

Guided nuclear exploration increases CTCF target search efficiency

Anders S. Hansen^{1,2,3,4,*}, Assaf Amitai^{5,*}, Claudia Cattoglio^{1,2,3,4}, Robert Tjian^{1,2,3,4}, Xavier Darzacq^{1,2,3}

1: Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, USA; 2: Li Ka Shing Center for Biomedical and Health Sciences; 3: CIRM Center of Excellence, University of California, Berkeley, Berkeley, USA; 4: Howard Hughes Medical Institute, University of California, Berkeley, Berkeley, USA; 5: Department of Chemical Engineering, MIT, Cambridge 02139, Massachusetts, USA

*: ASH and AA contributed equally. Correspondence: XD (darzacq@berkeley.edu) and RT (jmlim@berkeley.edu).

Mammalian genomes are enormous. For a DNA-binding protein, this means that the number of non-specific, off-target sites vastly exceeds the number of specific, cognate sites. How mammalian DNA-binding proteins overcome this challenge to efficiently locate their target sites is not known. Here through live-cell single-molecule tracking, we show that CCTC-binding factor, CTCF, is repeatedly trapped in small zones in the nucleus in a manner that is largely dependent on its RNA-binding region (RBR). Integrating theory, we devise a new model, Anisotropic Diffusion through transient Trapping in Zones (ADTZ), to explain this. Functionally, transient RBR-mediated trapping increases the efficiency of CTCF target search by ~2.5 fold. Since the RBR-domain also mediates CTCF clustering, our results suggest a “guided” mechanism where CTCF clusters concentrate diffusing CTCF proteins near cognate binding sites, thus increasing the local ON-rate. We suggest that local “guiding” may represent a general target search mechanism in mammalian cells.

INTRODUCTION

Mammalian nuclei are organized by a myriad of biophysical forces into sub-compartments with distinct functions. At the micron-scale, nuclear compartments such as nucleoli, speckles, and Cajal bodies, carry out specialized biochemical functions that are spatially segregated (Mao et al., 2011). Below the micron-scale and diffraction-limit of conventional optical microscopy, many proteins interact dynamically to form local high concentration clusters or hubs (Banjade and Rosen, 2014; Boehning et al., 2018; Chong et al., 2018). Thus, many proteins exhibit an organized and non-random nuclear distribution. At the same time, several proteins display anomalous and non-Brownian diffusion inside the nucleus (Höfling and Franosch, 2013; Metzler et al., 2014; Rhodes et al., 2017), which has been proposed to be due to molecular crowding and transient interactions (Bancaud et al., 2012; Woringer and Darzacq, 2018). It is still unclear what mechanisms allow proteins to form stable or transient clusters and control their diffusion and target search mechanism *in vivo*. It has been proposed that some proteins with disordered domains are capable of forming hubs that behave as distinct liquid droplets or condensates (Feric et al., 2016). In such structures, protein-RNA interactions have been shown to be important (Lin et al., 2015). The most

obvious example is the nucleolus (Weber and Brangwynne, 2015), which is rich in ribosomal RNA. However, the functional role of these structures in loading, guiding and tightly regulating the functional search kinetics of associated proteins has yet to be elucidated *in vivo*.

Since the kinetics of a reaction or binding event depend on the diffusive properties and nuclear organization of the reactants (Bénichou et al., 2010; Rice, 1985), understanding the molecular interactions that control a nuclear protein’s target search mechanism and its overall distribution is essential to understanding its function. Target search mechanisms have been extensively probed in prokaryotes (Hammar et al., 2012; Kapanidis et al., 2018), where current models (Bauer and Metzler, 2012) suggest that proteins find their target sites through facilitated diffusion (Slutsky and Mirny, 2004). In this family of models, proteins are suggested to move according to two principal modes: (1) sliding in 1D along the DNA strand; (2) occasionally disassociating from the strand, diffusing in 3D, only to bind at a proximal DNA site. It is thought that a mixture of these two modes of motion, with specific rates of transition between the states, would serve as an optimal search strategy (Lomholt et al., 2005). However, how DNA-binding proteins find their targets in mammalian nuclei with enormous genomes and nucleosomes that would

seemingly obstruct 1D sliding remains poorly understood. One problem with applying the 1D+3D view in mammalian cells is that the number of off-target sequences and non-specific DNA binding sites is orders of magnitude greater in mammals than in prokaryotes (Iwahara and Clore, 2006). The much higher ratio of non-specific to specific binding sites in mammalian nuclei compared to prokaryotes suggests that mammalian cells may have evolved distinct mechanisms for guiding DNA-binding proteins towards their specific, cognate target sites. Indeed, recent work has suggested that non-specific binding to DNA can be a strong determinant of the distribution and nuclear exploration of DNA-binding proteins (McSwiggen et al., 2018).

Here we investigate how mammalian DNA-binding proteins find their nuclear targets and focus on the nuclear target-search mechanism of CTCF. CTCF, together with the cohesin complex, folds mammalian genomes into spatial domains known as Topologically Associating Domains (TADs), which regulate enhancer-promoter contacts and gene expression (Hansen et al., 2018a; Merkenschlager and Nora, 2016; Rada-Iglesias et al., 2018; Rowley and Corces, 2018). CTCF is an 11-Zinc Finger DNA-binding protein with unstructured N- and C-terminal domains, whose function remains poorly understood (Ghirlando and Felsenfeld, 2016). Cohesin is thought to be a ring-shaped multi-subunit complex, which holds together CTCF-demarcated TADs as chromatin loops (Hassler et al., 2018). Although CTCF and cohesin have emerged as key regulators of genome organization and function, their nuclear target search mechanisms have not been studied.

Using Single-Particle Tracking (SPT) and theoretical modeling, here we show that both CTCF and cohesin exhibit unusual nuclear target search mechanisms, where anisotropic diffusion arises from being transiently trapped in specific zones with a characteristic size of ~ 200 nm. Our results indicate that these zones correspond to CTCF clusters and, surprisingly, we find that trapping inside them is largely due to CTCF's RNA-binding region (RBR). Functionally, transient RBR-mediated trapping in zones increases the efficiency of CTCF's DNA-target search mechanism by about 2.5-fold. More generally, we suggest that RNA-mediated "guiding" could be an effective way to control and regulate the local concentration of proteins in the nucleus around specific sites.

RESULTS

CTCF exhibits highly anisotropic nuclear diffusion

Our initial imaging analysis revealed that CTCF exhibits highly anomalous diffusion inside the nucleus (Figure 1-Figure Supplement 1A; Video 1-4). Motivated by this observation, we decided to systematically investigate how CTCF, cohesin and other nuclear proteins explore the mammalian nucleus using single-particle tracking (SPT). Taking advantage of our established mouse embryonic stem cell (mESC) and human osteosarcoma (U2OS) cell lines where CTCF and the cohesin subunit Rad21 have been endogenously HaloTagged (Hansen et al., 2017), we used stroboscopic photo-activation SPT (spaSPT; Figure 1A; (Hansen et al., 2017, 2018b)) to track single protein molecules over time in live cells. spaSPT overcomes common biases in SPT by using stroboscopic excitation to minimize "motion-blur" bias and by using photo-activation to track at low densities of ~ 0.5 -1.0 molecules per nucleus per frame, which minimizes tracking errors (Elf et al., 2007; Grimm et al., 2016; Hansen et al., 2017, 2018b; Manley et al., 2008). Most nuclear proteins exist in either a "bound" chromatin-associated state (e.g. trajectory A in Figure 1B) or a seemingly "free" diffusing state (e.g. trajectory B in Figure 1B). To explore the nuclear diffusion mechanism, we need to analyze exclusively the diffusive trajectory segments. We achieved this using a Hidden-Markov Model (HMM) (Persson et al., 2013) to classify trajectories into bound and free segments and then removed the bound segments and calculated the angle (Burov et al., 2013; Izeddin et al., 2014; Liao et al., 2012) between 3 consecutive localizations provided that both displacements making up the angle were much longer than our localization error of ~ 35 nm (Figure 1B; full details in Materials and Methods). Finally, to comprehensively analyze the data at multiple spatiotemporal scales, we generated a very large data set by imaging 1,693 single cells at three different frame-rates (223 Hz, 134 Hz, 74 Hz; see Data Availability for raw data).

