Hidden complexity in the isomerization dynamics of Holliday junctions

Changbong Hyeon[1]★, Jinwoo Lee[2], Jeseong Yoon[1], Sungchul Hohng[2] and D. Thirumalai[3]★

A plausible consequence of the rugged folding energy landscapes inherent to biomolecules is that there may be more than one functionally competent folded state. Indeed, molecule-to-molecule variations in the folding dynamics of enzymes and ribozymes have recently been identified in single-molecule experiments, but without systematic quantification or an understanding of their structural origin. Here, using concepts from glass physics and complementary clustering analysis, we provide a quantitative method to analyse single-molecule fluorescence resonance energy transfer (smFRET) data, thereby probing the isomerization dynamics of Holliday junctions, which display such heterogeneous dynamics over a long observation time ($T_{obs} \approx 40$ s). We show that the ergodicity of Holliday junction dynamics is effectively broken and that their conformational space is partitioned into a folding network of kinetically disconnected clusters. Theory suggests that the persistent heterogeneity of Holliday junction dynamics is a consequence of internal multiloops with varying sizes and flexibilities frozen by Mg^{2+} ions. An annealing experiment using Mg^{2+} pulses lends support to this idea by explicitly showing that interconversions between trajectories with different patterns can be induced.

Although challenged occasionally[1–3], it has long been considered a general principle that the folded states of biomolecules are uniquely determined by their sequences and environmental conditions[4]. The observation of multiple folding paths and the dependence of folding routes on initial conditions[5–8] illustrate the ruggedness of the folding landscapes[9–13] and the complexity of the process. These findings, in themselves, do not challenge the notion of a unique native state. However, recent data from single-molecule experiments on several biomolecular systems explicitly show persistent heterogeneities in the time traces (or molecule-to-molecule variations) generated under identical folding conditions[11,12,14–17]. Phenotypic cell-to-cell variabilities among genetically identical cells, visualized using microscopy, seems plausible and hence is well appreciated[18]. In contrast, heterogeneities among individual biomolecules on much smaller length scales are rather surprising because they are difficult to reconcile with the conventional notion that the functional states of proteins and RNAs are unique or that various native basins of attraction easily interconvert. For example, in docking–undocking transitions of surface-immobilized hairpin ribozymes[11] and *Tetrahymena* group I intron ribozymes[16], each time trace of individual molecules displays very different dynamic patterns, with long memories and without apparent compromise in catalytic efficiency. It has therefore been suggested that these ribozymes have multiple native states[16]. Control experiments with vesicle encapsulation demonstrated heterogeneities similar to those with surface immobilization, indicating that the heterogeneities are intrinsic to the molecules being probed and are not instrumental artefacts[12,15,19].

Given the ubiquity of molecule-to-molecule variations in single-molecule experiments, it is important to devise analytical tools to quantify observations using rigorous theoretical treatments based on statistical mechanics. Furthermore, examination of a molecular system simpler than the structured RNAs mentioned above, but which still displays persistent conformational heterogeneity, might provide a glimpse into its molecular origin. To this end, we studied the dynamics of Holliday junctions, which globally undergo a simple two-state-like isomerization transition in the presence of Mg^{2+} ions, but reveal complex behaviour when examined in detail.

Time-series data from single-molecule measurements can reveal the rate of conformational space navigated by a molecule[20,21]. Thus, the manifestation of molecule-to-molecule variation in the measured trajectories implies that each molecule samples only a small subset of the entire conformational space on T_{obs}. However, this is often hard to quantify, because the time series of an observable in single-molecule measurements of biomolecules results from a projection of dynamics in high-dimensional space onto a lower dimension. The extent to which a trajectory samples the allowed conformational space depends on the length of observation time T_{obs}. Conversely, states hidden in multiple deep furrows of the folding landscapes at high dimensions restrict the dynamics of each molecule to one of many states that are non-interconvertible within T_{obs}.

Historically, the widely accepted notion of a unique native state in biomolecules was hypothesized based on bulk measurements, where an averaged property of a probe variable was obtained from an ensemble of snapshots. Such a conclusion assumes ergodicity, that is, the equivalence between time and the ensemble average of an observable. Although ergodicity is a necessary condition for equilibrium systems, it is difficult to realize in practice because of the unlikelihood that in a single time trace a molecule can sample the entire configurational space[22]. Furthermore, the situation is further exacerbated because, nominally, the observation time in practice (or relevant timescales for many biological phenomena) is limited. In a rugged landscape, a molecule with an initial conformation distinct from others would repeatedly sample distinct regions of the folding landscape over a long observation time, which could be longer than the 'biologically relevant timescale'. This scenario results in heterogeneous dynamics, and ensemble-averaging would obscure the complexity of the structural features of the underlying landscape. In this sense, the ergodicity of the system is therefore effectively broken. Now that molecular heterogeneities are clearly demonstrated in many single-molecule data for a

[1]School of Computational Sciences, Korea Institute for Advanced Study, Seoul 130-722, Korea, [2]Department of Physics and Astronomy, Seoul National University, Seoul 151-747, Korea, [3]Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, USA. *e-mail: hyeoncb@kias.re.kr; thirum@umd.edu

variety of unrelated systems it is a major theoretical challenge to devise practical tools to reconstruct rugged folding landscapes from time series.

Here, we perform single-molecule fluorescence resonance energy transfer (smFRET) experiments[23] and use concepts from glass physics[24,25] and complementary clustering algorithms[26,27] to carry out a systematic analysis of smFRET data from metal-ion-driven conformational changes in a Holliday junction with a DNA sequence that prohibits branch migration[28]. We show quantitatively that although the Holliday junction dynamics at the ensemble level can be pictured using a two-state model, the ergodicity of the system is effectively broken. The associated folding landscape of the Holliday junctions is visualized in terms of rarely interconverting multiple ordered states embedded in the two isoforms of Holliday junctions. Furthermore, the simplicity of the Holliday junction structure is explored to discover the structural origin of the heterogeneities in dynamics. The presence of internal multiloop topologies with various sizes and flexibilities at the junction, and the high local concentration of Mg^{2+} ions around it (as calculated using molecular dynamics simulations), lead us to propose that the heterogeneities in

Holliday junction isomerization are due to non-interconverting multiloop topologies 'pinned' by the complexed Mg^{2+} ions. We validate the structural explanation by showing that trajectories with different patterns interconvert by an annealing protocol involving cycling the Mg^{2+} ions from high to low to high concentrations.

Results and discussion

smFRET measurements of Holliday junction dynamics and analysis using ensemble averaging. In DNA recombination, Holliday junctions are essential intermediates for strand exchange (Fig. 1a)[29]. At Mg^{2+} concentrations exceeding $\sim$50 µM, Holliday junctions exist in two distinct isoforms (iso-I and iso-II), both of which have characteristic X-shaped architectures. According to previous studies[28,30], in the absence of divalent ions Holliday junctions have a stable open square structure with an apparent FRET efficiency value of $E \approx 0.3$, but at $[Mg^{2+}] > 50$ µM, conformational changes take place between two stable conformers, iso-I and iso-II, via the open square structure.

