Live-cell 3D single-molecule tracking reveals how NuRD modulates enhancer dynamics

Authors: Basu S[§1,2], Shukron O[§3], Ponjavic A[4], Parruto P[3], Boucher W[1], Zhang W[1], Reynolds N[2], Lando D[1], Shah D[1], Sober LH[1], Jartseva A[1], Ragheb R[2], Cramard J[2], Floyd R[2], Brown G[1], Gor K[2], Balmer J[1], Drury TA[1], Carr AR[4], Needham L-M[4], Aubert A[6], Communie G[6], Morey L[7,11], Blanco E[7], Barber MA[2], Mohorianu I[2], Bartke T[5], Di Croce L[7,8], Berger I[9], Schaffitzel C[9], Lee SF[4], Stevens TJ[10], Klenerman D*[4], Hendrich BD*[1,2] , Holcman D*[3], Laue ED*[1,2].

Affiliations:

[1] Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, United Kingdom

[2] Wellcome-MRC Cambridge Stem Cell Institute, Jeffrey Cheah Biomedical Centre, Puddicombe Way, Cambridge CB2 0AW, United Kingdom

[3] Department of Applied Mathematics and Computational Biology, Ecole Normale Supérieure, 75005 Paris, France

[4] Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, United Kingdom

[5] Helmholtz Zentrum München, German Research Center for Environmental Health, Institute of Functional Epigenetics, Ingolstädter Landstr. 1, 85764 Neuherberg

[6] The European Molecular Biology Laboratory EMBL, 71 Avenue des Martyrs, 38000 Grenoble, France

[7] Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Dr. Aiguader 88, Barcelona 08003, Spain

[8] Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluis Companys 23, Barcelona, Spain

[9] School of Biochemistry, University of Bristol, 1 Tankard's Close, Bristol BS8 1TD, United Kingdom

[10] MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge CB2 0QH

[11] Present addresses: Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Biomedical Research Building, 1501 NW 10th Avenue, Miami, Miami-Dade County 33136, USA

[§] These authors contributed equally to this work; * To whom correspondence should be addressed

Abstract

Enhancer-promoter dynamics are critical for the spatiotemporal control of gene expression, but it remains unclear how these dynamics are controlled by chromatin regulators, such as the nucleosome remodelling and deacetylase (NuRD) complex. Here, we use Hi-C experiments to show that the intact NuRD complex increases CTCF/Cohesin binding and the probability of the interaction of intermediate-range (~1Mb) genomic sequences. To understand how NuRD alters 3D genome structure in this way, we developed an approach to segment and extract key biophysical parameters from trajectories of the NuRD complex determined using live-cell 3D single-molecule imaging. Unexpectedly, this revealed that the intact NuRD complex decompacts chromatin structure and makes NuRD-bound sequences move faster, thus increasing the overall volume of the nucleus that these sequences explore. Interestingly, we also uncovered a rare fast-diffusing state of chromatin that exhibits directed motion. The intact NuRD complex reduces the amount of time that enhancers/promoters remain in this fast-diffusing state, which we propose would otherwise re-organise enhancer-promoter proximity. Thus, we uncover an intimate connection between a chromatin remodeller and the spatial dynamics of the local region of the genome to which it binds.

Introduction

3D genome organisation is thought to be critical for the spatiotemporal control of gene expression. However, little is known about the multi-scale chromatin dynamics of cis-regulatory elements or how they relate to genome organisation. To probe these dynamics at the single-cell level, one of two complementary methods are typically used: either live-cell imaging[1-4] or single nucleus versions of chromosome conformation capture experiments (such as Hi-C), which reveal the proximity of DNA sequences in different individual fixed cells[5-16]. For example, live-cell tracking of enhancers and promoters has revealed fast diffusion (or 'stirring') during transcription[4], but the mechanisms underlying these changes in enhancer dynamics or how they relate to Hi-C measurements of enhancer-promoter proximity remain unclear.

The nucleosome remodelling and deacetylase (NuRD) complex is a highly conserved 1 MDa multi-subunit protein complex that we have previously shown clusters in 3D space in the nucleus with active enhancers and promoters[16]. It combines two key enzymatic activities – nucleosome remodelling *via* its helicase-containing ATPase (predominantly CHD4 in ES cells) and lysine deacetylation *via* its HDAC1/2 subunits[16-22]. These activities are thought to be present in two sub-complexes (**Figure 1a**). HDAC1/2 associates, along with the histone chaperones RBBP4/7, with the core scaffold proteins MTA1/2/3 (metastasis tumour antigens) to form a stable sub-complex with deacetylase activity[23]. The nucleosome remodeller CHD4 interacts with chromatin by itself and may also form a second sub-complex with GATAD2A/B and DOC1 (CDK2AP1)[23-25]. The methyl-CpG DNA binding domain proteins MBD2/3 interact directly with both the deacetylase sub-complex and GATAD2A/B[23,26,27], and thus play a critical role in linking the CHD4 remodeller and HDAC sub-complexes together to assemble the holo-NuRD complex.

Assembly of the intact holo-NuRD complex is critical for controlling cell fate transitions. Knockout of *Mbd3*, which disrupts intact NuRD complex assembly, leads to moderate up- or down-regulation ('fine-tuning') of transcription levels, but this modulation is sufficient to prevent ES cell lineage commitment[28-31]. Nucleosome remodelling by CHD4 requires the MBD subunit (and presumably the deacetylase sub-complex) to evict transcription factors and RNA polymerase II from chromatin[22]. Whether this impacts enhancer-promoter dynamics has remained unclear. Here, to understand how enhancer/promoter proximity and dynamics are

regulated by this crucial chromatin remodeller, we combine Hi-C and live-cell single-molecule tracking. Specifically, we exploit the ability to unlink the chromatin remodelling and deacetylase subunits of intact NuRD by deleting *Mbd3* to explore the function of the intact NuRD complex.

Intact holo-NuRD complex assembly increases the proximity of intermediate-range genomic sequences

We initially set out to understand whether assembly of the intact NuRD complex, mediated by MBD3, might affect genome architecture by carrying out chromosome conformation capture (*in-nucleus* Hi-C) experiments on wild-type and *Mbd3* knockout ES cells. This strategy, which exploits the fact that MBD2 (and thus the MBD2-linked holo-NuRD complex) is only expressed at low levels in ES cells, allowed us to perturb the majority of NuRD complex function without killing the cells. We sequenced 178-181 million paired end reads and obtained a total of 50-60 million high-quality Hi-C contacts (**Extended Data Figure 1**). As previously observed, these contact maps show that the genome is segregated into: *1)* A and B compartments (broadly regions containing more or fewer genes respectively); *2)* megabase-scale topologically associating domains (TADs), which have a higher frequency of intra-domain chromatin interactions; and *3)* loops where specific genomic regions contact each other, such as loops mediated *via* CTCF/Cohesin binding[5-16].

Although there is little change in A/B compartments or TAD formation when comparing wild-type vs *Mbd3-ko* cells (**Extended Data Figures 2a,b**), analysis of the Hi-C contacts revealed that removal of MBD3 leads to a small increase in short-range contacts (at the scale of genes, < 250 kb), and to a more significant decrease in intermediate-range contacts at the scale of TADs (500 kb to 3 Mb) (**Figure 1b**). This trend occurred regardless of the contacts being within the A, the B, or between the A and B compartments, although most change occurred in the latter (**Extended Data Figure 2c**). The loss in intermediate-range interactions in the *Mbd3-ko* was also observed in enhancer and active promoter interactions that we had previously defined in wild-type mouse ES cells[16] (**Extended Data Figure 2d**) – with interactions of enhancers showing a greater change in contact frequency than interactions of active promoters (**Extended Data Figure 2e**).

Consistent with the decrease in intermediate-range enhancer-promoter interactions in *Mbd3-ko* cells, we also observed a decrease in the intensity of intermediate- to long-range ES cell-specific CTCF-Cohesin loops (**Extended Data Figure 2d**), which are thought to influence enhancer-promoter proximity[33-35]. In addition, there were fewer interactions between adjacent TADs in *Mbd3-ko* cells (**Extended Data Figure 2f**). We used ChIP-seq experiments to show that in *Mbd3-ko* cells there is a decrease in binding of CTCF and/or SMC3 (a subunit of the Cohesin complex) – two proteins that are known to be responsible for loop and TAD formation[12,33,36-40]. Notably, however, the decrease in levels of CTCF and SMC3 binding largely occurred without any change in

the sequences bound by these proteins (**Extended Data Figure 3**). In summary, the results suggest that formation of the intact NuRD complex increases intermediate-range chromatin interactions (**Figure 1c**), such as enhancer-promoter interactions within and between adjacent TADs, and this may be related to the increased CTCF/Cohesin binding seen in the presence of intact NuRD.

A novel trajectory segmentation algorithm for extracting NuRD complex dynamics and chromatin movement

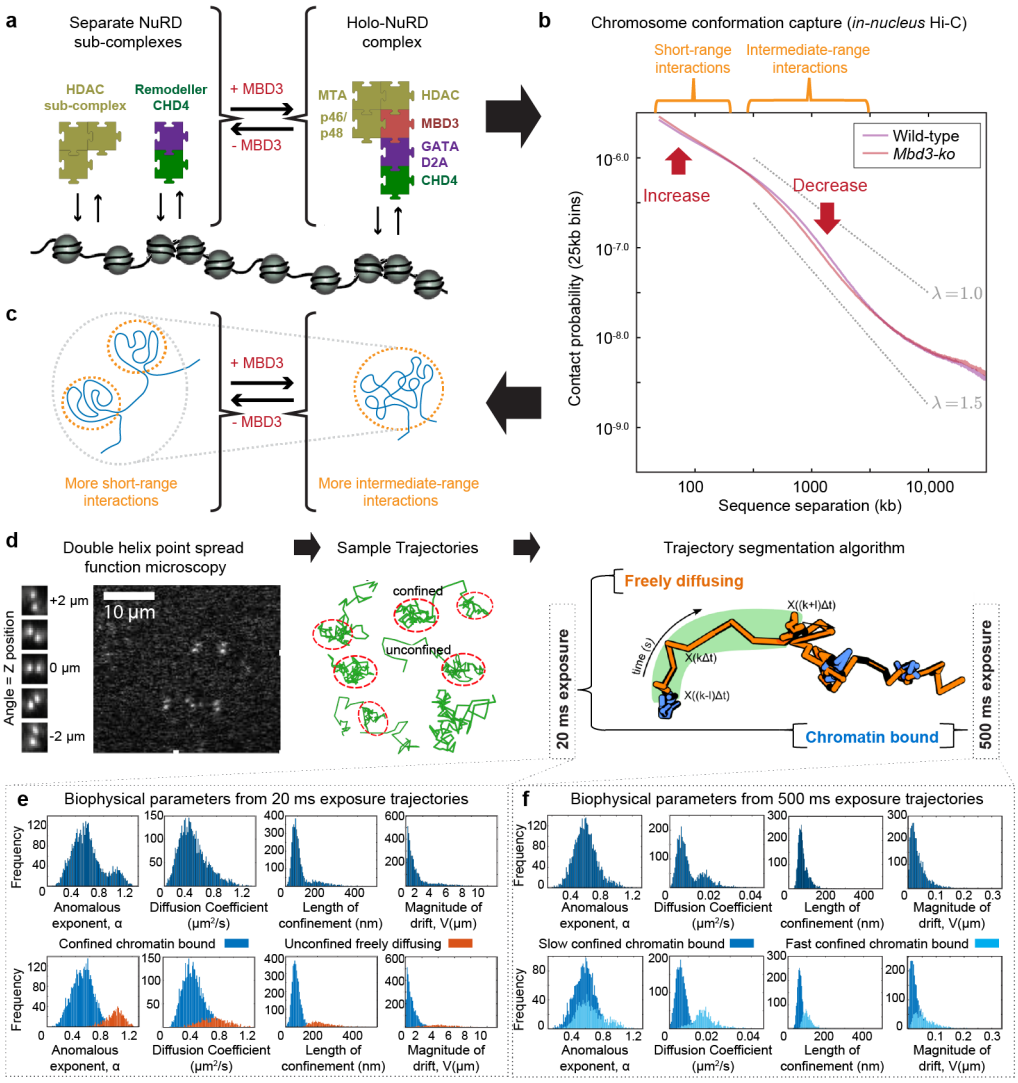
To provide a mechanistic understanding of how assembly of the intact NuRD complex increases mixing of intermediate-range genome sequences, we carried out live-cell 3D single-molecule tracking of NuRD complex subunits in the presence and absence of MBD3. We generated knock-in ES cell lines expressing HaloTag-CHD4, MBD3 and MTA2 fusions and confirmed that the tags did not affect NuRD complex expression or assembly (**Extended Data Figure 4**). We then used a double-helix point spread function microscope[24] to generate tracks of single JF_{549}-HaloTag-NuRD complexes in 3D over a 4 μm depth of focus, so that we could reliably determine biophysical parameters. (Tracking in 2D reduces the length of trajectories of fast-moving molecules because they can move out of the imaging plane during the experiment.) We recorded trajectories at two distinct temporal regimes, 20 ms and 500 ms (**Figure 1d**)[2]. Recording at 20 ms time resolution allows the detection of both freely diffusing and chromatin bound proteins[24], and can thus be used to extract the chromatin binding kinetics of NuRD complexes (**Figure 1e**). In contrast, at 500 ms time resolution, 'motion blurring' substantially reduces the detection of freely diffusing molecules, allowing us to focus on the slower sub-diffusive chromatin bound NuRD complexes[1] (**Figure 1f**).