We began by analyzing two independent mESC Halo-CTCF clones, C59 and C87. The distribution of angles from diffusing CTCF trajectories showed a large peak at $\sim 180^\circ$ (Figure 1C). This is surprising because it indicates that CTCF displays a directional bias: once CTCF has moved in one direction, it is significantly more likely to move backward in the opposite direction than it is to continue forward. This anisotropic diffusion behavior is not what one would expect for Brownian motion, which should have no directional bias. To quantify this effect, we use the fold anisotropy, $f_{180/0}$:

how many fold more likely is a step backward relative to a step forward, which for CTCF was 1.77 in mESCs. While confinement inside the nucleus is expected to cause some anisotropy due to e.g. collisions with the nuclear envelope, the free Halo-3xNLS protein was almost isotropic, $f_{180/0}=1.12$, ruling out this as an explanation (Figure 1C). Could this be a general effect of all DNA-binding proteins? To test this, we analyzed a Halo-Sox2 knock-in mESC line (Teves et al., 2016), but Sox2 was also much less anisotropic than CTCF ($f_{180/0}=1.27$; Figure 1C). Thus, CTCF exhibits unusual anisotropic diffusion behavior.

A previous study found that PTEFb exhibits scale-free anisotropic diffusion, with a magnitude of anisotropy that remains constant in space and time (Izeddin et al., 2014). To see if a similar mechanism holds for CTCF, we analyzed $f_{180/0}$ as a function of the lag time between frames (Figure 1D) and the mean displacement length (Figure 1E). After an initial decline, $f_{180/0}$ was relatively constant in time and somewhat noisy at long time-lags (Figure 1D). In contrast, CTCF showed a clear anisotropy peak at ~ 200 nm displacements (Figure 1E). This was surprising since this type of

diffusion has, to the best of our knowledge, not previously been observed. The spatial dependence was even clearer when we plotted $f_{180/0}$ as a function of the length of the first and second displacements: CTCF showed a prominent peak (Figure 1F), which was not seen for Halo-3xNLS (Figure 1G). Using simulations, we verified that the observed anisotropy was not due to our localization uncertainty of ~ 35 nm (Figure 1-Figure Supplement 1B-D). To further confirm that this result was also not due to tracking errors, we performed 2-color spaSPT control experiments: we labeled Halo-CTCF 1:1 with two distinguishable dyes, PA-JF₅₄₉ and PA-JF₆₄₆, which enabled us to identify any tracking errors (e.g. red \rightarrow green switches within the same trajectory) and analyzed how tracking errors depend on density and displacement lengths. As expected, tracking errors were almost non-existent at the low densities of our imaging (Figure 1-Figure Supplement 2). However, we found that tracking errors increase *exponentially* with displacement length reaching $\sim 5\%$ around ~ 800 nm displacements (Figure 1-Figure Supplement 2A,F), and we thus limited our analysis to displacements below this length (full discussion in Appendix 1). Finally, we asked whether this

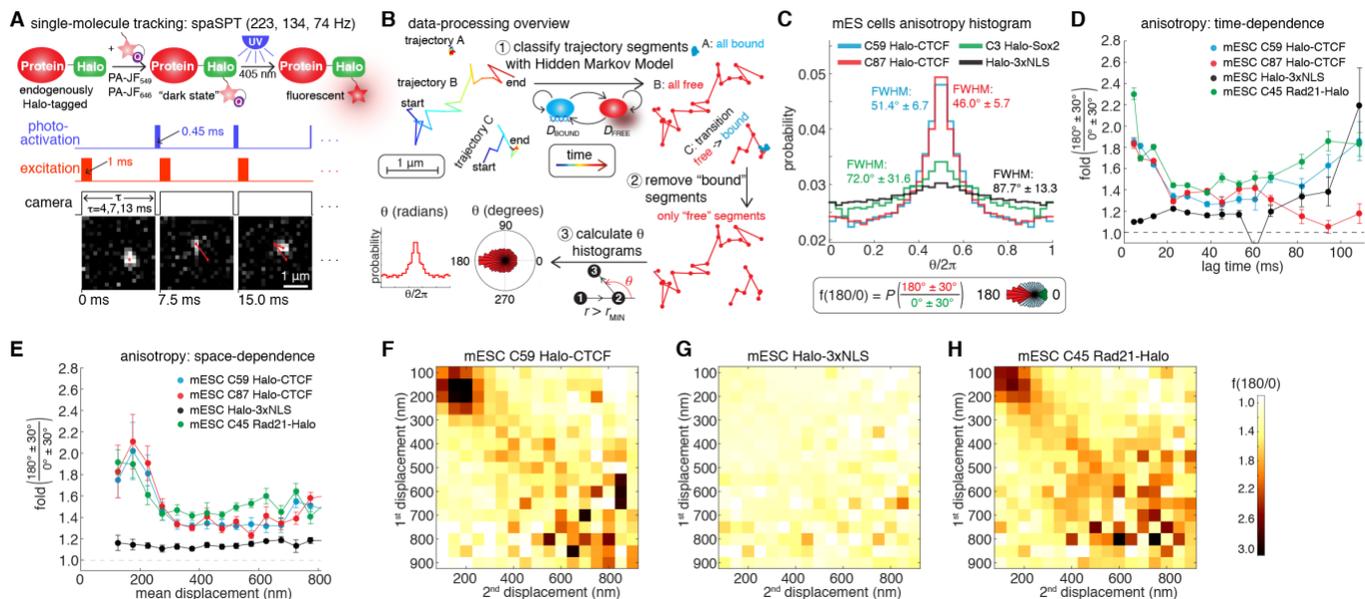


Figure 1. spaSPT reveals anisotropic CTCF diffusion in the nucleus.

(A) Overview of spaSPT. Sketch of Halo-labeling with PA-JF₅₄₉/PA-JF₆₄₆, photo-activation, and overview of laser pattern. Below, representative raw images with tracking. (B) Overview of data-processing. spaSPT trajectories are classified into “bound” and “free” segments using an HMM and angles are only calculated from “free” segments where both displacements are above 200 nm. (C) Bulk histograms of the distribution of angles for CTCF (C59, C87), Sox2 (C3) and Halo-3xNLS in mESCs. Definition of fold-anisotropy, $f_{180/0}$. (D) Plot of $f_{180/0}$ vs. lag time averaging over all displacement lengths. (E) Plot of $f_{180/0}$ vs. mean displacement length averaging over all lag times. (F-H) Anisotropy heatmaps showing $f_{180/0}$ vs. length of the first and second displacement for C59 Halo-CTCF (F), Halo-3xNLS (G), C45 mRad21-Halo (H); in S/G2 phase of the cell cycle, averaging over all lag times.

Figure 1-Figure Supplement 1. MSD-analysis and control Brownian motion simulations.

Figure 1-Figure Supplement 2. 2-color spaSPT experiments to estimate the misconnection probability during tracking.

Figure 1-Figure Supplement 3. Anisotropy for CTCF in human U2OS cells and cohesin in mESCs.

mechanism is conserved between species and analyzed CTCF diffusion in human U2OS cells. Like in mESCs, human CTCF showed highly anisotropic diffusion with a large peak at ~ 200 nm displacements (Figure 1-Figure Supplement 3A-D). Notably, although Sox2 did not, when we analyzed cohesin anisotropy, we found a similar anisotropic diffusion mechanism with a peak around ~ 200 nm displacements (Figure 1D,E,H and Figure 1-Figure Supplement 3E-J). In summary, we conclude that CTCF and cohesin exhibit a highly unusual mechanism of anisotropic nuclear diffusion with a clear spatial dependence that is specific and conserved between mouse stem cells and differentiated human cells.