We performed smFRET experiments for a range of Mg^{2+} ion concentrations to monitor the Mg^{2+}-dependent isomerization
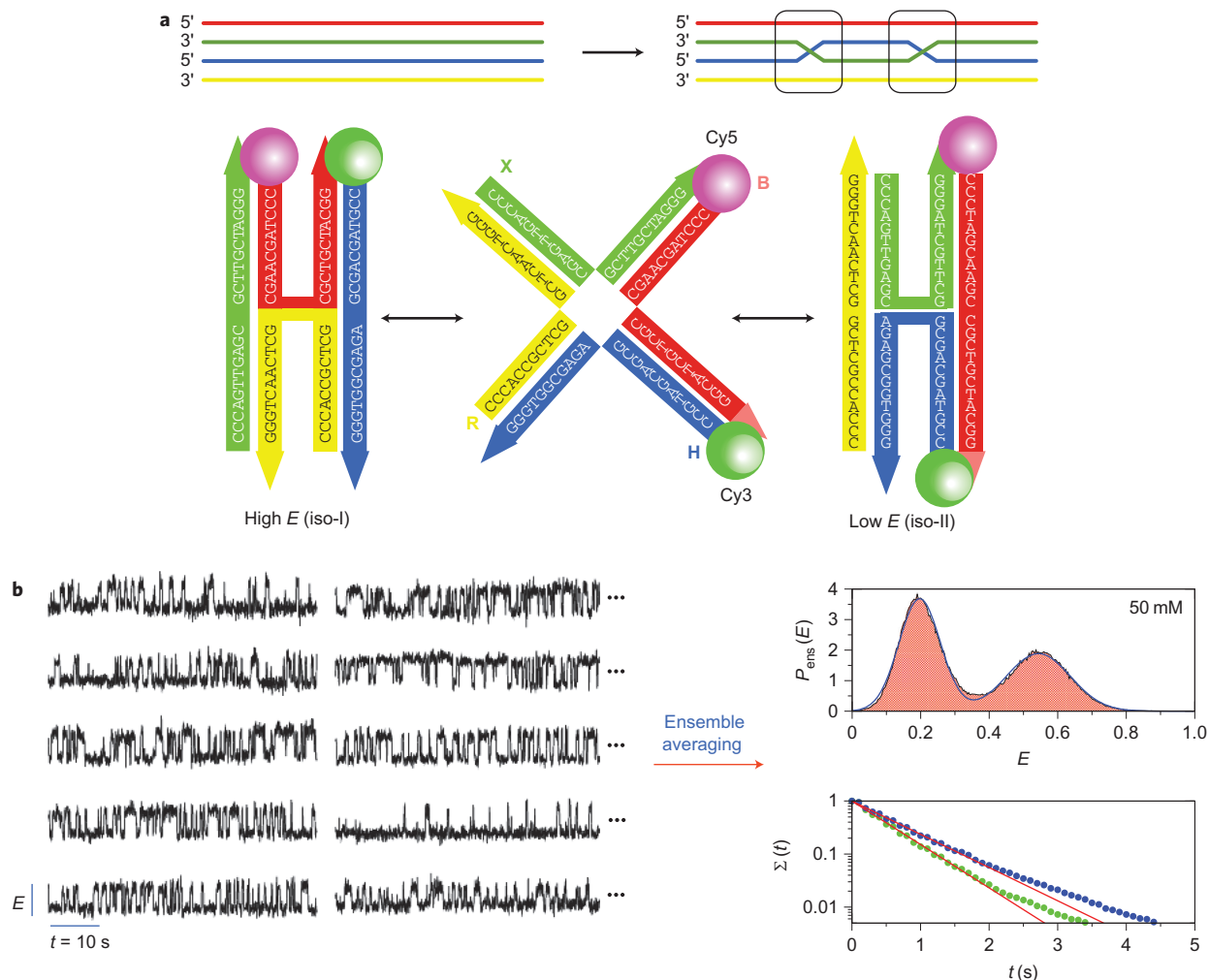
Figure 1 | Holliday junction dynamics probed with smFRET experiments and their analysis using a conventional ensemble averaging method. a, Schematic of strand exchange in DNA recombination (top) and the two isoforms connected by the open square structure (bottom). The Cy5 (magenta) and Cy3 (green) dyes attached to the B and H branches for the smFRET measurement are depicted as spheres. **b,** Part of FRET time traces ($\{E_i(t)\}$ with $i = 1, 2, \ldots, N$ and $N = 315$) obtained for individual Holliday junction molecules at $[Mg^{2+}] = 50$ mM. Similar to the findings in the literatures, conventional analysis using the ensemble of these time traces without deliberating the molecule-to-molecule variation gives rise to an interpretation of the apparent two-state behaviour for Holliday junctions. The ensemble-averaged histogram of the FRET efficiency E, that is, $P_{ens}(E)$, is nicely fit to a double Gaussian curve (blue line), and the dwell time distribution (bottom panel) for low (data in green) and high (data in blue) FRET states are approximately fit to single exponential functions (red lines).

dynamics of surface-immobilized Holliday junctions by attaching Cy3 and Cy5 dyes to the termini of the X and R branches (Fig. 1a; Supplementary Section S1). The time-dependent FRET efficiency $E_i(t)$ for the ith molecule ($i = 1, 2, \ldots, N$) was calculated by taking the ratio between the emission signals ($I_{A,i}(t)$ and $I_{D,i}(t)$) from acceptor and donor dyes using $E_i(t) = I_{A,i}(t)/(I_{A,i}(t) + I_{D,i}(t))$ (Supplementary Fig. S3). The individual time trajectories of $E_i(t)$ monitored for $\mathcal{T}_{obs} \approx 40$ s at $[Mg^{2+}] \geq 0.5$ mM show multiple transitions between high ($E \approx 0.5$) and low FRET ($E \approx 0.2$) values, corresponding to the iso-I and iso-II conformers, respectively. Histograms of FRET values collected over the entire observation time and population, $P_{ens}(E)$, reveal bimodal distributions for $[Mg^{2+}] \geq 0.5$ mM, all of which can be fit nicely to double Gaussian functions, and a unimodal distribution for $[Mg^{2+}] = 0.0$ mM (Fig. 1b and Supplementary Fig. S4). With increasing $[Mg^{2+}]$, the two peaks on $P_{ens}(E)$ separate more clearly, indicating that the transitions are becoming increasingly cooperative. In addition, the ensemble average of FRET efficiency $\langle E \rangle = \int_0^1 E P_{ens}(E) dE$ is invariant at $[Mg^{2+}] \geq 0.5$ mM (Supplementary Fig. S4), which is in agreement with a previous finding[28] that showed that the relative population of the two isoforms is entirely controlled by the junction sequences and not by the ionic conditions.