To extract dynamic biophysical parameters, we developed a machine learning method (a Gaussian mixture model) to segment the single molecule trajectories into different classes by studying their behaviour over a sliding window of 11 consecutive images (**Figure 1d** and **Extended Data Figure 5**; further details of the approach, and of the simulations we carried out to test the algorithm, are described in the Online Methods and in **Extended Data Figures 6,7**). To understand how NuRD affects chromatin structure, we estimated from each sub-trajectory not just the diffusion coefficient but also the anomalous exponent α, the localisation length Lc, and the drift magnitude V[41]. The anomalous exponent α, defined as the power of the mean squared

displacement against time curve [Mean squared displacement $\propto Time^{\alpha}$], is particularly informative: diffusing proteins are characterised by an anomalous exponent close to 1 whereas chromatin bound proteins exhibit a lower anomalous exponent[4,41,42]. In addition, for chromatin bound proteins, lower anomalous exponents represent a more condensed chromatin state[43] and higher values can represent energy-dependent directed motion. The localisation length Lc of chromatin bound proteins is also informative as it reflects the spatial scale that the molecule explores within the nucleus. The localisation length Lc is dependent on a range of parameters such as nucleosome density and CTCF/Cohesin-mediated looping interactions – for example, Lc is larger at lower nucleosome densities when the linker length between nucleosomes is longer[43,44].

Analysis of the 20 ms exposure trajectories of single CHD4 molecules using our approach revealed two diffusion states (**Figure 1e**). We identified a fast unconfined state that was freely diffusing with an α of 0.94 ± 0.12 and a diffusion coefficient of 0.7 ± 0.2 $\mu m^2 s^{-1}$, matching previous observations[23,24], and a confined chromatin bound state characterised by sub-diffusive motion with an α of 0.51 ± 0.15 and an apparent diffusion coefficient of 0.43 ± 0.17 $\mu m^2 s^{-1}$. Similar results were obtained when segmenting the trajectories of the other NuRD complex components, MBD3 and MTA2 (**Extended Data Figure 8**). By determining similar parameters for fixed dye molecules, we found that the apparent diffusion coefficient of the chromatin bound state was within our precision limit and so represents molecules that are essentially immobile (**Extended Data Figure 8**). We conclude that we can use 20 ms exposure trajectories to distinguish unconfined freely diffusing molecules from confined chromatin bound proteins.

Figure 1. MBD3-dependent holo-NuRD complex assembly increases intermediate-range genome interactions. (a) Schematic representation of the NuRD complex interacting with chromatin in the presence and absence of MBD3. (b) *In-nucleus* Hi-C of wild-type (purple) and *Mbd3* knockout ES cells (red), represented as log-scale plots of contact probability as a function of genomic sequence separation, averaged across the genome. (c) Schematic interpretation of the results of the Hi-C experiment, showing that in the *Mbd3* knockout there are more short range and fewer intermediate range chromatin interactions. (d) Approach for studying NuRD complex binding kinetics and function: (Left) Single JF_{549}-HaloTagged NuRD complex molecules are tracked in 3D using a double-helix point spread function microscope – two puncta are recorded for each fluorophore with their midpoint providing the lateral x, y position and the angle between them representing the axial position in z relative to the nominal focal plane. (Middle) Examples of extracted single particle trajectories from 20 ms exposure imaging show periods of free unconfined and confined diffusion. (Right) Example trajectory of a single CHD4 molecule from 20 ms exposure imaging showing a sliding window of 11 frames (shaded green) from which the biophysical parameters are calculated. The resulting segmentation, into confined chromatin bound (blue) and unconfined freely diffusing (orange) sub-trajectories, has been used to colour the trajectory. 500 ms exposure imaging, on the other hand, can be used to extract slow and fast sub-diffusive motion of chromatin-bound CHD4 molecules. (e, f) Distribution of the four biophysical parameters extracted from sliding windows within the single-molecule trajectories of CHD4 taken with (e) 20 ms and (f) 500 ms exposures – (top) before and (bottom) after classification based on the anomalous exponent α, the effective diffusion coefficient D, the length of confinement Lc, and the drift magnitude, norm‖V‖ of the mean velocity. Each dataset was produced by imaging 30 cells, and the parameters were extracted from 5,557 and 15,528 trajectories, respectively (see **Extended Data Figure 5** and the Online Methods for more detail).


a
Separate NuRD sub-complexes
HDAC sub-complex
Remodeller CHD4
+ MBD3
- MBD3
Holo-NuRD complex
MTA
p46/ p48
HDAC
MBD3
GATA D2A
CHD4
b
Chromosome conformation capture (in-nucleus Hi-C)
Short-range interactions
Intermediate-range interactions
Wild-type
Mbd3-ko
Increase
Decrease
λ = 1.0
λ = 1.5
Contact probability (25kb bins)
Sequence separation (kb)
c
+ MBD3
- MBD3
More short-range interactions
More intermediate-range interactions
d
Double helix point spread function microscopy
Sample Trajectories
Trajectory segmentation algorithm
Angle = Z position
+2 µm
0 µm
-2 µm
10 µm
confined
unconfined
Freely diffusing
Chromatin bound
20 ms exposure
500 ms exposure
time (s)
X(k∆t)
X((k+l)∆t)
X((k-l)∆t)
e
Biophysical parameters from 20 ms exposure trajectories
Frequency
Anomalous exponent, α
Diffusion Coefficient (µm²/s)
Length of confinement (nm)
Magnitude of drift, V(µm)
Confined chromatin bound
Unconfined freely diffusing
f
Biophysical parameters from 500 ms exposure trajectories
Slow confined chromatin bound
Fast confined chromatin bound

Chromatin association of the deacetylase sub-complex requires assembly into the intact holo-NuRD complex in ES cells

Having developed an approach to segment the 20 ms exposure trajectories of the NuRD complex into chromatin bound and freely diffusing molecules (**Figure 2a,b**), we investigated how the chromatin binding kinetics were affected by removal of MBD3, which disrupts the interaction between the HDAC- and CHD4-containing NuRD sub-complexes.

Single-molecule tracking of CHD4 in live ES cells revealed a small but significant increase in the diffusion coefficient of freely diffusing CHD4 in the absence of MBD3 (median increases 1.05-fold, $p = 0.009$), consistent with the formation of smaller CHD4 sub-complexes[22-24] (**Figure 2c, Extended Data Figure 8**). A larger increase might have been expected to arise from the disassembly of the holo-NuRD complex, but ES cells also contain a substantial amount of CHD4 that is not in NuRD complexes[45].

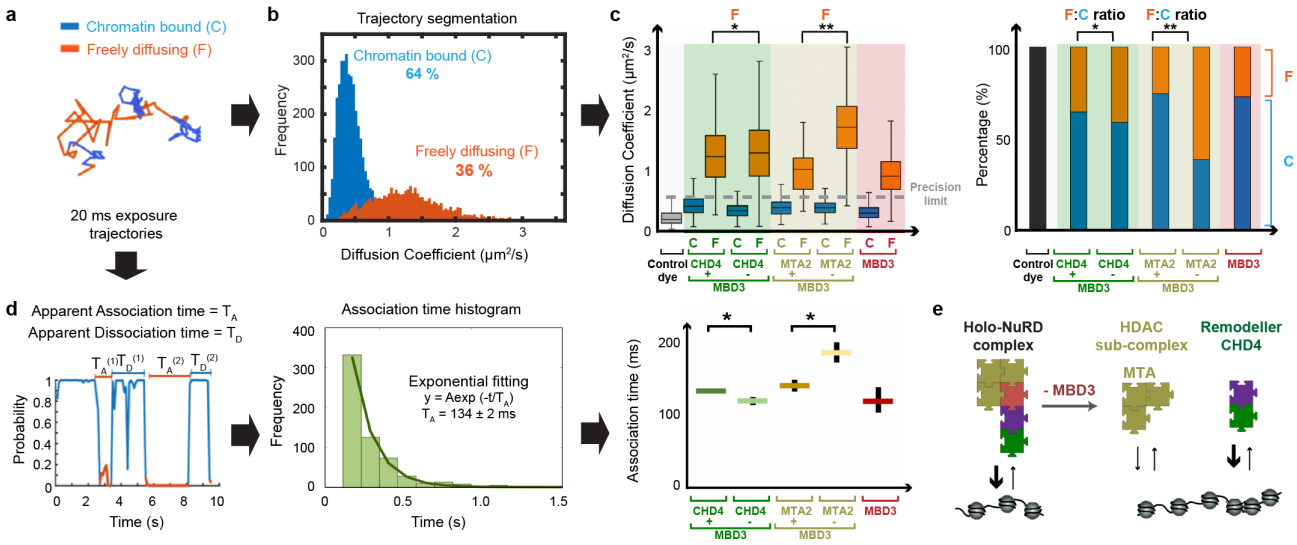
To explore whether the two sub-complexes are pre-assembled before binding to chromatin, we imaged the HDAC-containing sub-complex using tagged MTA2 molecules. This revealed a more substantial increase in diffusion coefficient for freely diffusing MTA2 in the absence of MBD3 (1.7-fold, $p < 10^{-5}$), consistent with the idea that the deacetylase subunit is normally associated with CHD4 in the intact holo-NuRD complex. We also imaged tagged MBD3 and showed that both freely diffusing MBD3 and MTA2 molecules have a similar diffusion coefficient, arguing that these two proteins are also normally associated (**Figure 2c**, **Extended Data Figure 8**). Finally, we showed that MBD3 does indeed interact with CHD4 via GATAD2A, both *in vitro* using purified GATAD2A in pull-down reconstitution experiments (**Extended Data Figure 8b**), and in ES cells where knock-down of GATAD2A and GATAD2B increased the diffusion coefficient of CHD4 (1.05-fold, $p < 10^{-4}$) (**Extended Data Figures 8c-e**). We conclude that NuRD normally exists as the intact complex *in vivo* and that, as with the removal of MBD3, depletion of GATAD2A/B is sufficient to unlink CHD4 from the HDAC-containing sub-complex in ES cells.

We then examined how the NuRD complex interacts with chromatin by comparing the percentage of freely diffusing versus chromatin bound CHD4 and MTA2 molecules in the presence and absence of MBD3. We observed a small decrease in the percentage of CHD4 molecules bound to chromatin in the absence of MBD3, but there was a much more significant decrease (2.4-fold) in the percentage of chromatin bound MTA2

molecules upon MBD3 depletion, suggesting that CHD4 rather than the deacetylase subunit is primarily responsible for the association of NuRD with chromatin (**Figure 2c**). This is consistent with the *in vitro* experiments which show that, in comparison to CHD4, the deacetylase subunit by itself does not bind strongly to nucleosomes (**Extended Data Figure 9a**).

To investigate the chromatin binding kinetics of the CHD4 remodeller and the MTA2 deacetylase sub-complex in the presence and absence of MBD3, we next determined association times from the time spent freely diffusing between confined chromatin-bound states. We also attempted to determine dissociation times from the time spent bound to chromatin between unconfined freely diffusing states (**Figure 2d** and **Extended Data Figure 5** – see also the Online Methods). The distribution of association and dissociation times were well approximated by a single exponential, suggesting a Poissonian process. Consistent with our conclusion that CHD4 is primarily responsible for recruitment of NuRD to chromatin, we found no increase in the association time of CHD4 upon removal of MBD3. (Instead, we found a small decrease consistent with the faster diffusion of the smaller CHD4 sub-complex.) However, we did find a significant increase in the association time of MTA2 upon MBD3 depletion (**Figure 2d**), consistent with the fact that the deacetylase subunit does not strongly associate with chromatin in the absence of CHD4 (**Extended Data Figure 9a**). Although no changes in dissociation time were observed (**Extended Data Figures 9b,c**), trajectories were likely truncated by photobleaching (our mean trajectory lengths were 8-11 frames). We therefore took advantage of 'motion blurring' when recording 500 ms trajectories (see below) to detect only chromatin bound proteins[1,46], and combined this with time-lapse imaging for different lengths of time, in an attempt to determine the dissociation time of the NuRD complex. To our surprise this showed that the dissociation times were much longer than we expected (greater than 100 seconds for MBD3, **Extended Data Figure 9d**), such that it proved impossible to track individual molecules for long enough in order to determine reliable rates. We conclude that, once bound to a target site the intact NuRD complex binds for unexpectedly long times.

Figure 2. Live-cell single-molecule tracking reveals that the assembly of the intact holo-NuRD complex increases chromatin association of its deacetylase sub-complex. (a) Segmentation of an example 20 ms trajectory into confined chromatin bound (C) (blue) and unconfined freely diffusing (F) states (orange). (b) Percentage of molecules and distribution of apparent diffusion coefficients for chromatin bound and freely diffusing CHD4 molecules. (c) (Left) Box plot of apparent diffusion coefficients for chromatin bound and freely diffusing CHD4 and MTA2 molecules compared to CHD4 and MTA2 in the absence of MBD3, and to MBD3 itself [*$p < 0.01$, **$p < 0.001$ (Kolmogorov-Smirnov test)]. The grey dotted line indicates the upper bound of the precision limit calculated at the 95 % confidence interval for immobilised JF_{549} dye molecules (Right) Percentage of freely diffusing and chromatin bound CHD4 and MTA2 molecules in the presence and absence of MBD3 [from Gaussian fitting, *$p < 0.01$, **$p < 0.001$ (2-way ANOVA)]. [The number of cells/trajectories used in the analysis were: 30/5,557 (CHD4), 25/2,337 (CHD4-MBD3), 10/336 (MTA2), 10/652 (MTA2-MBD3) and 30/2,224 (MBD3).] (d) (Left) A plot of the confinement probability allows determination of the association T_A and dissociation T_D times – defined respectively as the time a trajectory spends between periods of confined or unconfined motion. (Middle) A single exponential curve of rate lambda=$1/T_A$ is then fit to e.g. the distribution of association times. (Right) The association times extracted for CHD4 and MTA2 were then compared to those in the absence of MBD3, and to those for MBD3 itself [error bars show 95 % confidence intervals, *$p < 0.01$ (2-way ANOVA)]. (e) Schematic representation of a model in which MBD3-dependent holo-NuRD complex assembly increases the association rate of the deacetylase sub-complex.



