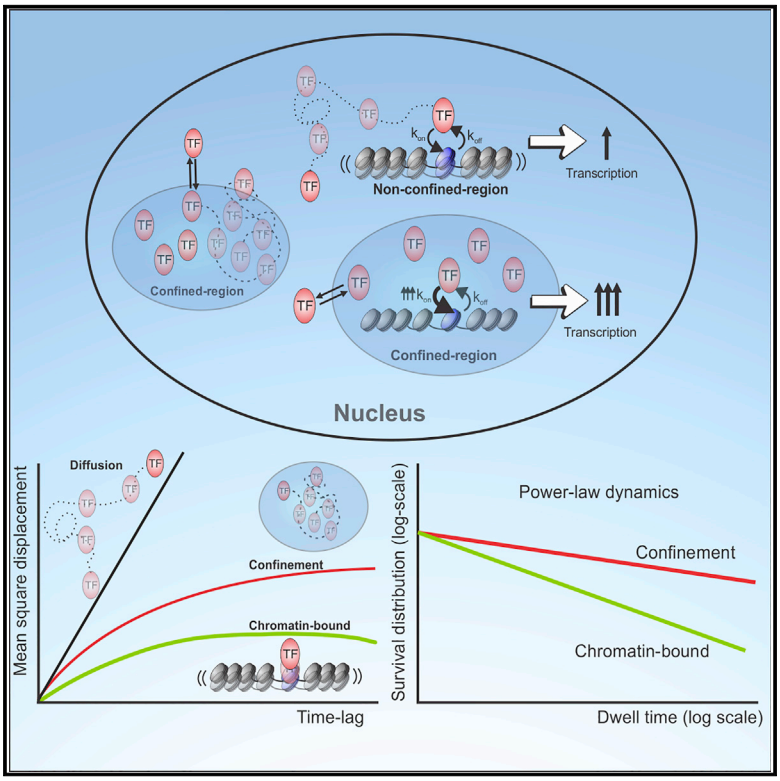
Molecular Cell

Article

An intrinsically disordered region-mediated confinement state contributes to the dynamics and function of transcription factors

Graphical Abstract



Authors

David A. Garcia, Thomas A. Johnson, Diego M. Presman, ..., Susanne Mandrup, Arpita Upadhyaya, Gordon L. Hager

Correspondence

arpitau@umd.edu (A.U.), hagerg@exchange.nih.gov (G.L.H.)

In Brief

Garcia et al. use a systems-level approach to analyze single-molecule tracks of the glucocorticoid receptor. In addition to the known chromatin-bound state, the authors characterized an IDR-mediated, long-lived confined state consistent with liquid condensates that can amplify transcriptional output by increasing the local concentration of TFs at enhancer sites.

Highlights

- TF dynamics were analyzed by combining SMT and a machine-learning-based approach
- GR exhibits two states with limited mobility: chromatin bound and confinement
- Dwell times of both states are distinguishable and follow a power-law distribution
- Confinement is mediated by IDRs and might amplify transcriptional output

Garcia et al., 2021, Molecular Cell *81*, 1484–1498
April 1, 2021
https://doi.org/10.1016/j.molcel.2021.01.013

CellPress

CellPress

Molecular Cell

Article

An intrinsically disordered region-mediated confinement state contributes to the dynamics and function of transcription factors

David A. Garcia,[1,2] Thomas A. Johnson,[1,7] Diego M. Presman,[3,7] Gregory Fettweis,[1] Kaustubh Wagh,[1,2] Lorenzo Rinaldi,[1] Diana A. Stavreva,[1] Ville Paakinaho,[4] Rikke A.M. Jensen,[1,5] Susanne Mandrup,[6] Arpita Upadhyaya,[2,6,*] and Gordon L. Hager[1,8,*]

[1]Laboratory of Receptor Biology and Gene Expression, National Cancer Institute, National Institutes of Health, Bethesda, MD 20893, USA
[2]Department of Physics, University of Maryland, College Park, MD 20742, USA
[3]Instituto de Fisiología, Biología Molecular y Neurociencias (IFIBYNE), CONICET-Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, C1428EGA Buenos Aires, Argentina
[4]Institute of Biomedicine, University of Eastern Finland, Kuopio, P.O. Box 1627, 70211 Kuopio, Finland
[5]Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark
[6]Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA
[7]These authors contributed equally
[8]Lead contact
*Correspondence: arpitau@umd.edu (A.U.), hagerg@exchange.nih.gov (G.L.H.)
https://doi.org/10.1016/j.molcel.2021.01.013

SUMMARY

Transcription factors (TFs) regulate gene expression by binding to specific consensus motifs within the local chromatin context. The mechanisms by which TFs navigate the nuclear environment as they search for binding sites remain unclear. Here, we used single-molecule tracking and machine-learning-based classification to directly measure the nuclear mobility of the glucocorticoid receptor (GR) in live cells. We revealed two distinct and dynamic low-mobility populations. One accounts for specific binding to chromatin, while the other represents a confinement state that requires an intrinsically disordered region (IDR), implicated in liquid-liquid condensate subdomains. Further analysis showed that the dwell times of both subpopulations follow a power-law distribution, consistent with a broad distribution of affinities on the GR cistrome and interactome. Together, our data link IDRs with a confinement state that is functionally distinct from specific chromatin binding and modulates the transcriptional output by increasing the local concentration of TFs at specific sites.

INTRODUCTION

The specific binding of transcription factors (TFs) to regulatory sites embedded within promoter-proximal elements and enhancers guides the assembly of the transcription apparatus and ensures the expression of target genes (Lazar, 2017). Fluorescent imaging of TFs and coactivators in living cells has revealed that they are dynamic and only transiently interact with chromatin targets (Hager et al., 2009).

Single-molecule tracking (SMT) has made it possible to observe individual TF molecules in live cells (Brouwer and Lenstra, 2019). A number of studies have used SMT to explore the kinetics of mammalian TFs (reviewed in Liu and Tjian [2018]) and their interactions with nuclear structures and the nuclear architecture as they search for specific binding sites on the genome (Bénichou et al., 2011; Izeddin et al., 2014; Kent et al., 2020; Normanno et al., 2015; Reingruber and Holcman, 2011). More recent work has revealed that, in addition to diffusion, TF kinetics are indicative of complex interactions beyond specific and nonspecific binding to DNA (Garcia et al., 2020; Hansen et al., 2020; Hipp et al., 2019; Lerner et al., 2020; Reisser et al., 2020; Stavreva et al., 2019). When traveling through the crowded nucleus, TFs are likely to interact with other proteins and coregulators, chromatin, and diverse RNA and species and may also be sequestered in various nuclear compartments. Indeed, previous experiments in mammalian cells have pointed to a plethora of possible interactions in the nucleus (Grünwald et al., 2006; Normanno et al., 2015), but their role in transcription remains unclear.

Nuclear receptors (NRs) are ligand-regulated TFs that recognize and bind their cognate regulatory sites throughout the genome upon activation. Interactions of all NRs are mediated by well-structured DNA-binding domains (DBDs) and one or more activation function domains (ADs), which bind to coactivators and corepressors via protein-protein interactions to regulate gene expression (Dasgupta et al., 2014). The ADs of the NRs

frequently contain intrinsically disordered regions (IDRs) of low-complexity amino acid sequences that assume multiple different conformations (Kumar and Litwack, 2009). The glucocorticoid receptor (GR) is a typical member of the NR family harboring an IDR (Voss and Hager, 2014). While the structures of the GR's DBD and ligand binding domain (LBD) have been elucidated by X-ray crystallography (Bledsoe et al., 2002; Luisi et al., 1991), the structure of its N-terminal domain activation function 1 (AF1) is not well understood, despite being a major region for control of GR's transcriptional activity (Khan et al., 2012; Simons and Kumar, 2013). The fact that ADs of NRs and other TFs are so poorly characterized limits our understanding of their interactions with the Mediator complex and coactivators (Allen and Taatjes, 2015; Reiter et al., 2017).

It was recently demonstrated that the ADs of diverse TFs, including the estrogen receptor, can form heterotypic condensates with the IDR of the MED1 subunit of the Mediator complex *in vitro*, and this process requires the TFs' IDRs (Boija et al., 2018). These IDR-IDR interactions can result in the formation of phase-separated condensates (Alberti, 2017; Banani et al., 2017; Hyman et al., 2014; Shin and Brangwynne, 2017). Such condensates are membraneless micron-scale compartments organized through liquid-liquid phase separation driven by multivalent macromolecular interactions. These compartments are prevalent in eukaryotic cells and are implicated in many biological processes (summarized by Banani et al., 2017).

The recruitment of the transcription machinery at genomic sites is also driven by liquid-liquid phase separation (LLPS) (Boehning et al., 2018; Chong et al., 2018; Hnisz et al., 2017; Lu et al., 2018; Sabari et al., 2018). Many TFs (e.g., FET family TFs, OCT4, SP1, including GR), co-activators (e.g., Mediator and BRD4), and RNAPII contain IDRs, which can drive their phase separation, leading to formation of discrete nuclear foci in mammalian cells sensitive to short-chain aliphatic alcohols, which can dissolve these membraneless structures (Boehning et al., 2018; Cho et al., 2018; Chong et al., 2018; Lu et al., 2018; Sabari et al., 2018; Stortz et al., 2020). Despite sustained interest in the role of IDRs in the formation of macromolecular condensates, their influence on TF diffusion and the consequent impact on transcription are unclear. Moreover, whether the kinetics of an IDR-containing and IDR-less TF are qualitatively different remains to be explored.

Advances in imaging and statistical analysis of TF dynamics are required to distinguish between the diverse diffusive properties of single molecules and link their kinetics to the underlying function(s). Here, we used a machine-learning (ML)-based method called perturbation expectation-maximization (pEM; Koo et al., 2015) to classify individual trajectories of TFs tracked using SMT. Our analysis uncovered two distinct states with limited mobility. One of these low-mobility states accounts for the expected specific GR binding to chromatin, while the other represents a novel confined state, mediated by the IDR regions of the receptor. The dwell times of both chromatin-bound and confined states follow a power-law distribution. In the case of the chromatin-bound population, the power law emerges as a result of the heterogeneity of binding to response elements with different motif strength. On the other hand, the power-law in the confined state likely emerges as a consequence of the broad distribution of effective binding affinities due to IDR-mediated protein-protein interactions. We propose that the confined state can amplify transcriptional output by increasing the local concentration of TFs at specific sites, thus providing a functional link between confinement and gene regulation.

RESULTS

The glucocorticoid receptor exhibits four distinct populations within the nucleus

Imaging studies have revealed a high degree of heterogeneity in nuclear architecture (Finn and Misteli, 2019; Lerner et al., 2020). How the diversity of interactions between TFs and the nuclear environment influences the dynamics of TFs within the nucleus and its subsequent effect on transcription is unclear.

The glucocorticoid receptor (GR) is a ligand-regulated transcription factor. In the absence of ligand, the GR remains inactive and only after ligand binding does the receptor bind chromatin and regulate gene expression (Presman and Hager, 2017), providing a platform to test the behavior of a physiologically relevant TF. We performed SMT experiments under highly inclined and laminated optical sheet (HiLO) illumination (Tokunaga et al., 2008) on HaloTag-fused wild-type GR (GRwt-Halo) (Paakinaho et al., 2017) conjugated with the photo-activatable PA-JF_{549} fluorophore (Grimm et al., 2016) and activated with dexamethasone (Dex), a synthetic glucocorticoid agonist. We imaged continuously using 12 ms exposure times for an optimal balance between fast acquisition, a good signal-to-noise ratio, and minimization of localization noise. The trajectories of localized particles from a representative cell are shown in Figure 1A.

For any particle trajectory, the mean squared displacement (MSD) is a measure of the movement of the particle (in our case a single TF molecule) over time and can be used to elucidate the type of motion that the particle undergoes (Levi and Gratton, 2007). When the MSD of an ensemble of particles is proportional to time [$\mathrm{MSD}(\tau) \equiv \langle |\mathrm{r}(\mathrm{t}+\tau) - \mathrm{r}(\mathrm{t})|^2 = 2\mathrm{n}\mathrm{D}\tau + \mathrm{c} \rangle$; where n is the dimensionality, D the diffusion coefficient and c a constant, Einstein Relation], the particles are said to follow Brownian motion (i.e., simple diffusion). When the MSD is non-linear ($\mathrm{MSD}(\tau) \propto \tau^b$) and scales faster ($b > 1$) or slower ($b < 1$) with time, the particles are said to undergo super-diffusion or sub-diffusion, respectively. Super-diffusion may indicate directed motion, while sub-diffusion implies restrictions to movement such as binding or confinement (Ben-Avraham and Havlin, 2000; Metzler et al., 2014).

