (F) The Mean frequency (s-1) plot from the fast flourier transform (FFT) calculation shows mainly two frequencies, 0.17 (s-1) and 0.4 (s-1). (G) A representation of mathematical modelling of the quasiperiodic oscillation. The blue circle is the schematic representation of quasi-periodic function was generated by the equation y= $\mathbf{A} * \mathbf{Sin}\ (\boldsymbol{\omega}\mathbf{t} + \boldsymbol{\phi}) + \frac{\mathbf{t}}{\mathbf{k}} + \mathbf{y_0}$, where A represents amplitude, $\boldsymbol{\omega}$= angular frequency, $\boldsymbol{\phi}$ = phase shift

and yo was the offset. When k = positive is representing the growth curve in the quasi-periodic oscillation and the negative value of k represents the decay curve. The black line is the connecting line of the cross corelation represents here as fitting line. (H-I) Representative plot of Quasiperiodic oscillation obtained from the cross corelation of donor and acceptor intensity trajectories.

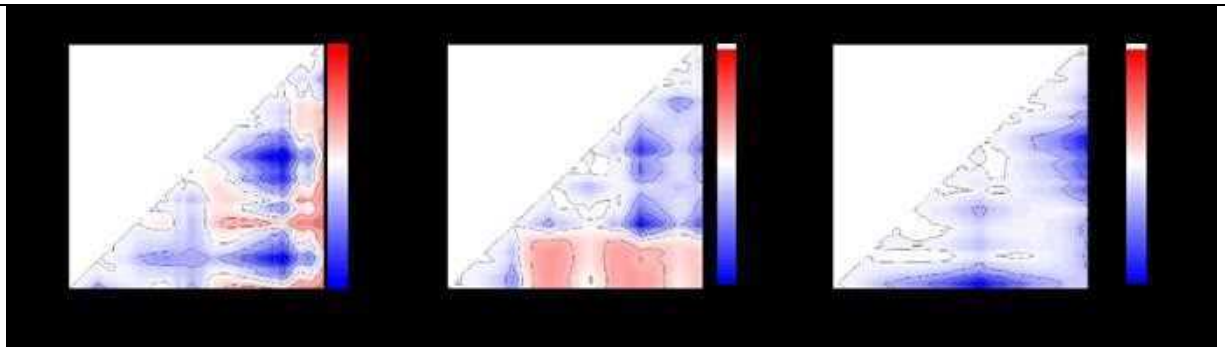


Figure 6.5 (A-C) 2D regional correlation analysis by two-dimensional cross-correlation function amplitude distribution mapping.

For these reasons, in the case of conventional correlation calculation considers the whole trajectory fluctuation equally and cannot able to identify the intermittent fluctuation origin.[266] Figure 6.4.A and 6.4.B represent the cross-correlation function of the trajectories, which show stochastic fluctuations without any pattern that can be defined by mathematical model. In Figure 6.4.C and Figure 6.5, we have calculated the two dimensions cross-correlation function by the equation 6.3, where we can be able to identify that the anticorrelated behavior was dominating. The cross-correlation function amplitudes of each time segment with separate starting and ending timings are generated in such an analysis, which involves scanning each of the smFRET D-A intensity trajectories with various starting and ending durations. To find the anti-correlated pattern a color bar is utilized to represent the amplitude of cross-correlation. Still, the molecule also behaves in a correlated pattern, which was an identification of thermal noise or faster function at that time point. The correlated pattern may be attributed to the local environment change in the DNA conformation also.

In Figures 6.4.(D-F), we have plotted the frequency plot calculated from the fast Fourier transformation from the cross-correlation function of Figures 6.4.(A-B) which shows a fluctuation of frequency in 0.1-0.8 s^{-1}. In the resonance, coherence or decoherence pattern, we have also observed that some of the trajectories follow a quasiperiodic oscillation, which was a dynamic behavior that follows the recurrence of a pattern at a smaller scale but was unpredictable in a larger time scale.