The Holo-NuRD complex modulates chromatin movement at enhancers and promoters

Analysis of the chromatin bound NuRD complex protein molecules in 500 ms exposure trajectories initially revealed two states of chromatin bound CHD4 (**Figure 1f**): both the slow and fast states were primarily sub-diffusive (with an anomalous exponent α of around 0.5), but were characterised by different confinement lengths of 62 ± 12 nm and 110 ± 30 nm and apparent diffusion coefficients of 0.006 ± 0.002 $\mu m^2 s^{-1}$ and 0.018 ± 0.006 $\mu m^2 s^{-1}$. These experiments thus reveal for the first time dynamics below the 200-250 nm length scale that had previously been observed[43,47]. By determining similar parameters for fixed dye molecules, we found that the slow-diffusing state was within our precision limit and so represents molecules that are essentially immobile (**Extended Data Figure 10**).

Interestingly, visualisation of the fast state trajectories revealed a proportion of molecules exhibiting periods of directed motion (**Figure 3a**). We used Gaussian fitting to characterise the distributions in anomalous exponent for both the slow- and fast-states of chromatin bound CHD4 (in separate calculations). This revealed a single slow state (S) with α_s of 0.59 ± 0.01 (67 % of sub-trajectories) and two fast states (F1 and F2) with different anomalous exponents: α_{F1} of 0.60 ± 0.01 (26 %) and α_{F2} of 0.89 ± 0.02 (7 %) (**Figure 3b, Extended Data Figure 10)**. Molecules in the fast F1 state have the same distribution of anomalous exponents as in the slow state, and therefore explore the same chromatin environment, but they diffuse faster and have a larger length of confinement (and thus move further within the nucleus). The molecules in the fast F2 state have a higher anomalous exponent and explore a larger area of the nucleus than those in both the slow and the fast F1 states. Moreover, they have higher drift, indicative of movement in a directed manner, which is also consistent with the higher anomalous exponent (**Figures 3b,c**). Importantly, visualisation of the trajectories identified individual molecules that switch between the three states; S, F1 and F2 (see **Figure 3a**) – thus, the three states are unlikely to represent NuRD complex molecules binding to different regions of chromatin in the cell.

We then compared the dynamics of the MBD3 component to that of chromatin bound CHD4 and found that it also exhibited the one slow and two fast states. Both chromatin bound MBD3 and CHD4 molecules exhibited motion in the fast F1 and F2 states in around 22-26% and 7-8 % of trajectories, respectively,

confirming that these states are a property of the intact NuRD complex and not just of CHD4 (**Figure 3c, Extended Data Figure 10**).

We next asked whether loss of MBD3 (and thus disconnecting the remodeller and deacetylase sub-complexes) affects chromatin condensation and the fast chromatin states. Surprisingly, we found that the anomalous exponent, length of confinement and apparent diffusion coefficient of chromatin bound CHD4 all decreased in the fast F1 and F2 states (**Figure 3c**). The decrease in anomalous exponent unexpectedly suggests that in the absence of the intact NuRD complex chromatin is more *condensed*, whilst the diffusion coefficient and length of confinement show that CHD4 molecules diffuse more slowly and explore a smaller nuclear volume. We had expected to find that in the absence of deacetylation by NuRD, that the more acetylated chromatin would be *decondensed*[48-52] and that CHD4 would explore a greater nuclear volume. Notably, however, in the absence of MBD3 we observed a significant increase in the proportion of CHD4 molecules in the fast F2 state exhibiting directed motion (18 % vs 7.4 %, see **Figure 3d**, **Extended Data Figure 10**).

The fast F1 and F2 states of chromatin bound NuRD could result from movement on DNA due to chromatin remodelling or, bearing in mind the long dissociation times we determined for CHD4 (see above), from movement of NuRD-bound chromatin. To distinguish between these possibilities, we used the CARGO-dCas9 system[4] to track the dynamics of a well-characterised NuRD-bound enhancer near the *Tbx3* gene. A 1 kb region of the enhancer was labelled by expressing 36 gRNAs to target GFP-tagged inactive dCas9 molecules to this site[4] (**Figure 4a, Extended Data Figure 11**). We imaged the enhancer in cells showing bright undivided foci to reduce the likelihood of using data from cells undergoing DNA replication (which exhibit blurred foci or doublets). Because the background fluorescence from freely diffusing dCas9-GFP prevented 3D tracking of specific genomic loci on our DH-PSF microscope, we tracked the *Tbx3* enhancer locus in 2D. Although this meant that we could not directly compare the parameters obtained in the 2D (CARGO-dCas9) and 3D (NuRD single molecule) tracking experiments, classification of the sub-trajectories using our Gaussian mixture model and subsequent Gaussian fitting analysis of the anomalous exponent distributions once again revealed one slow and two fast chromatin states. In particular, as was the case when tracking CHD4, removal of MBD3 significantly decreased the rate of diffusion of the *Tbx3* enhancer in the fast-diffusing F1 and F2 states and increased the proportion of sub-trajectories in the fast decondensed F2 state (**Figure 4b, Extended Data Figure 11**). Whilst we cannot rule out that some of the fast motions we observe when imaging the NuRD complex

might result from movement of the complex along DNA due to chromatin remodelling, given that NuRD associates predominantly with active enhancer and promoters in ES cells, and that similar movements were seen for NuRD-bound foci and the *Tbx3* enhancer, these results suggest that single molecule tracking of chromatin-bound NuRD largely monitors the spatial dynamics of active enhancers and promoters. Further, our data indicate that the assembly of the intact holo-NuRD complex increases the search space that enhancers/promoters can explore in the nucleus, changing the length scale at which interactions can occur (see **Figure 3c**). This increase in search space of enhancers and promoters in the presence of the intact NuRD complex is consistent with and provides a molecular explanation for the increase in intermediate-range Hi-C contacts that we observe in wild-type compared to *Mbd3-ko* ES cells (**Figure 1b**).

How might the fast sub-diffusive states of the bound NuRD complex (F1 and F2) relate to the previous enhancer/promoter tracking experiments of Gu *et al.* (2018), and their transcription-dependent enhancer 'stirring' model? The similarity in both the diffusion parameters and anomalous exponents of their slow/fast and our S/F1 states suggests that these correspond to either inactive (slow/S) or actively transcribed (fast/F1) enhancers and promoters. In our experiments, however, we additionally observe a fast F2 state which exhibits directed motion and we wondered whether this was due to chromatin movement resulting from transcriptional elongation. We therefore tracked bound CHD4 molecules after adding a small molecule inhibitor of transcriptional elongation (DRB[53]). Consistent with the results of Gu *et al.* (2018), premature termination by DRB led to a reduction in the proportion of bound CHD4 molecules exhibiting the fast transcription dependent F1 motion (from 26 % to 19 %). However, there was no significant change in the proportion of molecules in the fast decondensed F2 state nor in the chromatin environment (i.e. in the anomalous exponent or length of confinement) in the presence of a block on transcriptional elongation (**Extended Data Figures 10, 12a**). Finally, we tracked MBD3 molecules while blocking HDAC1/2 deacetylase activity with FK228[54]. As was the case when we prevented the association of the deacetylase subunit with CHD4 (see **Figure 3c**) we observed a decrease in the anomalous exponent (reflecting a more condensed state), but again there was no significant change in the proportion of molecules in the fast decondensed F2 state (**Extended Data Figures 10, 12b**). We conclude that the fast F2 state is not dependent on transcriptional elongation or deacetylation activity. Instead the results suggest that the assembly of the intact holo-NuRD complex (from both the chromatin remodelling

and deacetylase subunits) modulates the dynamics of active enhancers/promoters, affecting their conversion between the slow S, fast F1 and fast F2 states (**Figure 4c**).

Figure 3. Assembly of the intact NuRD complex reduces directed motion and decondenses chromatin. (a) (Left) Example trajectory of a chromatin bound CHD4 molecule showing periods of both slow (dark green) and fast sub-diffusive motion (light green). Two fast states (F1 and F2) are observed, with the F2 state showing periods of directed motion. (Right) The four biophysical parameters calculated along this trajectory with the fast F2 sub-trajectories showing a higher anomalous exponent, increased length of confinement and increased drift. (b) (Left) Gaussian fitting to the distribution of the anomalous exponents identifies a single α for slow confined NuRD complex molecules (dark blue), and two α values for molecules in the faster confined state (light blue). The resulting distribution of the lengths of confinement (Middle) and apparent diffusion coefficients (Right) are also shown. (c) Comparison of biophysical parameters for the CHD4 remodeller in the presence and absence of MBD3, and for MBD3 itself [the number of cells/trajectories used in the analysis were: 30/3,059 (CHD4), 15/2,111 (CHD4-MBD3), and 30/1,816 (MBD3)]. (Left) The anomalous exponents resulting from Gaussian fitting (error bars show 95 % confidence intervals, $*p < 0.01$, 2-way ANOVA). Boxplots of the lengths of confinement (Middle) and of the apparent diffusion coefficients (Right) are also shown, with the precision limit of the single molecule imaging indicated by a grey dotted line ($*p < 0.01$, Kolmogorov-Smirnov test). (We imaged JF_{549} dye fixed on a coverslip and found that all three biophysical parameters were significantly lower than the S, F1 and F2 states.) (d) (Left) Percentage of molecules in the slow or fast chromatin bound states (from the Gaussian fitting, $*p < 0.01$, 2-way ANOVA). (Right) Schematic representation of the three states of chromatin-bound NuRD: a slow sub-diffusive S state with small exploration radius, one fast sub-diffusive F1 state with a larger length of confinement (exploration radius) and another fast sub-diffusive F2 state exhibiting directed motion.

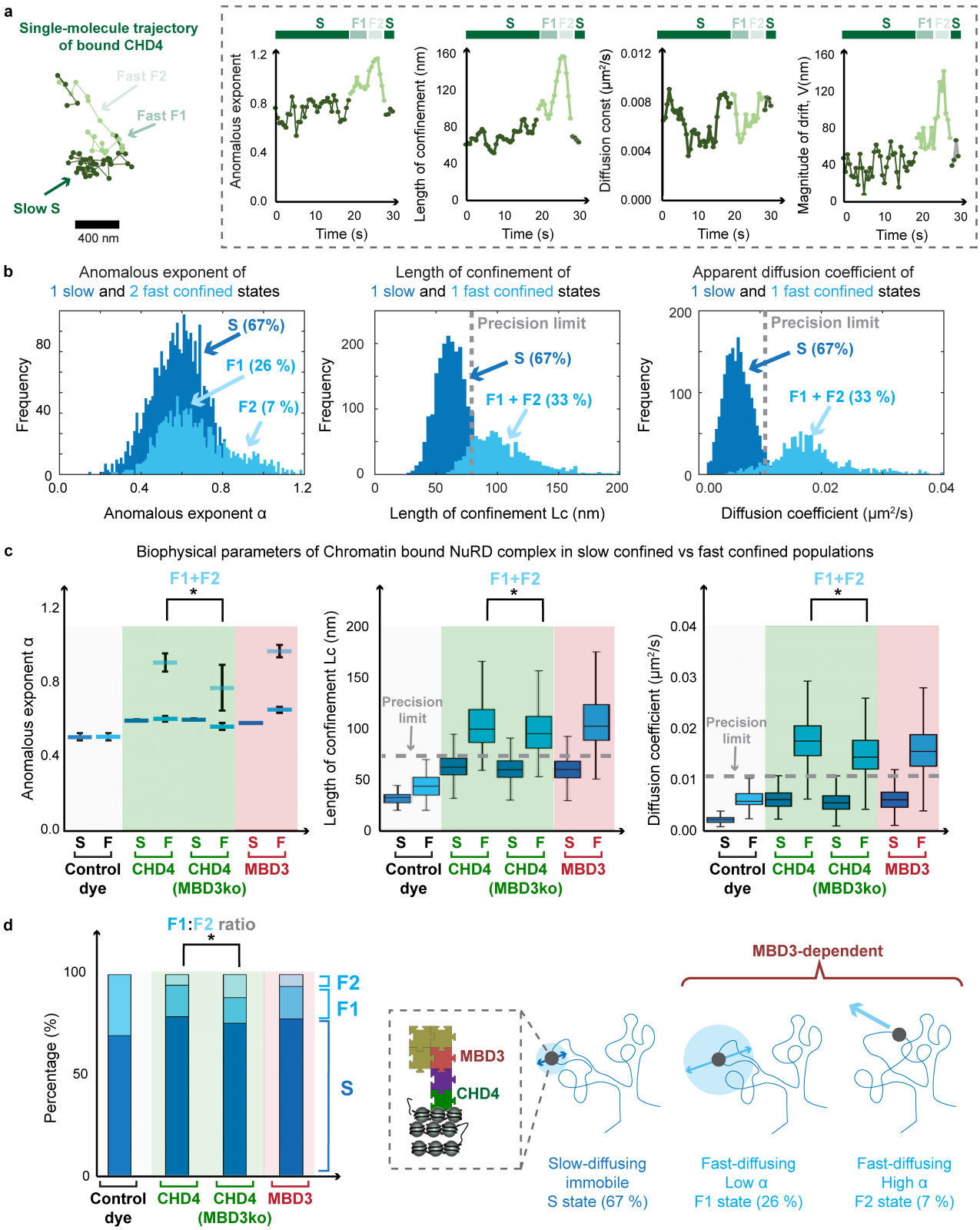
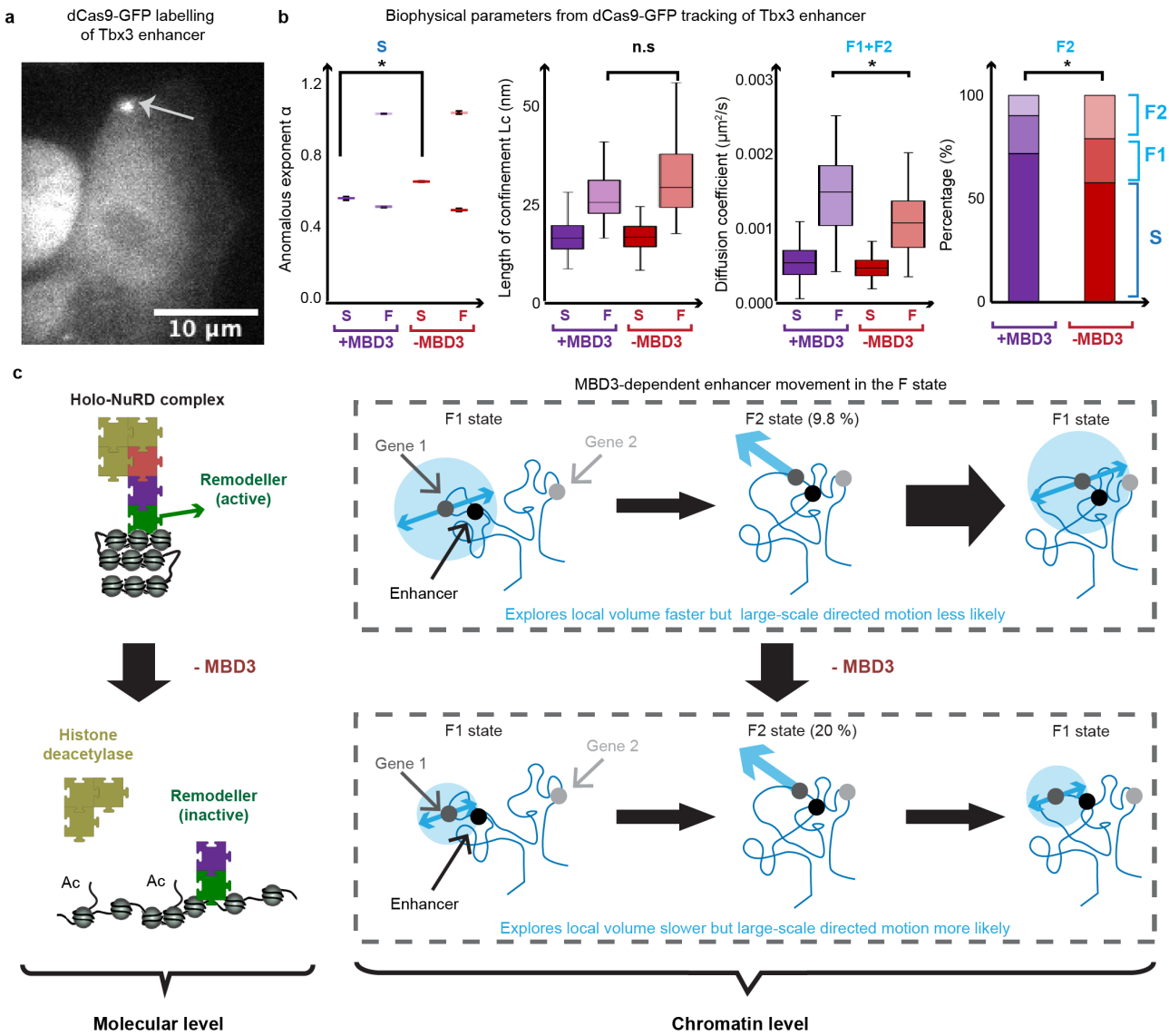