A model featuring transient trapping in domains/zones can explain CTCF's behavior

We next sought to understand the underlying mechanism of anisotropic diffusion by CTCF. A previous study explained anisotropic diffusion of the transcription complex PTEFb by suggesting that PTEFb moves along a fractal structure, most likely chromatin in the nucleus, and when it reaches a dead-end, must turn back (Izeddin et al., 2014). Since CTCF anisotropy is not uniform with displacement length (*i.e.*, it peaks at ~ 200 nm), a similar fractal model with scale-free behavior cannot explain CTCF's diffusion. We thus instead hypothesized that weak and transient interactions likely govern CTCF motion in the nucleus. If the probability of binding transiently at a given position was uniform throughout the nucleus, the resulting protein diffusion should be isotropic and Brownian, but with an effective (reduced) diffusion coefficient. However, if the source of CTCF anisotropic diffusion is transient localized trapping, such retention should occur in a domain/zone of a characteristic size (Figure 2A).

We tested this hypothesis by simulating chromatin as self-avoiding polymers confined inside the nucleus (Amitai et al., 2017) (Figure 2-Figure Supplement 1A) and simulated protein motion and interaction with chromatin (see (Amitai, 2018) and Materials and Methods for full details). The protein undergoes Brownian motion inside the nucleus but upon encountering a Transiently Trapping Zone (monomer; TTZ) has a certain probability P_{trap} of becoming absorbed and trapped. We model TTZs as spherical. These zones trap CTCF transiently with a time-scale comparable to our experimental frame rates and thus result in trajectories that appear as both free and bound, depending on the bound state of the protein. CTCF also binds strongly to Cognate DNA Binding Sites (CBSs)

(Figure 2A) with a residence time for specific binding of $\sim 1-4$ minutes (Hansen et al., 2017), which is effectively infinitely stable on the time-scale of our SPT experiments (milliseconds).

We thus considered three kinds of interactions. The first class, CBSs, accounts for stable binding of CTCF. The protein remains bound there for a characteristic time $\tau_{CBS} = 1$ min. The second class, TTZs, traps the protein for a much shorter characteristic time τ_{TB} ($\tau_{CBS} \gg \tau_{TTZ}$) (Figure 2A). While trapped, the particle is free to diffuse inside the spherical TTZs that have a radius ϵ_{TTZ} . When the diffusing protein hits the boundary, it can exit the TTZ with probability P_{exit} or be reflected. Upon exit, the particle is placed at a distance a_{TTZ} from the TTZ center and starts diffusing again (Figure 2-Figure Supplement 4B). The model parameters are chosen such that the protein has a probability to escape a zone upon release rather than rebinding immediately. Upon encountering another TTZ (arriving at a distance ϵ_{TTZ} from its center), it has a probability P_{trap} to yet again become trapped. While diffusing inside the TTZ, the protein can transiently bind at any position with Poissonian on-rate k_{on} and then be released with Poissonian off-rate k_{off} . We will discuss a third class of zones, Power-law distributed Trapping Zones (PTZs), in further details below.

We then tested if our model could reproduce our experimental observations. We simulated Brownian diffusion using Euler's scheme (Schuss, 2009) subject to cognate and transient interactions (full details in Materials and Methods). We chose the simulation parameters including diffusion coefficient, total bound fraction and localization error to approximately match our experimental measurements. In the experiments, a CTCF molecule stably bound to a CBS appears to be immobile throughout its trajectory and we therefore simulate CBS binding as immobile subject to 35 nm localization uncertainty. We note that stably bound CTCF will thus be filtered out by our analysis pipeline (Figure 1B), but we nevertheless simulate stably bound CTCF to match the experimental data. We take the size of the TTZs to be $\epsilon_{TTZ} = 200$ nm. Although protein diffusion is Brownian in our model, transient trapping of CTCF in the TTZs faithfully reproduced our experimental observations, including high anisotropy (Figure 2C-D) and a clear peak of anisotropy at ~ 200 nm mean displacements (Figure 2F) and a relatively constant trend in time (Figure 1D; Figure 1-Figure Supplement 3B). The relatively constant value of $f_{180/0}$ in time (Figure 2E) suggests that the interaction time of a

protein with a zone has a broad distribution. Interestingly, we note that the location of the peak in Figure 2G provides information about the size of the TTZ. Our simulations reproduced this behavior at high trapping probability – if we lowered the probability P_{trap} of trapping CTCF in these zones, the diffusion remained anisotropic, but the anisotropy peak largely disappeared (Figure 2F). While we considered a large variety of models and parameter combinations, only models with zones of a fairly well-defined size that trap CTCF with high probability could reproduce the experimental data. We refer to this model as Anisotropic

Diffusion through transient Trapping in Zones (ADTZ).

What is the underlying mechanism of ADTZ? When a protein transiently interacts with chromatin, its dynamics will be governed by rapid reattachment to the release site (Amitai, 2018). Indeed, if a protein is bound to a site for a given time, upon disassociation it is much more likely to re-attach to the same site rather than bind to another site. Protein diffusion therefore is biased toward its starting position. When the characteristic time to return is on the order of the frame-rate of our SPT experiments (milliseconds) the protein will appear to take a step back following a forward step. We call this

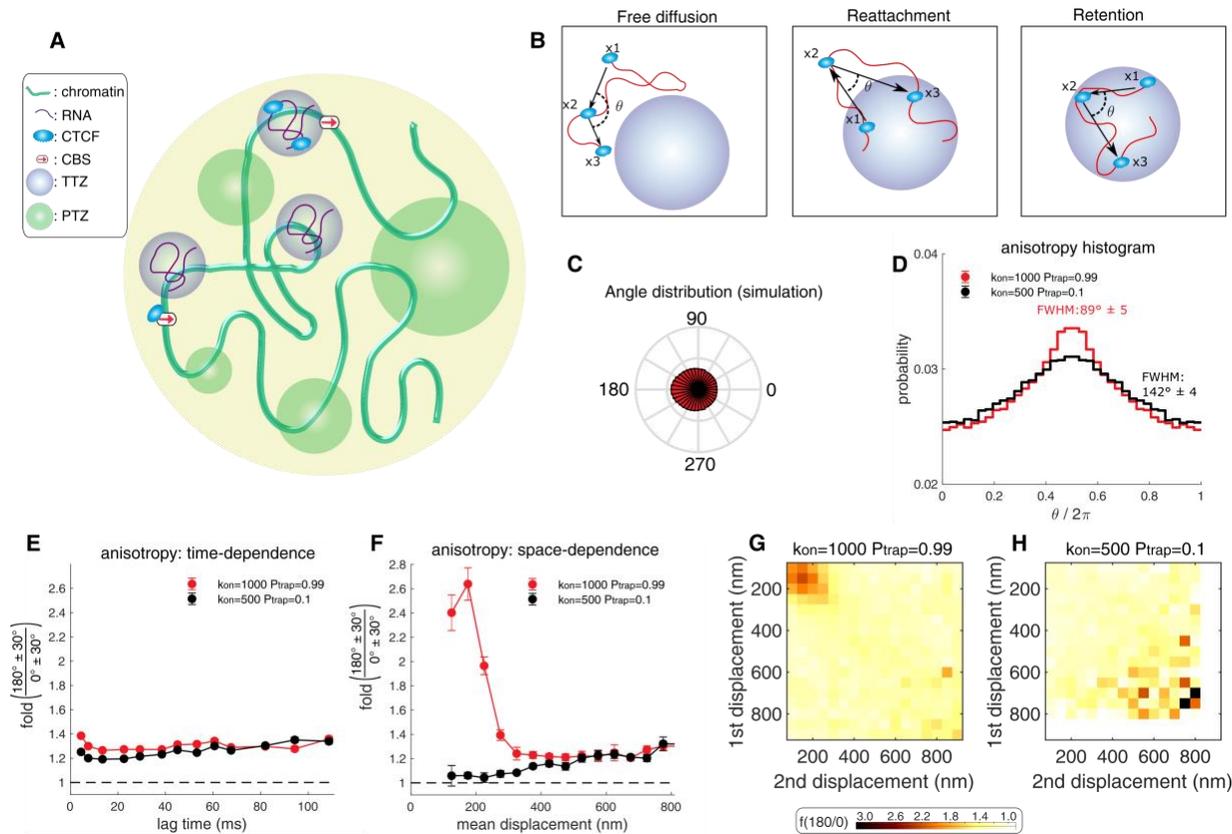


Figure 2. A model wherein CTCF diffusion in the nucleus is governed by its interaction with binding zones can explain the experimental data.