We have used the following function to describe this pattern,

$$y = A * \mathrm{Sin}\,(\omega t + \phi) + \frac{t}{k} + y_0 \qquad (6.20)$$

where A represents amplitude, ω= angular frequency, ϕ = phase shift, y_0 = the offset, k = modulating coefficient, when k has a positive value, then it shows an increasing pattern of amplitude. When k was negative, it showed a decreasing pattern, as described in Figure 6.4.G. Interestingly, we have found these two types of patterns in our case, with a frequency of 12.52 s^{-1} in the case of increasing amplitude of quasi-oscillation function (Figure 6.4.H) and 14.34 s^{-1} when this function shows decreasing amplitude (Figure 6.4.I).

6.3.5 Analysis of TCF Calculates the Rate of Conformational Changes and Activation Energy

We calculated all the kinetic information from the second-order TCF of FRET efficiency at different time intervals, which was the average product of two successive FRET efficiency at the time interval t_1, and t_2 by using the formula as per the previously reported method.[279]

$$C^2(\tau) = \langle E_{FRET}(0)E_{FRET}(\tau)\rangle - \langle E_{FRET}\rangle^2 \quad (6.21)$$

where $\langle E_{FRET}\rangle$ was the mean FRET value of the time trajectory, and the angle brackets denote the average value over all possible trajectory ranges Where $\tau = t_2 - t_1$.
In Figure 6.6.A, we plotted the second-order TCF vs the lag-time from the single molecule fluorescence time-trajectories. The data was fitted with equation 6.4 with a linear log scale with a biexponential decay where the $\tau_{fast} =$ 0.241 s (fastest component) and $\tau_{slow} =$ 5.2 s (slowest component) with an amplitude A_2 =0.449 and A_3 = 0.550. The second-order TCF shows two individual decay components which were well separated. According to the Markov model, the second-order TCF contains an (N-1) decay term, where N number of discrete and stochastic transitions were possible in an equilibrium kinetic chemical transition.[280] So, from the relevance of the number of decay terms in second-order TCF, we can clearly understand that in our system there was N $\geq$ 3 states were possible.