a
Single-molecule trajectory of bound CHD4
Fast F2
Fast F1
Slow S
400 nm
S F1 F2 S
Anomalous exponent
Length of confinement (nm)
Diffusion const (μm²/s)
Magnitude of drift, V(nm)
Time (s)
b
Anomalous exponent of 1 slow and 2 fast confined states
Length of confinement of 1 slow and 1 fast confined states
Apparent diffusion coefficient of 1 slow and 1 fast confined states
S (67%)
F1 (26 %)
F2 (7 %)
Precision limit
F1 + F2 (33 %)
Frequency
Anomalous exponent α
Length of confinement Lc (nm)
Diffusion coefficient (μm²/s)
c
Biophysical parameters of Chromatin bound NuRD complex in slow confined vs fast confined populations
F1+F2
*
Anomalous exponent α
Length of confinement Lc (nm)
Diffusion coefficient (μm²/s)
Precision limit
S F
Control dye
CHD4
CHD4 (MBD3ko)
MBD3
d
F1:F2 ratio
Percentage (%)
F2
F1
S
MBD3
CHD4
MBD3-dependent
Slow-diffusing immobile S state (67 %)
Fast-diffusing Low α F1 state (26 %)
Fast-diffusing High α F2 state (7 %)

Figure 4. The intact complex modulates the movement of NuRD-regulated enhancers. (a) Image of dCas9-GFP targeted to the *Tbx3* Enhancer locus via CARGO plasmids expressing 36 gRNAs. (b) The four biophysical parameters were calculated along the trajectories after tracking the position of the *Tbx3* enhancer in the presence and absence of MBD3 [the number of trajectories: 237 (+MBD3) and 287 (- MBD3) with 1-4 enhancers labelled per cell]. (Left) Gaussian fitting to the distribution of the anomalous exponents identifies a single α for slow confined *Tbx3* loci, and two α values for molecules in the faster confined state (error bars show 95 % confidence intervals from the Gaussian fitting, *$p < 0.01$, 2-way ANOVA). (Middle) Boxplots of localisation length and apparent diffusion coefficient (*$p < 0.01$, Kolmogorov-Smirnov test). (Right) Percentage of sub-trajectories of the *Tbx3* enhancer loci exhibiting slow or fast chromatin motion (from the Gaussian fitting, *$p < 0.01$, 2-way ANOVA). (c) Schematic model depicting: (Left) a molecular level view of how the deacetylase subunit is required for NuRD complex remodelling; and (Right) a chromatin level view of how NuRD increases the volume explored by an enhancer while at the same time reducing its likelihood of entering the F2 state in which the directed motion of chromatin might re-organise enhancer-promoter proximity.


a
dCas9-GFP labelling of Tbx3 enhancer
10 µm
b
Biophysical parameters from dCas9-GFP tracking of Tbx3 enhancer
S
*
Anomalous exponent α
1.2
0.8
0.4
0.0
n.s
Length of confinement Lc (nm)
50
25
0
F1+F2
*
Diffusion coefficient (µm²/s)
0.003
0.002
0.001
0.000
F2
*
Percentage (%)
100
50
0
F2
F1
S
S F S F
+MBD3 -MBD3
c
Holo-NuRD complex
Remodeller (active)
- MBD3
Histone deacetylase
Remodeller (inactive)
Ac
Ac
Molecular level
MBD3-dependent enhancer movement in the F state
F1 state
Gene 1
Gene 2
Enhancer
F2 state (9.8 %)
F1 state
Explores local volume faster but large-scale directed motion less likely
- MBD3
F1 state
Gene 1
Gene 2
Enhancer
F2 state (20 %)
F1 state
Explores local volume slower but large-scale directed motion more likely
Chromatin level

The NuRD complex modulates the dynamics of enhancers and promoters to control transcription

By tracking the dynamics of the *Tbx3* enhancer we were able to show that movement of the chromatin bound NuRD complex reflects enhancer/promoter dynamics. How might the NuRD complex influence transcription? Histone acetylation is thought to lead to a looser packing between nucleosomes and chromatin decompaction[48-52], but here, by studying compaction in the immediate vicinity of the stably bound NuRD complex, we unexpectedly find that recruitment of the deacetylase subunit leads to a more open chromatin structure where regulatory elements can diffuse faster and explore more of the nucleus (**Figure 3c**). In addition, however, we found that assembly of the intact chromatin-bound NuRD complex reduces the time NuRD-regulated enhancers/promoters spend in a decondensed chromatin state undergoing directed motion, and we propose that this may limit the re-organisation of enhancer/promoter interactions leading to their being maintained for longer. We also found that the intact NuRD complex leads to an increase in both CTCF and (more particularly) Cohesin binding, suggesting a possible mechanism for our findings. Our results suggest that the intact NuRD complex may increase or stabilise CTCF/Cohesin binding to promote a chromatin environment whereby enhancers and promoters can contact each other over longer distances. This increased CTCF/Cohesin binding might also increase cross-linking between different enhancer and promoter containing regions of chromatin, thereby reducing the fast F2 motions that our imaging approach has uncovered.

References (main)

1 Gebhardt, J. C. *et al.* Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. *Nature methods* **10**, 421-426, doi:10.1038/nmeth.2411 (2013).

2 Chen, J. *et al.* Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell* **156**, 1274-1285, doi:10.1016/j.cell.2014.01.062 (2014).

3 Liu, Z. & Tjian, R. Visualizing transcription factor dynamics in living cells. *J Cell Biol* **217**, 1181-1191, doi:10.1083/jcb.201710038 (2018).

4 Gu, B. *et al.* Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements. *Science* **359**, 1050-1055, doi:10.1126/science.aao3136 (2018).

5 Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F. & de Laat, W. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* **10**, 1453-1465, doi:10.1016/s1097-2765(02)00781-5 (2002).

6 Dostie, J. *et al.* Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome research* **16**, 1299-1309, doi:10.1101/gr.5571506 (2006).

7 Zhao, Z. *et al.* Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat Genet* **38**, 1341-1347, doi:10.1038/ng1891 (2006).

8 Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-293, doi:10.1126/science.1181369 (2009).

9 Ethier, S. D., Miura, H. & Dostie, J. Discovering genome regulation with 3C and 3C-related technologies. *Biochim Biophys Acta* **1819**, 401-410, doi:10.1016/j.bbagrm.2011.12.004 (2012).

10 Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380, doi:10.1038/nature11082 (2012).

11 Rao, S. S. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680, doi:10.1016/j.cell.2014.11.021 (2014).

12 Nora, E. P. *et al.* Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**, 930-944 e922, doi:10.1016/j.cell.2017.05.004 (2017).

13 Nagano, T. *et al.* Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59-64, doi:10.1038/nature12593 (2013).

14 Lando, D. *et al.* Combining fluorescence imaging with Hi-C to study 3D genome architecture of the same single cell. *Nature protocols* **13**, 1034-1061, doi:10.1038/nprot.2018.017 (2018).

15 Nagano, T., Wingett, S. W. & Fraser, P. Capturing Three-Dimensional Genome Organization in Individual Cells by Single-Cell Hi-C. *Methods Mol Biol* **1654**, 79-97, doi:10.1007/978-1-4939-7231-9_6 (2017).

16 Stevens, T. J. *et al.* 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* **544**, 59-64, doi:10.1038/nature21429 (2017).

17 Zhang, Y., LeRoy, G., Seelig, H. P., Lane, W. S. & Reinberg, D. The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell* **95**, 279-289, doi:10.1016/s0092-8674(00)81758-4 (1998).

18 Tong, J. K., Hassig, C. A., Schnitzler, G. R., Kingston, R. E. & Schreiber, S. L. Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. *Nature* **395**, 917-921, doi:10.1038/27699 (1998).

19 Xue, Y. *et al.* NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Mol Cell* **2**, 851-861, doi:10.1016/s1097-2765(00)80299-3 (1998).

20 Wade, P. A., Jones, P. L., Vermaak, D. & Wolffe, A. P. A multiple subunit Mi-2 histone deacetylase from Xenopus laevis cofractionates with an associated Snf2 superfamily ATPase. *Curr Biol* **8**, 843-846, doi:10.1016/s0960-9822(98)70328-8 (1998).

21 Whyte, W. A. *et al.* Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature* **482**, 221-225, doi:10.1038/nature10805 (2012).

22 Bornelov, S. *et al.* The Nucleosome Remodeling and Deacetylation Complex Modulates Chromatin Structure at Sites of Active Transcription to Fine-Tune Gene Expression. *Mol Cell* **71**, 56-72 e54, doi:10.1016/j.molcel.2018.06.003 (2018).

23 Zhang, W. *et al.* The Nucleosome Remodeling and Deacetylase Complex NuRD Is Built from Preformed Catalytically Active Sub-modules. *J Mol Biol* **428**, 2931-2942, doi:10.1016/j.jmb.2016.04.025 (2016).

24 Carr, A. R. *et al.* Three-Dimensional Super-Resolution in Eukaryotic Cells Using the Double-Helix Point Spread Function. *Biophysical journal* **112**, 1444-1454, doi:10.1016/j.bpj.2017.02.023 (2017).

25 Low, J. K. *et al.* CHD4 Is a Peripheral Component of the Nucleosome Remodeling and Deacetylase Complex. *J Biol Chem* **291**, 15853-15866, doi:10.1074/jbc.M115.707018 (2016).

26 Millard, C. J. *et al.* The structure of the core NuRD repression complex provides insights into its interaction with chromatin. *eLife* **5**, e13941, doi:10.7554/eLife.13941 (2016).

27 Gnanapragasam, M. N. *et al.* p66Alpha-MBD2 coiled-coil interaction and recruitment of Mi-2 are critical for globin gene silencing by the MBD2-NuRD complex. *Proc Natl Acad Sci U S A* **108**, 7487-7492, doi:10.1073/pnas.1015341108 (2011).

28 Kaji, K., Nichols, J. & Hendrich, B. Mbd3, a component of the NuRD co-repressor complex, is required for development of pluripotent cells. *Development* **134**, 1123-1132, doi:10.1242/dev.02802 (2007).

29 Kaji, K. *et al.* The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. *Nat Cell Biol* **8**, 285-292, doi:10.1038/ncb1372 (2006).

30 Reynolds, N., O'Shaughnessy, A. & Hendrich, B. Transcriptional repressors: multifaceted regulators of gene expression. *Development* **140**, 505-512, doi:10.1242/dev.083105 (2013).

31 Reynolds, N. *et al.* NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment. *Cell stem cell* **10**, 583-594, doi:10.1016/j.stem.2012.02.020 (2012).

32 O'Shaughnessy-Kirwan, A., Signolet, J., Costello, I., Gharbi, S. & Hendrich, B. Constraint of gene expression by the chromatin remodelling protein CHD4 facilitates lineage specification. *Development* **142**, 2586-2597, doi:10.1242/dev.125450 (2015).

