Protein-DNA Clusters Explain the Non-Exponential DNA Residence Time Distributions of Transcription Factors

Zafer Kosar [1] **and Aykut Erbas** [1,✉]

[1]UNAM-National Nanotechnology Research Center and Institute of Materials Science & Nanotechnology, Bilkent University, 06800 Ankara, Turkiye

The transcription process is regulated by temporal interactions of transcription factors with DNA. In the last decade, computational and experimental studies revealed the residence times of transcription factors on DNA correlate with transcriptional output. Biochemical studies suggest that transcription factor bindings exhibit bi-exponential dynamics, often explained by the binary affinity model composed of nonspecific and specific protein-DNA interactions. Recently, transcription factor residence times were shown to display a power law *in vivo* implicating effective protein-DNA interactions controlling the dissociation kinetics are rather more complex than suggested. One contribution that can cause such continuous residence-time distributions could be higher-order protein-DNA complexes or protein coacervates. Here, by using molecular dynamics simulations of a coarse-grained polymer model for bacterial chromosomes interacting with homodimeric transcription factors at physiologically relevant concentrations, we demonstrate that residence time distributions of dimeric proteins follow a multi-exponential pattern even when a single interaction describes the affinity between DNA and protein. Our simulations reveal that this emergent behavior is due to the formation of DNA-protein clusters of various sizes at a wide range of protein concentrations and affinities. These findings add another layer to transcriptional regulation and, consequently, to gene expression by connecting transcription factor concentrations and affinities, DNA-protein clusters, and DNA residence times of transcription factors.

Residence Time Patterns | Transcription Factor Dynamics | Transcriptional Regulation

Correspondence: ***aykut.erbas@unam.bilkent.edu.tr***

Introduction

Many aspects of cellular behavior, including cellular identity and response to physiochemical and even biomechanical stimuli, are dictated by gene expression patterns. Among the many regulators of gene expression, transcription factors (TFs) play one of the most prominent roles in inhibiting or activating the transcription of target genes at the single-binding site level (Browning, 2004; Seshasayee et al., 2011; Bintu et al., 2005; Gentry et al., 2023). TFs locate their specific DNA binding sequences through random search. While 3-dimensional (3D) diffusion and intersegmental transfer of TFs are possible (Halford and Marko, 2004), the search process is driven mostly by 1-dimensional (1D) sliding along nonspecific DNA sites, suggesting that TFs remain DNA-bound for most of their lifetimes (Stracy et al., 2021).

Regulation of a gene by a TF starts at the binding site once the TF binds to its specific DNA sequence and forms a relatively stable protein-DNA complex. This complex controls the repression or activation of the gene until the protein unbinds from the DNA binding site via thermal fluctuations or collision with other molecular species (de Jonge et al., 2022; Haberle and Stark, 2018). Accordingly, recent studies in TF dynamics revealed that the residence time of a TF on its DNA binding site (i.e., its unbinding or dissociation rate with the units of inverse time) is directly intertwined with transcriptional output (Clauß et al., 2017; Lickwar et al., 2012; de Jonge et al., 2020). Thus, the factors affecting the DNA-residence time of TFs emerge as key components in the gene expression patterns.

The duration of a TF on its DNA binding site can depend on various factors, including TF's affinity to the DNA (Connaghan-Jones et al., 2007; Schaaf and Cidlowski, 2003), temperature (Bonner et al., 1992; Zhu et al., 2023), and 3D DNA structure (Kim and Shendure, 2019; Inukai et al., 2017). Studies focusing on unbinding (or dissociation) kinetics of TFs from their single-binding sites have demonstrated a strong concentration dependency of TF unbinding rates Graham et al. (2011); Joshi et al. (2012); Kamar et al. (2017). In scarce concentrations, a TF remains bound to its target site for extended periods. In contrast, the abundance of unbound TFs in solution leads to a competition for the binding sites on DNA, resulting in much higher unbinding rates (i.e., shorter residence times) via the process referred to as facilitated dissociation (FD) (Kamar et al., 2017; Koşar et al., 2022; Erbaş and Marko, 2019; Joshi et al., 2012; Graham et al., 2011; Dahlke and Sing, 2017).

The impact of TF concentration is not limited to protein unbinding dynamics via FD. TFs can also contribute to 3-dimensional (3D) genome organization and form DNA-protein clusters in a concentration-dependent manner (Kim and Shendure, 2019; Noort et al., 2004; Skoko et al., 2006, 2004; Remesh et al., 2020; Koşar

and Erbaş, 2022; Arold et al., 2010; Dame et al., 2000; Winardhi et al., 2015). Such structural effects are more pronounced with bacterial Nucleoid-Associated Proteins (NAPs), a class of DNA-binding proteins often with dual functionality (i.e., regulatory and structural roles). NAPs are involved in chromosome organization, many of which also function as transcription factors (Dillon and Dorman, 2010; Dorman, 2014). Due to their common multivalent nature (i.e., multiple DNA binding domains) Lee (1992), DNA-binding proteins can drive the bridging of distinct DNA segments and the formation of DNA-protein clusters of various shapes and sizes and other forms of chromosome architectural changes (Skoko et al., 2004; Dillon and Dorman, 2010; Wang et al., 2011; Hammel et al., 2016; Dame, 2005; Verma et al., 2019). Molecular Dynamics (MD) simulations of model bacterial systems suggest that these events are highly dependent on protein concentration and nonspecific interaction potential (i.e., the affinity between the DNA-binding protein and nonspecific DNA sites) (Koşar et al., 2022; Brackley et al., 2013; Jung et al., 2024; Dahlke and Sing, 2019).

Recent experimental studies employing single molecule tracking (SMT) demonstrated that the residence duration of several TFs and chromatin-associated proteins dynamics follow a power-law pattern (Garcia et al., 2021; Garcia, 2021). These findings support a continuum model for TF dynamics and TF-DNA interaction affinities rather than a bi-exponential model attributed to binary specific (U_{sp}) and nonspecific (U_{ns}) affinities. In the bi-exponential model, TFs are considered to have longer residence times on their specific target DNA sequence and shorter residence times on nonspecific DNA, often resulting in the utilization of a bi-exponential decay equation to explain the distribution of residence times (Chen et al., 2014; Ball et al., 2016; Morisaki et al., 2014). Contrary to this suggestion, even single affinity ($U_{ns} = U_{sp}$) and binary affinity models ($U_{ns} < U_{sp}$), in principle, can generate a non-trivial dissociation energy landscape via many-body interactions via formation of protein-rich phases through liquid-liquid phase separation (Boija et al., 2018; Hnisz et al., 2017; Sabari et al., 2018; Wang et al., 2021; Nozawa et al., 2020; Larson et al., 2017; Alberti et al., 2019).

Here, we employed a topologically accurate large DNA polymer and state-of-the-art protein models with varying levels TF-DNA affinity and protein concentrations for extensive MD simulations. Analysis of our simulations revealed apparent multi-exponential distribution patterns even with the single and binary affinity models, hinting at emergent behavior. Consequently, we explored the driving factors of multi-exponential residence behaviors of TFs. In particular, high nonspecific affinity cases, which enabled the formation of much larger clusters, exhibited distinct RT distributions and required more exponents to be matched with decay curves. Investigation of TF-DNA cluster formations of different sizes revealed that cluster dissipation means lifetimes are coupled to their sizes. Therefore, concentration and affinity-dependent cluster formations could be the driving factors of the observed multi-exponential patterns.

We then explored the distributed affinity models with uniform or normal distributions to check whether the additional complexity via mimicking the continuum model would lead to power-law behavior. However, it is clear that discrete affinity distributions with limited resolutions again form multi-exponential decay patterns, and thus, they do not provide sufficient complexity to generate a power-law behavior.

In this study, we demonstrate that TF dynamics follow multi-exponential patterns even without multiple TF-DNA affinity levels in our model. In parallel with our findings, multi-exponential fits were previously used in several SMT studies to interpret the residence behaviors (Hipp et al., 2019; Reisser et al., 2020; Agarwal et al., 2017). We also show how TF residence patterns are impacted by DNA-protein cluster formations. Our model predicts the power-law behavior might be plausible even with discrete affinities with some additional complexities or with higher resolution of affinities. These findings may help establish a more advanced understanding of transcriptional regulation and, thus, of gene expression regulation.

Results

Residence time behaviors are multi-factorial

We employed multiple cases for nonspecific interaction affinities ranging from a very weak $1kT$ to a strong $4kT$ (i.e., $U_{ns} = U_{sp}$) per bead, where $U_{sp} = 4kT$ per bead in all cases. For each of these affinity levels, four different physiologically relevant (Verma et al., 2019; Azam et al., 1999; Ball et al., 1992) concentrations ranging from $10 - 60\mu M$ of TFs were employed in the simulations. We obtained residence time patterns, as shown in Figure 1 and as described in methods, in the form of occurrence versus duration, where occurrence is the number of times a duration was achieved or observed. We then tried fitting several equations to define TF residence patterns (Figure 1E). Fit equations included exponential decay with up to five exponents (Figure 2).

The projection of RT patterns in the log-log scale exhibited an apparent arching pattern, eliminating the possibility of a good power-law fit, which produces a straight-line pattern in that scale. The single exponential decay (ED) equation can only describe short-duration ($< 10a.u.$) distributions regardless of affinity and concentration (Figure S1), implying the full-spectrum residence time behavior is not dependent on a single parameter. This behavior was indeed expected due to distinct nonspecific and specific affinity levels leading to at least a bi-exponential behavior.

For most of the cases, our simulations utilized binary values where, for each simulation, there was only one nonspecific and one specific affinity where $U_{ns} < U_{sp}$. If there was no emergent factor affecting residence times,

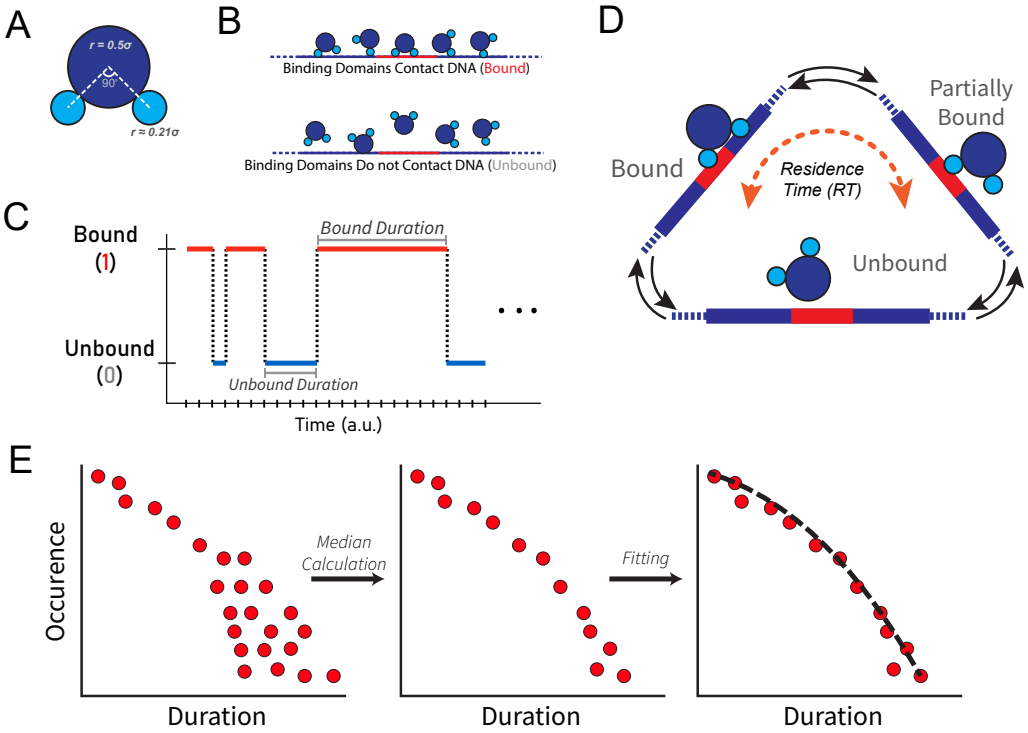


Figure 1. Graphical abstract of the study
(**A**) The coarse-grained model of a generic homodimeric transcription factor. The light blue parts represent the binding domains and the dark blue part is the hinge (non-binding) domain. Radii of the coarse-grained beads are given in Lennard-Jones (LJ) units and the angle between binding domains are shown in degrees. (**B**) Distinction between bound and unbound states. (**C**) Graphical representation of DNA-residence durations. (**D**) Possible binding states of a transcription factor and calculation of residence times (RT). A transcription factor could be fully or partially bound to DNA, or it could be in an unbound state in which it does not interact with the DNA. Note that red and dark blue regions represent specific binding sites and nonspecific binding sites of the DNA, respectively. Nevertheless, binding to either site is considered equivalent. (**E**) Collection, median calculation, and analysis of transcription factor residence patterns. Bound durations of each transcription factor throughout the simulations are collected and then minimized to their medians. Resulting patterns are analyzed by fitting several exponential equations.

a double exponential decay (DED) would be sufficient to interpret the residence time patterns obtained from the simulations. Respectively, we used a double exponential decay equation for characterizing RT patterns. DED provided much better fits compared to the ED equation but notably failed to generalize well for all durations (Figure 2A), which is intriguing considering simulations utilized a single type of TF and two distinct types of DNA sites.