(A) CTCF (light blue ellipse) diffuses inside the nucleus. It interacts with one of three zone types: (1) small, Cognate DNA Binding Sites (CBSs) (red arrow). (2) Transiently trapping zones (TTZs) of size 200nm (purple). (3) Trapping zones of different dimensions, whose size is derived from a power law distribution (Power-law distributed Trapping Zones – PTZs – green; for details in PTZs, please see Figure 3 and below). (B) The protein (light blue) is observed at 3 time points (x_1 , x_2 , x_3) along its stochastic trajectory (red curve). Left: Outside of the zones, CTCF performs free diffusion. The angle of two consecutive steps is uniformly distributed for this mode of motion. Middle: When CTCF leaves a TTZ (x_1), diffuses freely (x_2), and then reattaches back to the TTZ (x_3), the angle distribution is anisotropic. Right: While inside a zone, CTCF can be reflected from its boundary. The angle distribution, in this case, will be anisotropic (C) The angle distribution for a protein performing diffusion in a nucleus containing only CBSs and TTZs. See Table S1 for the simulation parameters. (D) Angle distribution of the trajectory computed from the simulation when the nucleus contains only zones of type 1 and 2 (CBSs and TTZs). CTCF is trapped with high probability ($P_{trap} = 0.99$) when touching the zones (red curve). While inside the zone, it can transiently arrest its motion with a high rate k_{on} (black curve). The protein can be trapped with low probability ($P_{trap} = 0.1$) when touching the zone. While inside the zone, it can transiently arrest its motion with smaller rate. (E) Plot of $f_{180/0}$ vs. lag time averaging over all displacement lengths of the simulation data shown in D. (F) Plot of $f_{180/0}$ vs. mean displacement length averaging over all lag times. (G-H) Anisotropy heatmaps showing $f_{180/0}$ vs. length of the first and second displacement for high binding probability ($P_{trap} = 0.99$, $k_{on} = 1000$) (G) and for smaller binding probability ($P_{trap} = 0.1$, $k_{on} = 500$) (H).

Figure 2-Figure Supplement 1. Overview of the model.

Figure 2-Figure Supplement 2. MSD from model simulations.

mechanism that will contribute to the angle anisotropy *reattachment* (Figure 2B). While the protein is trapped in the TTZ it is reflected from the domain boundary with some probability. We call this mechanism *retention* (Figure 2B). Hence, we suggest that the origin of anisotropy could be two-fold and originate from a combination of *reattachment* and *retention*.

Finally, diffusion is often analyzed using mean square displacement (MSD) analysis. Many nuclear proteins exhibit subdiffusive behaviors (Metzler et al., 2014; Saxton, 2007), where their MSD grows as a power law with time $MSD_i(\tau) = (r_i(t + \tau) - r_i(t)) \sim \tau^\alpha$, with $\alpha < 1$. By analyzing the trajectories, we found that CTCF exhibits anomalous diffusion with an exponent of $\alpha \sim 0.8$ (Figure 1-Figure Supplement 1A; value is sensitive to conditions). The observed anomalous diffusion has been attributed to motion in a crowded environment (Höfling and Franosch, 2013) or diffusion in a hypothetical fractal domain of the nucleus (Bancaud et al., 2009). This motion is often phenomenologically modeled as a continuous time random walk (CTRW), where a particle is trapped for a time drawn from a long tailed power-law distribution (Metzler and Klafter, 2000). To examine the consequences of the ADTZ model on the dynamics of a protein, we computed the MSD of our simulation. We find that the protein appears to perform subdiffusive motion with an anomalous exponent of about $\alpha \sim 0.89$ (Figure 2-Figure Supplement 2; value is sensitive to conditions) due to the rapid reattachment at the release site. The value of α is not universal and depends on model parameters. Interestingly, higher reattachment probability leads to a lower anomalous exponent. A freely diffusing particle confined inside the nucleus has an anomalous exponent $\alpha = 0.97$, which is smaller than 1 because of the nuclear confinement. Thus, the ADTZ model can mechanistically explain the phenomenology of subdiffusive protein motion in the nucleus without invoking power-law dissociation times as is assumed in CTRW (Metzler and Klafter, 2000).

Distinct interactions by CTCF's RBR- and ZF-domains control CTCF diffusion

Our simulations suggest an intriguing mechanistic model (ADTZ) where zones of an unknown nature transiently trap diffusing CTCF with high probability. What could be the nature of these TTZs? We previously demonstrated using super-resolution PALM imaging that CTCF and cohesin form small co-localizing clusters in the nucleus (Hansen et al., 2017). Here, we hypothesize that the TTZs could correspond

to CTCF clusters, since clustering is due to self-interaction and self-interaction might also explain transient CTCF trapping. In our companion paper (Hansen et al., 2018c), we show that CTCF self-association is largely RNA-mediated and DNA-independent consistent with a previous report (Saldaña-Meyer et al., 2014) and that CTCF clustering is substantially reduced after endogenous deletion of CTCF RNA-binding region (RBR; mESC C59D2 Δ RBR-Halo-CTCF). Thus, if the ADTZ model hypothesis is correct and TTZs correspond to CTCF clusters, it should be possible to abolish TTZ-mediated trapping of CTCF in two ways: a) by reducing the trapping probability of CTCF (Figure 2F) or b) by reducing the number of CTCF clusters.

To test these hypotheses and elucidate the molecular mechanism of CTCF diffusion we performed spaSPT for a series of CTCF mutants in mESCs (Figure 3A-F) and U2OS cells (Figure 3-Figure Supplement 1). Whereas a full deletion of CTCF 11 Zinc Finger domain (Δ ZF), which is required for DNA-binding (Hansen et al., 2017; Nakahashi et al., 2013), caused CTCF diffusion to become nearly isotropic, Δ RBR-CTCF remained almost as anisotropic as wt-CTCF at the bulk level (Figure 3A). However, analysis of the fold-anisotropy ($f_{180/0}$) as a function of time (Figure 3B) and space (Figure 3C) revealed two surprising results. First, the anisotropy peak largely, though not entirely, disappeared in Δ RBR-CTCF (Figure 3C-E). Our ADTZ theory (Figure 2) predicts that this can occur in two ways: a) by reducing the number of TTZs (Figure 3-Figure Supplement 2) and b) by reducing the trapping probability (Figure 2A-B), P_{trap} to the TTZs. Clustering (TTZs) is clearly reduced in Δ RBR-CTCF (Hansen et al., 2018c) and it is likely that P_{trap} is too. These results paint a mechanistic picture for the ADTZ model: the RBR domain mediates CTCF clustering, which serve as TTZs that transiently trap diffusing CTCF in zones of a defined size (~ 200 nm), resulting in highly anisotropic CTCF diffusion. However, this points to the second surprising result: Δ RBR-CTCF diffusion is still highly anisotropic (Figure 3A), but with a magnitude ($f_{180/0}$) that is approximately constant in both time and space (Figure 3B-C). Thus, without the RBR, CTCF exhibits approximately scale-free anisotropy similar to PTEFb (Izeddin et al., 2014).

Since Δ ZF-CTCF is essentially isotropic, these results suggest a more complicated model wherein CTCF anisotropy arises through a combination of 2 mechanisms. First, RBR mediated interactions

transiently trap CTCF in small TTZs, resulting in a peak of anisotropy at ~ 200 nm. Second, Zinc Finger (ZF) mediated interactions, perhaps due to transient DNA-interactions, generally trap CTCF without a clearly defined scale. Lending further support to this interpretation, we observed the same ΔZF - and ΔRBR -mutant and wild-type CTCF behavior in human U2OS cells (Figure 3-Figure Supplement 1).

Since our model with TTZs already reproduces anisotropy at a defined scale, we next studied whether adding additional zones could explain the scale-free anisotropy observed without the RBR (Figure 3A-E).