33 Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430-435, doi:10.1038/nature09380 (2010).

34 Symmons, O. *et al.* Functional and topological characteristics of mammalian regulatory domains. *Genome research* **24**, 390-400, doi:10.1101/gr.163519.113 (2014).

35 Symmons, O. *et al.* The Shh Topological Domain Facilitates the Action of Remote Enhancers by Reducing the Effects of Genomic Distances. *Developmental cell* **39**, 529-543, doi:10.1016/j.devcel.2016.10.015 (2016).

36 Wendt, K. S. *et al.* Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**, 796-801, doi:10.1038/nature06634 (2008).

37 Hadjur, S. *et al.* Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* **460**, 410-413, doi:10.1038/nature08079 (2009).

38 Zuin, J. *et al.* Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc Natl Acad Sci U S A* **111**, 996-1001, doi:10.1073/pnas.1317788111 (2014).

39 Mizuguchi, T. *et al.* Cohesin-dependent globules and heterochromatin shape 3D genome architecture in S. pombe. *Nature* **516**, 432-435, doi:10.1038/nature13833 (2014).

40 Ing-Simmons, E. *et al.* Spatial enhancer clustering and regulation of enhancer-proximal genes by cohesin. *Genome research* **25**, 504-513, doi:10.1101/gr.184986.114 (2015).

41 Amitai, A., Seeber, A., Gasser, S. M. & Holcman, D. Visualization of Chromatin Decompaction and Break Site Extrusion as Predicted by Statistical Polymer Modeling of Single-Locus Trajectories. *Cell reports* **18**, 1200-1214, doi:10.1016/j.celrep.2017.01.018 (2017).

42 Metzler, R. The future is noisy: the role of spatial fluctuations in genetic switching. *Phys Rev Lett* **87**, 068103, doi:10.1103/PhysRevLett.87.068103 (2001).

43 Shukron, O., Seeber, A., Amitai, A. & Holcman, D. Advances Using Single-Particle Trajectories to Reconstruct Chromatin Organization and Dynamics. *Trends in genetics : TIG* **35**, 685-705, doi:10.1016/j.tig.2019.06.007 (2019).

44 Verdaasdonk, J. S. *et al.* Centromere tethering confines chromosome domains. *Mol Cell* **52**, 819-831, doi:10.1016/j.molcel.2013.10.021 (2013).

45 O'Shaughnessy, A. & Hendrich, B. CHD4 in the DNA-damage response and cell cycle progression: not so NuRDy now. *Biochem Soc Trans* **41**, 777-782, doi:10.1042/BST20130027 (2013).

46 Etheridge, T. J. *et al.* Quantification of DNA-associated proteins inside eukaryotic cells using single-molecule localization microscopy. *Nucleic Acids Res* **42**, e146, doi:10.1093/nar/gku726 (2014).

47 Hauer, M. H. *et al.* Histone degradation in response to DNA damage enhances chromatin dynamics and recombination rates. *Nat Struct Mol Biol* **24**, 99-107, doi:10.1038/nsmb.3347 (2017).

48 Collepardo-Guevara, R. *et al.* Chromatin Unfolding by Epigenetic Modifications Explained by Dramatic Impairment of Internucleosome Interactions: A Multiscale Computational Study. *J Am Chem Soc* **137**, 10205-10215, doi:10.1021/jacs.5b04086 (2015).

49 Otterstrom, J., Garcia, A. C., Vicario, C., Cosma, M. P. & Lakadamyali, M. Nanoscale decompaction of nucleosomal DNA revealed through multi-color super-resolution microscopy. *bioRxiv*, 470823, doi:10.1101/470823 (2018).

50 Toth, K. F. *et al.* Trichostatin A-induced histone acetylation causes decondensation of interphase chromatin. *J Cell Sci* **117**, 4277-4287, doi:10.1242/jcs.01293 (2004).

51 Annunziato, A. T., Frado, L. L., Seale, R. L. & Woodcock, C. L. Treatment with sodium butyrate inhibits the complete condensation of interphase chromatin. *Chromosoma* **96**, 132-138, doi:10.1007/bf00331045 (1988).

52 Wang, X., He, C., Moore, S. C. & Ausio, J. Effects of histone acetylation on the solubility and folding of the chromatin fiber. *J Biol Chem* **276**, 12764-12768, doi:10.1074/jbc.M100501200 (2001).

53 Bensaude, O. Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? *Transcription* **2**, 103-108, doi:10.4161/trns.2.3.16172 (2011).

54 Furumai, R. *et al.* FK228 (depsipeptide) as a natural prodrug that inhibits class I histone deacetylases. *Cancer Res* **62**, 4916-4921 (2002).

Methods

A) In-nucleus chromosome conformation capture (Hi-C)

In-nucleus Hi-C was carried out on E14 wild-type and 7g9 *Mbd3-ko* ES cells as previously described[16,29]. 50 bp paired-end sequencing was carried out on a HiSeq4000 instrument. Sequences were mapped to the genome using NucProcess (https://github.com/tjs23/nuc_processing) and analysed further using NucTools (https://github.com/tjs23/nuc_tools). Enhancer-promoter interactions were analysed using active enhancers previously defined in mouse ES cells by the presence of H3K27Ac, H3K4me1, p300 and lack of H3K4me3, whilst active promoters were defined by the presence of H3K4me3[16]. CTCF/SMC3-mediated looping was analysed using loops previously defined in mouse ES cells by Pekowska et al[55].

B) CTCF and SMC3 ChIP-seq in wild-type and *Mbd3* knockout ES cells

ChIPseq was carried out as previously described[22] with DNA fragmentation in the presence of 0.6% SDS. The antibodies used were:

ANTIBODY	RAISED IN	COMPANY	CATALOGUE NUMBER	WESTERN BLOT	ChIP	Clonality
αCTCF	Rabbit	Millipore	07-729	-	5µl	Polyclonal
αSMC3	Rabbit	Abcam	ab9263	-	5µl	Polyclonal

50 bp single-end sequencing was carried out on a HiSeq4000 instrument – two biological replicates per sample were obtained with 38-45 million total and 20-32 million mapped reads, respectively, whilst the input samples had 38-40 million total and 11-14 million mapped reads, respectively. All ChIP-seq data was trimmed using trim_galore and then aligned using standard BWA parameters to the Mus Musculus reference genome (mm10) (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)[56]. Pie charts and_heatmaps of ChIP-seq enrichment were made using Deeptools v2.5.0[57]. ChIP-seq bigwig tracks were calculated from the bamCompare output with options --extendReads 200 --binSize 1 --ratio log2. The coverage was calculated with computeMatrix reference-point with options --binSize 10. The heatmap of standardised signal was then plotted

using plotHeatmap in R version 3.6.0. Peaks were called using MACS2 so as to give a False Discovery Rate of 5% and above 5-fold enrichment.

C) Mouse embryonic stem cell line generation

Mouse embryonic stem (ES) cell lines were cultured in standard serum and mouse leukemia inhibitory factor (mLIF) conditions: Glasgow minimum essential medium (Sigma-Aldrich G5154) containing 100 mM 2-mercaptoethanol (Life tech, cat. 21985023), 1x minimum essential medium non-essential amino acids (Sigma-Aldrich, M7145), 2 mM L-glutamine (Life tech, cat. 25030024), 1 mM sodium pyruvate (Sigma-Aldrich, S8636-100ML), 10 % fetal bovine serum (HyClone FBS, Lot nr SZB20006, GE Healthcare Austria SV30180.03) and 10 ng/ml mLIF (provided by the Biochemistry Department, University of Cambridge). They were passaged every two days by washing in PBS (Sigma-Aldrich, D8537), adding Trypsin-EDTA 0.25 % (Life tech, cat. 25200072) to detach the cells, and then washing in media before re-plating in fresh media. To help the cells attach to the surface, plates were incubated for 15 minutes at room temperature in PBS containing 0.1 % gelatin (Sigma Aldrich, G1890). The background E14tg2a ES cell lines (available from Sigma Aldrich, 08021401) were characterized by qPCR, RNA-seq, ChIP-seq, and potency assays, and they were routinely screened for mycoplasma contamination and tested negative.

ES cells expressing CHD4 tagged at the C-terminus with HaloTag were generated in the presence and absence of MBD3, as previously described[16,23,24]. Briefly, this was achieved by CRISPR/Cas9 based knock-in of a cassette containing mEos3.2-HaloTag-FLAG and a puromycin selection gene into one CHD4 allele of the ES cells (**Extended Data Fig 4a**). The puromycin cassette was then removed using Dre recombinase to generate the CHD4 allele with a C-terminal HaloTag fusion. Since knockout of CHD4 is lethal, we used cell viability assays to verify that the function of the tagged CHD4 was unaffected. We similarly generated knock-in ES cells in an E14Tg2a (XY) background expressing MBD3 and MTA2 tagged at the C-terminus with HaloTag (**Extended Data Fig 4b-c**). MTA2-HaloTag knock-in cell lines were generated in MBD3-inducible ES cells[22] (**Extended Data Fig 4c**), in which MBD3 is fused to oestrogen receptor at the N- and C-termini so that it initially localises at the cytoplasm but then translocates to the nucleus when induced for 48 hours with 4-hydroxytamoxifen added directly to the culture media to a final concentration of 0.4 nM. Western blots and

immunoprecipitation experiments were carried out on nuclear lysates, as previously described [22], to confirm the expression and assembly of the NuRD complex (**Extended Data Fig 4**). The antibodies used were:

ANTIBODY	RAISED IN	COMPANY	CATALOGUE NUMBER	WESTERN BLOT	Clonality
αCHD4	Mouse	Abcam	ab70469	1:5000	Monoclonal [3F2/4]
αFlag	Mouse	Sigma	F1804	1:5000	Monoclonal M2
αGATAD2A	Rabbit	Abcam	ab87663	1:2000	Polyclonal
αHDAC1	Rabbit	Abcam	ab7028	1:2000	Polyclonal
αMBD3	Rabbit	Abcam	ab157464	1:5000	Monoclonal [EPR9913]
αMTA2	Mouse	Abcam	ab50209	1:5000	Monoclonal [MTA2-276]
αPCNA	Mouse	Santa cruz	Sc56	1:2000	Monoclonal [PC10]

D) Mouse ES cell live-cell 3D single-molecule imaging

ES cells expressing HaloTag-tagged MBD3, CHD4 and MTA2 were passaged two days before imaging onto 35 mm glass bottom dishes No 1.0 (MatTek Corporation P35G-1.0-14-C Case) in serum/LIF imaging medium: Fluorobrite™ DMEM (Thermo Fisher Scientific, A1896701) containing 100 mM 2-mercaptoethanol (Life tech, cat. 21985023), 1x minimum essential medium non-essential amino acids (Sigma-Aldrich, M7145), 2 mM L-glutamine (Life tech, cat. 25030024), 1 mM sodium pyruvate (Sigma-Aldrich, S8636-100ML), 10 % fetal bovine serum (HyClone FBS, Lot nr SZB20006, GE Healthcare Austria SV30180.03) and 10 ng/ml mLIF (provided by the Biochemistry Department, University of Cambridge). Just before single-molecule imaging experiments, cells were labelled with 0.5 nM HaloTag®-JF_{549} ligand for 15 minutes, followed by two washes in PBS and a 30 minute incubation at 37 °C in imaging medium, before imaging the cells in fresh serum/LIF imaging medium. Cells were under-labelled to prevent overlap of fluorophores during single-molecule tracking experiments. The HaloTag dyes were a kind gift from Luke D. Lavis (HHMI).

Transcription elongation was inhibited using 100 μM 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) and deacetylase activity using 10 nM FK228 (TOCRIS Bioscience, UK) both for two hours prior to imaging[54,58]. qPCR experiments confirmed inhibition of transcription upon addition of DRB (data not shown).

A custom-made double-helix point spread function (DHPSF) microscope was then used for 3D single-molecule tracking as previously described[24]. The setup incorporates an index-matched 1.2 NA water immersion objective lens (Plan Apo VC 60×, Nikon, Tokyo, Japan) to facilitate imaging above the coverslip surface. The DHPSF transformation was achieved by the use of a 580 nm optimized double-helix phase mask (PM) (DoubleHelix, Boulder, CO) placed in the Fourier domain of the emission path of a fluorescence microscope (Eclipse Ti-U, Nikon). The objective lens was mounted onto a scanning piezo stage (P-726 PIFOC, PI, Karlsruhe, Germany) to calibrate the rotation rate of the DHPSF. A 4*f* system of lenses placed at the image plane relayed the image onto an EMCCD detector (Evolve Delta 512, Photometrics, Tucson, AZ). Excitation and activation illumination was provided by 561 nm (200 mW, Cobolt Jive 100, Cobolt, Solna, Sweden) and 405 nm (120 mW, iBeam smart-405-s, Toptica, Munich, Germany) lasers, respectively, that were circularly polarized, collimated, and focused to the back focal plane of the objective lens. Oblique-angle illumination imaging was achieved by aligning the laser off axis such that the emergent beam at the sample interface was near-collimated and incident at an angle less than the critical angle $\theta_c \sim 67°$ for a glass/water interface. The fluorescence signal was then separated from the excitation beams into the emission path by a quad-band dichroic mirror (Di01-R405/488/561/635-25x36, Semrock, Rochester, NY) before being focused into the image plane by a tube lens. Finally, long-pass and band-pass filters (BLP02-561R-25 and FF01-580/14-25, respectively; Semrock) placed immediately before the camera isolated the fluorescence emission. Using 561 nm excitation, fluorescence images were collected as movies of 60,000 frames at 20 ms or 4,000 frames at 500 ms exposure. A continuous 561 nm excitation beam at ~1 kW/cm^2 was used for 20 ms exposure imaging and at ~40 W/cm^2 for 500 ms exposure imaging. Each experiment was carried out with at least 3 biological replicates (3 fields of view, each containing around 3 cells).