The strategy employed here was to increase the number of exponents to better characterize TF residence behaviors. We gradually increased the number of exponents for the decay equation to up to five exponents. On top of the visual inspection of the fits, we quantitatively analyzed their accuracy via normalized Residual Sum of Squares (RSS), where the lower RSS indicates a better fit. The increase from single to double as well as from double to triple exponential decay resulted in the order of magnitudes lower RSS (Figure 2B). Additional increments in the number of decay exponents also reduced RSS for the fits, but the changes were not as drastic. These analyses demonstrate that RTs follow a multi-exponential decay pattern, suggesting RT patterns are shaped by additional factors besides binary DNA-protein binding affinities.

Residence time behaviors are dependent on the emergent behavior of TF concentration and binding energies

Residence time behaviors (Figure 2A, Figure S1) suggest a pattern beyond a bi-exponential system that cannot be simply explained nor attributed to the dual model of short-lived TF-DNA interactions on nonspecific sites and longer-lived, more stable interactions on specific binding sites. Therefore, an emergent behavior is required to explain such multi-exponential patterns. Specifically, patterns obtained from elevated concentration and high-affinity (kT) cases require more exponents (Figure2A, Figure S1), indicating such a behavior emerges as a result of concentration and energy levels. The relatively high binding energies and concentration of DNA-binding proteins were shown to lead to cluster formations as well as local and global condensations of the chromosome (Koşar et al., 2022; Lin et al., 2012; Brackley et al., 2013; Agback et al., 1998). Unsurprisingly, higher concentrations yield larger clusters (Figure 3C,E). Remarkably, global compaction of the chromosome and cluster sizes, even conformations, are mainly regulated by nonspecific interactions, and specific interactions have little effect (Koşar et al., 2022). This behavior can be attributed to the abundance of nonspecific DNA sites over specific sites. In this work, we reduced the impact of global chromosome compaction via minia-

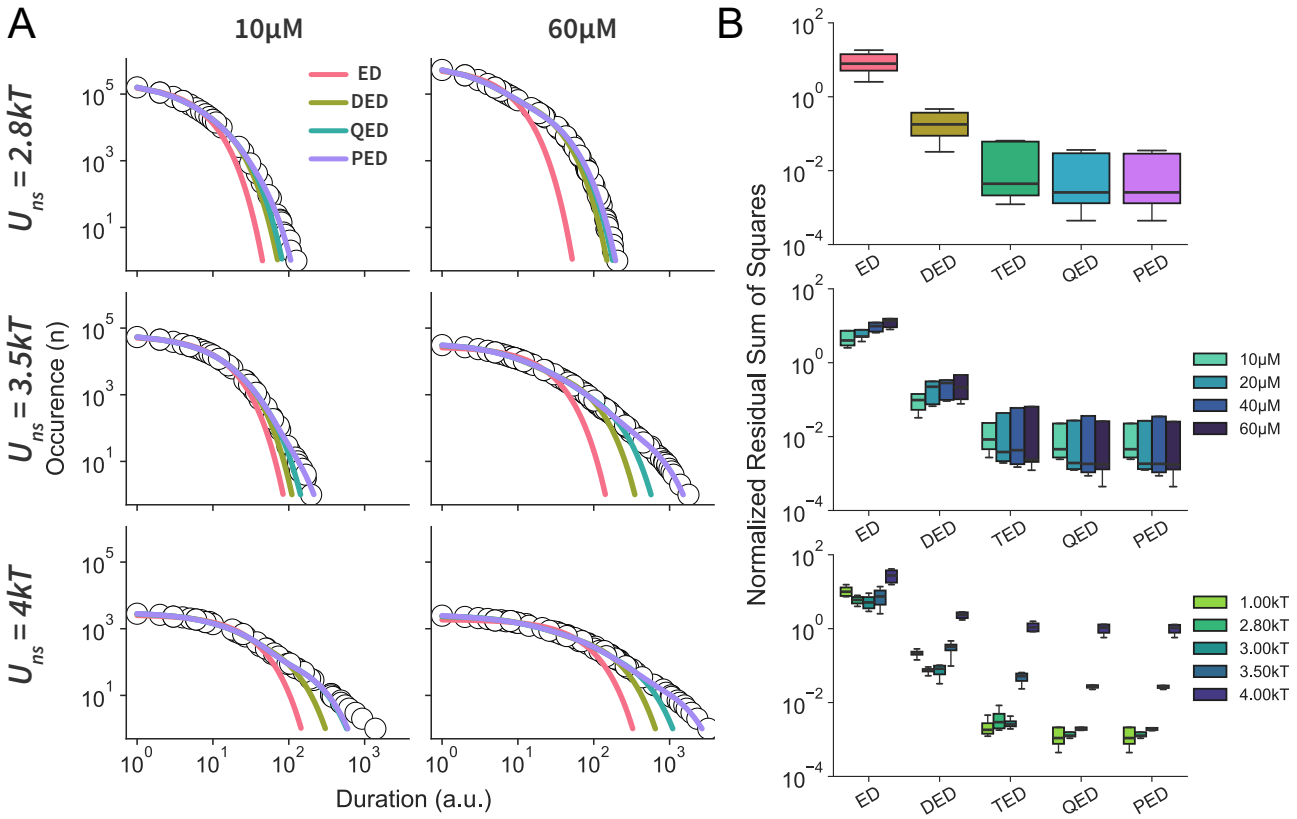


Figure 2. Analysis of transcription factor residence time patterns
(**A**) Distribution of transcription factor residence times in arbitrary unit (a.u.) at $10\mu M$ and $60\mu M$ concentrations for distinct U_{ns} levels and the fits of Exponential Decay (ED), Double Exponential Decay (DED), Quadruple Exponential Decay (QED), and Pentuple Exponential Decay (PED) equations to residence time data. Due to the extensive number of data points, only representative 20% are shown here. (**B**) Normalized Residuals Sum of Squares (RSS) of the fits depicting the scale of difference between observation and the fits. For all equations, overall RSS, RSS by concentrations, and by binding affinities are shown, respectively. Note that specific binding energies are $4kT$ for all the systems.

turizing TF binding domains (Figure 1A), which minimizes the bridging of multiple DNA segments. Moreover, the use of adequate binding energies eliminated the possibility of high chromosomal compaction even at high concentrations (Figure 3A). This strategy allowed TFs to roam relatively freely within the cellular confinement.

Our model depicts TFs residing for increasingly higher durations on DNA instead of freely roaming around with the increasing nonspecific potential with a fixed specific binding energy (Figure 3A,B). Consistent with the previous findings, at a very low nonspecific potential ($U_{ns} = 1kT$), bound proteins are sparsely distributed around the DNA polymer, mainly near binding sites (Figure 3A), but do not form DNA-polymer complexes (i.e., clusters) (Figure 3C). Relatively higher nonspecific potentials ($U_{ns} = 2.8kT$ and $U_{ns} = 3kT$) led to primarily small and globular clusters. In contrast, stronger nonspecific affinities ($U_{ns} = 3.5kT$ and $U_{ns} = 4kT$) enabled the formation of much larger clusters (Figure 3D) with filamentous conformation (Figure 3C). Therefore, we hypothesized the cluster sizes and conformations to play significant roles in TF residence patterns.

Cluster formations drive multi-exponential residence time patterns

To unravel the relationship between DNA-protein clusters and multi-exponent patterns, we initially considered tracking and collecting residence times of TFs for each individual cluster, as described in Figure 1. However, clusters are dynamic formations, and it is not rational to follow TFs that were initially part of a cluster because after they dissipate, they are free to bind anywhere on DNA. Therefore, we rather investigated the dissipation mean lifetimes of individual clusters (see methods). As opposed to RT pattern acquisitions, partial unbinding events were also counted. This modification was needed simply because at high nonspecific affinities, the time needed for obtaining cluster decay rates would exceed simulation lifetimes when only full unbinding events were counted.

At relatively low nonspecific energies ($U_{ns} = 2.8kT$ and $U_{ns} = 3kT$), correlation analysis of cluster size and their dissipation durations did not reveal any relation. Contrarily, higher nonspecific interaction potentials ($U_{ns} = 3.5kT$ and $U_{ns} = 4kT$) led to correlations with a coefficient of $r \geq 0.5$ (Figure 4), where larger clusters had longer mean lifetimes. These imply differently sized clusters could dissipate at diverse rates, contributing to multi-exponential behaviors that are more apparent for

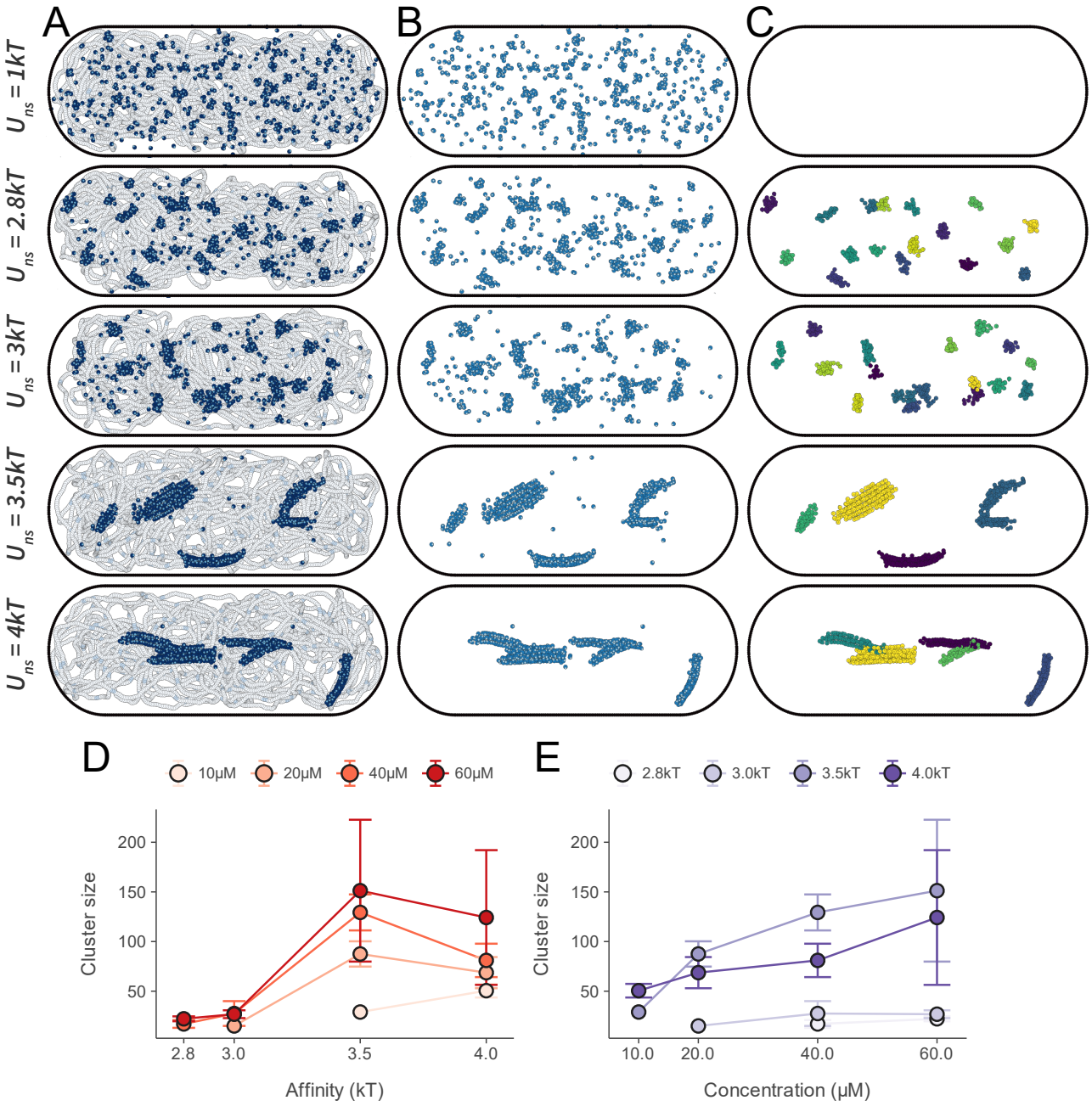


Figure 3. Visualizations of the system, protein distributions, and cluster formations of the coarse-grained bacterium model at various nonspecific binding energies

(**A**) Overview of the system DNA (white-light blue), and transcription factors. (**B**) Transcription factor distributions within the confinement. (**C**) Cluster formations of transcription factors. Here, the coloring only serves to distinguish distinct protein clusters. Snapshots were obtained from systems with transcription factor concentrations of $60\mu M$. Specific binding energies are $4kT$ for all the systems. (**D**) Distribution of cluster sizes (number of transcription factors) by nonspecific TF-DNA affinities or (**E**) by TF concentrations.

high nonspecific affinity cases (Figure 2A).

It should also be noted that at $2.8kT$ and $3kT$ nonspecific affinities, cluster sizes ranged from 20 to 45 and 20 to 70 TFs, respectively (Figure 4B). The sizes were significantly improved with higher nonspecific interactions. The cases of $3.5kT$ and $4kT$ nonspecific affinities allowed the formation of clusters of size ranging up to 250 TFs (Figure 4B). This increase, of course, provides a statistical advantage for the analysis of the relation between lifetimes and cluster sizes.