Because the data suggests a “scale-free” anisotropy, which is due to transient DNA-interactions, we added to the nucleus zones of different sizes (Figure 2A) which are sampled from a power law distribution (Figure 3H) ($P_{PTZ}(\epsilon) \sim \epsilon^{-\gamma}$) and term these zones *Power-law Trapping Zones* (PTZs). To reproduce our experiments, we choose the PTZs such that large zones are rarer and trap the protein less efficiently (specifically $P_{trap,PTZ}(\epsilon) \sim \epsilon^{-\delta}$; see Materials and Methods). For very small zones, the theory suggests that for δ of order 1, we would get scale-free anisotropy. We then performed Brownian

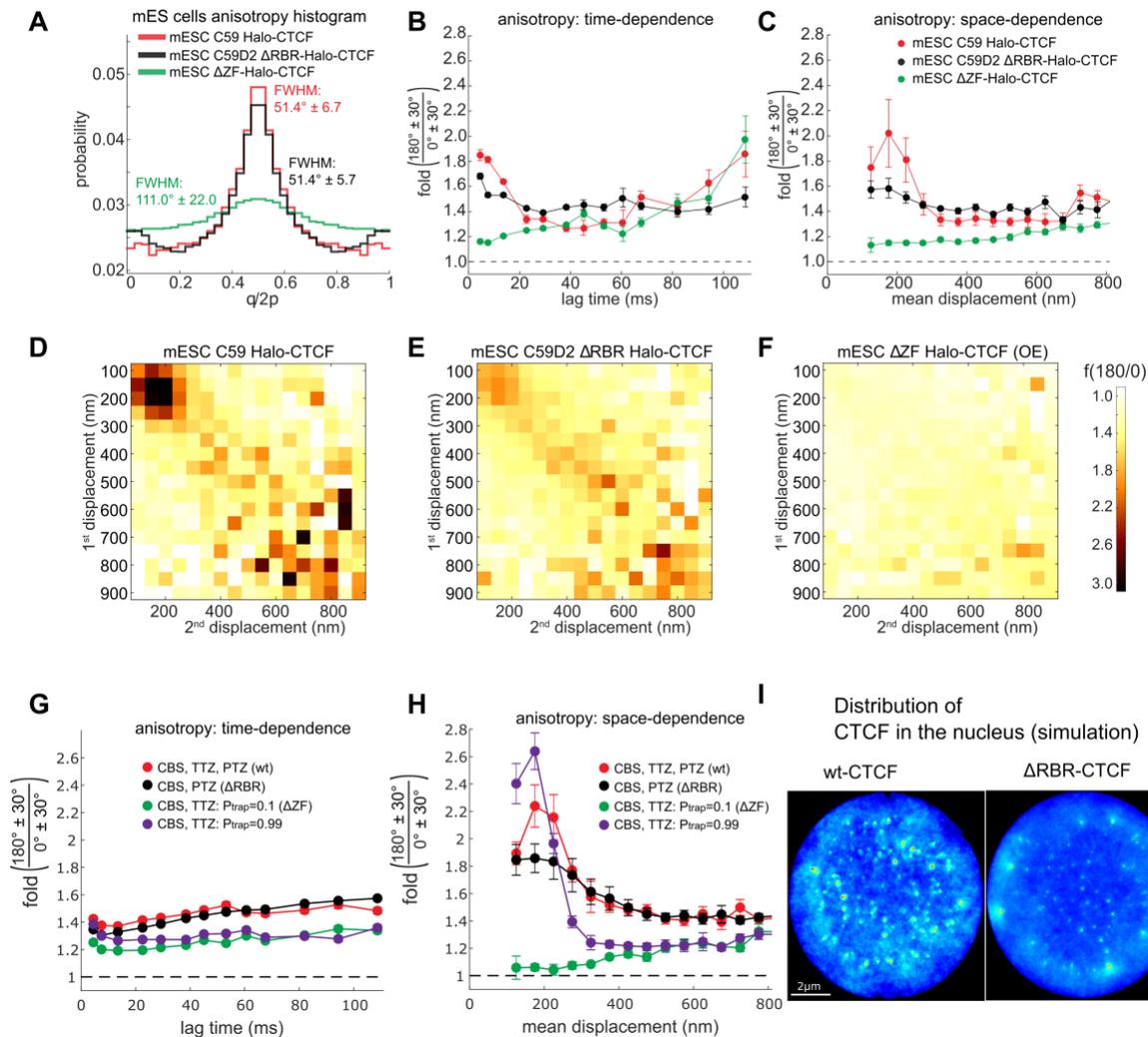


Figure 3. Anisotropy and nuclear distribution of ΔRBR -CTCF.

(A) Bulk angle distribution histograms for CTCF (C59) (red), CTCF mutant that does not have the RBR domain and does not bind RNA (C59D2) (black), CTCF with a deletion of CTCF’s 11 Zinc Finger domain (ΔZF) and which does not bind both DNA and RNA (ΔZF -Halo-CTCF) (green). (B) Plot of $f_{180/0}$ vs. lag time averaging over all displacement lengths. (C) Plot of $f_{180/0}$ vs. mean displacement length averaging over all lag times. (D-F) Anisotropy heatmaps showing $f_{180/0}$ vs. length of the first and second displacement for the three cell lines, averaging over all lag times. **Simulations.** (G-H) Plot of $f_{180/0}$ vs. lag time averaging over all displacement lengths (G) or vs. mean displacement length averaging over all lag times (H). The full model representing the wt, where CTCF interacts with high trapping probability with all three zones (CBS, TTZ, PTZ – see Figure 2A) (red). CTCF interacts with high trapping probability with CBS and PTZ but has a very weak affinity for the TTZ (black) (representing ΔRBR -CTCF). CTCF has a very low affinity to the CBSs and TTZs. There are no PTZs (green) (representing ΔZF -CTCF). CTCF has high affinity towards CBSs and TTZs. There are no PTZs (purple). (I) The distribution of CTCF in the nucleus estimated from the simulation (full model corresponding to wt-CTCF) (left), and the model where CTCF weakly interacts with TTZ (corresponding to ΔRBR -CTCF) (right).

Figure 3-Figure Supplement 1. Anisotropy wild-type and mutant CTCF in human U2OS cells.

Figure 3-Figure Supplement 2. Effect of TTZs abundance on anisotropy.

simulations of an interacting protein where chromatin was represented by a polymer where the majority of its monomers are TTZs (of size 200nm), a smaller fraction is PTZs (with a Power Law size distribution), and a third fraction consists of CBSs. The results recapitulate both the peak of anisotropy at ~ 200 nm and the large ($f_{180/0} \sim 1.5$) and scale-free anisotropy at longer displacements (Figure 3G-H). Thus, the ADTZ model can explain the behavior of wt-CTCF and when we “computationally mutate” the ADTZ-part, the model can also explain the behavior of Δ RBR-CTCF. Finally, we computationally estimated the distribution and clustering of wt-CTCF and Δ RBR-CTCF in the nucleus as it results from our simulations (Figure 3I). Notably, the computationally estimated difference in distribution between wt-CTCF and Δ RBR-CTCF looks qualitatively similar to the distribution observed in super-resolution PALM experiments (Hansen et al., 2018c).