In addition to the ensemble-averaged thermodynamic measure $P_{ens}(E)$, transition kinetics (which occur with an average time of $\tau_{I\leftrightarrow II} \approx 0.1\text{--}1$ s) between the two isomers provide a glimpse into how Holliday junction isomerizations occur with varying Mg^{2+} conditions. As suggested by the shapes of the Mg^{2+} ion-dependent $P_{ens}(E)$ (Supplementary Fig. S4), the transitions between the two conformers slow down with increasing Mg^{2+} ion concentration, presumably due to an increase in the free energy barrier at higher concentrations. The survival probabilities ($\Sigma(t)$) calculated from the dwell time distributions ($p_{dwell}(t)$) for both high and low FRET values $\Sigma(t)[= 1 - \int_0^t p_{dwell}(\tau)d\tau]$ can be approximately fit using a single exponential function, although 5% of the population deviates from the fit. Thus, it is tempting to surmise that an approximate two-state picture is adequate for describing the transition between the two means[28] (Fig. 1b, Supplementary Fig. S5).

Molecule-to-molecule variation in individual time traces. The two-state model for the Mg^{2+} ion-dependent isomerization dynamics of the Holliday junction, gleaned from averaging over an ensemble of molecules, ignores the intrinsic heterogeneities among the trajectories. However, an inspection of a few disparate trajectories makes clear the dramatic variations between individual molecules. In Fig. 2a one can immediately identify the differences between the three exemplary trajectories that maintain the characteristic pattern of dynamics on $\mathcal{T}_{obs}$. The multiple transitions in each time trace render the time average of $E_i(t)$, that is, $\varepsilon_i(t)[= \frac{1}{t}\int_0^t d\tau E_i(\tau)]$ stationary, which suggests that the conformational space associated with that particular trajectory is exhaustively sampled. Thus, using such a trajectory, it is legitimate to calculate a stationary distribution $p_s(E;i) = \lim_{t \to \mathcal{T}_{obs}} p(E,t;i)$ (ref. 31), which represents the population of microstates probed on $\mathcal{T}_{obs}$ in terms of FRET efficiency. Notably, despite a number of rapid transitions in each time trace, a finding that is nominally associated with canonical ergodic sampling of conformational space, there are qualitative differences between $p_s(E;i)$ and $P_{ens}(E)$, and among $p_s(E;i)$s with $i = 1, 2, 3$, indicating that the ergodicity of Holliday junction dynamics is effectively broken at $\mathcal{T}_{obs} \approx 40$ s.

In the literature, the overall variation in dynamics and the associated heterogeneities are often conveniently visualized in the form of scatter plot of the mean dwell times at low and high FRET signals for each molecule, as in Fig. 2b (refs 15,32,33). The mean dwell times for the three time traces in Fig. 2a are circled in Fig. 2b. Note that the scattered data points in Fig. 2b should in principle distribute in a single spot if the Holliday junction dynamics is truly ergodic.

At an abstract level, the following landscape picture can be implicated from the above findings. For a given time trace belonging to a dynamical pattern, corresponding to a specific molecule α, $\mathcal{T}_{obs} \approx 40$ s is long enough to observe multiple isomerization events, so that the timescale for single isomerization between iso-I and iso-II ($\tau_{I\leftrightarrow II}^{\alpha}$) is much smaller than $\mathcal{T}_{obs}$ ($\tau_{I\leftrightarrow II}^{\alpha} \ll \mathcal{T}_{obs}$). Thus, the Holliday junction explores the conformations in the α state exhaustively. However, this $\mathcal{T}_{obs}$ is not long enough for interconversion to
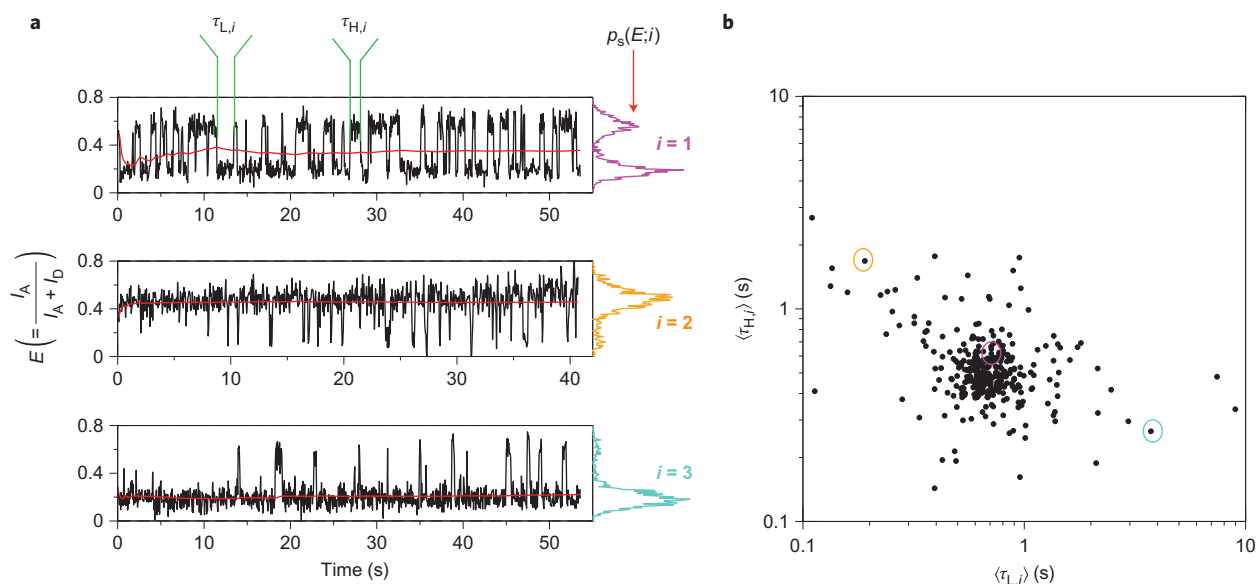
Figure 2 | Molecule-to-molecule variation (or molecular heterogeneity) manifested in the time traces of isomerization dynamics of Holliday junctions. a, Three disparate FRET time traces $E_i(t)$ at $[Mg^{2+}] = 50$ mM (blue), their time average $\varepsilon_i(t)$ (red), and the corresponding histograms on the right. Because each time trace can be considered stationary, as clearly indicated in the time independence of $\varepsilon_i(t)$, it is legitimate to build a histogram for each time trace and designate the histogram a stationary distribution, $p_s(E;i)$. **b**, Molecular heterogeneity revealed in the scatter plot of average dwell times at low and high FRET states for individual time traces ($\langle \tau_{L,i} \rangle, \langle \tau_{H,i} \rangle$). The circled data points correspond to the three time traces in **a**.
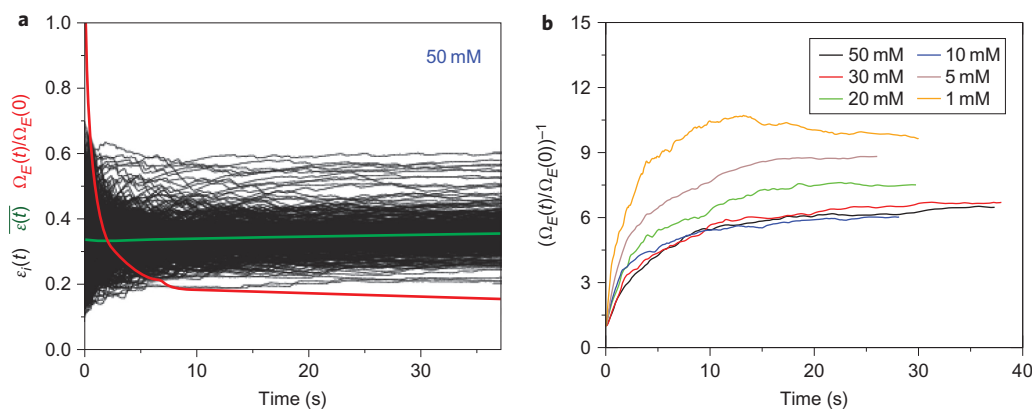
Figure 3 | Probing ergodicity breaking. a, Probes of ergodic behaviour in smFRET trajectories using the metric $\Omega_E(t)$ at $[Mg^{2+}] = 50$ mM. $\varepsilon_i(t)$, $\overline{\varepsilon(t)}$ and $\Omega_E(t)/\Omega_E(0)$ (defined in the text) are shown as black, green and red lines, respectively. **b**, The nonlinearity of $[\Omega_E(t)]^{-1}$ for $[Mg^{2+}] > 1$ mM shows that the Holliday junction dynamics is non-ergodic for all concentrations for observation time $\mathcal{T}_{obs}$.