TF residence times depend on their position relative to clusters

Upon establishing the relationship between cluster sizes and their lifetimes, we considered analyzing the residence times of TFs that are located at different regions of the clusters. Cluster-associated TFs can be classified as surface TFs and core TFs (Figure 5B). While the core TFs are rather trapped within the cluster, surface TFs are more exposed and could be expected to have shorter residence durations as they are free to unbind. Additionally, TFs that are not part of any cluster

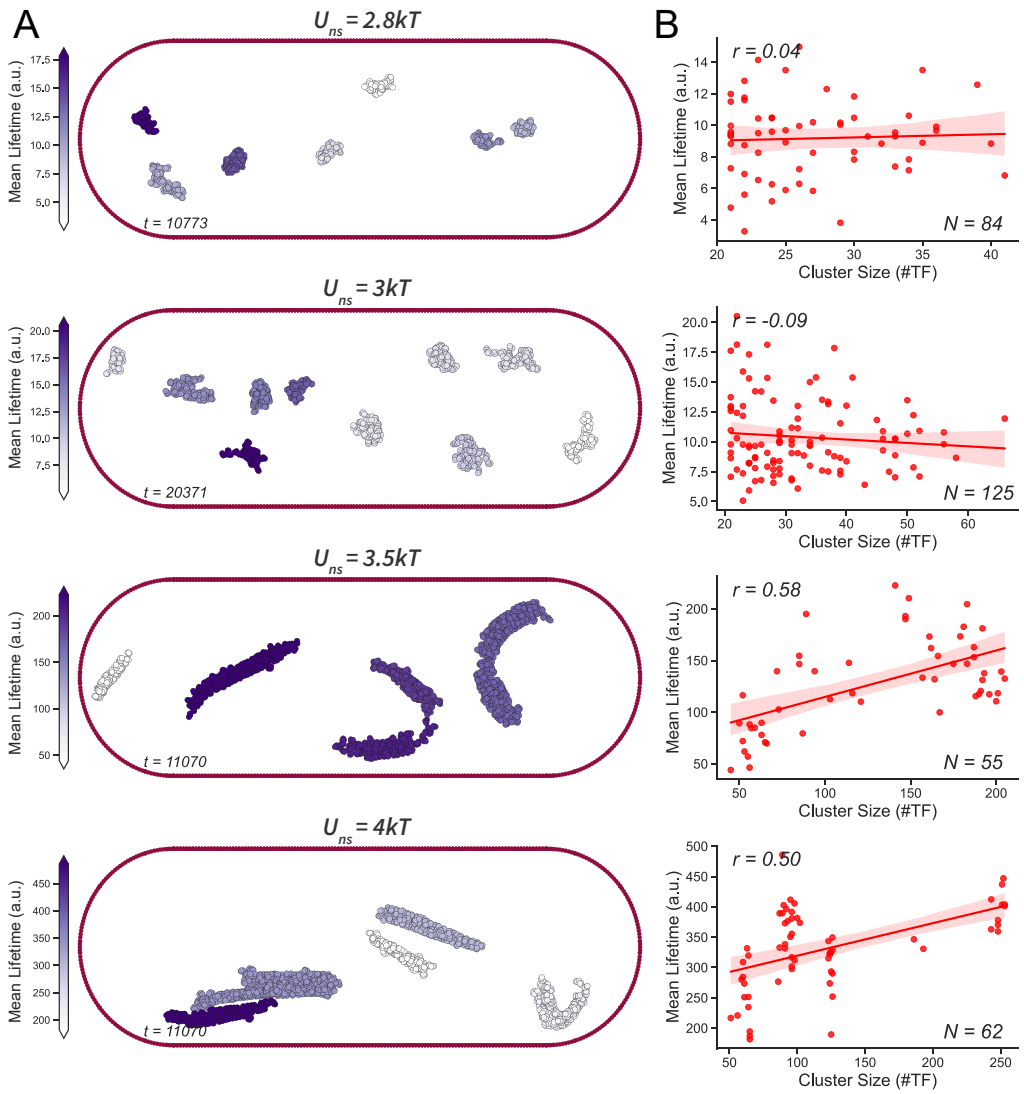


Figure 4. Correlation analysis of cluster sizes and dissipation times
(**A**) Snapshots of the transcription factor clusters at given nonspecific energies from random timesteps. Darker colors implicate a higher mean lifetime, as indicated by the color bars on the left. (**B**) Mean lifetimes in arbitrary units versus cluster sizes in the number of transcription factors they are composed of. Regression lines are used to determine correlations. Pearson correlation coefficients (r) are given on the upper-left side, and a number of the clusters in the analysis is given on the bottom-right side. Semi-transparent bands around the regression lines depict 95% confidence intervals. Transcription factor concentrations were $60\mu M$ for all cases. The threshold was set to a minimum of 20 TFs for cluster size for decay rate analysis for a more reliable statistical approach. 12 time points were used for sampling and the specific binding energy was $4kT$ for all simulations.

are classified as free TFs (Figure 5B), which are useful as references against cluster-associated TFs.

Visual investigation of individual proteins located at different positions with respect to the TF-DNA clusters did not show apparent distinction in their residence times (Figure 5A). However, in an analytic (i.e., Welch's *t-test*) and more comprehensive approach utilizing multiple time points, core TFs exhibited significantly longer residence times compared to surface and free TFs (Figure 5C). However, this behavior was limited to the high nonspecific affinity cases and was not observed at lower nonspecific potentials, which may add another layer of complexity to residence behaviors and, therefore, could explain the more prominent multi-exponential residence patterns at such energy levels. At $U_{ns} = 3.5kT$ case, surface TFs also had significantly higher residence times in comparison to free TFs (Figure 5C). This disparity indeed reveals that even located at the exposed regions of clusters, TFs could behave differently compared to freely roaming TFs. Thus, the variations among core, surface, and free TFs residence times could shape the overall residence distributions (Figure 5C-D) and further contribute to multi-exponential DNA-residence behaviors of TFs at single and binary affinity models.

Another factor that should be considered is the differing number of TFs of each type for distinct cases. Even though the concentration ($60\mu M$) and specific binding potentials ($U_{sp} = 4kT$) are fixed among the cases, the higher nonspecific energies drive clustering rather than scattering (Figure 3B), also leading to the formation of much bigger clusters (Figure 3C-E, Figure 4). Thus, more of the TFs are located within the clusters for high nonspecific affinity cases (Figure 5D). Moreover, the

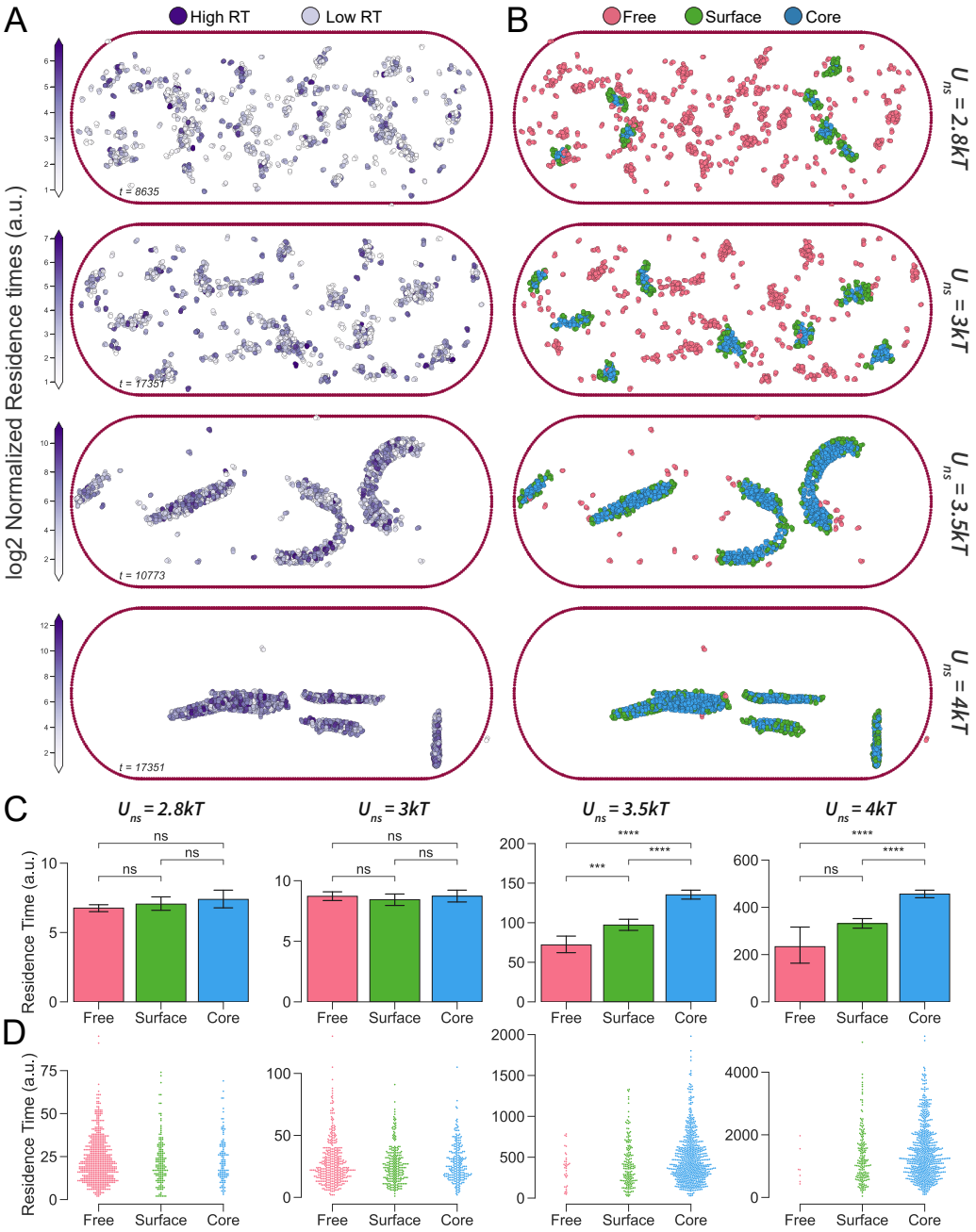


Figure 5. Residence time analysis based on the transcription factor localization with respect to clusters
(**A**) log2 normalized residence times of individual transcription factors at various nonspecific DNA-TF interaction affinities at $60\mu M$ TF concentration (**B**) Classification of TFs according to their positions with respect to the clusters. 'Free' transcription factors do not belong to any clusters. Transcription factors on the 'Surface' and in the 'Core' of the clusters were named accordingly. (**C**) Differences among residence times for free, surface, and core transcription factor types at given nonspecific affinities. Error bars represent 95% confidence intervals. Welch's *t-test* was used for significance analysis. (**D**) Distributions of TF residence times for the given transcription factor types at $60\mu M$ cases. 12 random timesteps were used for sampling, and the specific binding energy of $4kT$ was used for all simulations. Representative 12% of the data is shown to prevent over-crowding in swarm plots.

larger the clusters, the higher the core TF ratios were (Figure 5D), as expected due to the decrease in the surface-to-volume ratio.

Distributed affinity models do not lead to power-law behavior

Our MD simulations revealed that even single and binary affinity models could lead to multi-exponential residence time patterns. The next step was to mimic the continuum affinity model with a distributed affinity model, where we employed 13 distinct DNA site types and assigned them a range of DNA binding affinities. These affinities inclusively ranged between $1-4kT$ and $2-5kT$ with an increment of $0.25kT$ and the number of DNA sites for the given affinity assigned to provide either a normal (i.e., Gaussian) or uniform distribution.

Contrary to our expectations, distributed affinity models did not produce power-law behavior. Instead, the triple-exponential fit, an example of multi-exponential fit, better described the behaviors (Figure 6). Therefore,

we might speculate that mimicking the continuum model via discrete affinity distribution models is not sufficient for power-law behavior, and such behavior would require much higher complexity and resolution than our current systems for MD simulations could provide. On the other hand, the extraordinary complexity of biological systems or continuous affinity levels with higher resolutions (higher number of distinct affinity levels) could be sufficient to generate such behavior and might be the underlying reason behind the power-law behavior in the experimental approaches.

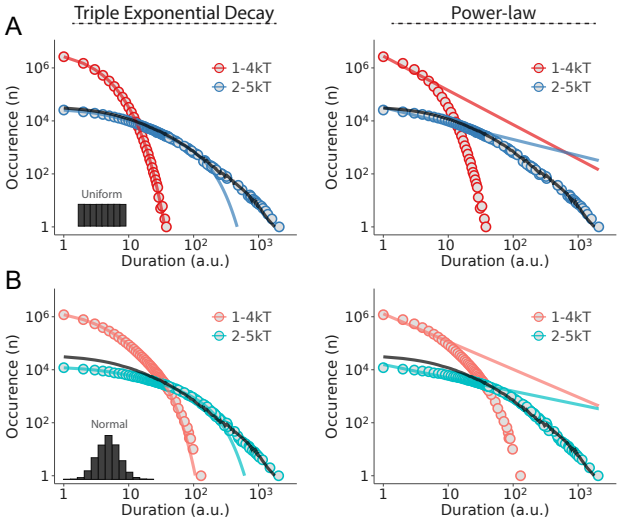


Figure 6. Transcription factor residence time behaviors at distributed affinities
(**A**) Uniform distribution and (**B**) Normal (Gaussian) distribution affinity models. Pattern analysis via fitting of triple exponential decay (left column) and power-law (right column) respectively. Example histograms of the uniform and the Gaussian distribution are shown in the bottom-left corner of each panel's first graph. Transcription factor concentrations were $60\mu M$ for all cases. Black lines are shown as a reference from the nonspecific affinity of $3.5kT$.

Discussion

The durations, therefore the distributions of durations of TFs on DNA have a significant impact on transcriptional regulation (Clauß et al., 2017). Using extensive MD simulations with a large-scale DNA polymer and state-of-the-art transcription factor coarse-grained models enclosed within rod-like confinement, we investigated the residence time distributions of the homodimeric generic TFs. In our investigations, we employed several biologically relevant TF concentrations (Ishihama et al., 2014) and distinct binding strengths between TFs and DNA. Our findings reveal an apparent multi-exponential behavior for all of the single, binary, and distributed affinity models of TF-DNA bindings. Single and binary affinities leading to multi-exponential patterns were beyond the prior suggestion of bi-exponential behavior due to specific and nonspecific binding sites. Here, we demonstrate how these multi-exponential behaviors are shaped by TF-DNA complexes, reveal the driving factors of such formations, and connect the residence times behavior to nonspecific binding affinities.