CTCF target search mechanism is RBR-guided

Having elucidated a surprisingly complicated mode of CTCF diffusion, we next asked what the function might be. We reasoned that cells might regulate the diffusion of DNA-binding proteins in order to regulate the efficiency of their cognate target site search. To study CTCF target-search efficiency, we performed spaSPT and Fluorescence Recovery After Photobleaching (FRAP) experiments. First, spaSPT experiments coupled with 2-state model-based analysis (Hansen et al., 2017, 2018b) revealed that the free diffusion coefficient (D_{FREE}) was apparently unaffected, whereas the total bound fraction decreased significantly from $\sim 62.4\%$ to 42.0% (Figure 4A-C). Since our spaSPT experiments capture both non-specifically “bound”

CTCF and specifically bound CTCF (characterized by a $\sim 1-4$ min residence time (Hansen et al., 2017)), we next asked using Fluorescence Recovery After Photobleaching (FRAP) experiments which population was affected. FRAP independently confirmed that the bound fraction was substantially reduced (Figure 4D), but model fitting revealed that the apparent residence time of specifically bound CTCF was not significantly affected, though the population size of specifically bound CTCF was reduced (Figure 4-Figure Supplement 1). Thus, these results demonstrate a surprising function of CTCF RBR: the strength of CTCF specific binding to cognate sites is unaffected by the RBR, but the amount of CTCF engaged in specific binding is strongly impacted. In a simplified 2-state model, the fraction of CTCF specifically bound to chromatin is controlled by the balance between its binding (ON) and dissociation (OFF) rates ($F_{\text{BOUND}} = \frac{k_{\text{ON}}^*}{k_{\text{ON}}^* + k_{\text{OFF}}}$). Within this framework and using our measurements, we find that the RBR increases k_{ON}^* by ~ 2.5 -fold (see Materials and Methods for calculation). The ON-rate is the inverse of the time it takes the protein to find its cognate binding site (search time). Thus, the RBR increases the frequency or rate of CTCF finding a specific DNA-binding site by ~ 2.5 -fold, without affecting CTCF affinity for specific binding sites. This suggests that the ADTZ-diffusion mechanism serves to increase the efficiency of CTCF target search mechanism. Additionally, this suggests that TTZs serve to “load” or “guide” CTCF towards cognate DNA binding sites. Hence, CTCF cognate binding sites may be located within or near TTZs. If this were the case, mechanistically, the TTZs would increase the local concentration of CTCF around its cognate DNA binding sites, thus increasing the local ON-rate.

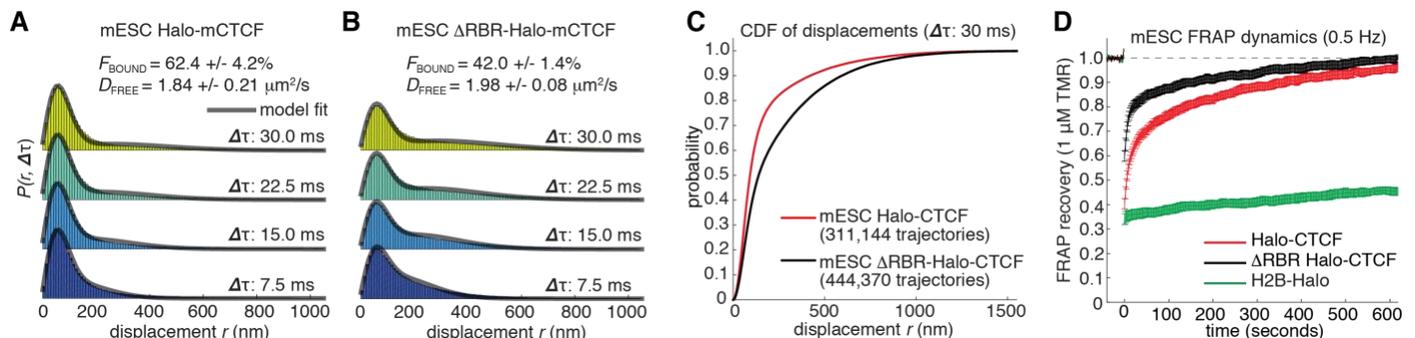


Figure 4. RBR-guided CTCF target search mechanism.

(A-B): spaSPT histograms of displacements for wt-CTCF (A) and Δ RBR-CTCF (B). Raw displacement data for four different lag times are shown with a 2-state Spot-On model fit (bound vs. free) overlaid (Hansen et al., 2017, 2018b). The inferred bound fraction and free diffusion coefficients are shown. Standard errors (+/-) are indicated. (C) Cumulative probability function of displacement lengths ($\Delta t=30$ ms) for the same data as in A-B. (D) Fluorescence Recovery After Photobleaching (FRAP) data for wt-CTCF, Δ RBR-CTCF, and histone H2B (control). Model-fits and inferred residence times are shown in Figure 4-Figure Supplement 1.

Figure 4-Figure Supplement 1. Model-fit to FRAP data.

DISCUSSION

Anisotropic diffusion through transient and weak interactions in the nucleus

It has long been known that diffusion in cells can be anomalous or non-Brownian (Metzler et al., 2014) and anomalous sub-diffusive behavior has previously been reported for DNA (Bronstein et al., 2009), RNA (Golding and Cox, 2006), and proteins (Izeddin et al., 2014). Here, we introduce two new concepts and a new approach for studying the diffusion of nuclear proteins in live cells. Technically, by largely eliminating tracking errors and motion-blur artifacts, spaSPT (Hansen et al., 2017, 2018b) overcomes common technical artifacts in SPT that can bias the analysis, and our analysis pipeline (Figure 1B) crucially filters out chromatin-associated trajectories. While we have focused here on analyzing the angle anisotropy, we anticipate that our analysis has only scratched the tip of the iceberg and we suggest that our approach can usefully be applied to many other nuclear proteins.

Conceptually, first, our work has revealed a new mode of nuclear exploration, which results in anisotropic motion: the Anisotropic Diffusion through transient Trapping in Zones (ADTZ) model (Figure 2). Unlike phenomenological models of sub-diffusion such as CTRW (Metzler et al., 2014), which assumes sub-diffusion is the result of power-law waiting time distributions at a binding site, our model offers a geometrical and mechanistic explanation for CTCF dynamics. Additionally, CTRW is not expected to result in anisotropic motion. We suggest that the primary contribution to wt-CTCF anisotropic dynamics is its interactions within RBR-mediated TTZs of a characteristic ~ 200 -nm size (Figure 5A). Within a zone,

CTCF is allowed to move and participate in transient interactions, which effectively reduce its in-zone diffusion coefficient. The ADTZ model can mechanistically explain sub-diffusion as coming from the retention and reattachment mechanisms (Figure 2). While inside the zone, CTCF is reflected from its boundary with a certain probability (*retention*), resulting in a characteristic residence time within the zone. Once CTCF escapes one zone, it is very likely to rebind to the same zone because of its proximity (*reattachment*), which results in anisotropic behavior. Thus, the ADTZ model suggests that CTCF anisotropy results from the non-uniform distribution of zones inside the nucleus. Mechanistically, our results suggest that the transient trapping zones (TTZs) likely correspond to clusters of CTCF and that the trapping is mediated by CTCF RBR domain. Speculatively, we suggest that if RNA(s) hold together CTCF clusters, weak CTCF-RNA interactions mediated by the RBR-domain may repeatedly trap CTCF. Identifying such putative RNA(s) will be an important future direction. It will be interesting to explore in the future if other proteins exhibit similar mechanisms – especially for proteins that also form clusters or hubs (Mir et al., 2018).

Second, Δ RBR-CTCF still displays anisotropic diffusion with a scale-free magnitude, but CTCF without its Zinc Finger domains exhibit isotropic diffusion (Figure 3). Conceptually, this suggests that protein diffusion is modular, tunable and programmable (Woringer and Darzacq, 2018). Specifically, some protein domains exhibit essentially isotropic diffusion (e.g. the N- and C-terminal domains of CTCF, Halo-3xNLS etc.). CTCF's RBR-domain mediates anisotropic diffusion through an ADTZ-type mechanism and CTCF's 11-Zinc Finger-domain mediates anisotropic diffusion with a scale-free magnitude (Figure 3). Since Δ RBR-CTCF does

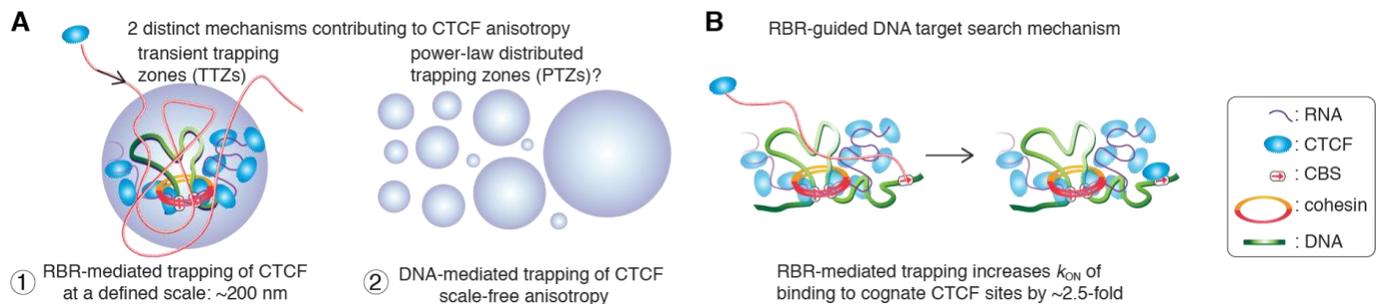


Figure 5. Model.