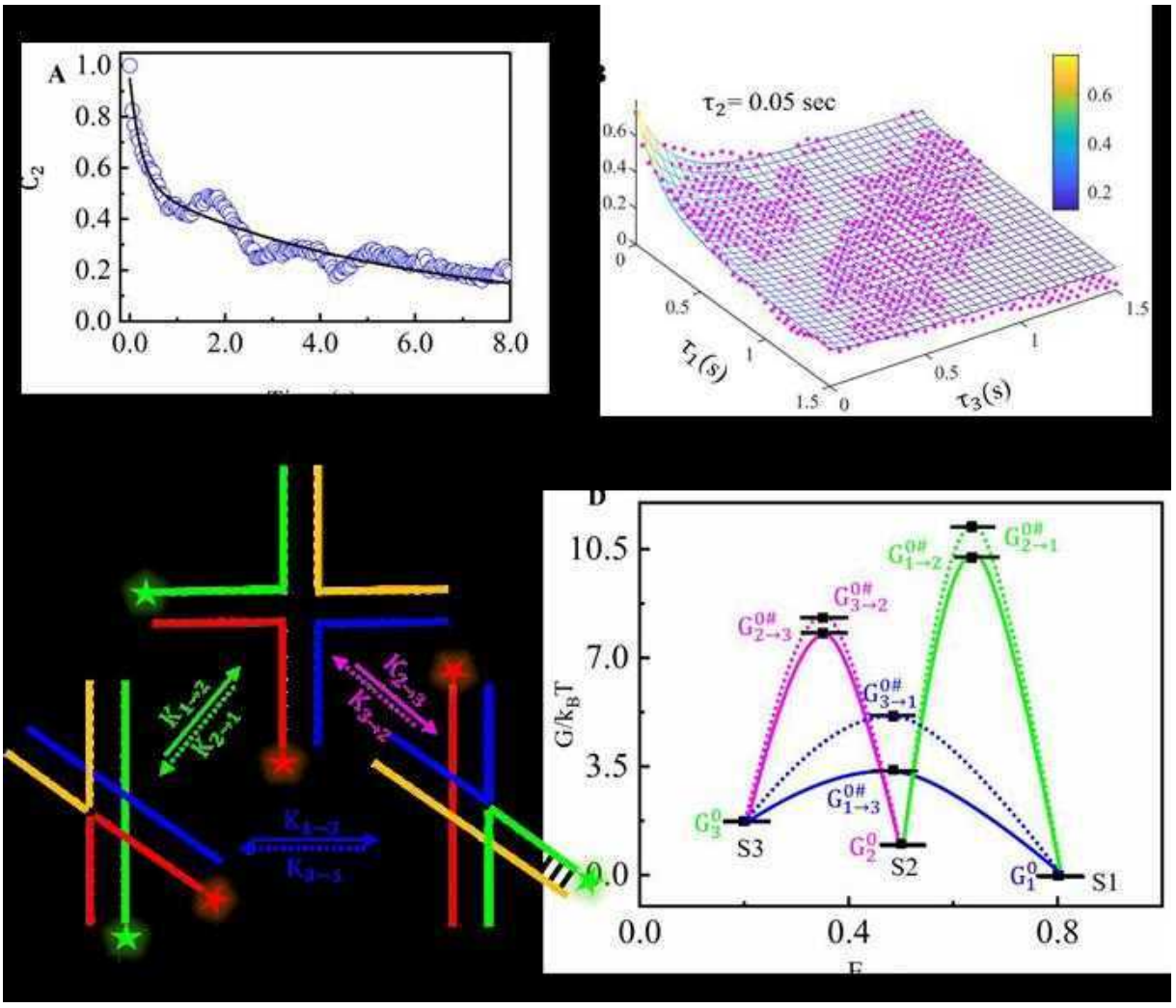


Figure 6.6 (A) Second-order TCF (blue circle) fitted with two exponential decay terms (black line) with two decay constants. (B) Three-dimension plot of fourth-order TCF at waiting time 0.05 s. (C) Schematic representation of 3 model states of HJ: State 1 (S1) (High FRET), State 3 (low FRET), and State 2 refers to the open state with

intermediate FRET. (D) Free Energy (k_BT) Landscape for DNA HJ calculated from the PDF (Table 6.1) and rate constant provided Table 6.2. The magenta line represents the free energy barrier between S2 and S3, the green line represents S1 to S2, and the blue line represents the free energy barrier of S1 to S3.

Table 6.1 Optimized parameter for PDF of three macrostate models fitted in bi-gaussian components as per equation 6.7.

A_1	$\langle E_{FRET}\rangle_1$	σ_1	P_1^{equ}	A_2	$\langle E_{FRET}\rangle_2$	σ_2	P_2^{equ}	A_3	$\langle E_{FRET}\rangle_3$	σ_3	P_3^{equ}
3.2	0.77	0.07	0.6	1.2	0.5	0.08	0.23	0.57	0.2	0.1	0.14

Table 6.2 Optimized forward and backward rate constant for DNA HJ N = 3 Macrostate. All the rate constants were derived from the 3 × 3 matrix as per equation 6.11-6.17.

k_{12} (ms^{-1})	k_{21} (ms^{-1})	k_{23} (ms^{-1})	k_{32} (ms^{-1})	k_{13} (ms^{-1})	k_{31} (ms^{-1})
0.00071	0.00026	0.008	0.005	0.66	0.117

Table 6.3 Optimized free energy minima and transition barrier for N=3 states of HJ. The free energy was derived as per the equation 6.18. (all the unit was given in kBT)