take place between the pattern α and another pattern, say β; that is, $\mathcal{T}_{obs} \ll \tau_{conv}^{\alpha \leftrightarrow \beta}$, where $\tau_{conv}^{\alpha \leftrightarrow \beta}$ is the interconversion time between α and β states, implying that a substantially high kinetic barrier separates states α and β. Therefore, the dynamics of Holliday junctions are effectively ergodic within each state on $\mathcal{T}_{obs}$, but $\mathcal{T}_{obs}$ is not long enough to ensure ergodic sampling of the entire conformational space—a situation that is reminiscent of ergodicity breaking in supercooled liquids[24] (see also Supplementary Section S2).

Assessing ergodicity from time series. To demonstrate quantitatively that interconversion between the multiple states of the Holliday junction implicit in smFRET trajectories is unlikely, we use concepts from glass physics[24,34]. If the entire conformational space of the Holliday junction is effectively sampled during $\mathcal{T}_{obs}$, which would establish ergodicity, then the time-averaged $E_i(t)$, that is, $\varepsilon_i(t)$, should converge to an ensemble average for all i. We use a time-dependent metric $\Omega_E(t) = \frac{1}{N} \sum_{i=1}^{N} (\varepsilon_i(t) - \overline{\varepsilon(t)})^2$ with $\overline{\varepsilon(t)} \equiv \frac{1}{N} \sum_{i=1}^{N} \varepsilon_i(t)$, introduced to probe the approach to equilibrium in the context of simulations of supercooled liquids and glasses[24,34], to analyse smFRET time-series data. For ergodic systems, $\Omega_E(t)$ converges to zero at $t \rightarrow \infty$. For finite time t with $\Omega_E(t) \neq 0$, it can be shown that $\Omega_E(t)$ decays as t^{-1} asymptotically; thus, $[\Omega_E(t)/\Omega_E(0)]^{-1} \approx D_E t$ (see Supplementary Section S3 for further details). Because the form of the metric $\Omega_E(t)$ is similar to the mean square displacement, the slope of $[\Omega_E(t)/\Omega_E(0)]^{-1} = D_E$ is interpreted either as an effective diffusion constant or an effective sampling rate of the conformational space projected onto the FRET efficiency coordinate.

For the entire ensemble of Holliday junction trajectories, however, we find that $[\Omega_E(t)]^{-1}$ is not linear in time, but converges to a finite value. The nonlinearity of $[\Omega_E(t)]^{-1}$ indicates that the conformations belonging to distinct states do not mix on the time scale of $\mathcal{T}_{obs}$, which implies effective ergodicity breaking in the Holliday junction dynamics at $\mathcal{T}_{obs}$ (Fig. 3b)[24]. Remarkably, this conclusion still holds for Holliday junction trajectories probed at an extended timescale of $\mathcal{T}_{obs} \approx 70$ s (Supplementary Fig. S6). Notably, the variance at $\mathcal{T}_{obs}$, quantified using $\Omega_E(\mathcal{T}_{obs})/\Omega_E(0)$, is greater at higher $[Mg^{2+}]$ (Fig. 3b), suggesting that interactions with Mg^{2+} ions are responsible for creating the heterogeneous environment for Holliday junction molecules.

Partitioning the conformational space using a complementary clustering algorithm. How many states that do not interconvert on $\mathcal{T}_{obs}$ are needed to fully account for the experimental data? To answer this question we complement the general approach discussed above for quantitatively assessing the ergodicity from time trajectories with a K-means clustering algorithm. This allows

us to partition the conformational space of the Holliday junction into multiple 'ergodic subspaces'. The K-means clustering algorithm partitions data into K clusters, so that the distance from data belonging to the cluster to the cluster mean is minimized and the intercluster distance is maximized. In our problem, the data correspond to the stationary distributions of individual time trajectories. We partitioned the set of stationary distributions $\{p_s(E;i)|i = 1, \ldots, N\}$ into K clusters, ensuring that $[\Omega_E(t)]^{-1} \approx t$ in each cluster (see Supplementary Sections S4, S5 and Fig. S7 for further details) so that within each cluster the dynamics is ergodic. The consequences of our analysis are summarized as follows: (i) for $[Mg^{2+}] = 50$ mM, $\{p_s(E;i)|i = 1, \ldots, N\}$ are partitioned into five clusters (ergodic subspaces) (Fig. 4a; see also Supplementary Section S5 and Fig. S7); (ii) the complementary clustering algorithm enables us to partition the set of time-averaged trajectories ($\{\varepsilon_i(t)\}$) (Fig. 4b) and scatter plot of dwell times (Fig. 4c), shown in Fig. 3a and Fig. 2b, respectively; (iii) the effective diffusion constant D_E in E-space associated with the conformational sampling of the Holliday junctions varies widely from one ergodic subspace to another (Fig. 4d). If we assume that the subspace $k = 2$ with the largest D_E (Fig. 4d) has a smooth landscape ($\epsilon_{k=2} = 0$), the roughness scale for subspace $k = 4$ is $\epsilon_{k=4} = \sqrt{[\log(0.5/0.07)]} \approx 1.4 k_B T$ from $D_E = D_E^0 e^{-\beta^2 \epsilon^2}$ (refs 35,36). A comparison of Fig. 4c and d shows that the large heterogeneities of dwell times in the scatter plot are mainly due to the molecule belonging to the ergodic subspace with small D_E (for example, $k = 1, 4, 5$). The physical criterion (namely that the dynamics in each subspace be ergodic) imposed in our clustering method makes our analysis unique, providing further glimpses into the details of folding landscapes, which are masked in ensemble-average quantities $P_{ens}(E)$ and $\Sigma(t)$.