To classify the single-molecule trajectories based on their diffusive properties, we applied a systems-level algorithm, perturbation expectation-maximization (pEM, Koo and Mochrie, 2016), which uses unsupervised ML together with Bayesian inference criterion (BIC) (Schwarz, 1978). We first segmented the tracks into 7-frame intervals to remove length effect artifacts and decrease the probability of transitions between diffusive states (see Figures S1A–S1D and STAR methods). This analysis revealed that activated GRwt exhibits four different types of diffusive motion (Figures 1B and 1C). Particle tracks that explore the most space are consistent with diffusive behavior (Figure 1B, GR high mobility states). Indeed, both states resemble 2D Brownian motion based on

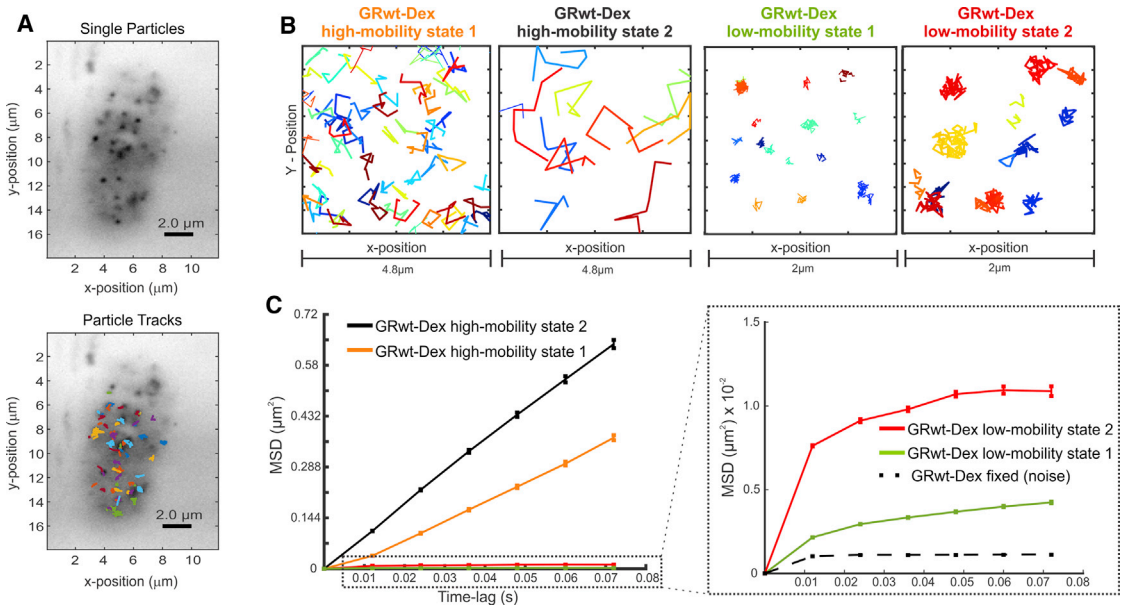


Figure 1. pEM-based MSD analysis reveals four types of GR movement within the nucleus

(A) Representative temporal projection image of an SMT experiment via HiLO imaging (top) with superimposed particle trajectories sampled over 84 ms with continuous acquisition (12 ms exposure, GRwt-Halo) (bottom).

(B) Representative examples of particle trajectories of the observed populations classified by pEM.

(C) MSD versus lag time for the four families of trajectories exhibited by GRwt-Halo conjugated with PA-JF_{549} and treated with 100 nM Dex (15–120 min prior to imaging). The right panel shows a zoomed-in section of the same plot. The noise floor was calculated by imaging GRwt-Halo in fixed cells (GR-fixed, black dotted line). MSDs are calculated from 7-frame tracks. The number of cells/tracks is 109/33,377. Error bars denote standard error measure (SEM).

See also Figure S1.

the linear behavior of their MSD (Figure 1C), with diffusion coefficients of 0.73 ± 0.02 and 2.11 ± 0.05 $\mu m^2/s$ for the high mobility states 1 and 2, respectively. Although these estimates fall within the range previously reported by other methods (Mikuni et al., 2007, 2017; Stasevich et al., 2010), it is surprising to find two distinct diffusive populations. However, as tracking TFs only in two dimensions (i.e., in a single focal plane) poses a limitation for accurate classification of faster diffusion modes since tracks may disappear from the focal plane faster than the tracking rate, we will not investigate these populations further in this work. The two rightmost panels in Figure 1B show trajectories that explore significantly restricted regions of space. MSD analyses revealed that these correspond to two distinct low-mobility states that appear sub-diffusive (Figure 1C, red/green lines), which we will explore further. Both states are distinguishable from GR fixed data under identical acquisition conditions (dashed black line), which represent the sensitivity threshold of the technique. We obtained similar results when tracks were segmented at 15 frames (Figure S1E).

Specific chromatin binding accounts for the first GR low-mobility state

We next analyzed the properties of the two low-mobility states and correlated the MSD analyses to the activity of GR. To confirm the existence of both states over a wider temporal window, we repeated the SMT experiments on GRwt-Halo conjugated with the JF_{549} fluorophore with a longer acquisition interval (200 ms) but similar exposure time (10 ms). As with the continuous 12 ms acquisition conditions, pEM analysis of tracks segmented at 7-frame intervals confirmed two low-mobility populations but only one higher mobility population (Figure S2A). The inability to discern the two high mobility states observed with the fast imaging is expected, as the sampling time of 200 ms is too slow to resolve them. Accordingly, the observed higher mobility state (blue curve in Figure S2A) represents only a small proportion of the tracks (Figure S2B) and will not be further explored. The two low-mobility populations show qualitatively different degrees of restricted mobility (Figure 2A), consistent with the quantification by MSD plots (Figure 2B).

We reasoned that the apparent sub-diffusive (low mobility) populations may arise from the binding of GR to its cognate binding sites on the DNA. To test this, we first took advantage of the mouse C428G (C440G in rat) GR mutation, in which the first zinc finger is disrupted by replacing one of the four key cysteine residues. This mutation produces a drastic conformational change in the receptor's DBD (Figure 2C), and this mutant is unable to bind DNA *in vitro* (Hollenberg and Evans, 1988). We confirmed this result *in vivo* by performing global chromatin immunoprecipitation sequencing (ChIP-seq) in a GRKO cell line (Paakinaho et al., 2019) stably expressing the GR-C428G mutant (see STAR methods). While we were able to detect thousands of binding sites in GRwt (+Dex), peak calling methods found no ligand dependent peaks in GR-C428G-expressing cells (Figure 2D).

MSD analysis of liganded GR-C428G revealed that the low-mobility state 1 found in GRwt was lost, suggesting that this state

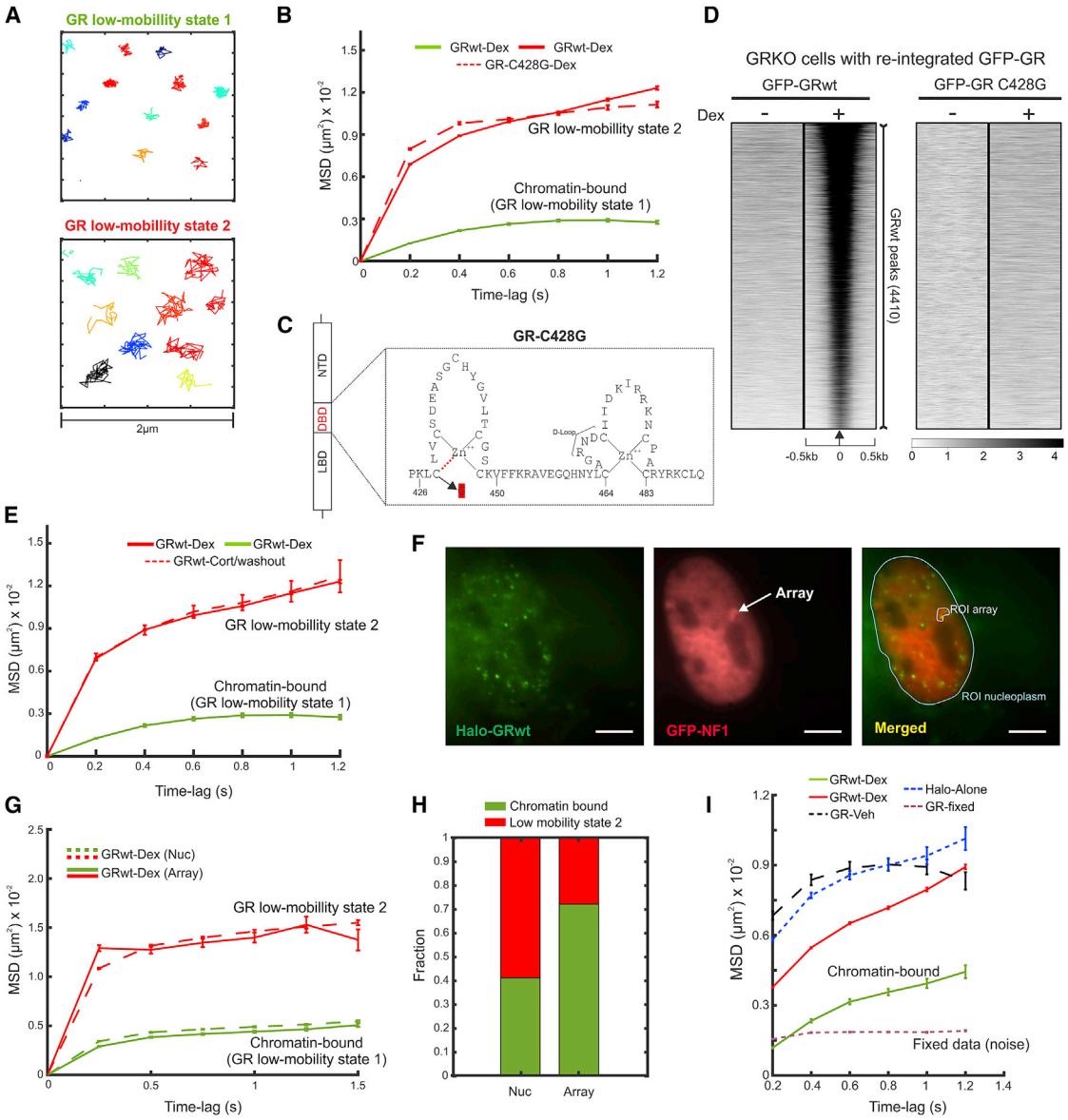


Figure 2. Chromatin binding accounts for one of the GR low-mobility states

(A) Randomly selected particle trajectories of the two low-mobility states of GRwt-Halo conjugated with JF_{549} and treated with 100 nM Dex (15–120 min prior to imaging) found by pEM analysis of 7-frame track segments, with a 200 ms acquisition interval, 10 ms exposure.

(B) MSD versus lag time of GRwt-Halo (solid lines, #cells/#tracks are 70/21,535) and GR-C428G (dashed lines, #cells/#tracks are 52/20,354). MSDs are calculated from 7-frame track segments, with a 200 ms acquisition interval, 10 ms exposure.

(C) Schematic of GR structural domains and location of the C428G mutation (arrow).

(D) Heatmap representation of ChIP-seq from the indicated cell lines, +/− 100 nM Dex for 1 h. Binding intensity is noted below on a linear scale. Heatmaps are sorted based on GRwt binding intensity and normalized for read depth and local tag density.

(E) MSD versus lag time as described in (B) with Dex-treated GRwt-Halo (solid lines, #cells/#tracks are 70/21,535) and 4 h Cort washout (dashed lines, #cells/#tracks are 60/32593).

(F) Representative projection image of Halo-GRwt. GFP-NF1 serves as a marker for the tandem array. ROI, region of interest. Scale bar 5 μm.

(G) MSD versus lag time as described in (B) for the nucleoplasm (#cells/#tracks are 82/7689) or the array (#cells/#tracks are 82/1866), with a 252 ms acquisition interval, 10 ms exposure.

(H) Proportions of two low-mobility states from (G) showing the relative fractions of tracks obtained from the nucleoplasm versus the array.

(I) Weighted MSD versus lag time for GRwt-Dex (solid lines, #cells/#tracks are 70/21,535), HaloTag-alone (blue dashed line, #cells/#tracks are 64/16,819), and GR vehicle (black dashed line, #cells/#tracks are 47/6236). The noise floor was calculated as in Figure 1C. In all cases, error bars denote SEM.

See also Figure S2.

corresponds to specific chromatin binding (Figure 2B). Consistent with this observation, the mouse GR-A465T/I634A monomeric mutant (GRmon) (Presman et al., 2014), which has almost no chromatin binding genome-wide (unpublished data), also lacks this "bound" sub-diffusive population (Figure S2C). Due to statistical limitations, if a small number of tracks (less than 5%) exhibit bound dynamics they will not be observed with the implemented classification. To rule out any artifacts of track segmentation, we analyzed tracks segmented at 30-frame intervals and found similar results (Figures S2D and S2E).

To further confirm the identity of GR low-mobility state 1, we analyzed the effect of inactivating GRwt by hormone withdrawal. We first activated the receptor with its natural ligand, corticosterone (Cort), and then inactivated it by culturing the cells with hormone-free media for several hours before imaging. Under these conditions, the receptor remains in the nucleus but is mostly inactive (Stavreva et al., 2015, 2009). While GR activated with Cort shows a similar MSD profile as the GR-Dex complex (Figures S2F and S2G), washing out the hormone from the GRwt removes the "bound" population (Figure 2E; Figure S2H).

The cell line used in these experiments harbors a tandem gene array that contains ~200 copies of a GR-responsive promoter structure (Stavreva et al., 2009), thus providing many specific sites for GR to bind at a discrete region of the nucleus. This array can be visualized using GFP-NF1 (Nuclear Factor 1) in live cells, while GR dynamics are tracked at the single-molecule level (Figure 2F). We defined a region of interest around the array using the GFP-NF1 fluorescence. We then assigned Halo-GRwt tracks that either originated in or visited this region as "array tracks." Tracks were terminated when the particle left the array. We applied our MSD analysis to the array tracks and compared these with tracks of Halo-GRwt particles throughout the rest of the nucleoplasm. We found that, while tracks from both the nucleoplasm and the array exhibited similar mobility states, the low-mobility state 1 (Figures 2G and 2H) was significantly enriched at the latter region. This observation supports our hypothesis that the low-mobility state 1 represents chromatin binding.

To further link GR low-mobility state 1 to specific chromatin binding, we performed two additional controls. As unliganded GR constantly shuttles between nucleus and cytoplasm (Vandevyver et al., 2012), the (small) nuclear population of unliganded GR molecules serves as the vehicle control (GR-Veh). Analysis of the posterior-weighted MSD (see STAR methods) shows that, as expected from its lack of DNA binding ability (GR-Veh in Figure 2I), unliganded GRwt lacks the GR "bound" state as does a control HaloTag (Figure 2I). In any SMT experiment, there will be a proportion of freely diffusive molecules that will appear as sub-diffusive because of the broad distribution of single-molecule mobility and the anisotropy of the nucleoplasm (Banaz et al., 2019; Mazza et al., 2012). Accordingly, for both the HaloTag-alone and the nuclear population of untreated GR, we detected a sub-diffusive population, but with a larger effective mobility compared to the low-mobility state 2 for activated GR (Figure 2I, cf. red line versus blue and black). In fact, by calculating the instantaneous velocity distribution for this state across the three different conditions (Figure S2I; STAR methods), we find significant differences between GRwt-Dex and HaloTag alone and between GRwt-Dex and GR-Veh but no differences between HaloTag alone and GR-Veh.