G_1^0	G_2^0	G_3^0	$G_{12}^{\#}$	$G_{21}^{\#}$	$G_{23}^{\#}$	$G_{32}^{\#}$	$G_{13}^{\#}$	$G_{31}^{\#}$
0	1.01	1.74	10.24	11.21	7.80	8.30	3.4	5.13

The PDF of FRET histogram reveals that three macrostates were present and describe in equation 6.5-6.9. State 1 contains the most dominant component with a mean FRET value of 0.77 and P_1^{equ}= 0.61, state 2 includes the majority of a peak at around 0.57 FRET efficiency with P_2^{equ} = 0.23. State 3 contains a very trace Gaussian component with P_3^{equ} = 0.14, where the FRET efficiency was around 0.2, as per Table 6.1.

Now, we consider here that a conformation change in the macrostates permits a stochastic transition from the i-th state to the j-th state, and the states were discrete and continuous. Then, the rate constant and the free energy of activation ($G_{ij}^{\#}$) were calculated as per the previously described method in equation 6.18, [281]We have also calculated the forward and backward rate constant (Table 6.2) for each macrostate from equation 6.11-6.17. It was worth noting that the interconversion process of State 1 ($\langle E_{FRET}\rangle_1$ = 0.77) to state 3 ($\langle E_{FRET}\rangle_3$ = 0.2) were much faster 0.66 ms^{-1} than that of backward state 3 and state 0.117 ms^{-1}, and the slowest interconversion happens in between state 1 to state 2 0.00071 ms^{-1} and for the backward reaction from state 2 to state 1 was 0.00026 ms^{-1}. The S2 to S3 state has an intermediate rate constant of 0.008 ms^{-1}, and for the backward process, the rate was 0.005 ms^{-1}.

To distinguish between different conformational changes of HJ as they connect independently or were interconnected, we have determined fourth order time correlation function; this was the product of four sequential data points of smFRET trajectories, separated by the time interval $\tau_1, (t_2 - t_1), \tau_2\ (t_3 - t_2)$, and $\tau_3(\ t_4 - t_3)$.

The fourth-order TCF can be calculated as[13]

$$C^{(4)}(\tau_1, \tau_2, \tau_3) = \langle E_{FRET}(0)E_{FRET}(\tau_1)E_{FRET}(\tau_2)E_{FRET}(\tau_3)\rangle \quad (6.22)$$

The fourth-order TCF illustrates the correlation between the interval at τ_1 and τ_3 at a fixed τ_2 value, depending on the weighted observation of each possible time. The fourth-order TCF was composed of $(N-1)^2$ terms, which were composed of four amplitude factors as [271, 273], where $\lambda_2 =$ 0.192 s^{-1} and $\lambda_3 = 4.14$ s^{-1}. The fourth-order TCF can be fitted by the equation[271, 279]

$$C^4(\tau_1, \tau_3)_{\tau_2 \text{ fixed}} = A_{22}(\tau_2)e^{\lambda_2(\tau_1+\tau_3)} + A_{23}(\tau_2)e^{\lambda_2\tau_1+\lambda_3\tau_3} + A_{32}(\tau_2)e^{\lambda_3\tau_1+\lambda_2\tau_3} + A_{33}(\tau_2)e^{\lambda_3(\tau_1+\tau_3)} \quad (6.23)$$