E) Residence time analysis from time-lapse 500 ms exposure imaging

Since photobleaching is related to the number of exposures, and the residence time is related to the time a molecule spends bound to chromatin, it is possible to change the time-lapse between exposures and use the data to extract both the residence time[1] and photobleaching rate. However, when we imaged at time intervals of 0.5 s, 2.5 s, 8 s and 32 s, we discovered that at the longest time lapse (32 s) we could see no decrease in the mean number of frames imaged before photobleaching, implying the residence time had no impact on the

measurement, which was thus dominated by photobleaching (**Extended Data Figure 8d**). To estimate the residence time would likely require imaging at much longer time-lapses, but because chromosomes and the cell itself move significantly during periods longer than this, it becomes unreliable to track individual chromatin bound NuRD complex subunits.

F) 3D single-molecule image processing and generation of trajectories

Single molecules were localized from 3D movies using the easy-DHPSF software[59] with a relative localization threshold of 100 in all 6 angles for 20 ms data and relative thresholds of 116, 127, 119, 99, 73 and 92 for the 500 ms data. Trajectories of individual molecules were then determined using custom Python code for connecting molecules in subsequent frames if they were within 800 nm for 20 ms trajectories and within 500 nm for 500 ms trajectories (https://github.com/wb104/trajectory-analysis).

G) Single-molecule trajectory analysis

In Section 1, below, we describe a switching model for a stochastic process between two or more states, characterized by different diffusion coefficients. We initially used this approach to justify the use of a two-state diffusion model to classify the distribution of displacements measured along single molecule trajectories. In later work we went on to develop an improved and novel method/algorithm to classify sub-trajectories into confined and unconfined states based on four physical parameters using a Gaussian mixture model. This approach is described in Section 2. However, the original classification based on diffusion coefficients provided a complementary analysis which justifies the results we obtained later and which we include here. The algorithm using the Gaussian mixture model was then tested on simulated trajectories before being applied to our live-cell single-molecule imaging data. In Section 3, we describe how we applied this classification algorithm to single particle trajectories to estimate the diffusion coefficients in confined and unconfined states for various molecular constructs.

1. Switching dynamics analysis

The motion of a particle can be described as switching between different states characterized by different diffusion coefficients[60,61]. We first used this description to develop a method to identify the best number of switching states, based on the distribution of instantaneous displacements $\Delta X = X(t+\Delta t) - X(t)$. First, we describe the switching dynamics and fitting procedure, and then how we used the Bayesian Information Criterion (BIC) to select the optimal model.

1.1 Diffusion model

The dynamics of a chromatin-binding protein complex such as NuRD, or some component of it, can be described by Langevin's equation[62,63] in the classical overdamped limit, where the position $X(t)$ satisfies:

$$\dot{\boldsymbol{X}} = \frac{\boldsymbol{f}(\boldsymbol{X})}{\gamma} + \sqrt{2D}\dot{\boldsymbol{\eta}}, \qquad (1)$$

The force f depends on the position X and the second term corresponds to steady-state diffusion in a crowded medium, characterized by a diffusion coefficient D. Here η is a Gaussian variable with mean 0 and variance 1. When the protein binds to its molecular partners or chromatin, the diffusion coefficient changes, and the motion can become restricted, e.g. to that along the path of the DNA.

Experimental trajectories consist of a series of points ($X(k\Delta t)$ $k = 1...$ and in the absence of any additional localization error noise, or an external force $f = 0$, the displacement dynamics at the sampling rate interval Δt, is given by:

$$\boldsymbol{X}(t+\Delta t) = \boldsymbol{X}(t) + \sqrt{2D\Delta t}\boldsymbol{\eta}, \qquad (2)$$

where η is a vector of Gaussian values of mean 0 and variance 1. In two dimensions, the distribution of displacements is:

$$Pr\left\{|\Delta \boldsymbol{X}| = u\right\} = \frac{u}{\sigma^2}\exp\left(-\frac{u^2}{2\sigma^2}\right), \qquad (3)$$

where $\sigma^2 = 2D\Delta t$ [5].

1.2 Modelling switching behaviour

A switching model for a chromatin-binding protein can be modelled by a Markov chain with several states i, characterized by a rate constant λ_{ij} between states i and j. For two states, it is characterized by the diffusion coefficients D_1 and D_2, and the associated jump process is:

$$\boldsymbol{X}(t+\Delta t) = \boldsymbol{X}(t) + \begin{cases} \sqrt{2D_1\Delta t}\boldsymbol{\eta} & \text{w. p. } 1-\lambda\Delta t \\ 0 & \text{w. p. } \lambda\Delta t \end{cases}, \tag{4}$$

$$\boldsymbol{X}(t+\Delta t) = \boldsymbol{X}(t) + \begin{cases} \sqrt{2D_1\Delta t}\boldsymbol{\eta} & \text{w. p. } 1-\mu\Delta t \\ 0 & \text{w. p. } \mu\Delta t \end{cases}, \tag{5}$$

where λ (resp. μ) are the transition rates from states 1 to 2:

$$1 \underset{\lambda}{\overset{\mu}{\rightleftharpoons}} 2. \tag{6}$$

The probability density function (pdf) p_d for the displacement of a molecule depends on the state of the process and $\kappa = \frac{\lambda}{\lambda+\mu}$ can be computed using Bayes' law:

$$Pr_d\{|\Delta X| = u\} = Pr\{|\Delta X| = u|X(t) \text{ and } X(t+\Delta t) \text{in state } 1\}q_1(t)$$
$$+Pr\{|\Delta X| = u|X(t) \text{ and } X(t+\Delta t) \text{ in state } 2\}q_2(t), \tag{7}$$

where using Bayes relation, for state $i = 1$ or 2,

$$q_i(t) = Pr\{X(t) \text{ and } X(t+\Delta t) \text{ in state i}\} = Pr\{X(t+\Delta t)\text{in state i}|X(t) \text{ in state i}\}$$
$$\times Pr\{X(t) \text{ in state i}\}.$$

Because $p_i(t) = Pr\{X(t) \text{ in state i}\}$ is the solution to the Master equation:

$$\frac{d}{dt}p_1 = -\lambda p_1 + \mu p_2 \tag{8}$$

$$\frac{d}{dt}p_2 = \lambda p_1 - \mu p_2. \tag{9}$$

And

$$p_1(t) = \kappa + Ce^{-(\lambda+\mu)t}$$
$$p_2(t) = 1 - \kappa - Ce^{-(\lambda+\mu)t},$$

where and C depend on the initial distribution and

$$Pr\{X(t+\Delta t)\text{in state } 1|X(t) \text{ in state } 1\} = 1 - \lambda\Delta t, \textit{ and}$$

$$Pr\{X(t+\Delta t)\text{in state } 2|X(t) \text{ in state } 2\} = 1 - \mu\Delta t.$$

We conclude that in two-dimensions, for a long-time t and a short-time Δt, using Eqs. 7 and 3, the pdf can be written as:

$$Pr_d\left\{|\Delta \boldsymbol{X}| = u\right\} = \kappa\frac{u}{\sigma_1^2}\exp\left(-\frac{u^2}{2\sigma_1^2}\right) + (1-\kappa)\frac{u}{\sigma_2^2}\exp\left(-\frac{u^2}{2\sigma_2^2}\right) \quad (10)$$

where $\sigma_k = 2D_k\Delta t$ and $k = 1,2$. In the case of a three-state model, the pdf for the displacement of a molecule is:

$$Pr_{\text{switch}}\left\{\Delta \boldsymbol{X} = u\right\} = \kappa_1\frac{u}{\sigma_1^2}\exp\left(-\frac{u^2}{2\sigma_1^2}\right) + \kappa_2\frac{u}{\sigma_2^2}\exp\left(-\frac{u^2}{2\sigma_2^2}\right)$$

$$+(1-\kappa_1-\kappa_2)\frac{u}{\sigma_3^2}\exp\left(-\frac{u^2}{2\sigma_3^2}\right), \quad (11)$$

where $\sigma_k = 2D_k\Delta t$ with $k = 1,\ 2,$ and 3 and $D_1,\ D_2,$ and D_3 are the three diffusion coefficients. We applied this method to analyse two sets of data acquired at either 500 ms or 20 ms, and came to the conclusion that a two-state diffusion model is optimal for both cases.

A steady-state displacement analysis of the 20 ms trajectories revealed that the data are best characterized by a two or three state model, consisting of a confined and either one or two unconfined states. To test the appropriateness of a one-, two- or three-state model, we extracted the diffusion coefficients from Eqs. 10-11 and the associated rate constants (**Extended Data Figure 6a**). We later also carried out a principle component analysis (PCA) after classifying the sub-trajectories using the four biophysical parameters as described in detail in Section 2, below. Although this showed that a three-state model could best describe the data, that analysis also revealed that the second unconfined state represented less than 20% of the data and thus we decided to develop a classification algorithm to segment trajectories into just two classes: confined and unconfined. Similarly, we evaluated 1, 2 and 3 state models for analysis of the 500 ms data (**Extended Data Figure 6b**).

In summary, the displacement analysis showed that a two-state diffusion model was appropriate to analyse NuRD complex dynamics, suggesting that the 20 ms trajectories can be characterised by confined (low diffusion

coefficient D) and unconfined (high diffusion coefficient D) states, whilst the 500 ms data can be characterised by slow and fast sub-diffusive chromatin bound states.

2. Classification of trajectories into two diffusion states

We describe here a novel method to classify a trajectory into two states – when analysing 20 ms data these sub-trajectories are labeled as confined (C) and unconfined (U). The method is based on a generalized Gaussian mixture model[64], where the input data is an ensemble of trajectories and the output is two ensembles of confined and unconfined sub-trajectories.

2.1 Input data

We used an ensemble of N trajectories $X_i(k\Delta t)$, $i = 1,...N$, $k = 0,1,...,n_i$, obtained at an acquisition time step Δt, such that each trajectory consists of n_i discrete points in three dimensions.

2.2 Statistical features are extracted in a sliding window along the trajectories

To classify sub-trajectories as confined C or unconfined U, we used four physical parameters, computed along single trajectories[41,65]. For a trajectory given by the successive points $X_i(k\Delta t)$, we used a sliding window W_k containing $2l + 1$ points, centered at $X(k\Delta t)$, and defined as:

$$W_k^{(d)}(m\Delta t) = \{X_i^{(d)}((k-l+m)\Delta t), m = 0, \ldots, 2l, \quad d = 1, 2 \text{ and } 3\}, \tag{12}$$

where d is the dimension. The sliding window $W_k(m\Delta t)$ is applied for $k = 1,...n_i$ along a trajectory containing n_i points. We now briefly discuss the four parameters we used for the classification:

Anomalous exponent α: The anomalous exponent characterizes the motion of a stochastic particle on a particular time scale[41]. It is computed from the Mean-Square-Displacement (MSD) $\langle|X(t + \Delta t) - X(t)|^2\rangle$ that behaves like At^α when Δt is small compared to the time of the process. A value of $\alpha = 1$ reflects Brownian motion, and an $\alpha > 1$ is called super-diffusion, which may represent dynamics containing an element of

deterministic (ballistic) directed motion. Finally, an $\alpha < 1$ is sub-diffusive motion[41,66,67]. To estimate α for each point X, we first compute the MSD S_k over the sliding window $W_k(m\Delta t)$ defined by:

$$S_k(m) = \langle \| W_k(m\Delta t) - W_k(0)) \|^2 \rangle, m = 0, \ldots, 2l, \quad (13)$$

Where $\langle . \rangle$ denotes the average where we use the intermediate point from 1 to m. To estimate the exponent $\alpha_i(k)$ for point $X(k\Delta t)$, we fitted the function $S_k(m)$ defined in Eq. (13) by

$$f_k(t) = \beta_i(k) t^{\alpha_i(k)} \quad (14)$$

with an additional parameter $\beta_i(k) > 0$. For the fit, we constrain the variable t in the ensemble $[0, \Delta t, \ldots, (2l+1)\Delta t]$.

Effective Diffusion coefficient: We estimate the effective diffusion coefficient[65,68] by computing the second statistical moment along the trajectories. We use the empirical estimator to estimate for each sub-trajectory in $W_k(m\Delta t)$[41]

$$D_i(k) = \frac{1}{2d\Delta t} \sum_{m=0}^{2l} (W_k((m+1)\Delta t)) - W_k(m\Delta t))^2 . \quad (15)$$

Length of confinement Lc. The length of confinement estimates the size of a domain where a trajectory is confined. We use the sub-trajectory $W_k^{(d)}(m\Delta t),\ m = 0, \ldots, 2l, \quad d = 1, 2, 3,$, located inside the window W_k [41]. The length of confinement Lc is then computed empirically in the window W_k by

$$Lc_i(k) = \frac{1}{3} \sum_{d=1}^{3} \sqrt{\frac{1}{2l} \sum_{m=0}^{2l} (W_k^{(d)}(m\Delta t) - \langle W_k^{(d)} \rangle)^2} = STD(W_k), \quad (16)$$

where

$$\langle W_k^{(d)} \rangle = \frac{1}{2l} \sum_{m=0}^{2l} W_k^{(d)}(m\Delta t), d = 1, 2 \text{ and } 3. \quad (17)$$

It is the standard deviation of the sub-trajectory W_k, where the average position is $(\langle W_k^{(1)} \rangle, \langle W_k^{(2)} \rangle, \langle W_k^{(3)} \rangle)$.