Thus far, our data indicate that GR exhibits two distinct sub-diffusive states. The most restricted state (termed the chromatin-binding state henceforth) is associated with specific binding of GR to chromatin. Accordingly, cellular conditions that preclude specific binding of GR exhibit only the sub-diffusive state with larger mobility.

IDRs account for the second GR low-mobility state

Recent studies have shown that protein-protein interactions mediated by IDRs form high-density aggregates, likely phase separated liquid-liquid droplets, in living cells (Cho et al., 2018; Chong et al., 2018; Sabari et al., 2018; Shin and Brangwynne, 2017). In the nucleus, these aggregates (also referred to as nuclear foci) might form through interactions with chromatin and have been linked to transcription (Boija et al., 2018; Lu et al., 2020). GRwt, upon ligand binding, distributes nonhomogeneously throughout the nucleus, forming regions with a higher concentration of receptor molecules that are compatible with liquid condensates (Stortz et al., 2020). Using live-cell confocal microscopy, we also observed fusion events between these structures, suggestive of an LLPS process (Figures S3A and S3B). We thus hypothesized that IDR interactions between GR molecules and other interacting proteins within the nucleus can create a local region of constrained motion or "confinement" where GR kinetics will be different from the rest of the nucleus and may explain our observations of the low-mobility state 2 for GR.

To test this hypothesis, we removed the entire N-terminal domain (NTD) of GR, which is enriched in IDRs (Figure 3A). This deletion mutant is referred to hereafter as GR407C, as it only has the last 407 C-terminal amino acids (Meijsing et al., 2007). MSD analyses on the SMT data collected from this mutant revealed a complete loss of the second low-mobility state (henceforth termed "confinement") while retaining low-mobility state 1, associated with chromatin binding (Figure 3B; Figures S3C and S3D). This observation suggests that "confinement" may be a result of protein-protein interactions associated with IDRs.

ChIP-seq analyses of the GR407C mutant in the presence of Dex shows a significant reduction in binding compared to GRwt (Figure 3C; Table S1, 898 versus 4,410 peaks, respectively). Representative genome browser track examples are shown in Figure S3E. This loss of chromatin occupancy together with the loss of the confined sub-diffusive state suggests that IDR-mediated confinement may facilitate GR binding and contribute to its activity. Interestingly, the GR407C mutant is able to bind both closed and pre-accessible chromatin based on ATAC-seq data from GRwt (Figure 3D), suggesting that the mutant receptor can still recruit remodeling factors (Fan et al., 2005; John et al., 2008), possibly through its LBD. Moreover, *de novo* motif analysis shows that the IDR-less GR mutant binds preferentially to more restrictive glucocorticoid response elements (GREs) motif sequences compared to the wild-type receptor (Figure 3E). In fact, the overall distribution of log-odds motif scores for GR407C is narrower compared to GRwt (Figure 3F), suggesting that GR407C binds a more homogeneous population of GREs in relation to its consensus sequence.

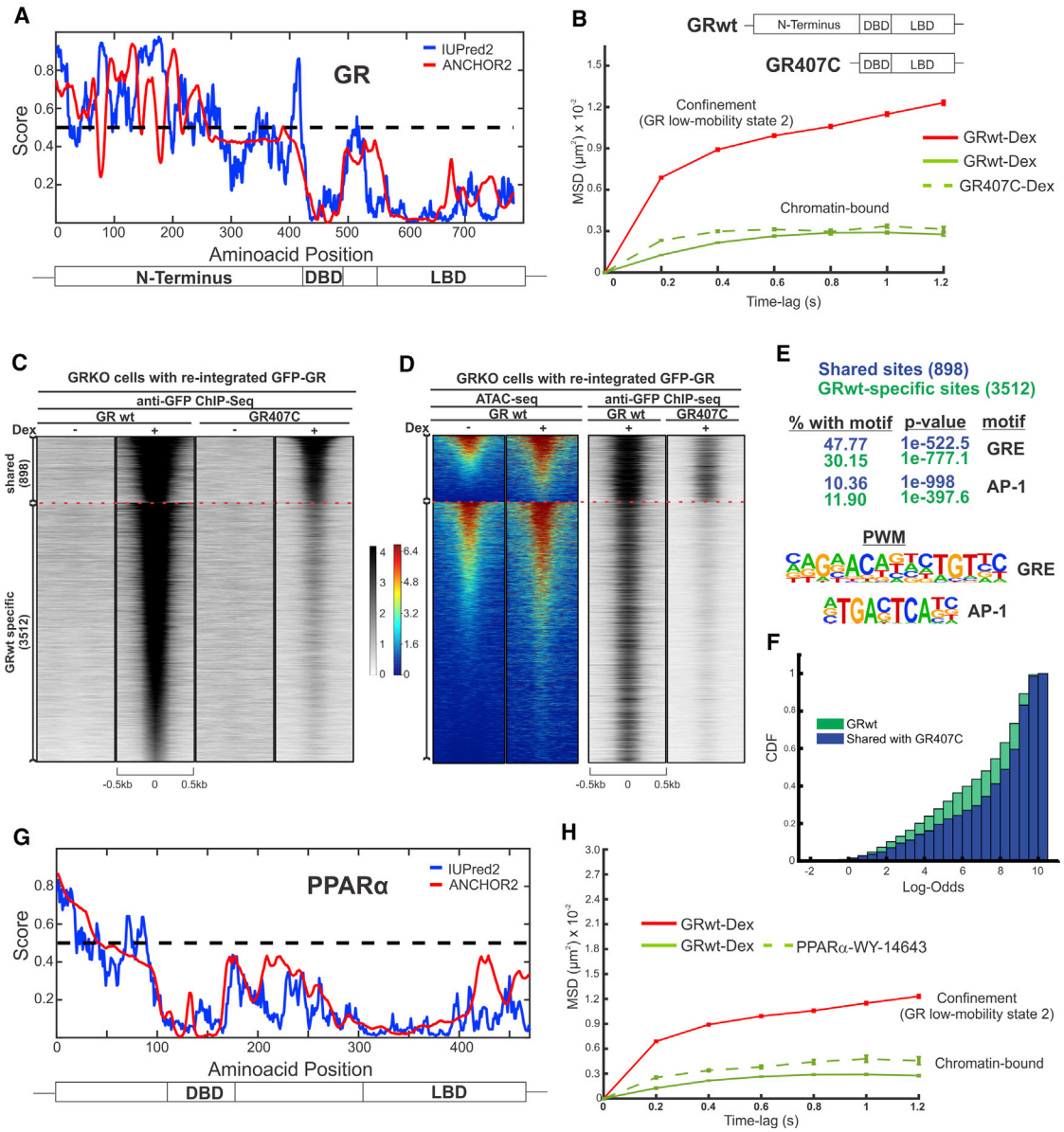


Figure 3. Interactions mediated by IDRs lead to confined diffusion of single TF molecules

(A and G) Plot of inherent protein disorder probability due to a lack of intrachain interactions as predicted by IUPred2A (blue) and ANCHOR (red) models for GR (A) and PPARα (G). The y axis denotes probability (0–1), and the x axis denotes amino acid position. Regions that have a score exceeding 0.5 (dashed line) are classified as disordered regions.

(B and H) MSD versus lag times of GRwt-Halo (solid lines, #cells/#tracks are 70/21,535), GR-407C (B, #cells/#tracks are 60/37,662), or PPARα (H, #cells/#tracks are 60/12,237, respectively) treated with 100 nM Dex or 10 μM WY-14643 (15–120 min prior to imaging). The plot shows MSD of 7-frame tracks, 200-ms acquisition interval, and 10 ms exposure. Error bars denote SEM.

(C) ChIP-seq heatmaps (top) and aggregate plots (bottom) of GFP-tagged GRwt and GR407C stably expressed in GRKO cells, +/− 100 nM Dex for 1 h. Heatmaps are sorted by GRwt binding intensity and clustered by GRwt-specific peaks and GRwt/GR407C-shared peaks, noted on the left. Heatmap binding intensity is noted to the right on a linear scale.

(D) ATAC-seq heatmap (left) and the same +Dex ChIP-seq data as shown in (C) (right) re-sorted within each cluster by No-Dex ATAC signal intensity. ATAC signal intensity is noted at the left of the heatmap on a linear scale.

(E) Motif analyses of each GR binding cluster at shared sites (blue) as compared to GRwt-specific sites (green). The position weight matrix (PWM) of the motifs is shown below.

(F) Distribution of log-odds of a GRE motif at shared sites (blue) and GRwt-specific sites (green). CDF, cumulative distribution function. The x axis represents bins of log-odds. Comparisons using the two sample Kolmogorov-Smirnov test ($p < 0.01$) are shown.

See also Figure S3 and Table S1.

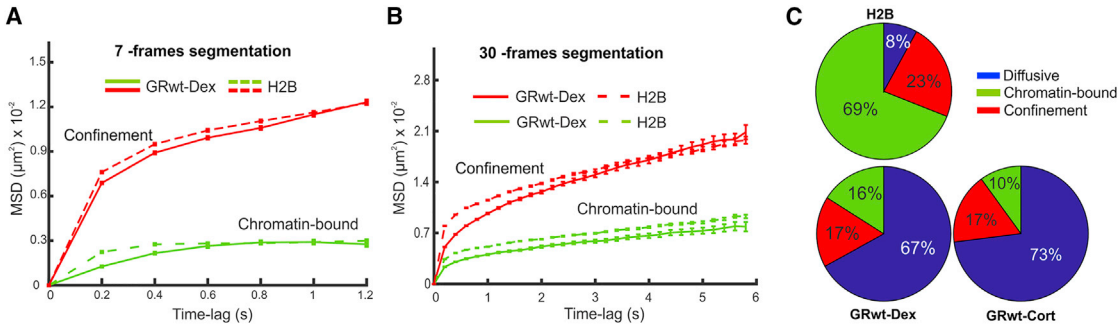


Figure 4. Histones exhibit both confined and chromatin-bound populations

(A and B) MSD versus lag time of GRwt-Halo treated with 100 nM Dex (15–120 min prior to imaging) (solid lines, #cells/#tracks are 70/21,535) and untreated HaloTag-H2B (dashed lines, #cells/#tracks are 70/27,218, respectively). Plots were obtained from 7-frame (A) or 30-frame (B) track segmentation, 200 ms acquisition interval, and 10 ms exposure. Error bars denote SEM.

(C) Pie charts showing percentage of the different diffusive states for GRwt (Dex), GRwt (Cort), and H2B. For 12 ms acquisition, #cells/#tracks are 100/20,000 for H2B, 109/33,377 for GRwt-Dex, and 101/22,182 for GRwt-Cort. For 200 ms acquisition, #cells/#tracks are 70/27,218 for H2B, 70/21,535 for GRwt-Dex, and 65/35,103 for GRwt-Cort.

See also Figure S4.

To further test the proposed relationship between IDRs and the ''confinement'' population, we analyzed peroxisome proliferator-activated receptor alpha (PPARα), another member of the nuclear receptor superfamily (Zhao et al., 2019). This protein naturally has a short IDR region (Figure 3G) and does not form foci (Figure S3F). Consistent with our IDR-mediated confinement hypothesis, diffusive state analysis on the SMT data collected with HaloTag-PPARα stimulated with the agonist WY-14643 only shows the chromatin-bound population and an absence of the confinement population (Figure 3H; Figure S3G).

Taken together, our data suggest that IDRs, which have been implicated in the formation of condensates that further compartmentalize the nucleoplasm (Shin and Brangwynne, 2017), also lead to sub-diffusive behavior of GR in the nucleus (confinement), which is not directly related to specific chromatin binding.

Histone H2B also exhibits both confinement and chromatin-bound populations

It has been proposed that chromatin itself exhibits LLPS (Gibson et al., 2019; Sanulli et al., 2019). In fact, core histones are disordered proteins and histone tails that are exposed in nucleosomes are classified as IDRs (Peng et al., 2012), suggesting that histones may also exhibit similar diffusive dynamics. Hence, we examined the dynamics of H2B using SMT. The MSD analysis of H2B (Figure 4A) shows two low-mobility states that are almost indistinguishable from the GR confinement and chromatin-bound states.

The lowest sub-diffusive state of H2B is remarkably similar to the state of the GR subpopulation arising from chromatin binding (Figure 4A). This is to be expected, as the dynamics of H2B incorporated into nucleosomes (i.e., chromatin) should mirror the dynamics of proteins directly bound to them. Furthermore, the second low-mobility state of H2B (Figure 4A, red line) is compatible with the sub-population of GR that showed IDR-mediated confinement. We obtained similar results with tracks segmented at 30-frame intervals (Figure 4B).

Using the population fraction estimates of the different diffusive states, we can calculate the relative proportions of tracks corresponding to confinement, chromatin binding, and diffusion (Figure S4; STAR methods). Fast acquisition SMT data (12 ms) can be used to estimate the relative proportions between the high mobility (diffusive and likely unbound) and the low-mobility or ''all bound'' trajectories (Koo and Mochrie, 2016). We note that these estimates will be influenced by the diffusion coefficient of the particle and the fraction of faster diffusing proteins will likely be underestimated, especially due to tracking only in two dimensions (z-plane ~300 nm). Furthermore, most diffusive particles will remain in the focal plane for less than 7 frames, and therefore the diffusion population is significantly underestimated. Despite these intrinsic limitations, for H2B, the immobile state dominates the dynamics as expected, in stark contrast with GRwt, which exhibits a larger diffusive population (Figure 4C). Longer-acquisition-interval SMT data (200 ms) provide the most reliable way to

Table 1. Fitting parameters for model selection

Protein	Evidence	Delta-BIC1	Delta-BIC2
PPARα	−22.73	−5.23	−54
GR407C-Dex	−23.39	−4376	−1.37
GR407C-Cort	−46.35	−20.12	−46.35
GRwt-Dex	2,118	985	204
GRwt-Cort	3,003	1,677	834
GRwt-Cortwash	800	417	106
GR-C428G-Dex	127	65	53
GRmon-Dex	1,596	744	174
GRwt-Dex confinement	N/A	N/A	107
GRwt-Dex chromatin binding	N/A	N/A	215

Values of the evidence and Delta-BIC1 and Delta-BIC2 for fitting comparison between the power-law and double-exponential models for all the different experimental conditions are shown. N/A, not applicable. Negative value indicates that the double exponential model is a better model for the data in comparison to the power-law model, while a positive value indicates the contrary.

calculate the relative proportions between chromatin binding and confinement (Figures S4B and S4C). We therefore used the 200 ms data to estimate the relative proportions of confined and chromatin-bound population for all the bound fractions identified with the fast acquisition data (Figure S4D). Using the proportions between bound/unbound and confinement/chromatin binding, we calculated the proportion estimates for the three different states—diffusive, confined, and chromatin bound (Figures S4C and S4D). As expected, H2B and GRwt-Dex have different fractions of chromatin-bound populations (Figure 4C, 69% and 16%, respectively) but surprisingly similar confined proportions (23% and 17%, respectively). GRwt-Cort shows a reduction in the chromatin-bound population compared to GRwt-Dex (Figure 4C, 10% versus 16%, respectively), consistent with previous results (Stavreva et al., 2019).