Structural origin of molecular heterogeneity. What is the structural origin of the multiple states in a Holliday junction that leads to dynamics heterogeneity? The simplicity of the Holliday junction structure allows us to infer the structural origin of the multiple states and their roles in complex dynamics. First, our calculations of the electrostatic potential and ion distribution using 100 ns molecular dynamics simulations show that Mg^{2+} ions are localized near the junction region and grooves that have high negative charge density (Fig. 5a). Second, the results from an m-fold algorithm[37] indicate that the open square form of the Holliday junction, representing the secondary structure of the Holliday junction that forms a transition state at the top of the path connecting the two isoforms at high $[Mg^{2+}]$ can have a spectrum of distinct internal multiloop topologies at the junction (Fig. 5b). In the absence of branch migration, which is ruled out
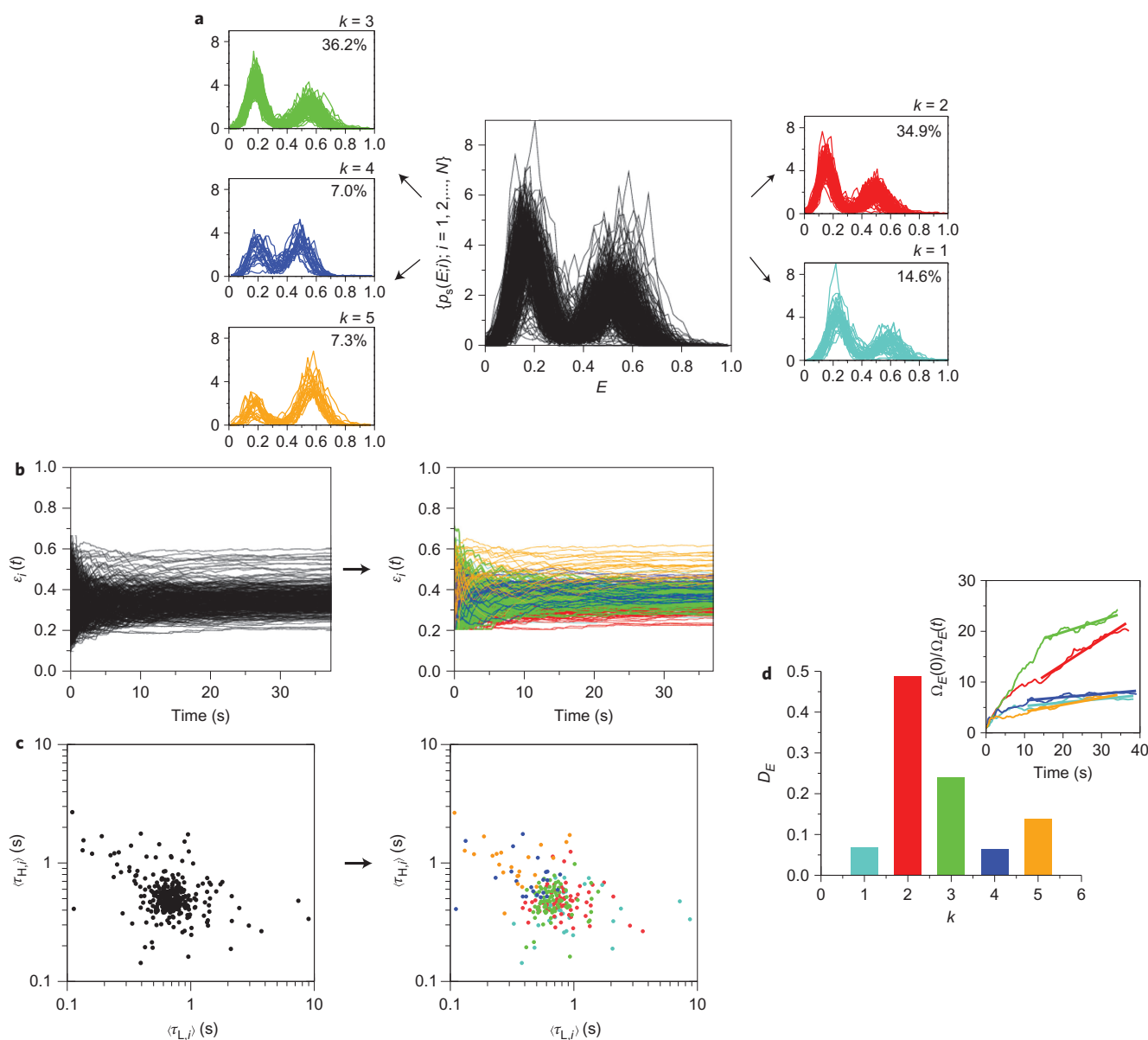
Figure 4 | Partitioning the molecules into distinct clusters. a, K-means clustering algorithm combined with the ergodic criteria partitions the set of stationary distributions $\{p_s(E;i)\}$ into five clusters for $[Mg^{2+}] = 50$ mM, and determines the list of time traces that belong to the clusters from $k = 1$ to $k = 5$. **b**, The list of time traces for each cluster determined in **a** is used to partition $\{\varepsilon_i(t)\}$ into $\{\varepsilon_i(t)|i \in k\}$ for $k = 1, 2, \ldots, 5$. **c**, D_E values calculated from the fits using $\Omega_E(0)/\Omega_E(t) \approx D_E t$ for each cluster ($k = 1, \ldots, 5$). **d**, Clustering of dwell time data as a result of the $\{p_s(E;i)\}$ clustering.

in our experiments, it is conceivable that the topology of the internal multiloop with varying sizes and flexibilities determines both the rate of Holliday junction isomerization and inter-dye distance. Variations in isomerization rate and inter-dye distance are reflected in $\{p_s(E;i)\}$, which is partitioned into five clusters. Taking these results together, we argue that the secondary structure of the Holliday junction, particularly at the internal multiloop (which mediates the conformational transition between the two isoforms), is pinned by Mg^{2+} ions. Consequently, the structural rearrangement needed for interconversion between two distinct states within a given isoform is prevented. Note that the transition state ensemble is quantized and that the actual free energy gaps in the spectrum of transition states in the presence of Mg^{2+} ions could be larger than the calculated values because the m-fold algorithm does not include the effect of the specific binding of multivalent ions (control experiments show that monovalent (Na^+) ions with high concentrations do not lead to