Magnitude of the drift vector kV_ik**.** To characterize the displacement of a trajectory between the beginning and the end of the sliding window W_k, we compute the magnitude kV_ik of the drift vector V_i for each dimension $d = 1,2,3,$ [68] using the formula:

$$V_i^{(d)}(k) = \frac{1}{2l\Delta t}\sum_{m=0}^{2l-1} W_k^{(d)}((m+1)\Delta t) - W_k^{(d)}(m\Delta t) \quad . \qquad (18)$$

The norm of the drift kV_ik is:

$$\|V_i\| = \sqrt{\sum_{d=1}^{3}(V_i^{(d)}(k))^2}. \qquad (19)$$

In summary, we compute four parameters: the anomalous exponent α, the diffusion coefficient D, the length of confinement Lc, and the magnitude of the drift vector kV_ik. These are all computed for each point along a trajectory using a sliding window. The sliding window of 11 was chosen by trial and error. Below this value, the anomalous exponent distributions for confined and unconfined molecules tended to merge suggesting that long trajectories are essential for reliable estimation of this parameter. Above this value, the diffusion coefficient histograms for confined and unconfined molecules tended to merge suggesting that transitions between these populations occur leading to averaging of the diffusion coefficients. In the next section, we use these four parameters for classification.

2.3 Classification of a point into either a confined (C) or an unconfined (U) state

To classify each time point $X((k-1)\Delta t)$ for $k = 1,...,n_i$ of all trajectories indexed by $i = 1,...,N$ into C and U classes, we first collected all the values from Eqs. 14-19, computed from each sliding window W_k (Eq. 12). This led to a total of $n_i \times 4$ parameters, that we organized in a matrix R_i associated with trajectory i:

$$R_i = [\alpha_i^T, D_i^T, Lc_i^T, \|V\|_i^T] = \begin{bmatrix} r_i(\Delta t) \\ r_i(2\Delta t) \\ . \\ . \\ r_i(n_i\Delta t) \end{bmatrix}, \qquad (20)$$

where T is the transpose operator, and $r_i(k\Delta t)$ are four dimensional vectors of the parameters (Eqs. 14, 15, 16, 19).

We then concatenate all feature matrices R_i (Eq. 20), $i = 1,...N$ into the $\left(\sum_i^N n_i\right) \times 4$ general matrices $\boldsymbol{R}$, defined by:

$$\boldsymbol{R} = \begin{bmatrix} R_1 \\ R_2 \\ \cdot \\ \cdot \\ R_N \end{bmatrix}. \qquad (21)$$

We normalize each column (feature) in $R^{(j)}, j = 1,...,4$ by subtracting its mean and dividing by its standard-deviation.

To separate the histograms of the four parameters into two independent classes, we constructed an unsupervised binary classifier using a two-component Gaussian mixture model in a four-dimensional space, corresponding to the four parameters Eqs. 14-19. The Gaussian mixture distribution p is a weighted sum of two multivariate Gaussian densities, defined as:

$$p(r|w_{C,U},\mu_{C,U},\Sigma_{C,U}) = w_C g(r|\mu_C,\Sigma_C) + w_U g(r|\mu_U,\Sigma_U), \qquad (22)$$

where w_C, w_U are the mixing weights, such that $w_C + w_U = 1$, $g(r|\mu,\Sigma)$ are the Gaussian densities

$$g(r|\mu,\Sigma) = \frac{1}{(2\pi)^2|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(r-\mu)^T\Sigma^{-1}(r-\mu)\right), \qquad (23)$$

and Σ_C, Σ_U and μ_C, μ_U are the four-dimensional covariance matrices and mean vectors for components C and U, respectively. We then find the values of the parameters $\mu_C, \mu_U, \Sigma_C, \Sigma_U$ and w_C, w_U of Eq. 22 which best separate the data as the maximal likelihood estimators of the density Eq. 23, given the observed statistics in $\boldsymbol{R}$ (Eq. 21), by using the Expectation-Maximization algorithm[69].

For each point $X_i(k\Delta t)$, with its associated feature vector $r_i(k\Delta t)$ (Eq. 21), we assign a label $n \in \{C,U\}$, based on the posterior probability P of the density Eq. 23, given by:

$$P(n|r_i(k\Delta t), w_n, \mu_n, \Sigma_n) = \frac{w_n g(r_i(k\Delta t)|\mu_n, \Sigma_n)}{w_{\mathcal{C}} g(r_i(k\Delta t)|\mu_{\mathcal{C}}, \Sigma_{\mathcal{C}}) + w_{\mathcal{U}} g(r_i(k\Delta t)|\mu_{\mathcal{U}}, \Sigma_{\mathcal{U}})} \qquad (24)$$

such that for each point:

$$X_i(k\Delta t) \in \begin{cases} \mathcal{C}, & Pr(\mathcal{C}|r_i(k\Delta t), w_{\mathcal{C}}, \mu_{\mathcal{C}}, \Sigma_{\mathcal{C}}) > Pr(\mathcal{U}|r_i(k\Delta t), w_{\mathcal{U}}, \mu_{\mathcal{U}}, \Sigma_{\mathcal{U}}) \\ \mathcal{U}, & \text{else.} \end{cases} \qquad (25)$$

In practice, if a point $X_i(k\Delta t) \in$ C or U and consists in an isolated class, e.g. C between two neighbouring U classes, then we relabel it as the class of its two immediate neighbours $(k-1)\Delta t, (k+1)\Delta t$.

2.4 Accuracy of the classification algorithm

We estimated the accuracy of the classification algorithm (subsections 2.2-2.3) using simulated (synthetic) trajectories, switching between two states C and U. The dynamics was described by an Ornstein-Uhlenbeck equation[65], and the generated trajectories were considered a ground-truth ensemble, which could be used to estimate the accuracy of the classification procedure developed in the previous section.

3. Simulations to validate the classification algorithm

3.1 Simulation of a ground-truth ensemble

We generated N trajectories $X_i(t)$, $i = 1,...,n_i$, which can switch between the confined C and unconfined U state at any time t. The stochastic switching process is defined by:

$$\frac{dX(t)}{dt} = \begin{cases} -\kappa(X(t) - X(\tau)) + \sqrt{2D_{\mathcal{C}}}\frac{d\boldsymbol{\omega}}{dt}, & X(t) \in \mathcal{C}; \\ \sqrt{2D_{\mathcal{U}}}\frac{d\boldsymbol{\omega}}{dt}, & X(t) \in \mathcal{U}, \end{cases} \quad (26)$$

where D_C, and D_U are the diffusion coefficients for the confined C and unconfined U states respectively, $d\omega/dt$ are standard three-dimensional Brownian motions, with a mean of zero and a standard-deviation of one, while κ is the strength of a potential well that attracts the trajectory $X(t) \in$ C to a fixed point $X(\tau)$, which is the last position, before time t where the trajectory was in state U prior to switching to C:

$$\tau = \sup_{s<t}\{X(s) \in \mathcal{U} | X(t) \in \mathcal{C}\}. \quad (27)$$

The transition probability between states (C, U) is defined by a Markov chain: The transient probability matrix P_{ij} between state j at time $s + \Delta s$ and state i at time s, is given by:

$$P_{ij} = \begin{cases} \lambda \Delta s, & i = \mathcal{C}, j = \mathcal{U}; \\ 1 - \lambda \Delta s, & i = \mathcal{C}, j = \mathcal{C}; \\ \mu \Delta s, & i = \mathcal{U}, j = \mathcal{C}; \\ 1 - \mu \Delta s, & i = \mathcal{U}, j = \mathcal{U}. \end{cases} \quad (28)$$

The probabilities $P_C(s)$ and ($P_U(s)$) satisfy[70]:

$$\begin{bmatrix} \dot{P}_{\mathcal{C}}(s) \\ \dot{P}_{\mathcal{U}}(s) \end{bmatrix} = \begin{bmatrix} -\lambda & \mu \\ \lambda & -\mu \end{bmatrix} \begin{bmatrix} P_{\mathcal{C}}(s) \\ P_{\mathcal{U}}(s) \end{bmatrix} \quad (29)$$

and the solution of Eq. 29 is:

$$\begin{bmatrix} P_{\mathcal{C}}(s) \\ P_{\mathcal{U}}(s) \end{bmatrix} = \frac{1}{\lambda + \mu} \begin{bmatrix} \lambda e^{-(\lambda+\mu)s} + \mu & \mu(1 - e^{-(\lambda+\mu)s}) \\ \lambda(1 - e^{-(\lambda+\mu)s}) & \lambda + \mu e^{-(\lambda+\mu)s} \end{bmatrix} \begin{bmatrix} P_{\mathcal{C}}(0) \\ P_{\mathcal{U}}(0) \end{bmatrix} \quad (30)$$

The simulation procedure was as follows: we initialized $X_i(0)$ in one of the states C or U. We used Euler's scheme to discretize Eq. 26 at a time step Δt. For each consecutive step $k\Delta t$, $k = 1,...n_i$, we determined the state of the trajectory from the probabilities $P_C(\Delta t)$,$P_U(\Delta t)$ (Eq. 30), using the previous state as the initial condition:

$$\begin{bmatrix} P_{\mathcal{C}}(0) \\ P_{\mathcal{U}}(0) \end{bmatrix} = \begin{cases} [1, 0]^T, & X_i((k-1)\Delta t) \in \mathcal{C}; \\ [0, 1]^T, & X_i((k-1)\Delta t) \in \mathcal{U}. \end{cases} \quad (31)$$

3.2 Comparing the classifier and ground-truth classes

To measure the accuracy of our classification (Sections 2.2-2.3), we used synthetic trajectories generated by the stochastic process described above. We generated $N = 100$ trajectories $Y_i(k\Delta t)$, $i = 1,...,N$, $k = 1,...,n_i$ from Eq. 27, where the length of a trajectory n_i was chosen randomly from a Poisson distribution with average length $n = 100$ time points, $\Delta t = 0.02s$, $\kappa = 2$, $D_C = 0.008\mu m^2/s$, $D_U = 0.1\mu m^2/s$, and for twenty equally spaced values of λ, $\mu \in [0,1]$ to obtain the ground-truth states $S_i(k\Delta t) \in$ [C,U] associated with each trajectory $Y_i(k\Delta t)$.

We used the four-parameter classifier (section 2.3) to classify the trajectories $Y_i(k\Delta t)$ and add a label $C_i(k\Delta t) \in$ [C, U]. To evaluate the similarity between the true class set S_i and the output of the classification classes C_i, we used an indicator function $\delta_i(k\Delta t)$, defined by:

$$\delta_i(k) = \begin{cases} 1, & S_i(k\Delta t) = C_i(k\Delta t); \\ 0, & \text{else}, \end{cases} \quad (32)$$

and then used Eq.32 to define a parameter M that measures the accuracy of the classification along all trajectories. It was defined by:

$$M = \frac{\sum_{i=1}^{N} \sum_{k=1}^{n_i} \delta_i(k)}{\sum_{i=1}^{N} n_i} \quad . \qquad (33)$$

We evaluated the accuracy matrix M (Eq. 32, computed for 20 values of the diffusion coefficient $D_C \in [0.002, 0.1]$ and $D_U \in [0.2, 1]$ $\mu m^2/s$, where we set the switching rates to be $\lambda = 0.8$, $\mu = 0.9$). We find that the classifier achieved above 90% accuracy in the majority of the parameter space, dropping to 76% for confined and unconfined diffusion constant where $D_C \in [0.02, 0.1]$, $D_U \in [0.1, 0.2]$. In addition, to test the robustness of the classification algorithm on the switching rates, we evaluated M (Eq. 33 for 20 values of the switching rates λ, μ ∈ [0, 1] for $D_C = 8 \times 10^{-3}$, $D_U = 0.01$ μm2/s and obtained a very good accuracy – above 90% for all values of λ, μ in that range (see **Extended Data Figure 6**).

In summary, when tested on simulated data, the classification algorithm was able to robustly classify trajectories into confined and unconfined states.

4. Estimation of association and dissociation times from trajectories

After classifying the trajectories into two classes C and U (see section 2.3), we estimated the association and dissociation time constants as follows. The association time τ_A for trajectory X_i is defined as the first time the trajectory reaches the state C given that it was previously in U

$$t_A = \inf_t (t > 0; X_i(t) \in \mathcal{C} | X_i(0) \in \mathcal{U}) \qquad (34)$$

The dissociation time τ_D is the first time the trajectory reaches a state U given it was initially in state C

$$t_D = \inf_t (t > 0; X_i(t) \in \mathcal{U} | X_i(0) \in \mathcal{C}) \qquad (35)$$

To construct the histogram of these times, we reset the origin of time to zero after each association or dissociation event. In practice, for each partition of a trajectory X_i in a given ensemble, we collected the sequence τ_A^i of association times as the consecutive time points for which the trajectory spends in class U between being $\langle \tau_A \rangle = \frac{1}{\mu}$ confined, as determined by the $\langle \tau_D \rangle = \frac{1}{\lambda}$ posterior probability P (Eq. 24). Equivalently,

dissociation times τ_D are those times in which the trajectory spends in confined state C between being unconfined. We fitted the histogram of association times and dissociation times with a single exponential[71]:

$$f_A(t) = \beta_D \exp(-\mu t), \text{ and} \quad (36)$$

$$f_D(t) = \beta_A \exp(-\lambda t). \quad (37)$$

The mean times are then obtained by $\langle \tau_A \rangle = \frac{1}{\mu}$ and $\langle \tau_D \rangle = \frac{1}{\lambda}$.

5. Estimation of other parameters

5.1 Estimation of rate and diffusion parameters from the displacements

The classification algorithm decomposes the ensemble of trajectories X into two sub-ensembles: C and U, of confined and unconfined trajectories, respectively. In both the two- and three-state diffusion model, when using Eq. 10, state 1 always corresponds to the distribution of confined displacements for which:

$$\kappa_1 = \frac{|X_C|}{|X_C| + |X_U|} \quad (38)$$

where |.| is the cardinal of the ensemble. The diffusion coefficient D_1 is obtained by fitting it to the data X_C. For the two-state diffusion model, we fitted the coefficient D_2 to X_U. For the three-state diffusion model, we considered only the distribution of displacements from X_U and fitted the two-state diffusion model to estimate the parameters κ_2, D_2 and D_3. The fitting procedure uses the fit function from MATLAB 9.2 (MathWorks), and the fits were repeated 50 times with randomly selected starting parameter values. The fit with the highest adjusted R^2 coefficient was kept.