Overall, our data support the idea that histones can exhibit constrained motion due to confinement in addition to their incorporation into chromatin itself.

Dissecting the origins of the power-law distribution in SMT dynamics

In addition to providing information on diffusive properties, SMT experiments allow us to calculate dwell-time distributions, which represent the time that a protein stays "bound" (Paakinaho et al., 2017; Presman et al., 2017). Most TFs have been described to exhibit a bi-exponential survival distribution, with the longer time constants representing specific binding (reviewed in Goldstein and Hager, 2018). However, alternative multi-exponential models (Hipp et al., 2019; Reisser et al., 2020) or power law may better describe survival distributions (Garcia et al., 2020; Normanno et al., 2015; Stavreva et al., 2019).

To determine how the different types of trajectories from TFs contribute to the dwell-time distribution, we performed long-exposure SMT experiments. We imaged HaloTag-GRwt, -GR-C428G, -GR407C, and -PPARα with 500 ms exposure/acquisition time (see STAR methods for details), using HaloTag-H2B as a reference for photobleaching correction and BIC for model fitting (Garcia et al., 2020) (see STAR methods). Upon either Dex (Figure 5A) or Cort (Figure S5A) stimulation, the dwell-time distribution of GRwt is better explained by power-law behavior compared to bi-exponential or multi-exponential models (see Table 1 for statistics). Interestingly, the survival distributions of GR-C428G (Figure 5B), GRmon (Figure 5C), or GRwt after washing out the hormone (Figure S5B) also show a power law, indicating that specific chromatin binding is not fully responsible for this type of distribution.

Strikingly, the absence of IDRs, either from PPARα (Figure 5D) or the IDR deletion mutant GR407C (Figure 5E; Figure S5C), results in a change in survival distributions to a bi-exponential behavior. The average residence times for the slowest component (Figure S5D) are similar to those previously reported by 3D orbital tracking (Stavreva et al., 2019). Hence, it appears that IDRs might be responsible for the power-law distribution. To test this hypothesis, we took advantage of our MSD analysis to split tracks into chromatin-bound and confined tracks (Figure 5F, see STAR methods), allowing us to independently analyze the dwell-time distribution of each low-mobility state.

If confinement were solely responsible for power-law behavior, then the chromatin-bound population should exhibit exponentially distributed dwell times consistent with chromatin binding (Garcia et al., 2020). However, contrary to our original hypothesis, both confinement and chromatin binding remain power-law distributed (Figure 5G), with the confined population exhibiting overall longer dwell times than chromatin binding for GRwt. We also found that the confined population of GRmon and GRwt have different survival distributions (Figure 5H), with GRwt exhibiting longer dwell times, likely reflecting the different interacting partners inside the confined region.

Overall, it appears that some properties of GRwt-chromatin interactions must account for the power-law behavior in contrast to the bi-exponential distribution seen for PPARα and GR407C. One plausible explanation could rely on the differences in heterogeneity of binding affinities between GRwt and PPARα/GR407C. The GR407C mutant binds to a narrower array of GRE motif sequences as illustrated by the distribution of log-odds motif scores as compared to GRwt (Figure 3F), suggesting that GR407C binds to less heterogeneous GREs. In the case of PPARα, it has been reported that it binds to relatively fewer sites compared to GRwt (Ratman et al., 2016), which might be indicative of less heterogeneity in binding. Taken together, our data suggest that, when the heterogeneity in binding affinity is low, chromatin binding will follow a bi-exponential distribution.

Finally, we wondered whether we could correlate our findings with transcriptional activity. The GRA465T mutant (GRdim) (Presman et al., 2016) is an extremely poor transcriptional activator, even though it binds to ~85% of GRwt's cistrome (Lim et al., 2015). Like GRwt, the GRdim SMT data show both confined and chromatin-bound populations (Figures S5E and S5F). By analyzing the dwell-time distributions, we discovered that, while confinement dynamics are almost identical for GRwt and GRdim, there is a difference in their chromatin-bound populations (Figure 5I). Specifically, our observations show that GRwt has longer binding events than GRdim, which could explain why the mutant receptor is not a good transcriptional activator.

DISCUSSION

Studying TF dynamics at the single-molecule level inside living cells is a sensitive approach that can unravel complex diffusion and binding kinetics of TFs as they locate and bind to their genomic targets. By analyzing single-molecule trajectories of TF using a ML-based method and classifying their residence times based on their kinetic profiles, we show that TF "binding" is composed of at least two distinct subpopulations, one reflecting chromatin binding, and a newly identified subpopulation that arises from IDR-IDR interactions and appears transiently confined.

In support of the notion that the most restricted sub-diffusive population arises from chromatin binding of TFs, we showed (1) loss of this population in GR mutants that do not bind to chromatin (Figures 2B and S2C), (2) its disappearance upon inactivation of the wild-type receptor by hormone withdrawal (Figure 2E), (3) an increase of this population at the tandem array that is expected to show enhanced binding (Figure 2H), and (4) the

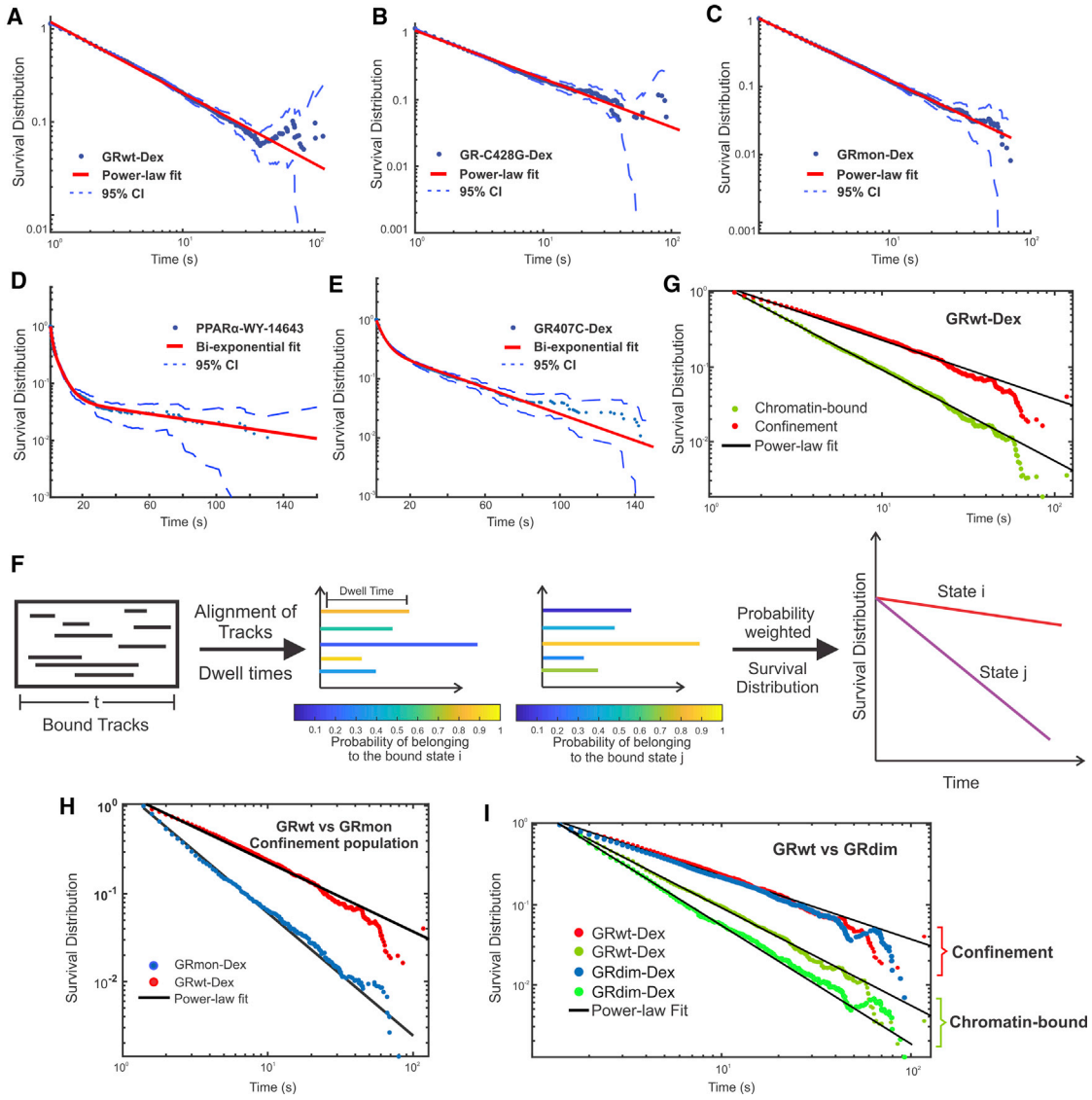


Figure 5. IDR-mediated interactions affect the distribution of dwell times

(A–E) Survival distribution fit to a power law for GRwt (Dex) (A, #cells/#tracks are 70/21,535), GR-C428G (B, #cells/#tracks are 52/20,354), and GRmon (Dex) (C, #cells/#tracks are 87/19,822). Survival distribution for PPARα (D, #cells/#tracks are 60/12,237) and GR-407C (E, #cells/#tracks are 60/37,662) fit to a bi-exponential. Fits are shown in red, 95% CI of the empirical survival distributions calculated using Greenwood's formula are indicated with dashed lines, and data points are shown as solid circles.

(F) Schematic pipeline for splitting chromatin-bound and confined tracks. Tracks are classified based on the posterior probability to belong to a particular state, which is then used to calculate the weighted dwell-time distribution for each binding state.

(G) GRwt (Dex) survival distribution for trajectories belonging to confinement (red) and chromatin binding (green) states fit to power laws (solid lines, #cells/#tracks are 70/21,535).

(H) Survival distributions of the confined population for GRwt (Dex) (red) and GRmon (Dex) (blue) fit to power laws (solid lines, #cells/#tracks are 87/19,822).

(I) Survival distributions of the confined (red) and chromatin-bound (green) population of GRwt (Dex) and confined (blue) and chromatin-bound population (bright green) of GRdim (Dex, #cells/#tracks are 80/30,794). Solid lines show power-law fits. In all cases, error bars denote SEM.

See also Figure S5.

identical constrained behavior of single H2B molecules likely incorporated into nucleosomes (Figures 4A and 4B).

The ML-based analysis identified a second novel population of trajectories that exhibits constrained but higher mobility when compared to chromatin binding (Figures 1 and 2). Surprisingly, we observe that this apparent confinement is sustained for longer time intervals than previous observations of transient confinement (Hansen et al., 2020). We showed that this confinement behavior is mediated by interactions between intrinsically disordered regions on the TFs. In support of this, confinement is lost when the IDR is removed as in the GR407C mutant (Figure 3B). Additionally, TFs such as PPARα that naturally present

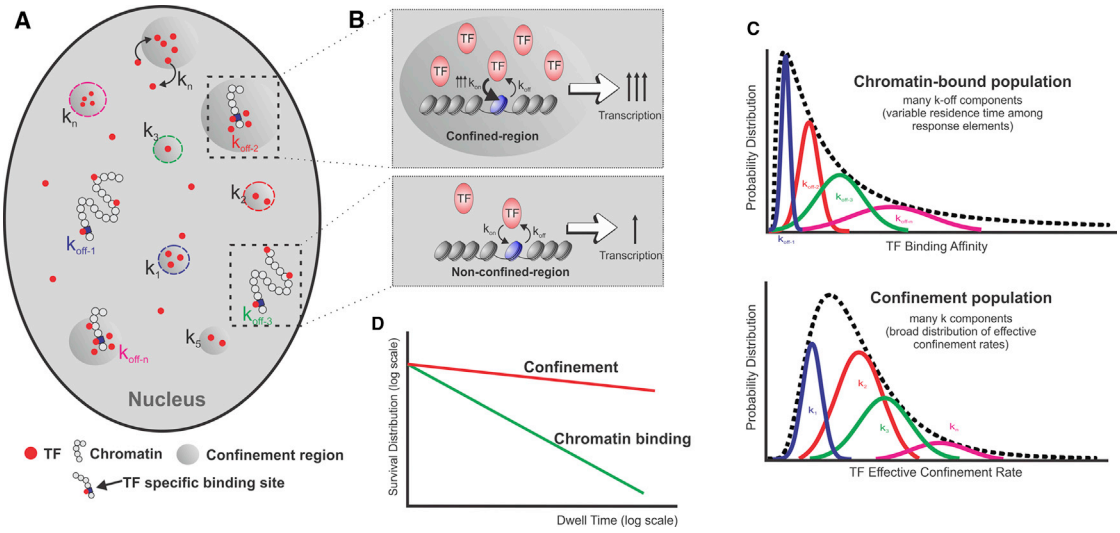


Figure 6. Proposed model for the modulation of gene expression by confinement and the emergence of power-law dwell-time distributions

(A) Transcription factors (TF, red spots) navigate the nucleoplasm until they find their targets. They can be freely diffusing in the nucleoplasm (isolated red spots), confined in high-density IDR-dependent hubs (shaded areas), or interacting with chromatin either specifically or non-specifically. k_{off}, dissociation rate from chromatin; k, dissociation rate from the confined region.