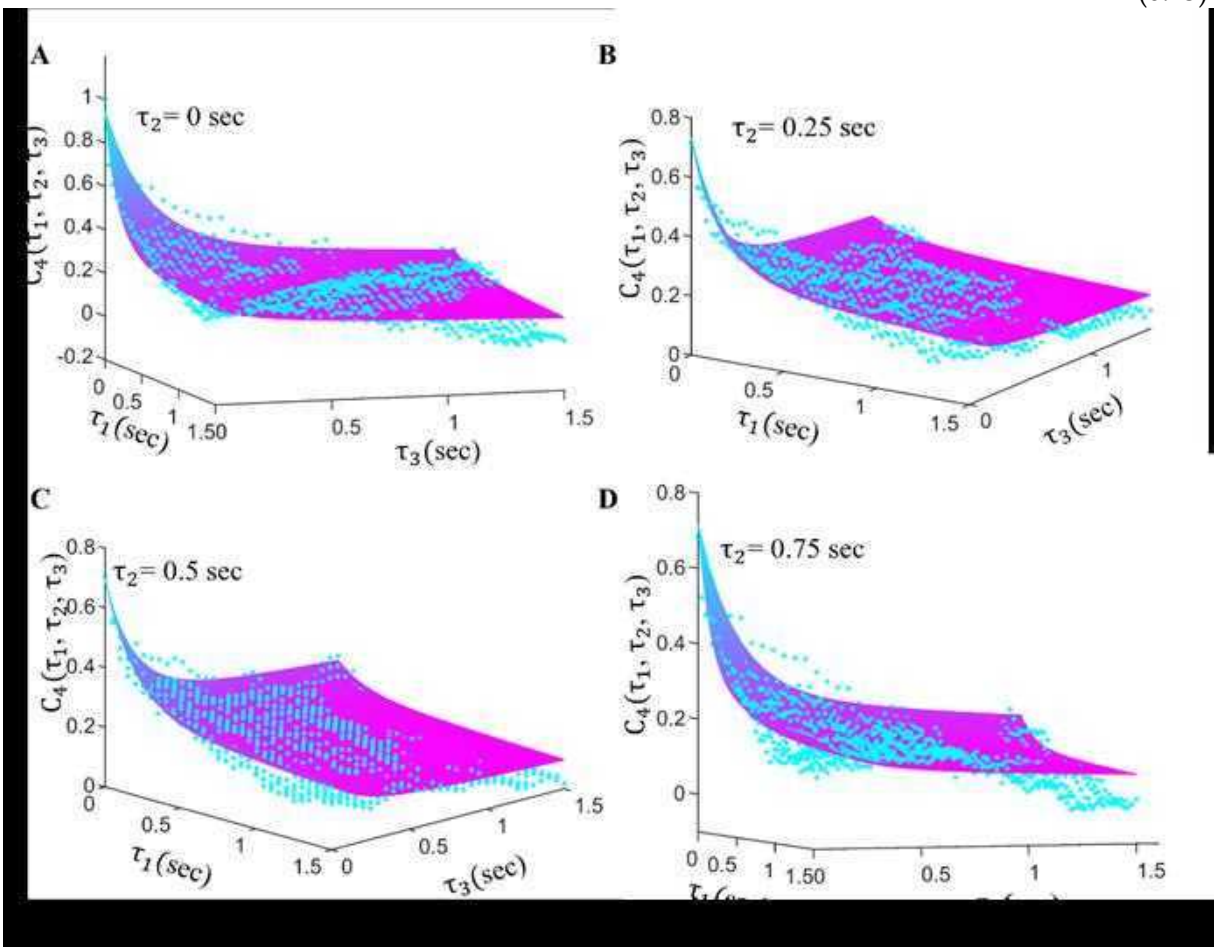


Figure 6.7 Three-dimension plot of fourth order TCF at waiting time (A) 0 sec (B) 0.25 sec (C) 0.5 sec and (D) 0.75 sec.

The fourth-order TCF at the waiting time $\tau_2 = 0.05$ s showed in Figure 6.6.B, reflects a positive cross amplitude of A_{23} and A_{32} = 0.05 for $\tau_2 = 0.05$ s. The one self-amplitude factor was A_{22} =0.2216 and another factor A_{33} = 0.4325 for $\tau_2 = 0.05$ s .The fourth-order TCF with various waiting time τ_2 were shown in Figure 6.7. The positive amplitude of cross amplitude term concludes that there was no kinetic bottleneck present in the intermediate steps, and the whole nature of the states was solely due to dynamic behavior.