(A) Two distinct mechanisms contribute to CTCF anisotropy. First, RBR-mediated trapping of CTCF in transiently trapping zones (TTZs) of a characteristic size (~ 200 nm). TTZs are likely to correspond to clusters of CTCF that form in a largely RBR-dependent manner and therefore, presumably, are held together by RNA(s). Cognate Binding Sites (CBSs) reside within the TTZ, as does a piece of DNA strand (green). Second, Δ RBR-CTCF still displays scale-free anisotropy, which is presumably due to non-specific interactions with DNA. We speculatively model this as arising from trapping in power-law distributed zones (PTZs). (B) Target-search mechanism. The RBR increases the apparent rate at which CTCF locates a cognate DNA-binding site. Given that the RBR also mediates CTCF clustering, we speculate that RNA-mediated CTCF clusters, perhaps near loop boundaries, help “guide” diffusing CTCF towards its cognate DNA binding sites (CBSs; DNA is shown in green).

not participate in transient interactions within the 200-nm zones, we suggest that the scale-free anisotropy observed for Δ RBR-CTCF is due to DNA-mediated interactions. This is because this anisotropy is lost when CTCF's DNA-binding domain is lost (Figure 3). We model scale-free DNA-mediated anisotropy as originating from interaction with zones of different sizes, where the larger zones are less sticky (Figure 5A). In these zones (PTZs), which we speculate have a power-law distribution, CTCF experiences transient interactions as in the TTZs. While the origin of the PTZs is unknown at this point, these might correspond to TAD or compartment structures, which vary both in range and size. Moreover, we note that diffusion in a disordered medium could also explain scale-free anisotropy (Havlin and Ben-Avraham, 2002).

We conclude that both RNA- and DNA-mediated interactions govern CTCF diffusion in a modular manner and these interactions result in anisotropy that manifests itself at different scales. Thus, by mixing and matching protein domains with defined diffusion mechanisms, it should in principle be possible to design a protein with a desired diffusion mechanism. This could both be exploited in synthetic biology approaches to engineer proteins with desired target search mechanisms or by the cell during evolution to fine-tune function.

Functional implications of RBR-mediated interactions

Although still speculative at this point, our results suggest a strong link between CTCF clustering, diffusion and target search mechanism (Figure 5B). Specifically, our results suggest that CTCF clusters correspond to TTZs. Since clusters largely form in an RBR-dependent manner (Hansen et al., 2018c), they are presumably held together by one or more species of RNA, as evidenced by co-IP experiments in our companion paper (Hansen et al., 2018c). Moreover, since the RBR increases the rate at which CTCF locates cognate DNA-binding sites by ~ 2.5 -fold, this suggests a model where the CTCF DNA-target search is guided by RNA and where CTCF clusters “guide” diffusing CTCF proteins towards nearby cognate DNA binding sites (Figure 5B). This model significantly changes the view on the mechanism of target search and protein localization in mammalian cells. Rather than the facilitated diffusion picture (alternating between 1D diffusion - e.g. by sliding along the DNA - and 3D diffusion), our model and experimental results suggest that protein localization and target search depend on the

formation of local high concentration zones or clusters, perhaps around cognate binding sites, mediated by RNA. It will be interesting to explore if other proteins exhibit similar “guided” target search mechanisms.

CTCF and cohesin regulate genome organization together (Rowley and Corces, 2018). It is therefore striking that both CTCF (Figure 1) and cohesin (Figure 1H and Figure 1-Figure Supplement 3) exhibit similar ADTZ-type anomalous diffusion. The similar diffusion mechanisms are especially interesting in light of the fact that how and where cohesin is loaded onto the genome remains poorly understood (Busslinger et al., 2017). It will be interesting to explore in the future if TTZs also contribute to topological loading of cohesin on chromatin. Consistent with an RBR-connection between CTCF, cohesin and genome organization, we show in a companion paper that a significant fraction of CTCF loops are lost in Δ RBR-CTCF mESCs (Hansen et al., 2018c). Thus, the same protein domain simultaneously regulates CTCF diffusion, clustering, target search mechanism, and function. Our results highlight the power of SPT, theory, and analysis, when coupled with genome-editing and mutations, as an approach to discover protein domains engaging in important interactions. It will be exciting to apply a similar approach to other nuclear proteins in the future.

Materials and Methods; Supplementary Figures

Detailed methods and Supplementary Figures and Tables have been uploaded as a separate Supplementary Information file.

Acknowledgements

We thank Luke Lavis for generously providing JF dyes, Maxime Woringer for insightful discussions and help with simSPT, Astou Tangara and Ana Robles for microscope assembly and maintenance, Gina Dailey for help and assistance with cloning, and Dr. Kartoosh Heydari at the Li Ka Shing Facility for flow cytometry assistance. We thank Andrew Seeber, Khanh Dao Duc, David McSwiggen, and other members of the Tjian and Darzacq labs for comments on the manuscript. This work was performed in part at the CRL Molecular Imaging Center, supported by the Gordon and Betty Moore Foundation. ASH is a postdoctoral fellow of the Siebel Stem Cell Institute. This work was supported by NIH grants UO1-EB021236 and U54-DK107980 (XD), the California Institute of Regenerative Medicine grant LA1-08013 (XD), by the Howard Hughes Medical Institute (003061, RT), and Massachusetts General Hospital grant 214931 (AA). We thank Sheila S Teves for providing the preprint template.