Within the context of their primary roles, TFs can inhibit and activate transcription. The formation of stable TF-DNA complexes extends the duration of a TF's residence time on its target DNA site. That, in turn, could significantly enhance their inhibitory or activatory effects. We may also speculate that larger and more compact clusters may decrease the accessibility by RNA polymerase, such as in heterochromatin-like domains (Amemiya et al., 2022), effectively reducing the transcriptional output. Contrarily, similar protein-DNA complexes, but in the form of transcriptional machinery, could initiate or enhance transcription.

The affinity of a transcription factor to its specific binding site is often stronger compared to its binding affinity towards a nonspecific DNA sequence. However, nonspecific DNA-TF interactions can contribute to the global chromosome organization due to their abundance. Collectively, specific and nonspecific interactions can lead to the formation of local DNA-protein complexes (i.e., clusters) (Koşar et al., 2022; Lin et al., 2012; Brackley et al., 2013; Agback et al., 1998). Moreover, these nucleic acid-protein complexes could significantly affect transcriptional regulation. For instance, transcription factories are multiprotein complexes formed by TFs, RNA polymerase, coactivators, etc. (Iborra et al., 1996; Cook, 2010; Melnik et al., 2011; Mitchell and Fraser, 2008). Therefore, multiprotein complexes or clusters are crucial regulators of transcriptional output. Furthermore, eukaryotic euchromatin regions, which are not densely packed as heterochromatin regions, are more accessible to transcription factors and thus more transcriptionally active (Amemiya et al., 2022; Elgin and Grewal, 2003; Penagos-Puig and Furlan-Magaril, 2020), emphasizing the significance of local DNA organization on residence times. In other words, the contribution of DNA-binding proteins to gene regulation is not restricted to their primary functions as TFs; their activity in domain-specific genome organization is also crucial.

Even though we demonstrated cluster size is an important factor for residence times, we also wanted to reveal any possible relationship between cluster shape (conformation) and dissipation rates. However, investigation of the conformations of the cluster formations on residence times was inconclusive due to cluster sizes being highly intertwined with nonspecific affinity levels. Thus, establishing a clear relation between cluster shape and RT requires eliminating affinity as a variable. In turn, that would require a separate MD model. Furthermore, clusters are not fixed formations; they are highly dynamic. Their sizes, conformations, orientations, and even the DNA segments they interact with change over time in the system. Moreover, clusters can merge and split, making it rather complicated to track individual clusters and make accurate analyses (Wang et al., 2011).

The utilization of dimeric TF models might have contributed to the multi-exponential behavior in two ways. First, our definition of unbound TF requires both binding

domains not to be in touch with any DNA molecule (Figure 1B). In other words, partial bindings to DNA (i.e., bindings with only one binding domain) are equivalent to full bindings in terms of being considered bound. The difference in stability of partial binding and full binding may indeed lead to different residence durations contributing to multi-exponential behavior. Moreover, divalent interactions are significantly more prone to facilitated dissociation (Kamar et al., 2017; Chen et al., 2018). Since FD is pronounced at high concentrations (Koşar et al., 2022), multi-exponential behavior could be partially attributed to FD as it can facilitate the unbinding of the exposed TFs (i.e., free or surface) more than the unexposed TFs (i.e., core).

Noise in gene expression is an important factor driving cellular heterogeneity (Liu et al., 2019). Gene expression noise could lead different cells in a homogeneous population to distinct phenotypes even under the same environmental conditions (Raser and O'Shea, 2005). This resulting cellular heterogeneity may provide evolutionary advantages to cells. Notably, the primary driving event of gene expression noise is considered to be TF binding (Parab et al., 2022). Also, infrequent or rare biochemical processes contribute to noise in gene expression (Raser and O'Shea, 2005). Such infrequency or lower occurrence could be seen for the higher duration residences in our systems, which are more apparent at high nonspecific affinity and high protein concentration cases. Therefore, such cases could lead to higher gene expression noise and yield a more heterogeneous population of cells. It should also be noted that longer residence times are more likely to cause transcriptional bursts (Raser and O'Shea, 2005), which is another notion that contributes to noise in gene expression. On the other hand, shorter DNA residence times are suggested to lower gene expression noise (Azpeitia and Wagner, 2020). Our study sheds light on DNA residence time distributions and may explain the diverse noise levels in gene expression responses.

One of the prominent features of our previous coarse-grained model was the global chromosome organization. Although that system allowed the modeling of the role of nonspecific affinities, selected bead size for TF binding domains led to multiple interactions facilitating the chromosomal collapse. Extreme compaction of the chromosome and resulting residence times of TFs would exceed simulation durations, making it unlikely to obtain RT patterns. By lowering the bead size of TFs, we minimized multiple binding and over-compaction. That also enabled the decoupling of 3D genome organization from RT patterns, increasing the control over other variables.

Similar to any MD study, this work has some limitations. Most prominently, the simplification of the cell for coarse-grained MD simulations removes most of the complexity possessed by the actual cellular systems. Additionally, we used a single type of TF for all MD simulations with discrete DNA binding affinities. In combination, reduced complexity may explain the lack of observation of power-law behavior, which was attributed to continuous TF-DNA affinities in experimental studies. Of course, the cellular complexity is also persistent in the experimental setup. Therefore, other factors, such as fluctuating TF levels and dynamic chromosomal landscape, might be crucial for power-law behavior. However, in our simulations TFs, even with discrete affinities, exhibited multi-exponential behaviors due to emergence, suggesting a power-law pattern is highly possible with some additional complexity even without continuous affinity levels. While this notion does not dismiss the continuum model, it provides an alternative interpretation for the observed power-law behavior.

On the other hand, MD simulations provide distinct advantages. Most prominently, experiments have to rely on approximations and assumptions due to current technological limitations such as indirect measurement techniques. Contrarily, simulations allow direct measurements, providing much higher precision and accuracy. Indeed, the lack of direct measurement necessitates a subsequent correction step, which can introduce biases and lead to inaccurate interpretations of the system or behavior. Correction of the observation data is often not required for MD simulations, resulting in bias-free or less-biased measurements. In fact, we might speculate that the power-law interpretation of residence times distributions in experimental studies might well be due to such biases introduced by corrections rather than the actual underlying TF residence behavior.

In this study, we demonstrate that TFs follow multi-exponential patterns with discrete affinities between DNA and TFs in our MD model. We demonstrate how binding affinity and concentration of TFs dictate these behaviors. There are notable implications of this behavior on biological systems such as cellular heterogeneity and noise in gene expression. Isolated from cellular complexity and continuum model, model homodimeric TFs exhibit multi-exponential patterns even with single and binary affinities. This type of behavior might be one of the underlying reasons for gene expression noise and consequent cellular heterogeneity and contribute to cellular differentiation and help overcome evolutionary bottlenecks. Moreover, the distributions of TF-DNA residence times may help explain discrete transcriptional bursts. Lastly, DNA-protein clusters in bacterial chromosomes could drive Topologically Associated Domain-like domain formations and further affect the regulation of gene expression. Overall, our findings on the TF residence time behaviors and their driving factors might contribute to a more comprehensive understanding of gene expression and regulation.

Methods

Establishing the system

We used a modified version of our previous coarse-grained model bacterial system mimicking an *Es-*

cherichia coli (*E. coli*) bacterium. The model system includes an adequately relaxed chromosome with uniformly distributed binding sites confined within cell wall-like boundaries resembling that of a rod-shaped bacterium. Generic homodimeric transcription factors with various concentrations $(10-60\mu M)$ were placed randomly in the volume created by confinement.

Modelling the DNA polymer

For coarse-grained modeling, we used 10 base pairs (bp) $\approx$ 1 bead "Kremer-Grest" (KG) model for DNA where 1 bead has a diameter of a σ in Lennard-Jones (LJ) units corresponding to $\sim 3.4nm$. We set the persistent length to 15 beads corresponding to $\sim 50nm$ consistent with a double helix DNA molecule. We established an $N = 12000$ KG bead DNA model where we maximized the number of binding sites and left sufficient spacing for DNA segmental flexibility to minimize the impacts by bridging the binding sites. The initial circular DNA structure was compacted using self-attractive forces to reduce its size to fit within the available volume, similar to our previous work (Koşar et al., 2022). The binding sites are composed of three beads (as opposed to two beads in our previous work), increasing the likelihood of maximum interaction with the TFs. 150 binding sites are placed uniformly along DNA with 80-bead spacing.

Constructing the confinement

To mimic the rod-like shape of the bacterial cell wall, we first built an open-ended cylinder (or simply a pipe) with a diameter of $R = 2r$ and a height of $2R$. Then, two semi-spheres with radiuses of r are used as caps to close the endings of the cylinder. The overall structure has a volume of $2/3 \times \pi \times R^3$. The radius (r) was set to provide a 1% DNA volume fraction to match that of *E. coli*. Therefore, r corresponds to $\sim 30\sigma$ in LJ units $(\sim 100nm)$ for a DNA polymer length of $N = 12000$. The fixed beads of the cell wall are placed dense enough to provide effective boundaries for DNA and proteins. Nevertheless, there were an insignificant number of TF leaks $(< 2\%)$ for extended simulations.

Designing the transcription factors

We designed a generic model of homodimeric TFs in which a TF has two identical binding domains and a hinge domain with no affinity to DNA. Binding domains were placed around the hinge domains with $90°$ angles in a semi-flexible fashion with 12 beads of persistent length. Radii were set to 0.5σ and 0.21σ for hinge and binding domains, respectively (Figure 1A). We used relatively small binding domains to minimize multiple interactions by single binding domains and prevent over-condensation of the nucleoid. The exact sizes and angles were used to ensure the well-fittings of dimeric TFs into three-bead binding sites. We employed four different concentration levels (i.e.,$10\mu M$, $20\mu M$, $40\mu M$, $60\mu M$). The corresponding number of TFs for the given concentrations was calculated using the volume provided by the confinement. TFs were distributed into the volume at random coordinates.

Modelling the TF-DNA affinities

We used a fixed specific interaction potential of $U_{sp} = 4kT$ per bead, energy high enough for robust binding and also low enough to allow unbinding in the time frame of our simulations, enabling us to extract residence time patterns. Varying nonspecific interaction potentials in the range of $1-4kT$ allowed tracking of residence times at distinct local compaction levels and diverging cluster sizes, as well as differing protein distribution over DNA. For the distributed affinity cases, there was no specific binding potential. Instead, affinities inclusively followed the given ranges of $1-4kT$ and $2-5kT$ with the increment of $0.25kT$, resulting in 13 distinct affinity levels throughout DNA to ensure either Gaussian or uniform distributions.

Retrieving the residence durations and generating the distributions

Transcription factors are considered bound under the condition that at least one of two binding sites is in direct contact with the DNA polymer (Figure 1B). For each time point, each protein is marked either bound or unbound (Figure 1C). Then, the duration of each uninterrupted bound state was calculated (Figure 1D). Each time a particular residence duration was encountered, the corresponding occurrence was incremented by 1. We then utilized pooled occurrences, containing the number of occurrences for each possible duration $(1-tmax)$, for analyzing residence patterns of transcription factors from simulations with distinct parameters. Obtained data is then minimized to ensure one occurrence value has only one corresponding duration by taking the median of the durations (Figure 1E). This step was necessary for equation fits and visualizations.

Fitting equations to the residence time distributions

We used several equations to interpret the behavior of transcription factors. Initially, we used single exponential decay $c \times \exp(-t \times k)$ and power-law $c \times t^{-k}$ where c stands for coefficient, t for duration in arbitrary units, k for decay rate (i.e., τ^{-1}). Due to sub-optimal fitting with these equations, we included double, triple, quadruple, and pentuple exponential decays in the form of

$$c_1 \times \exp(-t \times k_1) + \cdots + c_n \times \exp(-t \times k_n).$$

The resulting fits were then graphed and used for calculating the normalized Residual Sum of Squares (RSS). Normalization of RSS values was achieved in two steps. First, occurrence counts are divided by the maximum of occurrence values to normalize occurrences, preventing a higher number of events leading to higher RSS.

$$y_d^{\text{norm}} = \frac{y_d}{\max(y)}$$

Where y_d^{norm} is the normalized occurrence for duration d, y_d is the occurrence value for duration d, $\max(y)$ is the maximum occurrence value in the set of all occurrences.
Then, RSS is divided by the ratio of duration points with non-zero occurrences to prevent more observation points leading to higher total RSS.

$$\text{RSS}_{\text{norm}} = \frac{\text{RSS}}{D_{\text{nonzero}}/D_{\text{total}}}$$

Where, RSS_{norm} is the normalized residual sum of squares, RSS is the residual sum of squares, D_{total} is the total number of duration points, D_{nonzero} is the number of duration points where occurrences are not zero.

Clustering the transcription factors

Transcription factors within the threshold distance of each other are accepted to be part of the same cluster. We have used our own deterministic, non-centralized, and vectorized clustering algorithm. We set the threshold distance to 2.1σ in LJ units in agreement with visual inspections. For the overview of the system, the minimum number of transcription factors to form a cluster was 12, and for decay analysis, it was 20, ensuring a more accurate estimation of mean lifetimes.