(B) Confined regions concentrate TFs, reducing the search time (i.e., greater k_{on}, thicker arrow). Hence, transcriptional activity is potentiated compared to a gene whose enhancer element is not located in a confined region.

(C) Broadly distributed binding affinities of a TF (dashed line) are composed of binding distributions arising from different chromatin environments and/or motifs (solid lines, top graph). Similarly, a confined transcription factor can exhibit a broad distribution of effective binding affinities related to the time that it takes to escape the confinement region, which depends on the size and physical properties of the hub (solid lines, bottom graph).

(D) For a heavy-tailed distribution of binding affinities and confinement, the dwell-time distribution is expected to follow a power law. In the case of GR, the confinement dwell times are longer than for chromatin binding (as depicted). However, other TFs might present the opposite behavior if they have larger binding affinities.

smaller IDR regions, but otherwise have a full complement of motif preferences, do not show this confined state (Figure 3H). Furthermore, TFs that are not capable of binding their specific recognition sequences (GR-C428G and GRmon) still display a confined state. Finally, tracks within the MMTV array have a significantly higher fraction of the lowest mobility (bound) state. Taken together, these observations suggest that IDR-IDR interactions between GR molecules and GR with other nuclear proteins can create local regions where the receptor exhibits altered diffusivity resulting in the detection of a less constrained subdiffusive state. Our data suggest that IDR-mediated condensates (i.e., nuclear foci [Stortz et al., 2020]) are a good "microscopic" representation of this subpopulation. Whether IDR-IDR interactions are the only mechanism behind foci formation and/or confinement dynamics, or whether the foci are a functional homogeneous entity, remains to be proven.

SMT also enables the measurement of TF dwell times, which are indicative of their binding kinetics. We have recently reported that the dwell-time distributions of GRwt and many other TFs exhibit power-law behavior rather than bi-exponential, suggesting that non-specific and specific binding cannot be simply classified according to their residence times (Garcia et al., 2020). In fact, deviations from the bi-exponential model have been reported elsewhere (Hipp et al., 2019; Normanno et al., 2015; Reisser et al., 2020). Our ML-based techniques allowed us to classify the trajectories of bound proteins into chromatin-bound or confined sates and independently calculate their dwell-time distributions. We discovered that the dwell times of both these populations exhibit power-law behavior (Figure 5G). Theoretical considerations suggest that such distributions can emerge due to a broad distribution of effective binding affinities (Garcia et al., 2020), or from polymer models of chromatin with rapid rebinding of proteins (Amitai, 2018). Moreover, both models predict trapping of transcription factors in regions of the nucleus, consistent with the confinement population found in this study.

In the case of confinement, the broad distribution of effective binding affinities can originate from the heterogeneity in IDR-mediated protein-protein interactions. Equally plausible, TFs with IDRs could adopt a broad distribution of conformations due to the diversity of protein folding (Simons et al., 2014), hence accounting for a broad distribution of binding affinities (Brodsky et al., 2020). Alternatively, but not mutually exclusive, the heavy tailed size distribution of foci (Berry et al., 2015; Onuki, 2007; Ratke and Voorhees, 2011; Shakya et al., 2020) would produce effective binding energies consistent with power-law-distributed dwell times (Garcia et al., 2020).

In the case of chromatin binding, the power-law behavior likely emerges as a consequence of the broad affinity distribution (the k_{off}) of the TF's cistrome due to the heterogeneity of binding to response elements (Figure 6). This could explain why heterologous expression of the tetracycline receptor (TetR) in mammalian cells, where it does not bind specifically anywhere, follows power law throughout the genome but on an artificial array of single-response elements follows a single exponential (Normanno

et al., 2015). In support of this model, GR407C, which does not have a population of confined trajectories and binds less variable response elements (Figure 3F), behaves bi-exponentially (Figure 5E).

The dwell-time distributions of the confined populations of GR mutants can be either similar to (GRdim, Figure 5I) or different from (GRmon, Figure 5H) GRwt. More importantly, the survival distribution of the chromatin-bound population is significantly different between GRwt and GRdim (Figure 5I), which might explain the difference in their transcriptional activities (Jewell et al., 2012; Lim et al., 2015; Presman et al., 2014; Rogatsky et al., 2003). The approach used here thus provides a powerful means to correlate binding affinities of proteins to specific interactions within the cell nucleus.

We also found that H2B exhibits both confinement and chromatin-bound populations. Since free non-nucleosomal histones may electrostatically interact with different IDR-mediated liquid-liquid phase separated aggregates (Peng et al., 2012), a free histone that has not yet been incorporated into nucleosomes can have two different types of kinetics, normal diffusion, or confinement as described above. The kinetics under confinement will be determined by the diffusive properties of the histone in these high-density/high-viscosity regions, which in general will display larger MSD values than the incorporated histones due to the elastic properties of chromatin (Everaers and Schiessel, 2015; Koslover et al., 2010). The kinetics of a histone incorporated into the nucleosome will be dominated by the physical properties of chromatin (e.g., elasticity and thermodynamic properties) and thus set the scale for the mobility of the chromatin-bound fraction.

Our observations suggest that the subpopulation of TFs directly bound to chromatin could serve as "nucleators" for IDR-mediated condensates, consistent with a recent report (Stortz et al., 2020). While these nucleator molecules will exhibit a typical chromatin-bound behavior, their IDR-interacting partners will exhibit confined behavior (Figure 6). Considering the growing evidence for the role of phase separated condensates in transcription (Boehning et al., 2018; Chong et al., 2018; Hnisz et al., 2017; Lu et al., 2018; Sabari et al., 2018), the confined population herein described might be critical for efficient transcription initiation. A higher concentration of TF molecules found in these condensates should increase the on-rates (k_{on}) for TF binding to chromatin (Figure 6B), therefore improving the chances of successful activation of RNA polymerase II resulting in more frequent RNA bursting (Brouwer and Lenstra, 2019; Donovan et al., 2019; Stavreva et al., 2019). A prediction of this hypothesis is that, if one could measure TF binding at a single site within a confined region, then only the k_{on} and not the off-rate (k_{off}) (i.e., residence time) of the TF should change with respect to that in a non-confined region. In other words, confinement can only modulate the k_{on} of TFs, not their residence times, which may ultimately be a combination of the nature of the TF, the choice of ligand, the strength of the TF motif, and the chromatin landscape in which a response element is located (Coons et al., 2019; Johnson et al., 2018; Syed et al., 2020).

Taken together, our data suggest that IDR-mediated confinement is a natural mechanism that many TFs can use to regulate gene expression more efficiently (Figure 6). This is achieved by providing a higher local concentration of TFs at specific genomic sites (i.e., increasing k_{on}) while also effectively decreasing the exploration area within the complex nuclear environment. Although not essential for all TFs, dysregulated phase separation has been implicated in a number of disease conditions (Basu et al., 2020; Darling and Uversky, 2017; Innis et al., 2004; Muragaki et al., 1996; Nakamura et al., 2001). Our results suggest that the modulation of TF mobility by IDR-mediated interactions and the formation of condensates likely contribute to the regulation of transcriptional efficiency.

Limitations of study

Even though we found four distinct populations of GR within the nucleus, our implementation of HiLO only allows us to recover the dynamics of the population of particles with lower mobility. This is because rapidly diffusing, i.e., free TF molecules can quickly move away from the focal plane precluding accurate tracking and recovery of MSD information. While these populations might have some useful information regarding the properties of the nuclear environment, our implementation of HiLO does not allow us to report those. 3D tracking of TFs at very high frame rates is needed to better understand the behavior of the two freely diffusing populations found in this study.

A second limitation is that we do not consider possible transitions between the diffusive states within a single trajectory. This may lead to some tracks not being classified into one of the diffusive states described here (Figure S2B and S3D). Thus, it is possible that several biologically meaningful states with low probability of occurrence remain hidden to the classification scheme implemented in this work. A more robust classification algorithm using Bayesian non-parametrics might reveal these potential transitions and hidden states.

We have speculated that the confinement population described in this work is mediated by protein-protein interactions that could exhibit LLPS compartmentalization (Boehning et al., 2018; Chong et al., 2018; Hnisz et al., 2017; Lu et al., 2018; Sabari et al., 2018). However, a rigorous link between them has not been yet established and needs further investigation.

Finally, fluorophore stability is a limiting factor in all SMT experiments. To better quantify the dwell-time distribution of TFs in different diffusive states, experiments must be performed at low exposure/high laser power for optimal localization precision. However, these conditions limit the fluorescent lifetime of fluorophores to the order of seconds, which is not enough to reliably quantify long binding times in the tails of power-law-distributed dwell times.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Plasmids and cell lines

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2021.01.013.

ACKNOWLEDGMENTS

We thank Tatiana Karpova and David Ball from the Optical Microscopy Core at the NCI, NIH for assistance in imaging and data processing. We thank Luke Lavis (Janelia Research Campus) for providing HALO dyes. This research was supported (in part) by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. D.M.P. was supported by CONICET. V.P. was supported by the Academy of Finland, the Cancer Foundation Finland, and the Sigrid Jusélius Foundation. R.A.M.J. was supported by the Vissing Foundation, the William Demant Foundation, the Knud Højgaard Foundation, the Frimodt-Heineke Foundation, the Director Ib Henriksen Foundation, and the Ove and Edith Buhl Olesen Memorial Foundation. S.M. acknowledges support from the Danish Independent Research Council | Natural Sciences. A.U. acknowledges support from the NCI-UMD Cancer Technology Partnership and the awards NSF PHY 1607645 and NSF PHY 1806903.

AUTHOR CONTRIBUTIONS

D.A.G. conceived the initial project. D.A.G., D.M.P., G.F., K.W., L.R., V.P., and R.A.M.J. performed imaging experiments. D.A.G., D.M.P., G.F., and K.W. analyzed data. T.A.J. performed all ChIP-seq experiments. D.A.G. and L.R. analyzed the *in silico* IDR composition. D.A.G. wrote analysis scripts in MATLAB and implemented the ML method, with feedback and supervision from K.W. and A.U. S.M. supervised R.A.M.J. D.A.G., D.M.P., D.A.S., T.A.J., and A.U. wrote the original draft with subsequent input and editing from all authors. D.A.G., D.M.P., and T.A.J. prepared figures. A.U. supervised data analysis and theoretical models. G.L.H. supervised the entire project.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 31, 2020
Revised: December 8, 2020
Accepted: January 8, 2021
Published: February 8, 2021

REFERENCES

Alberti, S. (2017). The wisdom of crowds: regulating cell function through condensed states of living matter. J. Cell Sci. *130*, 2789–2796.

Allen, B.L., and Taatjes, D.J. (2015). The Mediator complex: a central integrator of transcription. Nat. Rev. Mol. Cell Biol. *16*, 155–166.

Amitai, A. (2018). Chromatin Configuration Affects the Dynamics and Distribution of a Transiently Interacting Protein. Biophys. J. *114*, 766–771.

Banani, S.F., Lee, H.O., Hyman, A.A., and Rosen, M.K. (2017). Biomolecular condensates: organizers of cellular biochemistry. Nat. Rev. Mol. Cell Biol. *18*, 285–298.

Banaz, N., Mäkelä, J., and Uphoff, S. (2019). Choosing the right label for single-molecule tracking in live bacteria: side-by-side comparison of photoactivatable fluorescent protein and Halo tag dyes. J. Phys. D Appl. Phys. *52*, 064002.

Basu, S., Mackowiak, S.D., Niskanen, H., Knezevic, D., Asimi, V., Grosswendt, S., Geertsema, H., Ali, S., Jerković, I., Ewers, H., et al. (2020). Unblending of Transcriptional Condensates in Human Repeat Expansion Disease. Cell *181*, 1062–1079.e30.

Ben-Avraham, Z., and Havlin, S. (2000). Diffusion and Reactions in Fractals and Disordered Systems (Cambridge University Press).

Bénichou, O., Chevalier, C., Meyer, B., and Voituriez, R. (2011). Facilitated diffusion of proteins on chromatin. Phys. Rev. Lett. *106*, 038102.

Berry, J., Weber, S.C., Vaidya, N., Haataja, M., and Brangwynne, C.P. (2015). RNA transcription modulates phase transition-driven nuclear body assembly. Proc. Natl. Acad. Sci. USA *112*, E5237–E5245.

Bledsoe, R.K., Montana, V.G., Stanley, T.B., Delves, C.J., Apolito, C.J., McKee, D.D., Consler, T.G., Parks, D.J., Stewart, E.L., Willson, T.M., et al. (2002). Crystal structure of the glucocorticoid receptor ligand binding domain reveals a novel mode of receptor dimerization and coactivator recognition. Cell *110*, 93–105.

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. *25*, 833–840.

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., et al. (2018). Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell *175*, 1842–1855.e16.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120, https://doi.org/10.1093/bioinformatics/btu170.

Brodsky, S., Jana, T., Mittelman, K., Chapal, M., Kumar, D.K., Carmi, M., and Barkai, N. (2020). Intrinsically Disordered Regions Direct Transcription Factor In Vivo Binding Specificity. Mol. Cell *79*, 459–471.e4.

Brouwer, I., and Lenstra, T.L. (2019). Visualizing transcription: key to understanding gene expression dynamics. Curr. Opin. Chem. Biol. *51*, 122–129.

Cho, W.K., Spille, J.H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, I.I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science *361*, 412–415.

Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G.M., Cattoglio, C., Heckert, A., Banala, S., Lavis, L., Darzacq, X., and Tjian, R. (2018). Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science *361*, eaar2555.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823, https://doi.org/10.1126/science.1231143.

Coons, L.A., Burkholder, A.B., Hewitt, S.C., McDonnell, D.P., and Korach, K.S. (2019). Decoding the Inversion Symmetry Underlying Transcription Factor DNA-Binding Specificity and Functionality in the Genome. iScience *15*, 552–591.

Darling, A.L., and Uversky, V.N. (2017). Intrinsic Disorder in Proteins with Pathogenic Repeat Expansions. Molecules *22*, 2027.

Dasgupta, S., Lonard, D.M., and O'Malley, B.W. (2014). Nuclear receptor coactivators: master regulators of human health and disease. Annu. Rev. Med. *65*, 279–292.