bimodal isomerization; see Supplementary Fig. S9). In principle, structural rearrangement could occur via transitions from high to low free energy transition states (Fig. 5b), which are quantized, thus allowing for interconversion between the five states. However, these quantized transition states (QTSs) are only transiently populated during the isomerization process, because the transition path times are considerably shorter than the time required to cross free energy barriers[38]. In other words, the lifetimes of the QTSs are very short, so that rearrangement of local bubbles at the transition state at high Mg^{2+} concentrations is unlikely. The transient nature of the QTSs implies that during multiple rounds of isomerization the local bubble structures are quenched. In addition to Fig. 3b, which shows a greater molecular heterogeneity under the high $[Mg^{2+}]$ condition, our idea of non-interconverting QTSs in the presence of Mg^{2+} ions also finds support in previous studies[29,39], which show that the obligatory open-square-form intermediates slow down DNA branch migration
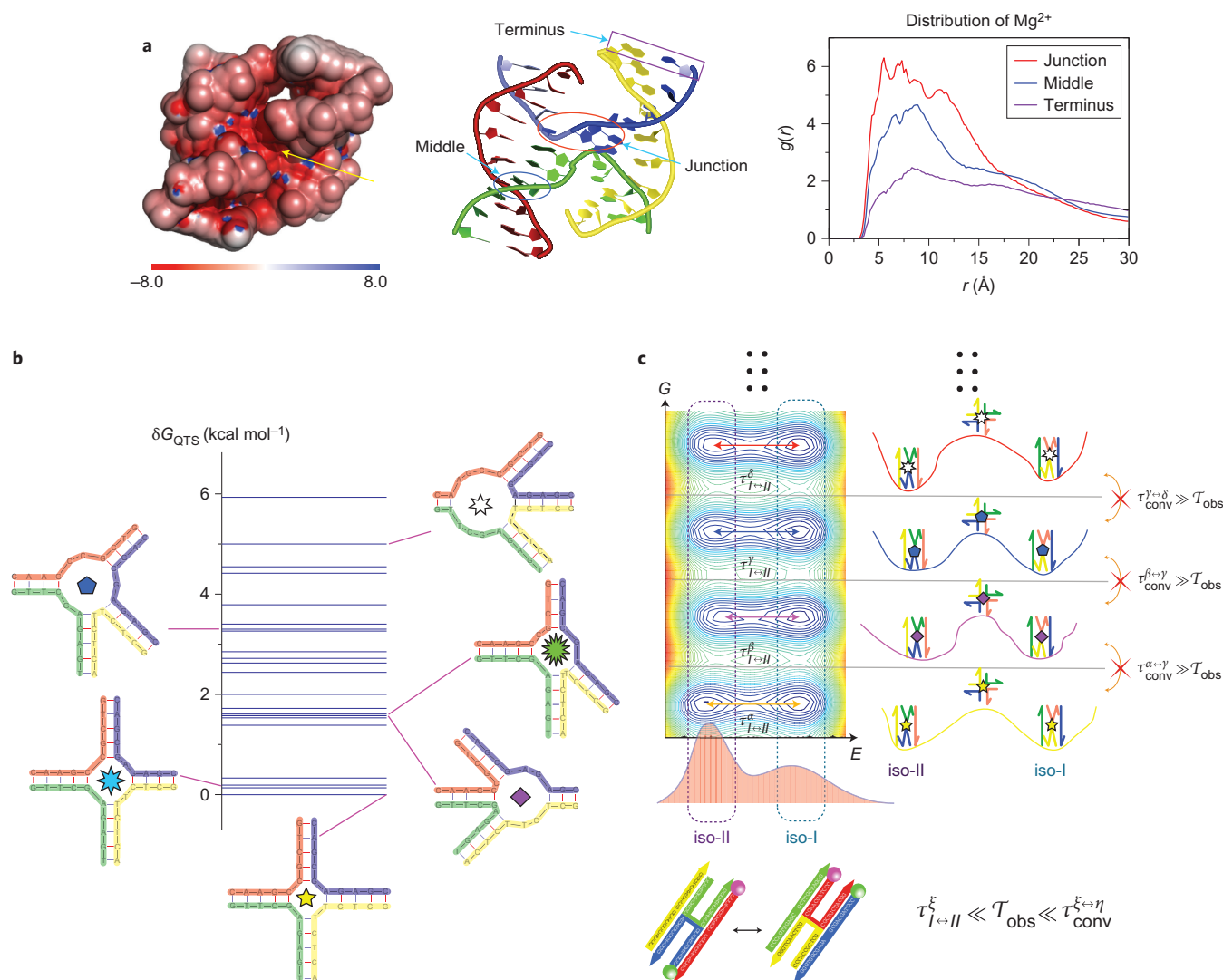
Figure 5 | Structural model to account for the origin of molecule-to-molecule variation in the Holliday junction dynamics. a, Left: electrostatic potential calculated at 200 mM monovalent ion condition using an X-ray structure of the Holliday junction (PDB code: 1DCW; Supplementary Section S6). The energy scale for the potential is in $k_B T/e$ units. Middle, right: 100 ns molecular dynamics simulation at $T = 310$ K (Supplementary Section S7), showing that Mg^{2+} ions are localized more at the junction and grooves than at the terminus. **b**, Holliday junctions with various topologies of internal multiloops, which are the putative QTSs connecting one state in iso-I and another in iso-II. **c**, Model for the dynamics of a Holliday junction constructed based on smFRET experiments and simulations. On the left are the free energy contours for various states. Two isoforms in each state are connected by a distinct open square form, the structures of which are shown in **b**. The ensemble-averaged distribution of the FRET efficiencies, $P_{ens}(E)$, is shown at the bottom. On the right, schematics of the free energy profiles are shown with cartoons of Holliday junction structures. The symbols (star, pentagon and so on) at the junction emphasize that the junction structure is intact during the isomerization process. Hence, $\tau^{\xi}_{I \leftrightarrow II}(\xi = \alpha, \beta, \ldots) \ll \mathcal{T}_{obs} \ll \tau^{\xi \leftrightarrow \eta}_{conv}(\xi, \eta = \alpha, \beta, \gamma, \ldots$ with $\xi \neq \eta$) is established.

by a factor of ~1,000 in the presence of Mg^{2+} ions, because this process requires the rupture and formation of base pairs.

The calculations summarized in Fig. 5 explain our experimental findings. In the folding landscape of the Holliday junction emerging from our analyses (Fig. 5c), transitions are only allowed between iso-I and iso-II via a band of QTSs within which the free energy gap is small enough to allow interconversion on $\mathcal{T}_{obs}$. A lack of transitions between two different states (say α and β) within a given isoform ($\mathcal{T}_{obs} \ll \tau^{\alpha \leftrightarrow \beta}_{conv}$) is explained by noting that the rupture of Mg^{2+}-stabilized base pairs is required for rearrangements from one multiloop topology to another. The conformational space connecting iso-I and iso-II is partitioned into a number of kinetically disjoint states ($\xi = \alpha, \beta, \gamma, \ldots$), reflecting the band structure of the QTS ensemble. In this sense, the persistent pattern of an smFRET trajectory is an imprint of specific disjoint states in the rugged folding landscape.