Cluster analyses

Cluster analyses included size, conformation, surface analysis, and decay rates. The size of a cluster is simply the number of transcription factors within that particular cluster. To classify the conformation (or shape) of a cluster, we first determined three possible formations, namely filamentous, globular, and semi-filamentous. We formed Rg tensor for each cluster and evaluated their eigenvalues. To distinguish surface and core cluster proteins, we employed the Convex-Hull algorithm included in the *SciPy* library. Decay rates (reverse of mean lifetimes) for clusters were calculated by fitting a single exponential decay equation to $N(t)/N_0$. Decay rates require statistically meaningful numbers for reliable analysis of lifetimes, which led us to select 20 as the threshold for minimum cluster size. We then used linear regression for correlation analysis between cluster dissipation mean lifetimes and their sizes.

Code Availability

The codes and the processed data necessary to reproduce the analyses and the graphs are available at https://github.com/Zaf4/residence2.

Acknowledgements

ZK thanks Alper Genceroglu, Irem Niran Çagıl, Leyla Yalçınkaya, Tutku Muratoglu, and Fatma Chafra for their review of the initial manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

ZK designed the study, created the models, ran the simulations, performed the data analysis and visualizations, constructed the figures, and wrote the manuscript. AE conceptualized and supervised the project, and revised the manuscript.

References

Agarwal, H., Reisser, M., Wortmann, C., and Gebhardt, J. C. M. (2017). Direct observation of cell-cycle-dependent interactions between ctcf and chromatin. *Biophysical Journal*, 112. doi: 10.1016/j.bpj.2017.04.018.

Agback, P., Baumann, H., Knapp, S., Ladenstein, R., and Härd, T. (1998). Architecture of nonspecific protein–dna interactions in the sso7d–dna complex. *Nature Structural Biology 1998 5:7*, 5:579–584. doi: 10.1038/836.

Alberti, S., Gladfelter, A., and Mittag, T. (2019). Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell*, 176(3):419–434. doi: 10.1016/j.cell.2018.12.035.

Amemiya, H. M., Goss, T. J., Nye, T. M., Hurto, R. L., Simmons, L. A., and Freddolino, P. L. (2022). Distinct heterochromatin-like domains promote transcriptional memory and silence parasitic genetic elements in bacteria. *The EMBO Journal*, 41:e108708. doi: 10.15252/EMBJ.2021108708.

Arold, S., Parkinson, G., Ladbury, J., and Leonard, P. (2010). H-ns forms a superhelical protein scaffold for dna condensation. *Proc. Natl Acad. Sci. USA*, 107:15728–15732. doi: 10.1073/pnas.1006966107.

Azam, T. A., Iwata, A., Nishimura, A., Ueda, S., and Ishihama, A. (1999). Growth phase-dependent variation in protein composition of the escherichia coli nucleoid. *Journal of Bacteriology*, 181:6361–6370. doi: 10.1128/jb.181.20.6361-6370.1999.

Azpeitia, E. and Wagner, A. (2020). Short residence times of dna-bound transcription factors can reduce gene expression noise and increase the transmission of information in a gene regulation system. *Frontiers in Molecular Biosciences*, 0:67. doi: 10.3389/FMOLB.2020.00067.

Ball, C. A., Osuna, R., Ferguson, K. C., and Johnson, R. C. (1992). Dramatic changes in fis levels upon nutrient upshift in escherichia coli. *Journal of Bacteriology*, 174:8043–8056. doi: 10.1128/JB.174.24.8043-8056.1992.

Ball, D. A., Mehta, G. D., Salomon-Kent, R., Mazza, D., Morisaki, T., Mueller, F., McNally, J. G., and Karpova, T. S. (2016). Single molecule tracking of ace1p in saccharomyces cerevisiae defines a characteristic residence time for non-specific interactions of transcription factors with chromatin. *Nucleic Acids Research*, 44. doi: 10.1093/nar/gkw744.

Bintu, L., Buchler, N. E., Garcia, H. G., Gerland, U., Hwa, T., Kondev, J., and Phillips, R. Transcriptional regulation by the numbers: Models, (2005). ISSN 0959437X.

Boija, A., Klein, I. A., Sabari, B. R., Dall'Agnese, A., Coffey, E. L., Zamudio, A. V., Li, C. H., Shrinivas, K., Manteiga, J. C., Hannett, N. M., Abraham, B. J., Afeyan, L. K., Guo, Y. E., Rimel, J. K., Fant, C. B., Schuijers, J., Lee, T. I., Taatjes, D. J., and Young, R. A. (2018). Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell*, 175(7):1842–1855.e16. doi: 10.1016/j.cell.2018.10.042.

Bonner, J. J., Heyward, S., and Fackenthal, D. L. (1992). Temperature-dependent regulation of a heterologous transcriptional activation domain fused to yeast heat shock transcription factor. *Molecular and Cellular Biology*, 12(3):1021–1030. doi: 10.1128/mcb.12.3.1021-1030.1992.

Brackley, C. A., Taylor, S., Papantonis, A., Cook, P. R., and Marenduzzo, D. (2013). Non-specific bridging-induced attraction drives clustering of dna-binding proteins and genome organization. *Proceedings of the National Academy of Sciences of the United States of America*, 110. doi: 10.1073/PNAS.1302950110.

Browning, S. B. D. (2004). The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.*, 2:57–65.

Chen, J., Zhang, Z., Li, L., Chen, B. C., Revyakin, A., Hajj, B., Legant, W., Dahan, M., Lionnet, T., Betzig, E., Tjian, R., and Liu, Z. (2014). Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell*, 156. doi: 10.1016/j.cell.2014.01.062.

Chen, T.-Y., Cheng, Y.-S., Huang, P.-S., and Chen, P. (2018). Facilitated unbinding via multivalency-enabled ternary complexes: New paradigm for protein–dna interactions. *Accounts of Chemical Research*, 51(4):860–868. doi: 10.1021/acs.accounts.7b00541.

Clauß, K., Popp, A. P., Schulze, L., Hettich, J., Reisser, M., Torres, L. E., Uhlenhaut, N. H., and Gebhardt, J. C. M. (2017). Dna residence time is a regulatory factor of transcription repression. *Nucleic Acids Research*, 45:11121–11130. doi: 10.1093/NAR/GKX728.

Connaghan-Jones, K. D., Heneghan, A. F., Miura, M. T., and Bain, D. L. (2007). Thermodynamic analysis of progesterone receptor–promoter interactions reveals a molecular model for isoform-specific function. *Proceedings of the National Academy of Sciences*, 104(7):2187–2192. doi: 10.1073/pnas.0608848104.

Cook, P. R. A model for all genomes: The role of transcription factories, (2010). ISSN 00222836.

Dahlke, K. and Sing, C. E. (2017). Facilitated dissociation kinetics of dimeric nucleoid-associated proteins follow a universal curve. *Biophysical Journal*, 112(3):543–551. doi: 10.1016/j.bpj.2016.11.3198.

Dahlke, K. and Sing, C. E. (2019). Influence of nucleoid-associated proteins on dna supercoiling. *The Journal of Physical Chemistry B*, 123(48):10152–10162. doi: 10.1021/acs.jpcb.9b07436.

Dame, R., Wyman, C., and Goosen, N. (2000). H-ns mediated compaction of dna visualised by atomic force microscopy. *Nucleic Acids Res.*, 28:3504–3510. doi: 10.1093/nar/28.18.3504.

Dame, R. T. (2005). The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Molecular Microbiology*, 56:858–870. doi: 10.1111/J.1365-2958.2005.04598.X.

de Jonge, W. J., Brok, M., Lijnzaad, P., Kemmeren, P., and Holstege, F. C. (2020). Genome-wide off-rates reveal how dna binding dynamics shape transcription factor function. *Molecular systems biology*, 16. doi: 10.15252/MSB.20209885.

de Jonge, W. J., Patel, H. P., Meeussen, J. V., and Lenstra, T. L. Following the tracks: How transcription factor binding dynamics control transcription, (2022). ISSN 15420086.

Dillon, S. C. and Dorman, C. J. (2010). Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nature Reviews Microbiology 2010 8:3*, 8:185–195. doi: 10.1038/nrmicro2261.

Dorman, C. J. (2014). Function of nucleoid-associated proteins in chromosome structuring and transcriptional regulation. *Microbial Physiology*, 24:316–331. doi: 10.1159/000368850.

Elgin, S. C. and Grewal, S. I. (2003). Heterochromatin: silence is golden. *Current Biology*, 13:R895–R898. doi: 10.1016/J.CUB.2003.11.006.

Erbaş, A. and Marko, J. F. (2019). How do dna-bound proteins leave their binding sites? the role of facilitated dissociation. *Current Opinion in Chemical Biology*, 53:118–124. doi: 10.1016/j.cbpa.2019.08.007.

Garcia, D. A. An intrinsic disordered region mediated confinement state contributes to dynamics and function of transcription factors .david a garcia et al, (2021).

Garcia, D. A., Fettweis, G., Presman, D. M., Paakinaho, V., Jarzynski, C., Upadhyaya, A., and Hager, G. (2021). Power-law behavior of transcription factor dynamics at the single-molecule level implies a continuum affinity model. *Nucleic Acids Research*, 49:6605–6620. doi: 10.1093/nar/gkab072.

Gentry, R. C., Ide, N. A., Comunale, V. M., Hartwick, E. W., Kinz-Thompson, C. D., and Gonzalez, R. L. (2023). The mechanism of mrna activation. *bioRxiv*. doi: 10.1101/2023.11.15.567265.

Graham, J. S., Johnson, R. C., and Marko, J. F. (2011). Concentration-dependent exchange accelerates turnover of proteins bound to double-stranded dna. *Nucleic Acids Research*, 39:2249–2259. doi: 10.1093/NAR/GKQ1140.

Haberle, V. and Stark, A. Eukaryotic core promoters and the functional basis of transcription initiation, (2018). ISSN 14710080.

Halford, S. E. and Marko, J. F. (2004). How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res.*, 32(10):3040–3052.

Hammel, M., Amlanjyoti, D., Reyes, F. E., Chen, J. H., Parpana, R., Tang, H. Y., Larabell, C. A., Tainer, J. A., and Adhya, S. (2016). Hu multimerization shift controls nucleoid compaction. *Science Advances*, 2. doi: 10.1126/SCIADV.1600650.

Hipp, L., Beer, J., Kuchler, O., Reisser, M., Sinske, D., Michaelis, J., Gebhardt, J. C. M., and Knöll, B. (2019). Single-molecule imaging of the transcription factor srf reveals prolonged chromatin-binding kinetics upon cell stimulation. *Proceedings of the National Academy of Sciences of the United States of America*, 116. doi: 10.1073/pnas.1812734116.

Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K., and Sharp, P. A. (2017). A phase separation model for transcriptional control. *Cell*, 169(1):13–23. doi: 10.1016/j.cell.2017.02.007.

Iborra, F. J., Pombo, A., Jackson, D. A., and Cook, P. R. (1996). Active rna polymerases are localized within discrete transcription 'factories' in human nuclei. *Journal of Cell Science*, 109. doi: 10.1242/jcs.109.6.1427.

Inukai, S., Kock, K. H., and Bulyk, M. L. Transcription factor–dna binding: beyond binding site motifs, (2017). ISSN 18790380.

Ishihama, A., Kori, A., Koshio, E., Yamada, K., Maeda, H., Shimada, T., Makinoshima, H., Iwata, A., and Fujita, N. (2014). Intracellular concentrations of 65 species of transcription factors with known regulatory functions in escherichia coli. *Journal of Bacteriology*, 196 (15):2718–2727. doi: 10.1128/jb.01579-14.

Joshi, C. P., Panda, D., Martella, D. J., Andoy, N. M., Chen, T. Y., Gaballa, A., Helmann, J. D., and Chen, P. (2012). Direct substitution and assisted dissociation pathways for turning off transcription by a merr-family metalloregulator. *Proceedings of the National Academy of Sciences of the United States of America*, 109:15121–15126. doi: 10.1073/PNAS.1208508109.

Jung, Y., Sadeghi, A., and Ha, B.-Y. (2024). Modeling the compaction of bacterial chromosomes by biomolecular crowding and the cross-linking protein h-ns. *Scientific Reports*, 14(1). doi: 10.1038/s41598-023-50355-2.

Kamar, R. I., Banigan, E. J., Erbas, A., Giuntoli, R. D., Cruz, M. O. D. L., Johnson, R. C., and Marko, J. F. (2017). Facilitated dissociation of transcription factors from single dna binding sites. *Proceedings of the National Academy of Sciences of the United States of America*, 114:E3251–E3257. doi: 10.1073/PNAS.1701884114.

Kim, S. and Shendure, J. (2019). Mechanisms of interplay between transcription factors and the 3d genome. *Molecular Cell*, 76:306–319. doi: 10.1016/J.MOLCEL.2019.08.010.

Koşar, Z. and Erbaş, A. (2022). Can the concentration of a transcription factor affect gene expression? *Frontiers in Soft Matter*, 0:6. doi: 10.3389/FRSFM.2022.914494.

Koşar, Z., Attar, A. G., and Erbaş, A. (2022). Facilitated dissociation of nucleoid-associated proteins from dna in the bacterial confinement. *Biophysical journal*, 121:1119–1133. doi: 10.1016/J.BPJ.2022.03.002.

Larson, A. G., Elnatan, D., Keenen, M. M., Trnka, M. J., Johnston, J. B., Burlingame, A. L., Agard, D. A., Redding, S., and Narlikar, G. J. (2017). Liquid droplet formation by hp1α suggests a role for phase separation in heterochromatin. *Nature*, 547(7662):236–240. doi: 10.1038/nature22822.

Lee, K. A. Dimeric transcription factor families: It takes two to tango but who decides on partners and the venue?, (1992). ISSN 00219533.

Lickwar, C. R., Mueller, F., Hanlon, S. E., McNally, J. G., and Lieb, J. D. (2012). Genome-wide protein–dna binding dynamics suggest a molecular clutch for transcription factor function. *Nature 2012 484:7393*, 484:251–255. doi: 10.1038/nature10985.