Donovan, B.T., Huynh, A., Ball, D.A., Patel, H.P., Poirier, M.G., Larson, D.R., Ferguson, M.L., and Lenstra, T.L. (2019). Live-cell imaging reveals the interplay

between transcription factors, nucleosomes, and bursting. EMBO J. *38*, e100809.

Everaers, R., and Schiessel, H. (2015). The physics of chromatin. J. Phys. Condens. Matter *27*, 060301.

Fan, H.Y., Trotter, K.W., Archer, T.K., and Kingston, R.E. (2005). Swapping function of two chromatin remodeling complexes. Mol. Cell *17*, 805–815.

Finn, E.H., and Misteli, T. (2019). Molecular basis and biological function of variability in spatial genome organization. Science *365*, eaaw9498.

Garcia, D.A., Fettweis, G., Presman, D.M., Paakinaho, V., Jarzynski, C., Upadhyaya, A., and Hager, G.L. (2020). Power-law behaviour of transcription factor dynamics at the single-molecule level implies a continuum affinity model. bioRxiv. https://doi.org/10.1101/637355.

Gibson, B.A., Doolittle, L.K., Schneider, M.W.G., Jensen, L.E., Gamarra, N., Henry, L., Gerlich, D.W., Redding, S., and Rosen, M.K. (2019). Organization of Chromatin by Intrinsic and Regulated Phase Separation. Cell *179*, 470–484.e21.

Goldstein, I., and Hager, G.L. (2018). Dynamic enhancer function in the chromatin context. Wiley Interdiscip. Rev. Syst. Biol. Med. *10*, e1390.

Grimm, J.B., English, B.P., Chen, J., Slaughter, J.P., Zhang, Z., Revyakin, A., Patel, R., Macklin, J.J., Normanno, D., Singer, R.H., et al. (2015). A general method to improve fluorophores for live-cell and single-molecule microscopy. Nat. Methods *12*, 244–250, 3, 250.

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., and Lavis, L.D. (2016). Bright photoactivatable fluorophores for single-molecule imaging. Nat. Methods *13*, 985–988.

Grünwald, D., Spottke, B., Buschmann, V., and Kubitscheck, U. (2006). Intranuclear binding kinetics and mobility of single native U1 snRNP particles in living cells. Mol. Biol. Cell *17*, 5017–5027.

Hager, G.L., McNally, J.G., and Misteli, T. (2009). Transcription dynamics. Mol. Cell *35*, 741–753.

Hansen, A.S., Amitai, A., Cattoglio, C., Tjian, R., and Darzacq, X. (2020). Guided nuclear exploration increases CTCF target search efficiency. Nat. Chem. Biol. *16*, 257–266.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

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. Proc. Natl. Acad. Sci. USA *116*, 880–889.

Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., and Sharp, P.A. (2017). A Phase Separation Model for Transcriptional Control. Cell *169*, 13–23.

Hollenberg, S.M., and Evans, R.M. (1988). Multiple and cooperative trans-activation domains of the human glucocorticoid receptor. Cell *55*, 899–906.

Hyman, A.A., Weber, C.A., and Jülicher, F. (2014). Liquid-liquid phase separation in biology. Annu. Rev. Cell Dev. Biol. *30*, 39–58.

Innis, J.W., Mortlock, D., Chen, Z., Ludwig, M., Williams, M.E., Williams, T.M., Doyle, C.D., Shao, Z., Glynn, M., Mikulic, D., et al. (2004). Polyalanine expansion in HOXA13: three new affected families and the molecular consequences in a mouse model. Hum. Mol. Genet. *13*, 2841–2851.

Izeddin, I., Récamier, V., Bosanac, L., Cissé, I.I., Boudarene, L., Dugast-Darzacq, C., Proux, F., Bénichou, 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 *3*, e02230.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2017). An Introduction to Statistical Learning: With Applications in R (Springer).

Jaynes, E.T., and Bretthorst, G.L. (2019). Probability Theory: The Logic of Science (Cambridge University Press).

Jewell, C.M., Scoltock, A.B., Hamel, B.L., Yudt, M.R., and Cidlowski, J.A. (2012). Complex human glucocorticoid receptor dim mutations define glucocorticoid induced apoptotic resistance in bone cells. Mol. Endocrinol. *26*, 244–256.

John, S., Sabo, P.J., Johnson, T.A., Sung, M.H., Biddie, S.C., Lightman, S.L., Voss, T.C., Davis, S.R., Meltzer, P.S., Stamatoyannopoulos, J.A., and Hager, G.L. (2008). Interaction of the glucocorticoid receptor with the chromatin landscape. Mol. Cell *29*, 611–624.

Johnson, T.A., Chereji, R.V., Stavreva, D.A., Morris, S.A., Hager, G.L., and Clark, D.J. (2018). Conventional and pioneer modes of glucocorticoid receptor interaction with enhancer chromatin in vivo. Nucleic Acids Res. *46*, 203–214.

Kent, S., Brown, K., Yang, C.H., Alsaihati, N., Tian, C., Wang, H., and Ren, X. (2020). Phase-Separated Transcriptional Condensates Accelerate Target-Search Process Revealed by Live-Cell Single-Molecule Imaging. Cell Rep. *33*, 108248.

Khan, S.H., Awasthi, S., Guo, C., Goswami, D., Ling, J., Griffin, P.R., Simons, S.S., Jr., and Kumar, R. (2012). Binding of the N-terminal region of coactivator TIF2 to the intrinsically disordered AF1 domain of the glucocorticoid receptor is accompanied by conformational reorganizations. J. Biol. Chem. *287*, 44546–44560.

Koo, P.K., and Mochrie, S.G. (2016). Systems-level approach to uncovering diffusive states and their transitions from single-particle trajectories. Phys. Rev. E *94*, 052412.

Koo, P.K., Weitzman, M., Sabanaygam, C.R., van Golen, K.L., and Mochrie, S.G. (2015). Extracting Diffusive States of Rho GTPase in Live Cells: Towards In Vivo Biochemistry. PLoS Comput. Biol. *11*, e1004297.

Koslover, E.F., Fuller, C.J., Straight, A.F., and Spakowitz, A.J. (2010). Local geometry and elasticity in compact chromatin structure. Biophys. J. *99*, 3941–3950.

Kumar, R., and Litwack, G. (2009). Structural and functional relationships of the steroid hormone receptors' N-terminal transactivation domain. Steroids *74*, 877–883.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lazar, M.A. (2017). Maturing of the nuclear receptor family. J. Clin. Invest. *127*, 1123–1125.

Lerner, J., Gomez-Garcia, P.A., McCarthy, R.L., Liu, Z., Lakadamyali, M., and Zaret, K.S. (2020). Two-Parameter Mobility Assessments Discriminate Diverse Regulatory Factor Behaviors in Chromatin. Mol. Cell *79*, 677–688.e6.

Levi, V., and Gratton, E. (2007). Exploring dynamics in living cells by tracking single particles. Cell Biochem. Biophys. *48*, 1–15.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079, https://doi.org/10.1093/bioinformatics/btp352.

Lim, H.W., Uhlenhaut, N.H., Rauch, A., Weiner, J., Hübner, S., Hübner, N., Won, K.J., Lazar, M.A., Tuckermann, J., and Steger, D.J. (2015). Genomic redistribution of GR monomers and dimers mediates transcriptional response to exogenous glucocorticoid in vivo. Genome Res. *25*, 836–844.

Liu, Z., and Tjian, R. (2018). Visualizing transcription factor dynamics in living cells. J. Cell Biol. *217*, 1181–1191.

Lu, H., Yu, D., Hansen, A.S., Ganguly, S., Liu, R., Heckert, A., Darzacq, X., and Zhou, Q. (2018). Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. Nature *558*, 318–323.

Lu, Y., Wu, T., Gutman, O., Lu, H., Zhou, Q., Henis, Y.I., and Luo, K. (2020). Phase separation of TAZ compartmentalizes the transcription machinery to promote gene expression. Nat. Cell Biol. *22*, 453–464.

Luisi, B.F., Xu, W.X., Otwinowski, Z., Freedman, L.P., Yamamoto, K.R., and Sigler, P.B. (1991). Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. Nature *352*, 497–505.

Mazza, D., Abernathy, A., Golob, N., Morisaki, T., and McNally, J.G. (2012). A benchmark for chromatin binding measurements in live cells. Nucleic Acids Res. *40*, e119.

Mazza, D., Ganguly, S., and McNally, J.G. (2013). Monitoring dynamic binding of chromatin proteins in vivo by single-molecule tracking. Methods Mol. Biol. *1042*, 117–137.

Meijsing, S.H., Elbi, C., Luecke, H.F., Hager, G.L., and Yamamoto, K.R. (2007). The ligand binding domain controls glucocorticoid receptor dynamics independent of ligand release. Mol. Cell. Biol. *27*, 2442–2451.

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.

Mikuni, S., Tamura, M., and Kinjo, M. (2007). Analysis of intranuclear binding process of glucocorticoid receptor using fluorescence correlation spectroscopy. FEBS Lett. *581*, 389–393.

Mikuni, S., Yamamoto, J., Horio, T., and Kinjo, M. (2017). Negative Correlation between the Diffusion Coefficient and Transcriptional Activity of the Glucocorticoid Receptor. Int. J. Mol. Sci. *18*, 1855.

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. Nat. Commun. *5*, 4456.

Muragaki, Y., Mundlos, S., Upton, J., and Olsen, B.R. (1996). Altered growth and branching patterns in synpolydactyly caused by mutations in HOXD13. Science *272*, 548–551.

Nakamura, K., Jeong, S.Y., Uchihara, T., Anno, M., Nagashima, K., Nagashima, T., Ikeda, S., Tsuji, S., and Kanazawa, I. (2001). SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. Hum. Mol. Genet. *10*, 1441–1448.

Normanno, D., Boudarène, L., Dugast-Darzacq, C., Chen, J., Richter, C., Proux, F., Bénichou, O., Voituriez, R., Darzacq, X., and Dahan, M. (2015). Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. Nat. Commun. *6*, 7357.

Onuki, A. (2007). Phase Transition Dynamics (Cambridge University Press).

Paakinaho, V., Presman, D.M., Ball, D.A., Johnson, T.A., Schiltz, R.L., Levitt, P., Mazza, D., Morisaki, T., Karpova, T.S., and Hager, G.L. (2017). Single-molecule analysis of steroid receptor and cofactor action in living cells. Nat. Commun. *8*, 15896.

Paakinaho, V., Johnson, T.A., Presman, D.M., and Hager, G.L. (2019). Glucocorticoid receptor quaternary structure drives chromatin occupancy and transcriptional outcome. Genome Res. *29*, 1223–1234.

Peng, Z., Mizianty, M.J., Xue, B., Kurgan, L., and Uversky, V.N. (2012). More than just tails: intrinsic disorder in histone proteins. Mol. Biosyst. *8*, 1886–1901.

Presman, D.M., and Hager, G.L. (2017). More than meets the dimer: What is the quaternary structure of the glucocorticoid receptor? Transcription *8*, 32–39.

Presman, D.M., Ogara, M.F., Stortz, M., Alvarez, L.D., Pooley, J.R., Schiltz, R.L., Grøntved, L., Johnson, T.A., Mittelstadt, P.R., Ashwell, J.D., et al. (2014). Live cell imaging unveils multiple domain requirements for in vivo dimerization of the glucocorticoid receptor. PLoS Biol. *12*, e1001813.

Presman, D.M., Ganguly, S., Schiltz, R.L., Johnson, T.A., Karpova, T.S., and Hager, G.L. (2016). DNA binding triggers tetramerization of the glucocorticoid receptor in live cells. Proc. Natl. Acad. Sci. USA *113*, 8236–8241.

Presman, D.M., Ball, D.A., Paakinaho, V., Grimm, J.B., Lavis, L.D., Karpova, T.S., and Hager, G.L. (2017). Quantifying transcription factor binding dynamics at the single-molecule level in live cells. Methods *123*, 76–88.

Qian, H., Sheetz, M.P., and Elson, E.L. (1991). Single particle tracking. Analysis of diffusion and flow in two-dimensional systems. Biophys. J. *60*, 910–921.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842, https://doi.org/10.1093/bioinformatics/btq033.

Ratke, L., and Voorhees, P.W. (2011). Growth and Coarsening: Ostwald Ripening in Material Processing (Springer).

Ratman, D., Mylka, V., Bougarne, N., Pawlak, M., Caron, S., Hennuyer, N., Paumelle, R., De Cauwer, L., Thommis, J., Rider, M.H., et al. (2016). Chromatin recruitment of activated AMPK drives fasting response genes co-controlled by GR and PPARα. Nucleic Acids Res. *44*, 10539–10553.

Reingruber, J., and Holcman, D. (2011). Transcription factor search for a DNA promoter in a three-state model. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. *84*, 020901.

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. Sci. Rep. *10*, 1758.

Reiter, F., Wienerroither, S., and Stark, A. (2017). Combinatorial function of transcription factors and cofactors. Curr. Opin. Genet. Dev. *43*, 73–81.

Rogatsky, I., Wang, J.C., Derynck, M.K., Nonaka, D.F., Khodabakhsh, D.B., Haqq, C.M., Darimont, B.D., Garabedian, M.J., and Yamamoto, K.R. (2003). Target-specific utilization of transcriptional regulatory surfaces by the glucocorticoid receptor. Proc. Natl. Acad. Sci. USA *100*, 13845–13850.

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., et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. Science *361*, eaar3958.

Sanulli, S., Trnka, M.J., Dharmarajan, V., Tibble, R.W., Pascal, B.D., Burlingame, A.L., Griffin, P.R., Gross, J.D., and Narlikar, G.J. (2019). HP1 reshapes nucleosome core to promote phase separation of heterochromatin. Nature *575*, 390–394.

Schwarz, G. (1978). Estimating the dimension of a model. Ann. Stat. *6*, 461–464.

Shakya, A., Park, S., Rana, N., and King, J.T. (2020). Liquid-Liquid Phase Separation of Histone Proteins in Cells: Role in Chromatin Organization. Biophys. J. *118*, 753–764.

Shin, Y., and Brangwynne, C.P. (2017). Liquid phase condensation in cell physiology and disease. Science *357*, eaaf4382.