Mg^{2+} pulse annealing experiments. An immediate prediction of our model (Fig. 5c) is that interconversion between states α and β should be facilitated by an annealing protocol, enabling the release of Mg^{2+} ions from frozen internal multiloop structures. To validate this prediction we performed single-molecule experiments using a Mg^{2+} pulse sequence $[Mg^{2+}] = 50$ mM $\rightarrow$ 0 mM $\rightarrow$ 50 mM to induce transitions between multiple states (Supplementary Fig. S2). The annealing experiments confirmed that washing Mg^{2+} ions from the Holliday junction molecules indeed facilitates interconversion between trajectories with distinct patterns (compare the trajectories or two $p_s(E;i)$s shown on the side of each panel in Fig. 6a, calculated from the blue and red intervals of the trajectories corresponding to the moment before and after the Mg^{2+} pulse). We also calculated the Euclidean distance of $p_s(E;i)$ to the centroid of the five clusters in Fig. 4a before and after the Mg^{2+} pulse annealing
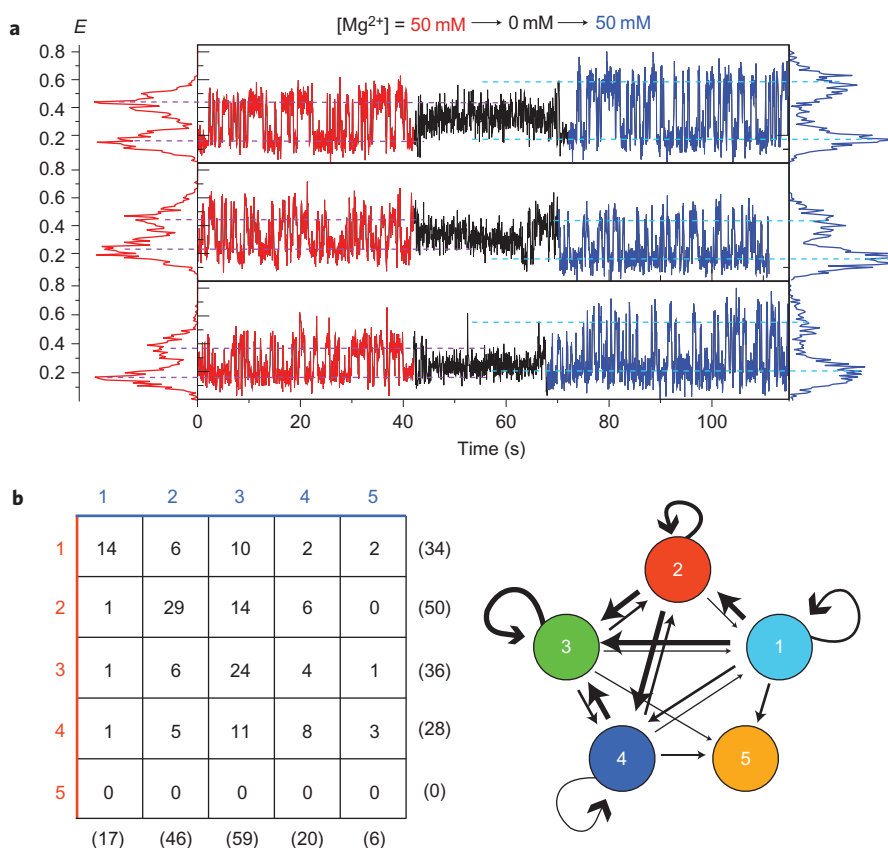
Figure 6 | Mg^{2+} pulse experiments to reset the molecular population in conformational space. a, Effect of Mg^{2+} pulse experiment. Shown are three representative trajectories that change their pattern ($E(t)$ or $p_s(E;i)$) in response to the Mg^{2+} pulse sequence. The dashed lines depict the peak positions of $p_s(E;i)$, underlying the differences in $p_s(E;i)$ before (red) and after (blue) washing off the Mg^{2+} ions. **b**, Mg^{2+} pulse-induced transition frequency matrix and diagram among kinetically disjoint states based on 148 FRET trajectories. The indices at the sides of the matrix and in the nodes denote the cluster number, $k = 1, 2, \ldots, 5$. The numbers in parentheses are the occupation number in each cluster, which can be obtained by summing the transition frequency from one cluster to another. The diagram on the right is the kinetic network describing the Holliday junction transition under the Mg^{2+} pulse. The widths of the arrows are in proportion to the number of transitions.

experiments. Here, the centroid is the arithmetic mean of $\{p_s(E;i)|i \in k\}$ with $k = 1, 2, \ldots, 5$ (Fig. 4a). The distances of $p_s(E;i)$ to the cluster means change after the Mg^{2+} pulse, which is also reflected in the reshuffling of the population of the Holliday junction molecules among the five clusters. Consequently, transitions that are prohibited on the timescale of 40 s are induced, as seen by a redistribution of the population among the five distinct clusters (Fig. 6b). Resetting the initial memory of each Holliday junction molecule is achieved by temporarily removing Mg^{2+} from the solutions, enabling conformational interconversions in the otherwise non-interconverting Holliday junction time traces. The observation of facilitated interconversion using a Mg^{2+} pulse corroborates the idea that Mg^{2+} ions and their interaction with the Holliday junction structure are the major cause of the persistent conformational heterogeneity in the Holliday junction, ruling out possible contributions to molecular heterogeneity from experimental artefacts such as the aforementioned surface immobilization[15], heterogeneous dye stacking with DNA bases[40] or chemical modifications[33]. (Supplementary Fig. S10 shows the non-responsiveness of donor-only tagged Holliday junctions to the Mg^{2+} pulse, corroborating that the Mg^{2+} ions affect the internal conformational dynamics, not the dye stacking or dye-to-surface interaction.) To recapitulate, in Holliday junctions, Mg^{2+} ions play a key role, not only in slowing down isomerization by stabilizing the native states in two isoforms (Supplementary Figs S4, S5), but also by reducing the probability of interconversion

between two different Holliday junction trajectories by enlarging the free energy gap of the QTS ensemble (Fig. 5b).

Conclusions

Single-molecule measurements provide a new means with which to probe the dynamics of biological systems, providing information not available in ensemble experiments. However, most of the current experimental studies build a distribution of observables or carry out dwell time analysis by ensemble-averaging, thus overlooking molecular heterogeneity. Our novel theoretical analysis makes use of the concept of ergodicity breaking, providing a practical framework to analyse single-molecule data and to decipher complex folding landscapes of biological systems. Only by quantitatively analysing each trajectory individually, without succumbing to the temptation to average, can the dynamical complexity of biological molecules be fully revealed. Indeed, as noted recently, quantifying and understanding the consequences of non-ergodic behaviour of RNA molecules is a major challenge[41], and the present work provides the framework for meeting it. Finally, it is worth noting that although we have used the Holliday junction as an example with which to explore quantitatively the concept of heterogeneity, our conclusions are far reaching. We expect similar behaviour in biological systems spanning spatial scales from nanometres to several micrometres (cell dimensions) and timescales from microseconds to minutes and longer.