Lin, J., Chen, H., Dröge, P., and Yan, J. (2012). Physical organization of dna by multiple non-specific dna-binding modes of integration host factor (ihf). *PLOS ONE*, 7:e49885. doi: 10.1371/JOURNAL.PONE.0049885.

Liu, J., François, J.-M., and Capp, J.-P. (2019). Gene expression noise produces cell-to-cell heterogeneity in eukaryotic homologous recombination rate. *Frontiers in Genetics*, 10. doi: 10.3389/fgene.2019.00475.

Melnik, S., Deng, B., Papantonis, A., Baboo, S., Carr, I. M., and Cook, P. R. (2011). The proteomes of transcription factories containing rna polymerases i, ii or iii. *Nature Methods*, 8. doi: 10.1038/nmeth.1705.

Mitchell, J. A. and Fraser, P. (2008). Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes and Development*, 22. doi: 10.1101/gad.454008.

Morisaki, T., Müller, W. G., Golob, N., Mazza, D., and McNally, J. G. (2014). Single-molecule analysis of transcription factor binding at transcription sites in live cells. *Nature Communications*, 5. doi: 10.1038/ncomms5456.

Noort, J. V., Verbrugge, S., Goosen, N., Dekker, C., and Dame, R. T. (2004). Dual architectural roles of hu: Formation of flexible hinges and rigid filaments. *Proceedings of the National Academy of Sciences of the United States of America*, 101:6969–6974. doi: 10.1073/PNAS.0308230101.

Nozawa, R., Yamamoto, T., Takahashi, M., Tachiwana, H., Maruyama, R., Hirota, T., and Saitoh, N. (2020). Nuclear microenvironment in cancer: Control through liquid-liquid phase separation. *Cancer Science*, 111(9):3155–3163. doi: 10.1111/cas.14551.

Parab, L., Pal, S., and Dhar, R. (2022). Transcription factor binding process is the primary driver of noise in gene expression. *PLoS genetics*, 18:e1010535. doi: 10.1371/JOURNAL.PGEN.1010535.

Penagos-Puig, A. and Furlan-Magaril, M. Heterochromatin as an important driver of genome organization, (2020). ISSN 2296634X.

Raser, J. M. and O'Shea, E. K. (2005). Noise in gene expression: Origins, consequences, and control. *Science*, 309:2010–2013. doi: 10.1126/science.1105891. Original paper for "gene expression NOISE".

Reisser, M., Hettich, J., Kuhn, T., Popp, A. P., Große-Berkenbusch, A., and Gebhardt, J. C. M. (2020). Inferring quantity and qualities of superimposed reaction rates from single molecule survival time distributions. *Scientific Reports*, 10. doi: 10.1038/s41598-020-58634-y.

Remesh, S. G., Verma, S. C., Chen, J. H., Ekman, A. A., Larabell, C. A., Adhya, S., and Hammel, M. (2020). Nucleoid remodeling during environmental adaptation is regulated by hu-dependent dna bundling. *Nature Communications 2020 11:1*, 11:1–12. doi: 10.1038/s41467-020-16724-5.

Sabari, B. R., Dall'Agnese, A., Boija, A., Klein, I. A., Coffey, E. L., Shrinivas, K., Abraham, B. J., Hannett, N. M., Zamudio, A. V., Manteiga, J. C., Li, C. H., Guo, Y. E., Day, D. S., Schuijers, J., Vasile, E., Malik, S., Hnisz, D., Lee, T. I., Cisse, I. I., Roeder, R. G., Sharp, P. A., Chakraborty, A. K., and Young, R. A. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. *Science*, 361(6400). doi: 10.1126/science.aar3958.

Schaaf, M. J. M. and Cidlowski, J. A. (2003). Molecular determinants of glucocorticoid receptor mobility in living cells: the importance of ligand affinity. *Molecular and Cellular Biology*, 23(6):1922–1934. doi: 10.1128/mcb.23.6.1922-1934.2003.

Seshasayee, A. S. N., Sivaraman, K., and Luscombe, N. M. (2011). An overview of prokaryotic transcription factors: A summary of function and occurrence in bacterial genomes. *Sub-Cellular Biochemistry*, 52:7–23. doi: 10.1007/978-90-481-9069-0_2/FIGURES/2_2_209873_1_EN.

Skoko, D., Wong, B., Johnson, R. C., and Marko*, J. F. (2004). Micromechanical analysis of the binding of dna-bending proteins hmgb1, nhp6a, and hu reveals their ability to form highly stable dnaprotein complexes†. *Biochemistry*, 43:13867–13874. doi: 10.1021/BI048428O.

Skoko, D., Yoo, D., Bai, H., Schnurr, B., Yan, J., McLeod, S. M., Marko, J. F., and Johnson, R. C. (2006). Mechanism of chromosome compaction and looping by the escherichia coli nucleoid protein fis. *Journal of Molecular Biology*, 364:777–798. doi: 10.1016/J.JMB.2006.09.043.

Stracy, M., Schweizer, J., Sherratt, D. J., Kapanidis, A. N., Uphoff, S., and Lesterlin, C. (2021). Transient non-specific DNA binding dominates the target search of bacterial DNA-binding proteins. *Mol. Cell*, 81(7):1499–1514.e6.

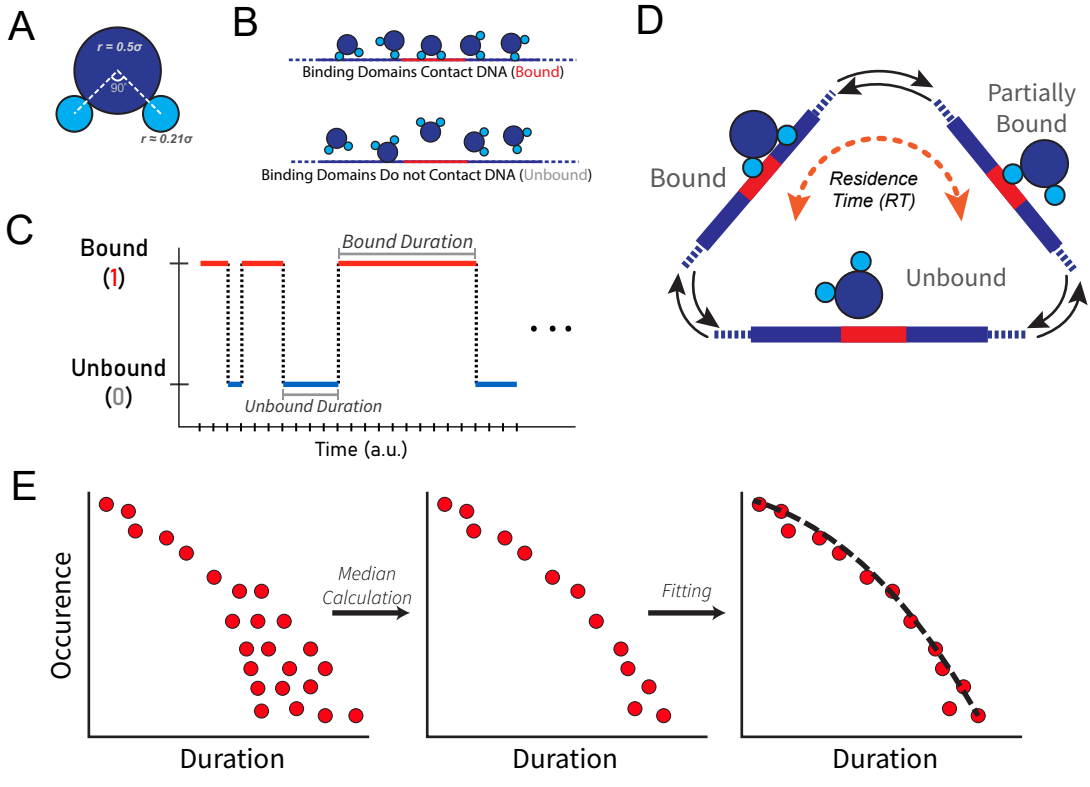
Verma, S. C., Qian, Z., and Adhya, S. L. (2019). Architecture of the escherichia coli nucleoid. *PLOS Genetics*, 15:e1008456. doi: 10.1371/JOURNAL.PGEN.1008456.

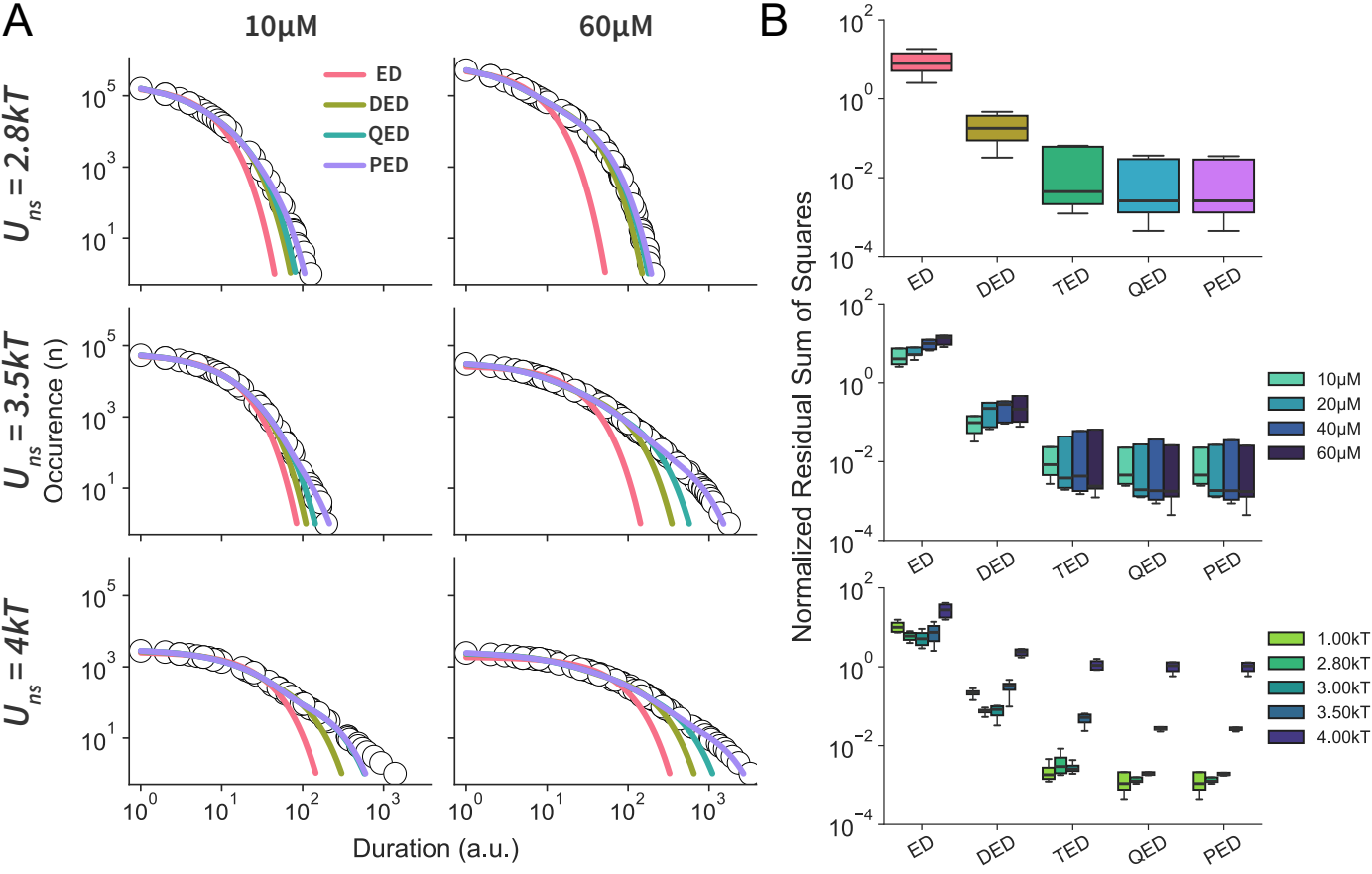
Wang, B., Zhang, L., Dai, T., Qin, Z., Lu, H., Zhang, L., and Zhou, F. (2021). Liquid–liquid phase separation in human health and diseases. *Signal Transduction and Targeted Therapy*, 6(1). doi: 10.1038/s41392-021-00678-1.