Simons, S.S., Jr., and Kumar, R. (2013). Variable steroid receptor responses: Intrinsically disordered AF1 is the key. Mol. Cell. Endocrinol. *376*, 81–84.

Simons, S.S., Jr., Edwards, D.P., and Kumar, R. (2014). Minireview: dynamic structures of nuclear hormone receptors: new promises and challenges. Mol. Endocrinol. *28*, 173–182.

Stasevich, T.J., Mueller, F., Michelman-Ribeiro, A., Rosales, T., Knutson, J.R., and McNally, J.G. (2010). Cross-validating FRAP and FCS to quantify the impact of photobleaching on in vivo binding estimates. Biophys. J. *99*, 3093–3101.

Stavreva, D.A., Wiench, M., John, S., Conway-Campbell, B.L., McKenna, M.A., Pooley, J.R., Johnson, T.A., Voss, T.C., Lightman, S.L., and Hager, G.L. (2009). Ultradian hormone stimulation induces glucocorticoid receptor-mediated pulses of gene transcription. Nat. Cell Biol. *11*, 1093–1102.

Stavreva, D.A., Coulon, A., Baek, S., Sung, M.H., John, S., Stixova, L., Tesikova, M., Hakim, O., Miranda, T., Hawkins, M., et al. (2015). Dynamics of chromatin accessibility and long-range interactions in response to glucocorticoid pulsing. Genome Res. *25*, 845–857.

Stavreva, D.A., Garcia, D.A., Fettweis, G., Gudla, P.R., Zaki, G.F., Soni, V., McGowan, A., Williams, G., Huynh, A., Palangat, M., et al. (2019). Transcriptional Bursting and Co-bursting Regulation by Steroid Hormone Release Pattern and Transcription Factor Mobility. Mol. Cell *75*, 1161–1177.e11.

Stortz, M., Pecci, A., Presman, D.M., and Levi, V. (2020). Unraveling the molecular interactions involved in phase separation of glucocorticoid receptor. BMC Biol. *18*, 59.

Syed, A.P., Greulich, F., Ansari, S.A., and Uhlenhaut, N.H. (2020). Anti-inflammatory glucocorticoid action: genomic insights and emerging concepts. Curr. Opin. Pharmacol. *53*, 35–44.

Tarantino, N., Tinevez, J.Y., Crowell, E.F., Boisson, B., Henriques, R., Mhlanga, M., Agou, F., Israël, A., and Laplantine, E. (2014). TNF and IL-1 exhibit distinct ubiquitin requirements for inducing NEMO-IKK supramolecular structures. J. Cell Biol. *204*, 231–245.

Tokunaga, M., Imamoto, N., and Sakata-Sogawa, K. (2008). Highly inclined thin illumination enables clear single-molecule imaging in cells. Nat. Methods *5*, 159–161.

Vandevyver, S., Dejager, L., and Libert, C. (2012). On the trail of the glucocorticoid receptor: into the nucleus and back. Traffic *13*, 364–374.

Voss, T.C., and Hager, G.L. (2014). Dynamic regulation of transcriptional states by chromatin and transcription factors. Nat. Rev. Genet. *15*, 69–81.

Zhao, L., Zhou, S., and Gustafsson, J.Å. (2019). Nuclear Receptors: Recent Drug Discovery for Cancer Therapies. Endocr. Rev. *40*, 1207–1249.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-GFP	Abcam	ab290; RRID: AB_303395
Chemicals, peptides, and recombinant proteins		
Corticosterone	Sigma-Aldrich	Cat# 27840; CAS: 50-22-6
Dexamethasone	Sigma-Aldrich	D1756; CAS: 50-02-2
16% Paraformaldehyde (formaldehyde) aqueous solution	Electron Microscopy Sciences	15710; CAS: 30525-89-4
8% Aqueous Solution Glutaraldehyde	Electron Microscopy Sciences	Cat# 16000
Tetracycline	Sigma-Aldrich	T7660; CAS: 60-54-8
JF549 HaloTag ligand	Grimm et al., 2015	N/A
PA-JF549 HaloTag ligand	Grimm et al., 2016	N/A
Critical commercial assays		
TruSeq Chip Sample Prep Kit	Illumina	IP-202-1012
Nextera DNA Library Prep Kit	Illumina	FC-121-1030
Magna ChIP Protein A+G Magnetic Beads	Millipore-Sigma	16-663
Deposited data		
Raw sequencing data generated	This paper	GEO: GSE154771
Sample of raw datasets of SMT imaging	This paper	https://doi.org/10.17632/k6g3xtr5k5.1
GR ChIP-seq and Atac-seq from 3617GRKO-GPFGR cells	Paakinaho et al., 2019	GEO: GSE108634
Experimental models: cell lines		
3617 cells	Paakinaho et al., 2017	N/A
GRKO cells with stably integrated HaloTag-GR	Paakinaho et al., 2017	N/A
GRKO cells with stably integrated GFP-GRwt	Paakinaho et al., 2019	N/A
GRKO cells with stably integrated GFP-GR407C	This paper	N/A
GRKO cells with stably integrated GFP-GR428G	This paper	N/A
Recombinant DNA		
gRNA-pX330 SpCas9	(Cong et al., 2013)	Addgene #42230
Donor-Rosa_Puro_CMV_01292016	Paakinaho et al., 2019	N/A
pHaloTag-GR	Morisaki et al., 2014	N/A
pHaloTag-H2B	Mazza et al., 2012	N/A
pHalo-PPARα	Promega	FHC03156
pHaloTag-GRdim	This paper	N/A
pHaloTag-GRmon	This paper	N/A
pHaloTag-GR407C	This paper	N/A
pHaloTag-GR440C	Paakinaho et al., 2017	N/A
pGFP-GRwt	Paakinaho et al., 2019	N/A
pGFP-GR407C	This paper	N/A
pGFP-GRC428G	This paper	N/A
pGPF-NF1	Presman et al., 2017	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Trimmomatic 0.36	Bolger et al., 2014	http://www.usadellab.org/cms/?page=trimmomatic; RRID: SCR_011848
SAMTools	Li et al., 2009	https://github.com/samtools/samtools
BEDTools	Quinlan and Hall, 2010	https://github.com/arq5x/bedtools2; RRID: SCR_006646
Bowtie2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml; RRID:SCR_005476
HOMER 4.9	Heinz et al., 2010	http://homer.ucsd.edu/homer/; RRID: SCR_010881
MATLAB	The MathWorks, Inc	https://www.mathworks.com/products/matlab.html
TrackRecord	Mazza et al., 2013 and Garcia et al., 2020	https://github.com/davidalejogarcia/PL_HagerLab
Custom MATLAB code	Garcia et al., 2020	https://github.com/davidalejogarcia/PL_HagerLab

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Gordon L. Hager (hagerg@dce41.nci.nih.gov).

Materials availability

All cell lines and plasmids generated in this study are available upon request.

Data and code availability

The GRC428G and GR407C ChIP-seq datasets are submitted to the NCBI Gene Expression Omnibus (GEO: GSE154771). The previously published GFP-GRwt ChIP and ATAC datasets are archived under GEO: GSE108634.

Tracking was performed in MATLAB (version 2016a) with custom scripts (Garcia et al., 2020). The code for pEM analysis is freely available at GitHub: https://github.com/p-koo/pEMv2. A sample of raw datasets of SMT imaging data have been deposited in Mendeley Data: https://doi.org/10.17632/k6g3xtr5k5.1. Other relevant data and materials that are not explicitly included in this article will be made available by the lead contact upon reasonable request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Plasmids and cell lines

The pHaloTag-GR expresses rat GR with HaloTag (Promega, Madison, WI, USA) fused in the C-terminal domain under the CMVd1 promoter (Morisaki et al., 2014). The pHalo-GRA477T (GRdim), pHalo-GRA477T/I646A, and pHalo-GRC440G (C428G in mouse) were generated by using a QuikChange II XL Site Directed Mutagenesis Kit according to the manufacturer's instructions (Stratagene, La Jolla, CA, USA), as previously described (Paakinaho et al., 2017). The pHaloTag-H2B expresses histone H2B fused with the HaloTag through its N-terminal end (Mazza et al., 2012). The pHalo-PPARα was purchased from Promega (Cat# pfn21ab9549) and expresses the human PPARα fused with the HaloTag through its N-terminal end. The pHalo-GR407C was generated by 1) PCR amplification of rGR's coding amino acids 407-794 from the pHaloTag-GR template using the Herculase II fusion DNA polymerase system (Agilent Technologies, Santa Clara, CA, USA), introducing XhoI and PvuI restriction sites; and 2) swapping the PCR product into the pHaloTag-GR using the same restriction enzymes. The pGFP-NF1 expresses Nuclear factor-1 fused to GFP (Presman et al., 2017). The mouse GFP-GR407C plasmid was generated by circular PCR with high-fidelity polymerase from the GFP-GRwt plasmid (Paakinaho et al., 2019) using primers to exclude the N terminus and sequence verified prior to use in making the cell line. The GFP-GRC428G plasmid was generated with the same PCR/verification methods as above using primers to incorporate a point mutation at residue 428 of mouse GR.

The 3617 mouse mammary adenocarcinoma cell line and its derivatives were routinely cultured in high glucose DMEM supplemented with 10% fetal bovine serum and 2 mM L-glutamine at 37°C in a CO_2-controlled humidified incubator. The GRKO cell lines with stably integrated forms of GFP-GR has been described previously (Paakinaho et al., 2019). Briefly, they were generated by integrating the designated form of GR into the GT(Rosa)26Sor locus. GFP-GR integrated cells were selected with puromycin and FACS

sorted for similar levels of GFP expression and size uniformity. A GR knockout subclone expressing Halo-GR was also used in this study and described elsewhere (Paakinaho et al., 2017).

For transfections, 5 million cells were electroporated using BTX T820 Electro Square Porator (Harvard Apparatus, Holliston, MA, USA) in 100 μL of DPBS with 2.5 μg of plasmid. 25 ms pulses of 120 V were used and cells were resuspended in fresh media. Single-molecule imaging experiments were set up as follows: 100,000 electroporated cells were seeded onto each well of a 2-well Lab-Tek chamber (1.5 German borosilicate coverglass, Thermo Fisher, Waltham, MA, USA) in high glucose DMEM supplemented with 10% FBS (Life Technologies), 2mM L- glutamine, 5 μg/ml tetracycline, and cultured overnight. The media was then replaced with high glucose DMEM supplemented with 10% charcoal stripped FBS (Life Technologies), 2mM L- glutamine, 5 μg/ml tetracycline, and incubated at 37°C O.N. before labeling.

Fluorescent labeling of Halo-tagged molecules and hormone treatments

Transfected cells were incubated with 5 nM JF_{549}-HaloTag (Grimm et al., 2015) or 50 nM photo-activatable PA-JF_{549}-HaloTag (Grimm et al., 2016) ligand for 20 min at 37°C. PA-JF_{549} was used for the fast acquisition SMT experiments (12ms) by stroboscopic activation of the fluorophore by a 473nm laser. Free ligand was depleted by washing three times with phenol red free DMEM media (supplemented with 10% charcoal stripped FBS and 5 μg/ml tetracycline) in 15 min intervals at 37°C. Next, cells were treated with 600 nM Corticosterone (Cort) (Sigma-Aldrich) or 100 nM Dexamethasone (Dex) (Sigma-Aldrich), and incubated for 20 min at 37°C before imaging. For "noise" control in Figure 1C and Figure 2F, after activating GR for 30 min with 100 nM dexamethasone, the cells were fixed with 4% (wt/vol) paraformaldehyde (Electron Microscopy Sciences) and 0.2% (wt/vol) glutaraldehyde (Electron Microscopy Sciences) for 35 min at room temperature. After fixation, cells were incubated with ProLong Gold Antifade Mountant for 24 hours. Cells treated with 10 μM WY-14643 (PPARα activation) were incubated for 1 hour at 37°C before imaging. For wash-out experiments, cells were washed with media three times for 4 different intervals (every 15 minutes for 1 hour or every hour for 4 hours) after 20 minutes of hormone treatment and finally imaged.

METHOD DETAILS

Image acquisition for single-molecule tracking

A custom HiLO microscope was used as previously described in detail elsewhere (Presman et al., 2017), with an objective heater to reduce drift. Briefly, the custom-built microscope from the CCR, LRBGE Optical Microscopy Core facility is controlled by μManager software (Open Imaging, Inc., San Francisco, CA.), equipped with an Okolab state top incubator for CO_2 (5%) and temperature control (37°C), a 150X 1.45 numerical aperture objective (Olympus Scientific Solutions, Waltham, MA), a 561nm and 473nm lasers (iFLEX-Mustang, Excelitas Technologies Corp., Waltham, MA), and an acousto-optic tunable filter (AOTFnC- 400.650, AA Optoelectronic, Orsay, France). Images were collected on an EM-CCD camera (Evolve 512, Photometrics). The imaging conditions were as follow: 12ms exposure and 12ms acquisition time for fast acquisition SMT data with a laser power of 0.96 mW, 10ms exposure and 200ms acquisition time with a laser power of 0.96 mW for confinement and chromatin binding analysis and 500ms exposure and 500ms acquisition time with laser power of 0.16 mW for survival distribution analysis. For the array experiments (Figures 2F–2H), the imaging conditions were 10ms exposure with sequential acquisition of GFP (for NF-1) and JF549 (for GR-halo) channels, giving a 252ms acquisition time due to limitations in the speed of the AOTF.

Image acquisition by Airyscan

Single plane images were acquired on a Zeiss LSM 880 point scanning confocal microscope using the Airyscan detector, a 100x Plan-Apochromat 1.4NA DIC oil immersion objective (Zeiss) and 488 nm and 561 nm laser lines. The Zeiss Zen 2.3 (black edition) software was used to control the microscope, adjust spectral detection for the emission of EGFP and JF_{549}-HaloTag fluorophores and for processing of the Airyscan raw images. Zeiss Definite Focus was used at each time point for time-lapse imaging.