REFERENCES

- Amitai, A. (2018). Chromatin Configuration Affects the Dynamics and Distribution of a Transiently Interacting Protein. *Biophys. J.*
- Amitai, A., Seeber, A., Gasser, S.M., and Holcman, D. (2017). Visualization of Chromatin Decompaction and Break Site Extrusion as Predicted by Statistical Polymer Modeling of Single-Locus Trajectories. *Cell Rep.* 18.
- Bancaud, A., Huet, S., Daigle, N., Mozziconacci, J., Beaudouin, J., and Ellenberg, J. (2009). Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. *EMBO J.* 28, 3785–3798.
- Bancaud, A., Lavelle, C., Huet, S., and Ellenberg, J. (2012). A fractal model for nuclear organization: current evidence and biological implications. *Nucleic Acids Res.* 40, 8783–8792.
- Banjade, S., and Rosen, M.K. (2014). Phase transitions of multivalent proteins can promote clustering of membrane receptors. *Elife* 3, e04123.
- Bauer, M., and Metzler, R. (2012). Generalized facilitated diffusion model for DNA-binding proteins with search and recognition states. *Biophys. J.* 102, 2321–2330.
- Bénichou, O., Chevalier, C., Klafter, J., Meyer, B., and Voituriez, R. (2010). Geometry-controlled kinetics. *Nat. Chem.* 2, 472.
- Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A.S., Yu, T., Marie-Nelly, H., McSwiggen, D.T., Kokic, G., Dailey, G.M., Cramer, P., et al. (2018). RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.*
- Bronstein, I., Israel, Y., Kepten, E., Mai, S., Shav-Tal, Y., Barkai, E., and Garini, Y. (2009). Transient Anomalous Diffusion of Telomeres in the Nucleus of Mammalian Cells. *Phys. Rev. Lett.* 103, 18102.
- Burov, S., Tabei, S.M.A., Huynh, T., Murrell, M.P., Philipson, L.H., Rice, S. a, Gardel, M.L., Scherer, N.F., and Dinner, A.R. (2013). Distribution of directional change as a signature of complex dynamics. *Proc. Natl. Acad. Sci. U. S. A.* 110, 19689–19694.
- Busslinger, G.A., Stocsits, R.R., van der Lelij, P., Axelsson, E., Tedeschi, A., Galjart, N., and Peters, J.-M. (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* 544, 503–507.
- Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G.M., Cattoglio, C., Heckert, A., Banala, S., Lavis, L., Darzacq, X., et al. (2018). Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* (80-.).
- Elf, J., Li, G.-W., and Xie, X.S. (2007). Probing transcription factor dynamics at the single-molecule level in a living cell. *Science* 316, 1191–1194.
- Feric, M., Vaidya, N., Harmon, T.S., Mitrea, D.M., Zhu, L., Richardson, T.M., Kriwacki, R.W., Pappu, R. V., and Brangwynne, C.P. (2016). Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* 165, 1686–1697.
- Ghirlando, R., and Felsenfeld, G. (2016). CTCF: Making the right connections. *Genes Dev.* 30, 881–891.
- Golding, I., and Cox, E.C. (2006). Physical nature of bacterial cytoplasm. *Phys. Rev. Lett.*
- Grimm, J.B., English, B.P., Choi, H., Muthusamy, A.K., Mehl, B.P., Dong, P., Brown, T.A., Lippincott-Schwartz, J., Liu, Z., Lionnet, T., et al. (2016). Bright photoactivatable fluorophores for single-molecule imaging. *Nat. Methods* 066779.
- Hammar, P., Leroy, P., Mahmutovic, A., Marklund, E.G., Berg, O.G., and Elf, J. (2012). The ⚬lac⚬ Repressor Displays Facilitated Diffusion in Living Cells. *Science* (80-.). 336, 1595 LP-1598.
- Hansen, A.S., Pustova, I., Cattoglio, C., Tjian, R., and Darzacq, X. (2017). CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *Elife*.
- Hansen, A.S., Cattoglio, C., Darzacq, X., and Tjian, R. (2018a). Recent evidence that TADs and chromatin loops are dynamic structures. *Nucleus*.
- Hansen, A.S., Wroninger, M., Grimm, J.B., Lavis, L.D., Tjian, R., and Darzacq, X. (2018b). Robust model-based analysis of single-particle tracking experiments with Spot-On. *Elife* 7, e33125.
- Hansen, A.S., Hsieh, T.-H.S., Cattoglio, C., Pustova, I., Darzacq, X., and Tjian, R. (2018c). An RNA-Binding Region Regulates CTCF Clustering and Chromatin Looping. *BioRxiv*.
- Hassler, M., Shaltiel, I.A., and Haering, C.H. (2018). Towards a Unified Model of SMC Complex Function. *Curr. Biol.* 28, R1266–R1281.
- Havlin, S., and Ben-Avraham, D. (2002). Diffusion in disordered media. *Adv. Phys.* 51, 187–292.
- Höfling, F., and Franosch, T. (2013). Anomalous transport in the crowded world of biological cells. *Reports Prog. Phys.* 76.
- Iwahara, J., and Clore, G.M. (2006). Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature* 440, 1227–1230.
- Izeddin, I., Recamier, V., Bosanac, L., Cisse, I.I., Boudarene, L., Dugast-Darzacq, C., Proux, F., Benichou, O., Voituriez, R., Bensaude, O., et al. (2014). Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. *Elife* 2014, 1–27.
- Kapanidis, A.N., Uphoff, S., and Stracy, M. (2018). Understanding Protein Mobility in Bacteria by Tracking Single Molecules. *J. Mol. Biol.* 430, 4443–4455.
- Liao, Y., Yang, S.K., Koh, K., Matzger, A.J., and Biteen, J.S. (2012). Heterogeneous single-molecule diffusion in one-, two-, and three-dimensional microporous coordination polymers: Directional, trapped, and immobile guests. *Nano Lett.* 12, 3080–3085.
- Lin, Y., Protter, D.S.W., Rosen, M.K., and Parker, R. (2015). Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol. Cell* 60, 208–219.
- Lomholt, M., Ambjörnsson, T., and Metzler, R. (2005). Optimal Target Search on a Fast-Folding Polymer Chain with Volume Exchange. *Phys. Rev. Lett.* 95, 260603.
- Manley, S., Gillette, J.M., Patterson, G.H., Shroff, H., Hess, H.F., Betzig, E., and Lippincott-Schwartz, J. (2008). High-density mapping of single-molecule trajectories with photoactivated localization microscopy. *Nat. Methods* 5, 155–157.
- Mao, Y.S., Zhang, B., and Spector, D.L. (2011). Biogenesis and function of nuclear bodies. *Trends Genet.* 27, 295–306.
- McSwiggen, D.T., Hansen, A.S., Marie-Nelly, H., Teves, S., Heckert, A., Dugast-Darzacq, C., Hao, Y., Umemoto, K.K., Tjian,

- R., and Darzacq, X. (2018). Transient DNA Binding Induces RNA Polymerase II Compartmentalization During Herpesviral Infection Distinct From Phase Separation. *BioRxiv*.
- Merkenschlager, M., and Nora, E.P. (2016). CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. *Annu. Rev. Genomics Hum. Genet.* 1–27.
- Metzler, R., and Klafter, J. (2000). The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.* 339, 1–77.
- Metzler, R., Jeon, J.-H., Cherstvy, A.G., and Barkai, E. (2014). Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Phys. Chem. Chem. Phys.* 16, 24128–24164.
- Mir, M., Stadler, M.R., Ortiz, S.A., Harrison, M.M., Darzacq, X., and Eisen, M.B. (2018). Dynamic multifactor hubs interact transiently with sites of active transcription in *Drosophila* embryos. *BioRxiv*.
- Nakahashi, H., Kwon, K.R.K., Resch, W., Vian, L., Dose, M., Stavreva, D., Hakim, O., Pruett, N., Nelson, S., Yamane, A., et al. (2013). A Genome-wide Map of CTCF Multivalency Redefines the CTCF Code. *Cell Rep.* 3, 1678–1689.
- Persson, F., Lindén, M., Unoson, C., and Elf, J. (2013). Extracting intracellular diffusive states and transition rates from single-molecule tracking data. *Nat. Methods* 10, 265–269.
- Rada-Iglesias, A., Grosveld, F.G., and Papanonis, A. (2018). Forces driving the three-dimensional folding of eukaryotic genomes. *Mol. Syst. Biol.* 14.
- Rhodes, J., Mazza, D., Nasmyth, K., and Uphoff, S. (2017). *Scc2/Nipbl* Hops Between Chromosomal Cohesin Rings After Loading. *Elife* doi: 10.7554/eLife.30000.
- Rice, S.A. (1985). Diffusion-Limited Reactions. *Compr. Chem. Kinet.* 25, 3–46.
- Rowley, M.J., and Corces, V.G. (2018). Organizational principles of 3D genome architecture. *Nat. Rev. Genet.*
- Saldaña-Meyer, R., González-Buendía, E., Guerrero, G., Narendra, V., Bonasio, R., Recillas-Targa, F., and Reinberg, D. (2014). CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, *Wrap53*. *Genes Dev.* 28, 723–734.
- Saxton, M.J. (2007). A Biological Interpretation of Transient Anomalous Subdiffusion. I. Qualitative Model. *Biophys J.* 92, 1178–1191.
- Schuss, Z. (2009). *Diffusion and Stochastic Processes. An Analytical Approach* (New York, NY: Springer-Verlag).
- Slutsky, M., and Mirny, L. a (2004). Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential. *Biophys. J.* 87, 4021–4035.
- Teves, S.S., An, L., Hansen, A.S., Xie, L., Darzacq, X., and Tjian, R. (2016). A Dynamic Mode of Mitotic Bookmarking by Transcription Factors.
- Weber, S.C., and Brangwynne, C.P. (2015). Inverse size scaling of the nucleolus by a concentration-dependent phase transition. *Curr. Biol.* 25, 641–646.
- Woringer, M., and Darzacq, X. (2018). Protein motion in the nucleus: from anomalous diffusion to weak interactions. *Biochem. Soc. Trans.*