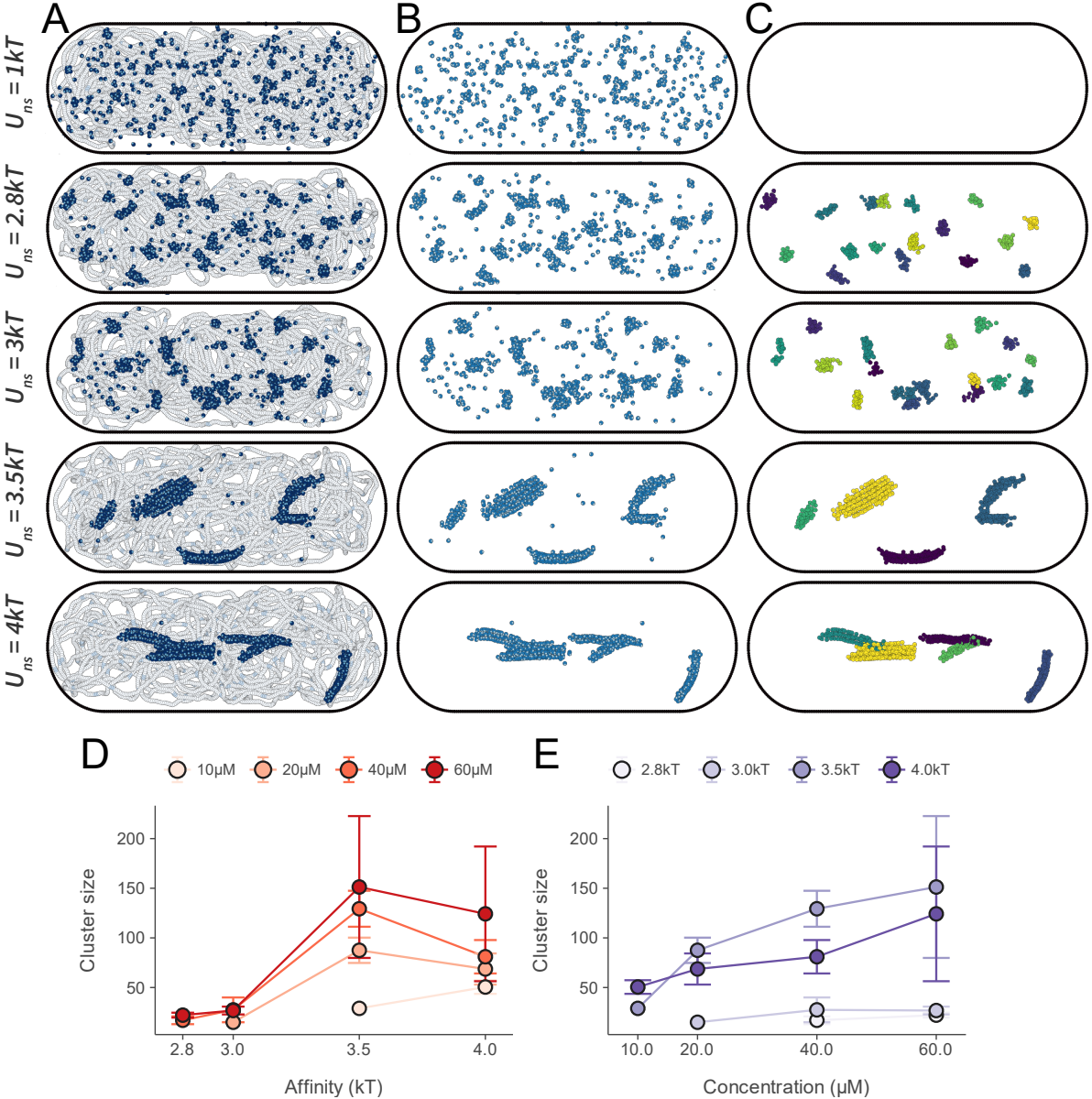
Wang, W., Li, G. W., Chen, C., Xie, X. S., and Zhuang, X. (2011). Chromosome organization by a nucleoid-associated protein in live bacteria. *Science*, 333:1445–1449. doi: 10.1126/SCIENCE.1204697.

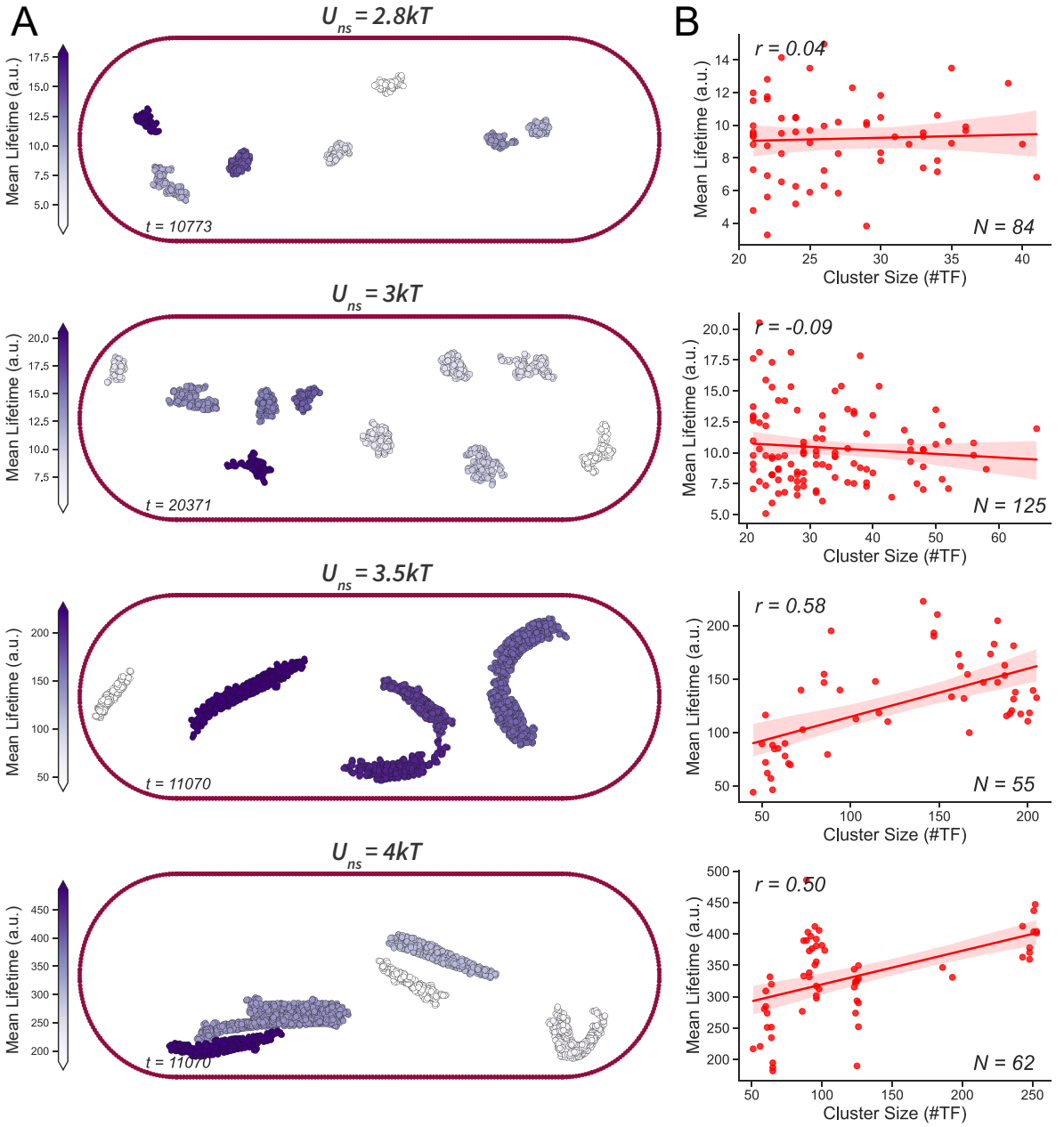
Winardhi, R. S., Yan, J., and Kenney, L. J. (2015). H-ns regulates gene expression and compacts the nucleoid: Insights from single-molecule experiments. *Biophysical Journal*, 109:1321–1329. doi: 10.1016/J.BPJ.2015.08.016.

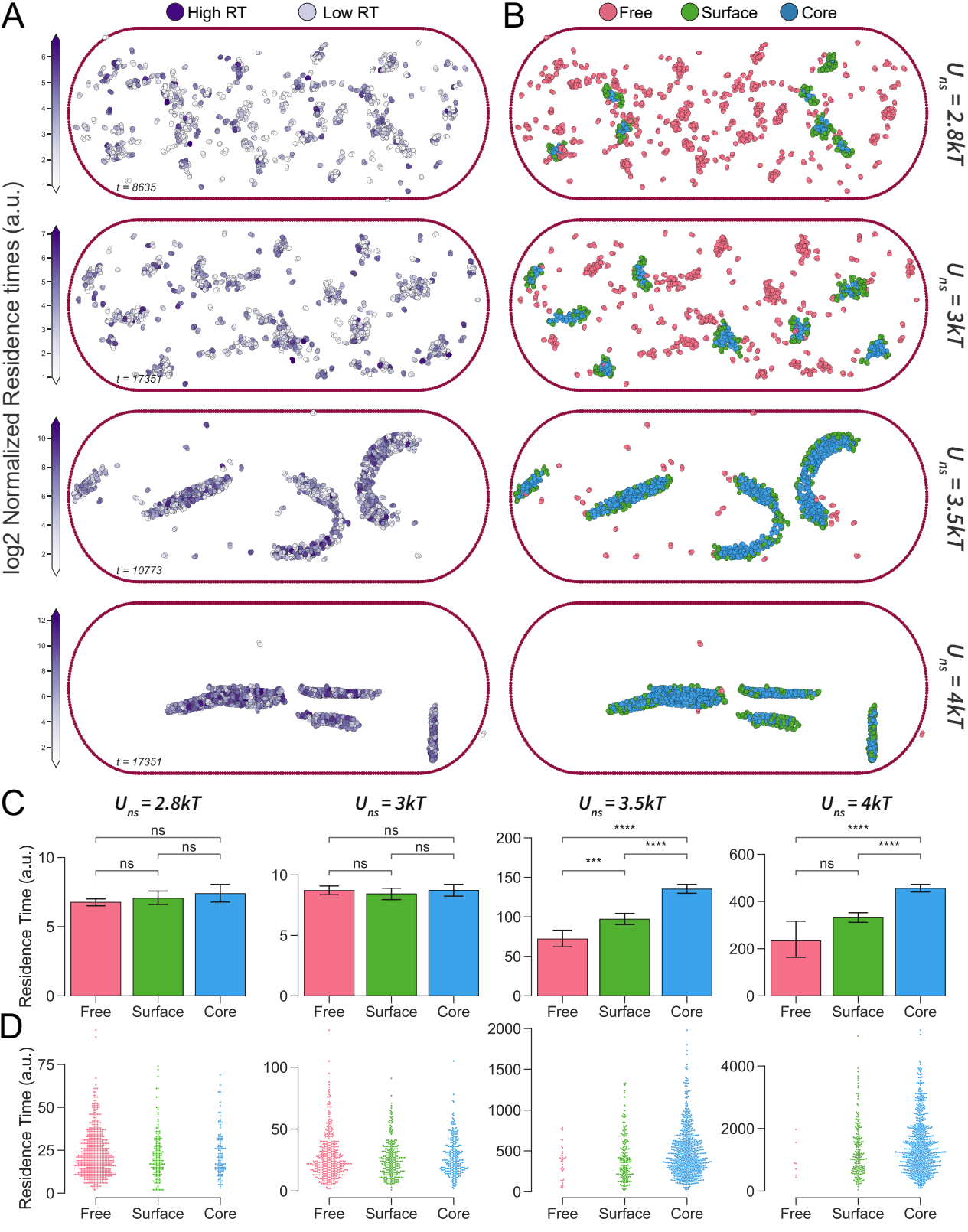
Zhu, J., Cao, X., and Deng, X. (2023). Epigenetic and transcription factors synergistically promote the high temperature response in plants. *Trends in Biochemical Sciences*, 48 (9):788–800. doi: 10.1016/j.tibs.2023.06.001.

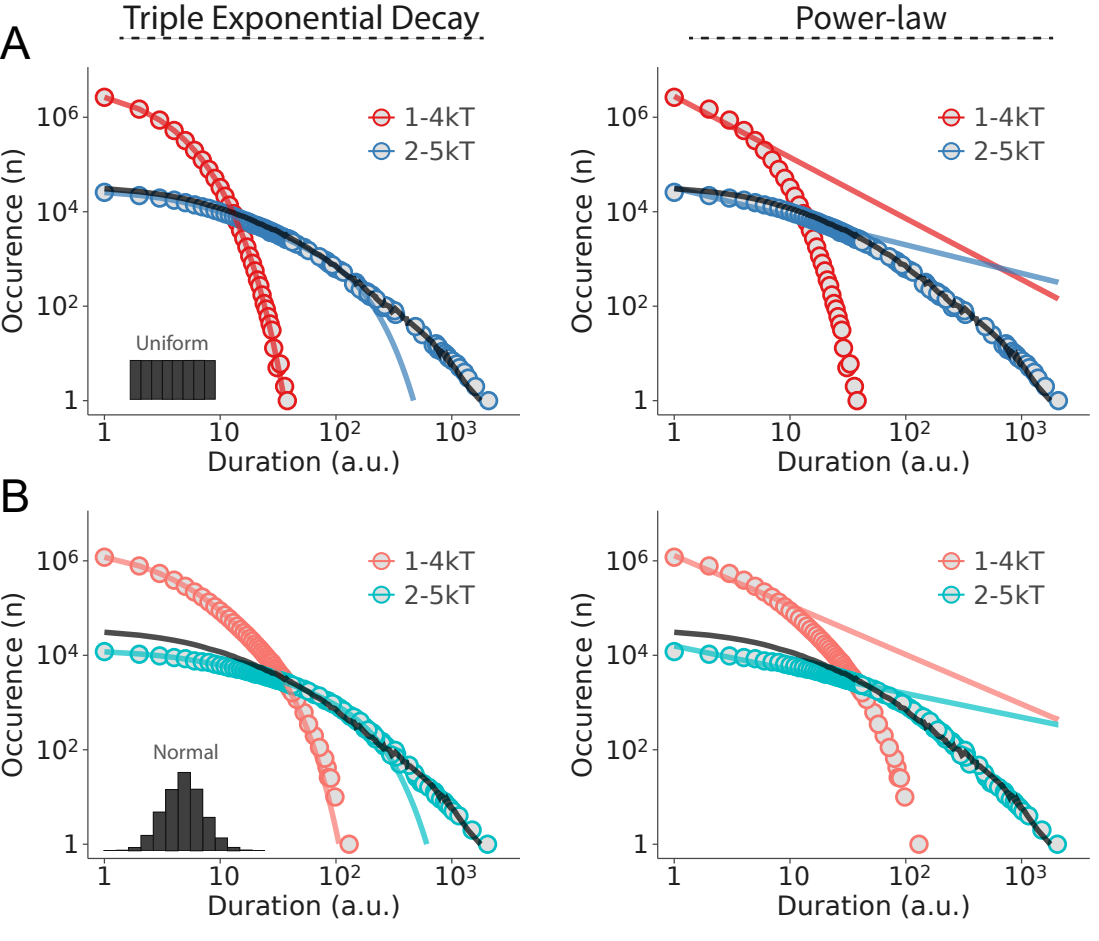

A
r = 0.5σ
90°
r = 0.21σ
B
Binding Domains Contact DNA (Bound)
Binding Domains Do not Contact DNA (Unbound)
C
Bound
(1)
Unbound
(0)
Bound Duration
Unbound Duration
Time (a.u.)
D
Bound
Partially
Bound
Residence
Time (RT)
Unbound
E
Occurence
Median
Calculation
Fitting
Duration
Duration
Duration