Single-molecule tracking analysis

We used the custom-made software TrackRecord (Mazza et al., 2013) in MATLAB (The MathWorks, Inc., Natick, MA). Briefly, to analyze each time series, data were filtered using top-hat, Wiener, and Gaussian filters. A region of interest (ROI) was defined to encompass the nucleus (and the array using the GFP-NF1 fluorescence when necessary), then particles were detected, fitted to two dimensional Gaussian function for subpixel localization, and finally tracked using a nearest neighbor algorithm (Presman et al., 2017). The tracking parameters were as follows: window size for particle detection 7 pixels, maximum frame to frame displacement of 6 pixels, shortest track 2 frames, and gaps to close 1.

The average single molecule localization precision (SMLP) was estimated as $\sigma_{SMLP} = \left(\sigma_{PSF} \,/\, \sqrt{N_{Photons}}\right)$, where σ_{PSF} corresponds to the standard deviation of the 2D Gaussian used for particle localization, $N_{Photons}$ the number of photons and $\langle\,\rangle$ indicates the ensemble average over all the detected particles in an experiment. The number of photons were estimated by converting the gray values (camera offset corrected) to electrons by using the conversion gain of the camera adjusted and quantum efficiency (QE) for digital gain as indicated by the manufacturer. The average localization precision for 12ms exposure – 12ms interval experiments is 23.8 nm and for 10ms exposure – 200ms interval experiments is 31.7 nm. The higher localization precision in the 12ms experiments is due to significantly higher laser power used.

Tracks were classified into distinct diffusive states using perturbation-Expectation Maximization [pEM v2, (Koo and Mochrie, 2016)]. Prior assumptions on the type of diffusion of the tracked particles are not needed with pEM and the number of diffusive states can be deduced from the analysis. pEM analysis requires all analyzed tracks to be of the same length. Tracks were split into 7 frame segments and the pEM classification analysis was performed on the set of all these track segments. For instance, a track of length 30 is segmented into 4 subtracks of length 7. Short tracks minimize the probability of transitions between diffusive states on the same tracks. To confirm the results, pEM analysis was also performed on tracks split into 15 or 30 frame segments. The minimum number of states for the system to converge to was set at 2 and the maximum at 7. If the optimal number of states that the analysis converged to was 7, the algorithm was rerun with a higher number of maximum states. The number of reinitializations was set to 30 with 50 perturbation trials. The maximum number of iterations was 10000 with a convergence criterion for the change of log-likelihood of 10^{-7}. The number of features for the covariance matrix was set to 5 for tracks of length 7 or 15 and 3 for tracks of length 30. Motion blur coefficient was calculated as $(1/6)(\Delta e/\Delta t)$, where Δe corresponds to the exposure time and Δt the acquisition interval.

Mean Squared Displacement calculation (MSD)

Ensemble MSD was calculated using custom scripts and routines previously published (Tarantino et al., 2014). All the calculations assume that the stochastic process of TFs diffusing in the nucleus or interacting with chromatin is wide sense stationary. SMT experiments provide us with tracks that can be described as a series of positions $\{\vec{r}i(t)\}$, acquired every Δt sec. For any stationary stochastic process, the MSD of a trajectory *i* at time lag τ can be calculated as:

$$MSD_i(\tau) = \frac{1}{N}\sum_n \left[\vec{r_i}(n\Delta t + \tau) - \vec{r_i}(n\Delta t)\right]^2$$

After pEM classification, states with a population fraction less than 5% are removed. For a particular state *i*, let us define the set $I \equiv \{k | P_k(i) \geq 0.6\}$ where *k* is a track of length 7 and $P_k(i)$ corresponds to the probability of *k* to belong to the state *i*. If the population of a set *I* (i.e., the ratio of the number of tracks assigned to state *I* to the total number of tracks) is less than 0.05, the set *I* is removed. The ensemble average MSD for a particular state *i* is given by:

$$eMSD_i(\tau) = \frac{\sum_{j\in I} MSD_j(\tau)}{|I|}$$

where $|I|$ is the number of tracks assigned to state *I*. In other words, after tracks of length 7 (or 15) are classified into different diffusive states, the ensemble average MSD is calculated for each particular diffusive state from tracks of length 7 (or 15) that have a posterior probability higher than 0.6 of belonging to that particular state.

Alternately, we can also calculate the weighted ensemble averaged MSD for a particular state by using the posterior probability of a track to belong to a state *I* as the weight function:

$$WMSD_i(\tau) = \frac{\sum_{j=1}^{N} P(i)\cdot MSD_j(\tau)}{N}$$

Where $WMSD_i(\tau)$ is the weighted ensemble average MSD for state *i*, and N is the total number of tracks.

Standard error was calculated as $\sigma_w/\sqrt{N_f}$, where σ_w corresponds to the weighted standard deviation and N_f the number of degrees of freedom in the weighted mean. Weights correspond to the number of points averaged to generate the mean square displacement value at the given time lag.

Sample tracks for illustration were selected as follows: sets composed of all tracks from the different diffusive states with a posterior probability larger than 0.9 were generated. For low-mobility tracks, tracks were selected randomly in a 2 μm X 2 μm window for illustration. For high-mobility states, a 4.8 μm X 4.8 μm window was selected instead.

Diffusion coefficients were estimated from tracks with a posterior probability larger than 0.6 to belong to the particular diffusive state for the 12ms acquisition interval experiments. The estimation was done from the variance of the instantaneous velocity vector *v* by $v^2 = 4D/\Delta t$, D corresponds to the diffusion coefficient and Δt the acquisition interval (Qian et al., 1991).

Survival distribution calculation

Survival distributions were calculated from particle tracks as in (Garcia et al., 2020) using the Kaplan-Meier estimate. The 95% confidence interval was estimated using Greenwood's Formula. HaloTag-H2B data was acquired with identical imaging conditions as the transcription factors for the different acquisition conditions. The survival distribution of this was fitted to a double and triple exponential model to extract the photobleaching rate, and model selection was used to determine the best predictive model (See Fitting and model selection). The survival distribution of the different transcription factors was corrected for photobleaching as follows ($S(t) = e^{\gamma t}S_E(t)$, where *S(t)* corresponds to the survival distribution after photobleaching correction, $S_E(t)$ the empirical survival distribution and γ the photobleaching rate).

Weighted survival distribution calculation

For the diffusivity analysis, we divide every track into subtracks of length 7 frames. These subtracks are then classified into different diffusive states using pEM and states that are not representative are removed as explained above. Suppose a track *k* is divided into *n*

subtracks of length 7. After the pEM analysis, each subtrack i is assigned a posterior probability $P_k^i(j)$ to belong to a diffusive state j. Here, $i \in \{1,2,\ldots,n\}, j \in \{1,2,\ldots,m\}$, where m is the number of diffusive states to which pEM converges. From this state assignment, we would like to calculate the survival distribution of a particular diffusive state. To do so, we first calculate a posterior probability for the reconstructed track k:

$$P_k(j) = \frac{\sum_i P_k^i(j)}{\sum_{ij} P_k^i(j)}$$

We will use these probabilities to calculate a weight function for each bin of the dwell time histogram. All the dwell times can be distributed into N bins such that each bin Ω_i contains n_i tracks. Mathematically, $\Omega_i \equiv \{t \in \mathbb{R} | (i-1)\Delta t \leq t < i\Delta t\}$, where Δt is the acquisition interval.

We can define a weight function W_i^j for each of the bins i, per state j as

$$W_i^j = \sum_{k=1}^{n_i} P_k(j)$$

Previously survival distributions were calculated by assigning a unit weight to each observed bound track. Instead, we can define the statistic

$$\widehat{p}_i^j = \frac{W_i^j}{n}$$

as the unbiased estimator (i.e., $E[p_i^j] = \widehat{p}_i^j$) of the weighted survival distribution.

Let $P_k(j)$ be the probability assigned to a track k as defined previously. The proportion $(M(i))$ of a particular diffusive state i is given by:

$$M(i) = \frac{\sum_k P_k(i)}{\sum_i \sum_k P_k(i)}$$

Fitting and model selection

All fits performed to the data were implemented with the nonlinear least square method using bisquare weights due to the noise on the tail of the survival distribution.

Graphical inspection was used to qualitatively determine if a straight line was observed for multiple decades in the case of a power-law fit in a log-log plot. Three different metrics were used to determine the difference between exponential models and power-law models. The first metric corresponds to Bayesian information criterion (BIC) using the probability distribution function (PDF) corrected for photobleaching as the likelihood function. The PDF was normalized between the minimum and maximum observation range of TFs dwell time (BIC1). BIC is a criterion for model selection that penalizes for model complexity (number of free parameters in the model). BIC1 is given by (James et al., 2017):

$$BIC1(M) = kln(n) - 2ln\left(P\left(D|\widehat{\theta}, M\right)\right)$$

where M corresponds to the model (Power Law, Double Exponential and Triple Exponential), k corresponds to the number of parameters of the model, $\widehat{\theta}$ corresponds to the model parameters found by fitting, D the observed data and n the number of observations. $P(x|\widehat{\theta}, M)$ corresponds to the realization probability of x given the model PDF with parameters $\widehat{\theta}$. For SMT, D is the set of independent and identically distributed discrete experimental events and $P(D|\widehat{\theta}, M)$ is calculated as follows:

$$P\left(D|\widehat{\theta}, M\right) = \prod_{x \in D} \int_{x-\frac{\delta t}{2}}^{x+\frac{\delta t}{2}} P\left(x|\widehat{\theta}, M\right) dx$$

Where $P(x|\widehat{\theta}, M)$ corresponds to the PDF of the model after photobleaching correction. For instance, the double exponential PDF is given by:

$$p\left(x|\widehat{\theta} = (\alpha, \beta), M_{DE}\right) = C\left(f_1 \gamma\alpha e^{-(\gamma+\alpha)t} + (1-f_1)\gamma\alpha e^{-(\gamma+\alpha)t}\right)$$

where α, β are the exponential parameters, γ the photobleaching rate and C a normalization constant.

The second metric, the evidence in decibels (Db) for a particular model given the observed data and priors, was calculated to compare the alternative models explored. The evidence measures the probability of a particular model being the best predicting model in comparison with all the other models. For instance, for the power law model (M_{PL}) the evidence versus the double exponential model (M_{DE}) is given by (Jaynes and Bretthorst, 2019):

$$E\left(M_{PL}|D,\widehat{\theta},A\right) = E(M_{PL}|A) + 10log_{10}\left[\frac{P\left(D|M_{PL},\widehat{\theta}\right)}{P\left(D|M_{DE},\widehat{\theta}\right)}\right]$$

$$E(M_{PL}|A) = 10\log_{10}\left(\frac{P(M_{PL}|A)}{P(M_{DE}|A)}\right)$$

where A corresponds to the priors; P,D and $\widehat{\theta}$ as defined for BIC1. Uniform priors were used throughout the model comparison and therefore $E(M_{PL}|A)$ is set to 0. For instance, an evidence of 30 Db corresponds to a probability higher than 0.999 that the power law model better describes the data in comparison with the alternative models tested. If the evidence was not high enough to reach a conclusion between the different models, more data was acquired until the evidence reached a satisfactory value.

A final metric using BIC and the residual sum of squares (RSS) as a likelihood function was used (BIC2). For each functional model *f*, BIC2 was calculated as:

$$BIC2(M) = klog(N) + Nlog\left(\frac{RSS}{N}\right)$$

where *k* corresponds to the number of parameters estimated and *N* the number of observations in the survival distribution (James et al., 2017).

Delta-BIC1 (Delta-BIC2) is defined as the difference between BIC1 (BIC2) calculated for double exponential and power-law (DeltaBIC1 = BIC1(DE)-BIC1(PL)). A negative value indicates that the double exponential model is a better model for the data in comparison to the power-law model, while a positive value indicates that a power-law better describes the data. A negative value of evidence (as calculated above) corresponds to a higher probability for a double exponential model being the best model for the data compared to the power-law model. Table 1 shows the values of the evidence and Delta-BIC1 and Delta-BIC2 for fitting comparison between the power law and double-exponential models for all the different experimental conditions.

GR chromatin immunoprecipitation (ChIP) and ChIP-seq

GFP-GR wild-type and mutant expressing cells were left untreated or treated with 100 nM of Dex (Sigma) for 1 h. For ChIP, after cross-linking with 1% paraformaldehyde (5 min) and cell collection the chromatin was sonicated (Bioruptor, Diagenode) to an average DNA length of 200–500 bp. For immunoprecipitation, 600 μg of chromatin was incubated with anti-GFP antibody (Abcam #ab290) coupled onto Protein A/G magnetic beads (Millipore) with rotation overnight at 4°C. After stringent washes, the antibody-bound chromatin fragments' cross-linking was reversed, and the remaining proteins digested. Immunoprecipitated DNA was extracted from the samples with phenol-chloroform-isoamyl alcohol and ethanol precipitation. ChIP-seq libraries were generated using Illumina TruSeq Chip Sample Prep Kit (Illumina # IP-202-1012) according to manufacturer's instructions.

ChIP-sequencing data analyses

Two biological duplicate ChIP samples each for GRC428G and GR407C cell lines were sequenced using Illumina NextSeq 500 with single-end reads. The data were aligned to the mouse reference mm10 genome using Bowtie2 (Langmead and Salzberg, 2012). Subsequent downstream analysis was performed using HOMER (Heinz et al., 2010). Replicate datasets were merged and peaks in each dataset were called using findPeaks with style factor for TF, FDR 0.001, 4-fold enrichment of normalized reads to the control and 4-fold enrichment over local background; however, individual replicates correlated well with each other for all called peaks (Table S1). The mergepeaks command was used to determine the shared/unique peaks from the GRwt and GR407C. Data matrices for heatmaps/aggregate plots for the ChIP-seq and the ATAC-seq (Paakinaho et al., 2019) were generated using the annotatePeaks.pl command with a 20bp sampling window.

QUANTIFICATION AND STATISTICAL ANALYSIS

For statistical analysis, all measured quantities are reported as ensemble averages with standard error and number of observations. At least three biological replicates of SMT experiments were performed for each condition. Two sample K-S tests on the survival distribution were performed between replicates to confirm statistical reproducibility. At least 20 cells were imaged per SMT replicate for each condition for slow acquisition intervals (200 ms and 500 ms) and 60-100 cells were imaged for fast acquisition intervals (12 ms). The exact number of tracks and cells are specified in figure legends.