In Figure 6.6.C, we depicted the HJ in three state schematic diagrams, where S1 represent too high FRET ($\langle E_{FRET}\rangle_1 = 0.77$), representing a fully folded state in the presence of 50 mM Mg^{2+}, where S3 represents the low FRET structure where the mean FRET value was 0.2. The S2 represents the open state which refers to the FRET efficiency of 0.5. In Figure 6.6.D, we have plotted the free energy surface as a function of the FRET efficiency for each macrostate contribution. Here, we have considered G_1^0=0 as this state was taken as the ground state, the '0'

superscript refers here that we considered state 1 as minimum as a reference for all other energies, and the excited state energy was calculated by[271]

$$G_i^0 = -k_B T \ln \frac{P_i^{eq}}{\sigma_i \sqrt{2\pi}} - G_1 \quad (6.24)$$

the activation energy was determined by equation 6.18. We have also plotted the activation energy diagram for the transition from one state to another (Table 6.3). The highest activation energy was required for the transition S1 to S2 transition; the activation energy was 10.24 $k_B T$, whereas the backward activation energy (11.21 $k_B T$) was higher than forward. The forward transition barrier for S2 to S3 was 7.80 $k_B T$, whereas the backward transition barrier was 8.30 $k_B T$, which was lower than for S1 to S2. The lowest activation energy was required for S1 to S3, and the forward activation energy was 3.4 $k_B T$, whereas the backward transition was 5.13 $k_B T$. The parabolic behavior of the free energy surface vs the FRET efficiency indicates that we have considered a Gaussian distribution dependency for representing the macrostates.

Here, we would like to summarize the discussion based on the experimental results. (1) We have reported the conformational changes of the HJ induced by the Mg^{2+}, which shows an oscillation pattern in the correlation function analysis. To elucidate the information about the pattern, we have developed a mathematical model for the resonance, coherence and decoherence pattern. The coherence pattern shows the reoccurrence of the conformational states which can be related to the memory effect in DNA HJ dynamics. The decoherence pattern reflects that after some time, there was no oscillation found, which may be due to the stabilization of one conformation upon another conformation, or it may be the consequence of another environmental effect. (2) We have also found that the correlation function can be a stochastic frequency fluctuation or can be in a semi periodic motion, which was mostly common in a biological system. (3) We have also calculated 2D regional correlation analysis to get the exact information about the dynamic functional (anticorrelation) under a non-correlated noise background which may be induced by thermal fluctuation or from some environmental effect. In smFRET data, if this type of fluctuation occurs in a similar timescale in which FRET occurs, then there was a high chance that the exact information was not extracted, which can be distinguished by the 2D regional correlation analysis. (4) We have also interpreted the data based on the TCF to find out the information about the conformational state, rate constant and activation energy which modulates the rate constant of the conformational states. As the second-order TCF has fitted with biexponential decay, as per the Markov model, there were three possible states: High FRET (S1), intermediate FRET (S2) and low FRET (S3). The fourth-order TCF reflects that there were no kinetic bottlenecks present in the present, which reflects the independent dynamic behavior of each state. Among the states, the fastest rate conversation occurs between the High FRET state and to low FRET state (S1 to S3), also confirmed by the lowest activation energy of the system, 3.4 $k_B T$. The highest activation energy was required when the conversation occurred from the intermediate state (S2) to the High FRET state (S1), with the lowest rate constant of 0.00026 ms^{-1}.

6.4 Conclusion

In summary, we have observed an oscillation pattern including resonance, coherence and decoherence motion changes of DNA HJ under the influence of Mg^{2+}. The frequency fluctuation in the motion comes from the stochastic fluctuation of three conformational states. There might be a possibility that many complex wave patterns exist in a biological system which directly relate to the conformational change with time. We have also found the three macrostate populations in the DNA HJ. However, many intermediate states can be present which is above the experimental resolution.