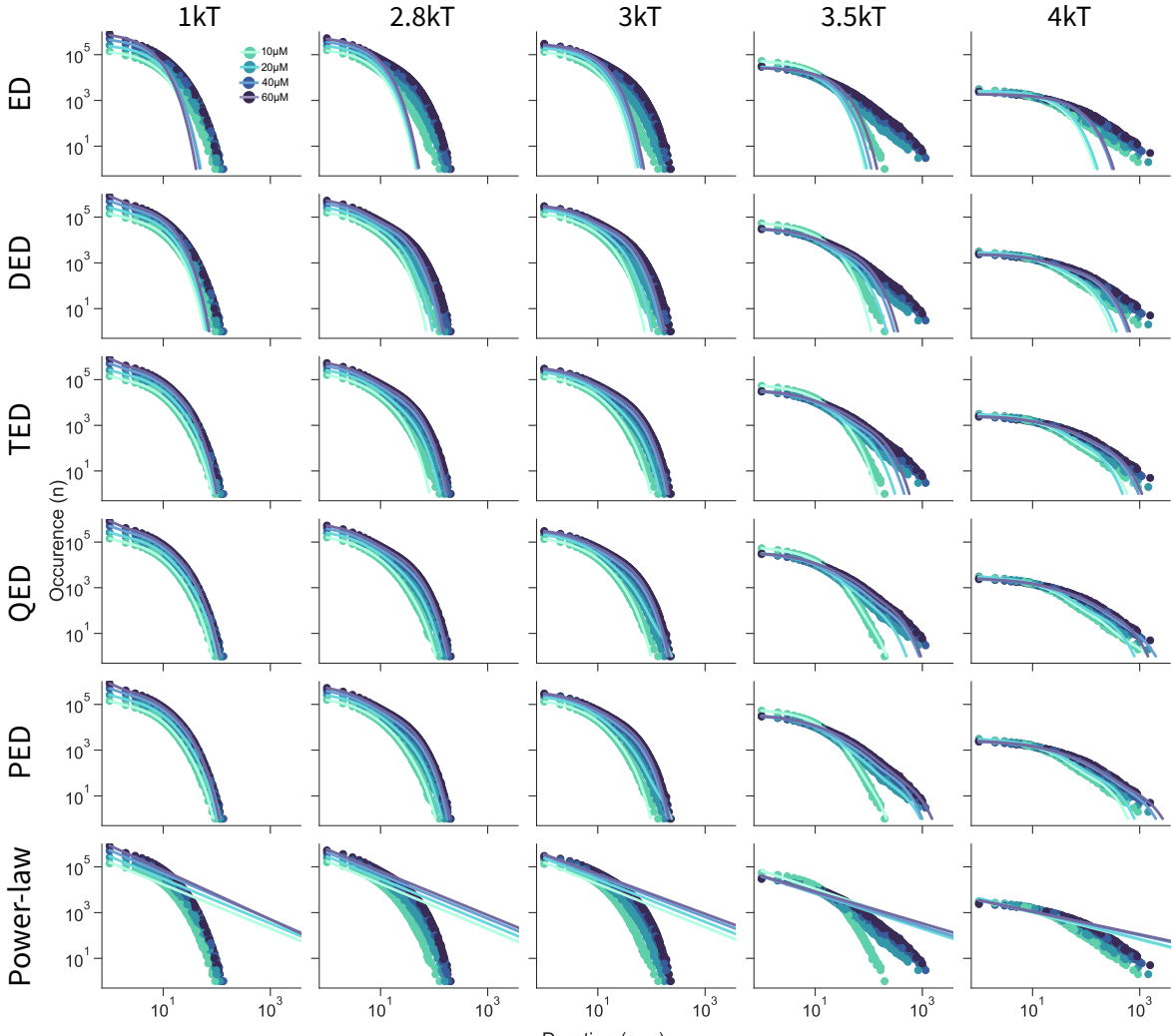

A
10μM
60μM
ED
DED
QED
PED
$U_{ns} = 2.8kT$
$U_{ns} = 3.5kT$
$U_{ns} = 4kT$
Occurence (n)
Duration (a.u.)
B
Normalized Residual Sum of Squares
ED
DED
TED
QED
PED
10μM
20μM
40μM
60μM
1.00kT
2.80kT
3.00kT
3.50kT
4.00kT

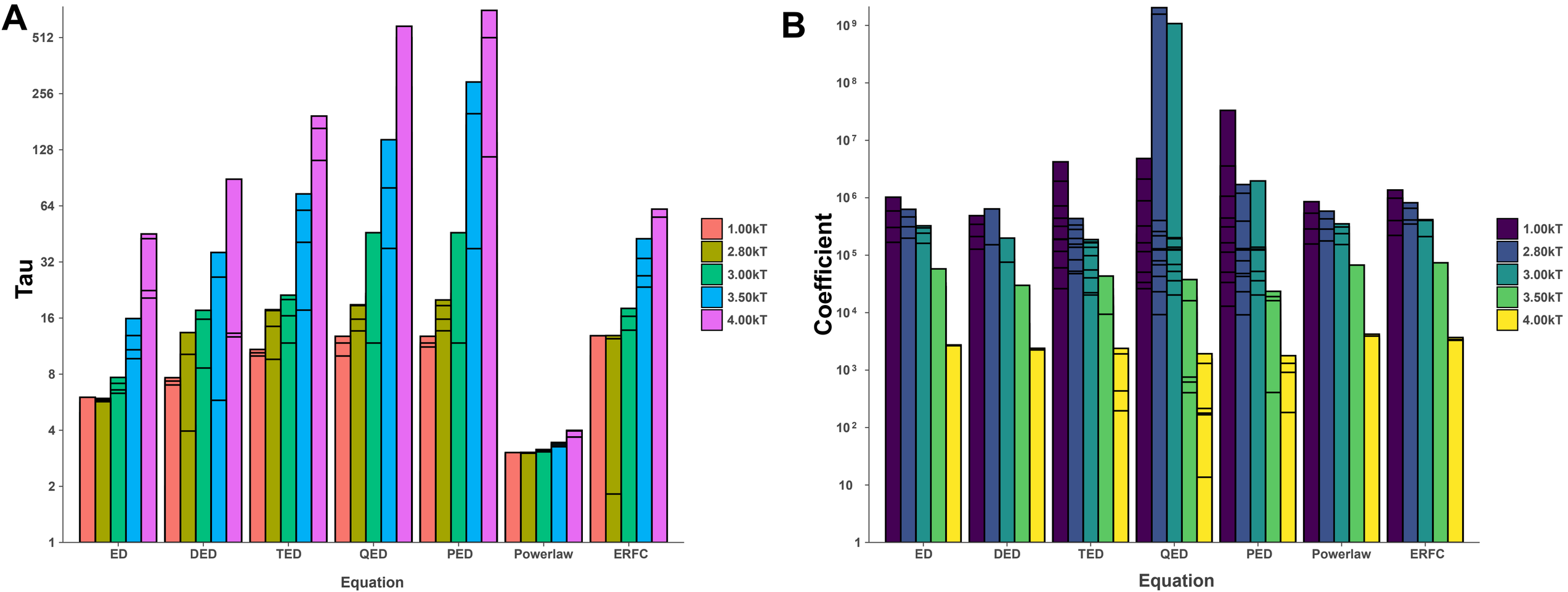

A
B
C
$U_{ns} = 1kT$
$U_{ns} = 2.8kT$
$U_{ns} = 3kT$
$U_{ns} = 3.5kT$
$U_{ns} = 4kT$
D
10μM
20μM
40μM
60μM
Cluster size
200
150
100
50
2.8
3.0
3.5
4.0
Affinity (kT)
E
2.8kT
3.0kT
3.5kT
4.0kT
Cluster size
200
150
100
50
10.0
20.0
40.0
60.0
Concentration (μM)

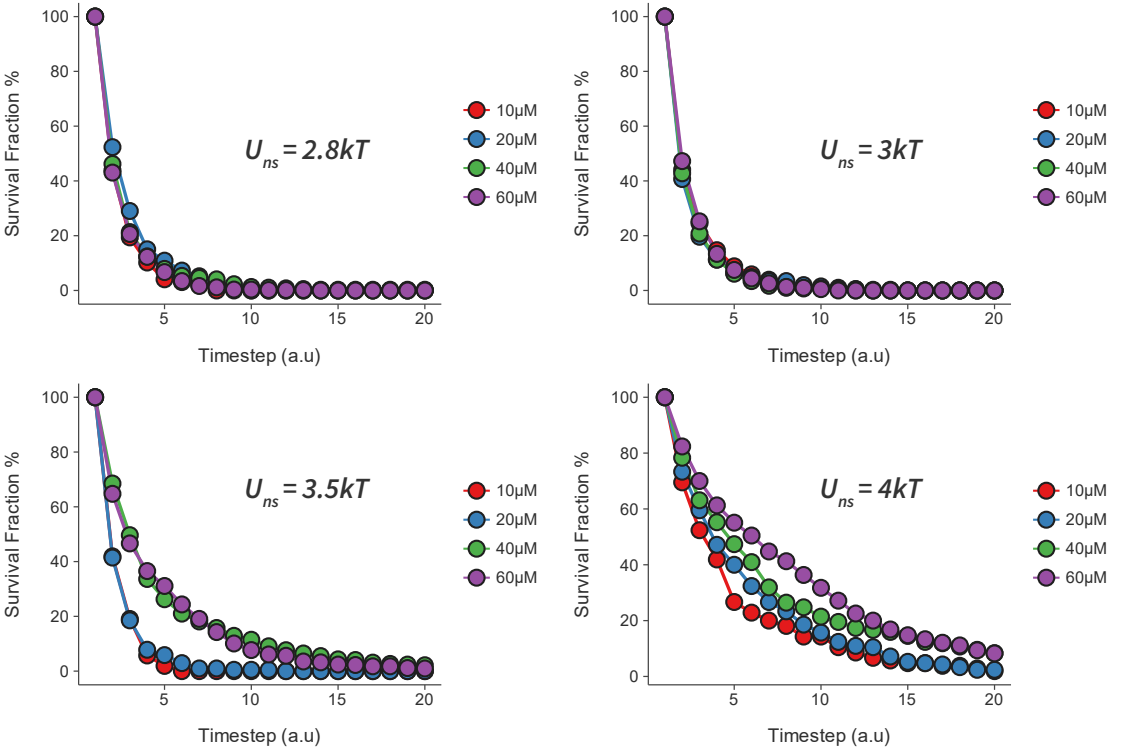

A
B
U_{ns} = 2.8kT
Mean Lifetime (a.u.)
17.5
15.0
12.5
10.0
7.5
5.0
t = 10773
r = 0.04
N = 84
Cluster Size (#TF)
U_{ns} = 3kT
20.0
t = 20371
r = -0.09
N = 125
U_{ns} = 3.5kT
200
150
100
50
t = 11070
r = 0.58
N = 55
U_{ns} = 4kT
450
400
350
300
250
200
t = 11070
r = 0.50
N = 62


A
High RT
Low RT
B
Free
Surface
Core
log2 Normalized Residence times (a.u.)
t = 8635
t = 17351
t = 10773
t = 17351
U_{ns} = 2.8kT
U_{ns} = 3kT
U_{ns} = 3.5kT
U_{ns} = 4kT
C
U_{ns} = 2.8kT
U_{ns} = 3kT
U_{ns} = 3.5kT
U_{ns} = 4kT
Residence Time (a.u.)
ns
***
****
Free
Surface
Core
D
Residence Time (a.u.)
Free
Surface
Core


Triple Exponential Decay
Power-law
A
B
1-4kT
2-5kT
Uniform
Normal
Occurence (n)
Duration (a.u.)
10^6
10^4
10^2
1
10
10^2
10^3


1kT
2.8kT
3kT
3.5kT
4kT
ED
DED
TED
QED
PED
Power-law
10μM
20μM
40μM
60μM
Occurence (n)
Duration (a.u.)


A
Tau
512
256
128
64
32
16
8
4
2
1
ED
DED
TED
QED
PED
Powerlaw
ERFC
Equation
1.00kT
2.80kT
3.00kT
3.50kT
4.00kT
B
Coefficient
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10
1
ED
DED
TED
QED
PED
Powerlaw
ERFC
Equation
1.00kT
2.80kT
3.00kT
3.50kT
4.00kT


Survival Fraction %
Timestep (a.u)
$U_{ns} = 2.8kT$
$U_{ns} = 3kT$
$U_{ns} = 3.5kT$
$U_{ns} = 4kT$
10μM
20μM
40μM
60μM
0
20
40
60
80
100
5
10
15
20