References

1. Frieden, C. Slow transitions and hysteretic behavior in enzymes. *Annu. Rev. Biochem.* **48**, 471–489 (1979).
2. Schmid, F. X. & Blaschek, H. A. Native-like intermediate on the ribonuclease A folding pathway. *Eur. J. Biochem.* **114**, 111–117 (1981).
3. Honeycutt, J. D. & Thirumalai, D. Metastability of the folded states of globular proteins. *Proc. Natl Acad. Sci. USA* **87**, 3526–3529 (1990).
4. Anfinsen, C. B. & Scheraga, H. A. Experimental and theoretical aspects of protein folding. *Adv. Protein Chem.* **29**, 205–300 (1975).
5. Silverman, S. K., Deras, M. L., Woodson, S. A., Scaringe, S. A. & Cech, T. R. Multiple folding pathways for the P4–P6 RNA domain. *Biochemistry* **39**, 12465–12475 (2000).
6. Treiber, D. K. & Williamson, J. R. Exposing the kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.* **9**, 339–345 (1999).
7. Russell, R. *et al.* Exploring the folding landscape of a structured RNA. *Proc. Natl Acad. Sci. USA* **99**, 155–160 (2002).
8. Thirumalai, D. & Hyeon, C. RNA and protein folding: common themes and variations. *Biochemistry* **44**, 4957–4970 (2005).
9. Xie, Z., Srividya, N., Sosnick, T. R., Pan, T. & Scherer, N. F. Single-molecule studies highlight conformational heterogeneity in the early folding steps of a large ribozyme. *Proc. Natl Acad. Sci. USA* **101**, 534–539 (2004).
10. Onoa, B. *et al.* Identifying kinetic barriers to mechanical unfolding of the *T. thermophila* ribozyme. *Science* **299**, 1892–1895 (2003).
11. Zhuang, X. *et al.* Correlating structural dynamics and function in single ribozyme molecules. *Science* **296**, 1473–1476 (2002).
12. Ditzler, M. A., Rueda, D., Mo, J., Hakansson, K. & Walter, N. G. A rugged free energy landscape separates multiple functional RNA folds throughout denaturation. *Nucleic Acids Res.* **36**, 7088–7099 (2008).
13. Mickler, M. *et al.* Revealing the bifurcation in the unfolding pathways of GFP by using single-molecule experiments and simulations. *Proc. Natl Acad. Sci. USA* **104**, 20268–20273 (2007).
14. Frauenfelder, H., Parak, F. & Young, R. Conformational substates in proteins. *Annu. Rev. Biophys. Biophys. Chem.* **17**, 451–479 (1988).
15. Okumus, B., Wilson, T. J., Lilley, D. M. J. & Ha, T. Vesicle encapsulation studies reveal that single molecule ribozyme heterogeneities are instrinsic. *Biophys. J.* **87**, 2798–2806 (2004).
16. Solomatin, S. V., Greenfeld, M., Chu, S. & Herschlag, D. Multiple native states reveal persistent ruggedness of an RNA folding landscape. *Nature* **463**, 681–684 (2010).
17. Borman, S. The more the merrier: new evidence proves that RNAs can take on multiple folded forms, just as proteins do. *Chem. Eng. News* **88**, 36–37 (2010).
18. Pelkmans, L. Using cell-to-cell variability—a new era in molecular biology. *Science* **336**, 425–426 (2012).
19. Rasnik, I., McKinney, S. A. & Ha, T. Surfaces and orientations: much to FRET about? *Acc. Chem. Res.* **38**, 542–548 (2005).
20. Flomenbom, O. & Silbey, R. J. Utilizing the information content in two-state trajectories. *Proc. Natl Acad. Sci. USA* **103**, 10907–10910 (2006).
21. Li, C.-B., Yang, H. & Komatsuzaki, T. Multiscale complex network of protein conformational fluctuations in single-molecule time series. *Proc. Natl Acad. Sci. USA* **105**, 536–541 (2008).
22. Ma, S. K. *Statistical Mechanics* (World Scientific, 1985).
23. Schuler, B. & Eaton, W. A. Protein folding studied by single-molecule FRET. *Curr. Opin. Struct. Biol.* **18**, 16–26 (2008).
24. Thirumalai, D., Mountain, R. D. & Kirkpatrick, T. R. Ergodic behavior in supercooled liquids and in glasses. *Phys. Rev. A* **39**, 3563–3574 (1989).
25. Kirkpatrick, T. & Thirumalai, D. Random solution from regular density functional Hamiltonian—a static and dynamical theory for the structural glass transition. *J. Phys. A* **22**, L149–L155 (1989).
26. Tamayo, P. *et al.* Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc. Natl Acad. Sci. USA* **96**, 2907–2912 (1999).
27. Sturn, A., Quackenbush, J. & Trajanoski, Z. Genesis: cluster analysis of microarray data. *Bioinformatics* **18**, 207–208 (2002).
28. Joo, C., McKinney, S. A., Lilley, D. M. J. & Ha, T. Exploring rare conformational species and ionic effects in DNA Holliday junctions using single-molecule spectroscopy. *J. Mol. Biol.* **341**, 739–751 (2004).
29. Lushnikov, A. Y., Bogdanov, A. & Lyubchenko, Y. L. DNA recombination: Holliday junctions dynamics and branch migration. *J. Biol. Chem.* **278**, 43130–43134 (2003).
30. Duckett, D., Murchie, A. & Lilley, D. The role of metal ions in the conformation of the four-way DNA junction. *EMBO J.* **9**, 583–590 (1990).
31. Gardiner, C. W. *Handbook of Stochastic Methods* 2nd edn (Springer-Verlag, 1985).
32. Tan, E. *et al.* A four-way junction accelerates hairpin ribozyme folding via a discrete intermediate. *Proc. Natl Acad. Sci. USA* **100**, 9308–9313 (2003).
33. Greenfeld, M., Solomatin, S. V. & Herschlag, D. Removal of covalent heterogeneity reveals simple folding behavior for P4–P6 RNA. *J. Biol. Chem.* **286**, 19872–19879 (2011).
34. Thirumalai, D. & Mountain, R. Ergodic convergence properties of supercooled liquids and glasses. *Phys. Rev. A* **42**, 4574–4578 (1990).
35. Zwanzig, R. Diffusion in rough potential. *Proc. Natl Acad. Sci. USA* **85**, 2029–2030 (1988).
36. Hyeon, C. & Thirumalai, D. Can energy landscape roughness of proteins and RNA be measured by using mechanical unfolding experiments? *Proc. Natl Acad. Sci. USA* **100**, 10249–10253 (2003).
37. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415 (2003).
38. Chung, H., McHale, K., Louis, J. & Eaton, W. Single-molecule fluorescence experiments determine protein folding transition path times. *Science* **335**, 981–984 (2012).
39. Panyutin, I. G. & Hsieh, P. The kinetics of spontaneous DNA branch migration. *Proc. Natl Acad. Sci. USA* **91**, 2021–2025 (1994).
40. Iqbal, A. *et al.* Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids. *Proc. Natl Acad. Sci. USA* **105**, 11176–11181 (2008).
41. Al-Hashimi, H. & Walter, N. RNA dynamics: it is about time. *Curr. Opin. Struct. Biol.* **18**, 321–329 (2008).

Author contributions

J.L. and S.H. carried out smFRET measurements on Holliday junctions under varying Mg^{2+} concentrations and Mg^{2+} pulse. C.H. carried out the smFRET data analysis. J.Y. carried out all-atom molecular dynamics simulations to determine the radial distribution of Mg^{2+} ions around the Holliday junctions. C.H. and D.T. conceived and directed the project, and prepared the manuscript.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permission information is available online at http://www.nature.com/reprints. Correspondence and requests for materials should be addressed to C.H. and D.T.

Competing financial interests

The authors declare no competing financial interests.