5.2 Quantification of the biophysical model using the BIC index

We estimate the quality of fit of the different models by using the Bayesian Information Criterion (BIC) defined as:

$$BIC = (p+1)n(\ln n)\left[\ln\left(\frac{2\pi \mathrm{RSS}}{n}\right) + 1\right] \quad (39)$$

where p is the number of parameters of the model from Eq. 10, n is the number of data points, and RSS is the residual sum of squares between the model and the data.

6. Application of the classification algorithm to NuRD single particle trajectories

To understand how well the four-parameter algorithm performs, we used the displacement histograms and determined diffusion coefficients either before or after (**Extended Data Figure 6b, c**) applying the classification algorithm. The algorithm provides a precise segmentation along trajectories, and necessarily has to discard short segments, but interestingly we found that the diffusion coefficients extracted either before or after using the algorithm are not that different. When we assume a single diffusion coefficient (left column in **Extended Data Figure 7**), we can see that the four-parameter algorithm discards some data. However, in both cases, the estimated diffusion coefficients are quite similar. More interestingly, when using the two-state diffusion model (Confined *vs.* Unconfined), we found that the slow diffusion coefficients are higher after the use of the algorithm – probably because short segments were discarded and the fraction confined was higher, and thus more of the trajectories were considered to be confined. As a consequence, the diffusion coefficient in the unconfined state was lower, as more displacements were probably used in the computation of the confined state.

7. Classification performance tested on ground truth simulated data

To evaluate the performance of the parameter estimation method, described in Section 2, we use the displacement histogram, as explained in Section 1. We find that the displacement histogram can indeed be used to classify the confined and unconfined trajectories for various ranges of parameters. Diffusion coefficients and the confined and unconfined rate constants after analysis of the 20 ms and 500 ms data are shown in **Extended Data Figure 7**. Finally, we summarize the parameters estimated from Δt = 500 ms trajectories after applying the classification algorithm (**Extended Data Figure 6c**), confirming that it can be applied robustly for the classification of trajectories.

8. Gaussian fitting of 500 ms exposure anomalous exponents

Although the algorithm we designed segments 500 ms exposure trajectories into slow-diffusing and fast-diffusing components, the distribution of the anomalous exponents in the fast-diffusing component showed that two peaks can be identified by Gaussian fitting. Indeed, when we fitted 1, 2 or 3 Gaussians to the anomalous exponent distributions for chromatin bound NuRD complex subunits such as CHD4, MBD3 and MTA2 in wild-type ES cells, we found that 2 Gaussians were the best minimal model to account for the data – having the lowest BIC value (**Extended Data Figure 9c**).

Next, we assessed the reproducibility of analysing the anomalous exponent distributions of the fast-diffusing chromatin bound CHD4. Data was collated from 3 fields of view with around 6 cells in each field of view imaged, leading to a total of around 18 cells per condition. To ensure enough data was collected to account for cell-to-cell heterogeneity, an additional 3 fields of view containing around 18 cells were collected for chromatin bound CHD4 and shown to have a similar anomalous exponent distribution (data not shown). In contrast, CHD4 in the absence of MBD3 showed a different distribution from CHD4 in wild-type ES cells when compared to either the first or second day CHD4 was imaged ($p = 0.005$ and $p = 0.02$ respectively), thus showing the reproducibility of the changes observed.

H) *In vitro* biochemical assays of NuRD complex with and without nucleosomes

Drosophila PMMR, Human CHD4 and Human GATAD2A-MBP were expressed in Sf21 cells and purified as described[23] with the following modifications: Cells were resuspended in 50 mM Tris-HCl pH 7.5, 1 M NaCl, 5 mM DTT, 1× complete EDTA-free protease inhibitor cocktail (Roche), lysed by sonication and cleared by centrifugation at 50 000 *g* for 1 h. The supernatant was applied to amylose resin pre-equilibrated with lysis buffer and incubated for 2 h with rotation at 4 °C. The resin was washed with 20 X column volumes (CV) of lysis buffer and then eluted with 10mM maltose in lysis buffer. Fractions containing hGATAD2A-MBP protein were concentrated and further purified by size exclusion chromatography using a Superose 6 Increase 3.2/300 column (GE Healthcare) equilibrated with 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 % glycerol, 1 mM DTT.

For pull-down experiments, purified protein was immobilised on MBP-Trap resin (ChromoTek) which was pre-equilibrated in pull-down buffer (50 mM HEPES, pH7.5 300 mM NaCl, 1 mM DTT, 5 % glycerol) followed by incubation for 1 h with rotation at 4 °C. A sample of the 6% protein:bead mixture was retained as

'Input'. The resin was washed 3 times with pull-down buffer, then a washed 'beads' sample was retained for analysis on a 4-12% NuPAGE gel (-Invitrogen).

Electrophoretic mobility shift assays (EMSAs) were performed on n3-Widom-78bp DNA or recombinant nucleosomes with this template[72,73] in 10 µL of binding buffer (20 mM Hepes pH7.5, 2mM $MgCl_2$, 5% Glycerol, 1mM TCEP) with varying concentrations of the indicated proteins. The reaction mixtures were incubated at 30^0C for 30 min followed by centrifugation at 1000 x g. The resulting reaction mixtures were loaded onto 5% native polyacrylamide gels and run in 0.2 × TBE. Gels were stained with SYBR Gold (Invitrogen) and imaged using a Typhoon FLA 9000 (GE healthcare).

I) CARGO-dCas9-GFP imaging of *Tbx3* enhancer loci

ES cells expressing dCas9 tagged with GFP were generated as previously described[4]. Briefly, MBD3-inducible ES cells were transfected by the PB-TRE3G-dCas9-eGFP-WPRE-ubcp-rtTA-IRES-puroR vector containing a dual promoter backbone, with a TRE3G (Tet-on) promoter expressing GFP-tagged inactive dCas9 and the ubiquitin C promoter expressing the reverse tetracycline-controlled transactivator, rtTA and a puromycin cassette via an IRES sequence. Puromycin-resistant ES cells were then selected for 7 days and doxycycline added for 24 hours to induce expression of dCas9-GFP (through activation of the rtTA). Stable transfectants were then sorted for low levels of GFP expression (to ensure that only a few copies of the plasmid were integrated stably into the genome).

Prior to imaging, Doxycycline was added to ES cells for 24 hours to induce expression of low levels of dCas9-GFP. CARGO vectors expressing 36 gRNAs targeting the *Tbx3* enhancer were then transfected using lipofectamine 2000 (Invitrogen) into these ES cells. Cells were transfected during passaging straight onto imaging dishes in Fluorobrite imaging medium as described above. After 24 hours, fresh media was added and after 48 hours, imaging was carried out. The CARGO and dCas9-GFP expressing plasmids were kind gifts from the J. Wysocka lab.

2D tracking of genomic loci was carried out using oblique illumination on a custom built 2D single-molecule tracking microscope as previously described[74]. Briefly, an IX73 Olympus inverted microscope was used with circularly polarized laser beams aligned and focused at the back aperture of an Olympus 1.40 NA 100× oil objective (Universal Plan Super Apochromat, 100×, NA 1.40, UPLSAPO100XO/1.4). A 561 nm laser

was used as a continuous wavelength diode laser light source. Oblique-angle illumination imaging was achieved by aligning the laser off axis such that the emergent beam at the sample interface was near-collimated and incident at an angle less than the critical angle $\theta_c \sim 67°$ for a glass/water interface. This generated a ~50 µm diameter excitation footprint. The power of the collimated 488 nm beam at the back aperture of the microscope was 100 W/cm^2. The lasers were reflected by dichroic mirrors which also separated the collected fluorescence emission from the TIR beam (Semrock, Di01- R405/488/561/635). The fluorescence emission was collected through the same objective and then further filtered using a combination of long-pass and band-pass filters (BLP01-561R and FF01-587/35). The emission signal was projected onto an EMCCD (Photometrics, Evolve 512 Delta) with an electron multiplication gain of 250 ADU/photon operating in a frame transfer mode. The instrument was automated using the open-source software micro-manager (https://www.micro-manager.org) and the data displayed using the ImageJ software[75,76].

For image processing, PeakFit[77] was used to localise genomic loci from these images using the filter settings: "shiftFactor":1.0, "signalStrength":5.0, "minPhotons":30.0, "precisionThreshold":40.0, "minWidthFactor":0.5, "maxWidthFactor":0.5, "precisionMethod":"MORTENSEN". Trajectories were then tracked in 2D using custom python code for connecting molecules in subsequent frames if they were within 500 nm (https://github.com/wb104/trajectory-analysis). Molecules were again classified using a Gaussian mixture model (see **Single molecule trajectory analysis**, above).

J) Software and code

1. Data collection

Microscope image acquisition	Micro-manager (https://www.micro-manager.org)
	ImageJ software[75,76]

2. Data analysis

Hi-C analysis	Python
	NucProcess (https://github.com/tjs23/nuc_processing)
	NucTools (https://github.com/tjs23/nuc_tools)
ChIP-seq analysis	R version 3.6.0

	(https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)
	Read alignment[56].
	Deeptools v2.5.0[57].
2D single-molecule peak fitting	PeakFit[77]
3D single-molecule peak fitting	easy-DHPSF software[59]
Trajectory analysis	(https://github.com/wb104/trajectory-analysis)
Gaussian mixture classification/fitting	MATLAB v2016

References (methods)

1 Gebhardt, J. C. *et al.* Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. *Nature methods* **10**, 421-426, doi:10.1038/nmeth.2411 (2013).

4 Gu, B. *et al.* Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements. *Science* **359**, 1050-1055, doi:10.1126/science.aao3136 (2018).

16 Stevens, T. J. *et al.* 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* **544**, 59-64, doi:10.1038/nature21429 (2017).

22 Bornelov, S. *et al.* The Nucleosome Remodeling and Deacetylation Complex Modulates Chromatin Structure at Sites of Active Transcription to Fine-Tune Gene Expression. *Mol Cell* **71**, 56-72 e54, doi:10.1016/j.molcel.2018.06.003 (2018).

23 Zhang, W. *et al.* The Nucleosome Remodeling and Deacetylase Complex NuRD Is Built from Preformed Catalytically Active Sub-modules. *J Mol Biol* **428**, 2931-2942, doi:10.1016/j.jmb.2016.04.025 (2016).

24 Carr, A. R. *et al.* Three-Dimensional Super-Resolution in Eukaryotic Cells Using the Double-Helix Point Spread Function. *Biophysical journal* **112**, 1444-1454, doi:10.1016/j.bpj.2017.02.023 (2017).

29 Kaji, K. *et al.* The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. *Nat Cell Biol* **8**, 285-292, doi:10.1038/ncb1372 (2006).

41 Amitai, A., Seeber, A., Gasser, S. M. & Holcman, D. Visualization of Chromatin Decompaction and Break Site Extrusion as Predicted by Statistical Polymer Modeling of Single-Locus Trajectories. *Cell reports* **18**, 1200-1214, doi:10.1016/j.celrep.2017.01.018 (2017).

54 Furumai, R. *et al.* FK228 (depsipeptide) as a natural prodrug that inhibits class I histone deacetylases. *Cancer Res* **62**, 4916-4921 (2002).

55 Pekowska, A. *et al.* Gain of CTCF-Anchored Chromatin Loops Marks the Exit from Naive Pluripotency. *Cell Syst* **7**, 482-495 e410, doi:10.1016/j.cels.2018.09.003 (2018).

56 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).

57 Ramirez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165, doi:10.1093/nar/gkw257 (2016).

58 Matsubara, H. *et al.* Involvement of extracellular signal-regulated kinase activation in human osteosarcoma cell resistance to the histone deacetylase inhibitor FK228 [(1S,4S,7Z,10S,16E,21R)-7-ethylidene-4,21-bis(propan-2-yl)-2-oxa-12,13-dithia-5,8 ,20,23-tetraazabicyclo[8.7.6]tricos-16-ene-3,6,9,19,22-pentone]. *The Journal of pharmacology and experimental therapeutics* **328**, 839-848, doi:10.1124/jpet.108.147462 (2009).

59 Lew, M., et al. Easy-DHPSF open-source software for three-dimensional localization of single molecules with precision beyond the optical diffraction limit. *Protocol Exchange*, doi:10.1038/protex.2013.026 (2013).

60 Reingruber, J. & Holcman, D. Gated narrow escape time for molecular signaling. *Phys Rev Lett* **103**, 148102, doi:10.1103/PhysRevLett.103.148102 (2009).

61 Holcman, D. S., Z. Stochastic narrow escape in molecular and cellular biology. *Springer-Verlag New York*, doi:10.1007/978-1-4939-3103-3 (2015).

62 Schuss, Z. Diffusion and Stochastic Processes. An Analytical Approach. *Springer-Verlag New York* **170**, doi:10.1007/978-1-4419-1605-1 (2009).

63 Holcman, D. & Schuss, Z. 100 years after Smoluchowski: stochastic processes in cell biology. *Journal of Physics A: Mathematical and Theoretical* **50**, 093002, doi:10.1088/1751-8121/50/9/093002 (2017).

64 Hastie, H. T., R.; Friedman, J. The elements of statistical learning. *Springer-Verlag New York* **1**, doi:10.1007/978-0-387-84858-7 (2001).

65 Schuss, Z. Theory and applications of stochastic processes: an analytical approach. *Springer-Verlag New York* **170**, doi:10.1007/978-1-4419-1605-1 (2010).

66 Kepten, E., Bronshtein, I. & Garini, Y. Improved estimation of anomalous diffusion exponents in single-particle tracking experiments. *Physical review. E, Statistical, nonlinear, and soft matter physics* **87**, 052713, doi:10.1103/PhysRevE.87.052713 (2013).

67 Dion, V. & Gasser, S. M. Chromatin movement in the maintenance of genome stability. *Cell* **152**, 1355-1364, doi:10.1016/j.cell.2013.02.010 (2013).

68 Hozé, N. H., D. Statistical methods for large ensembles of super-resolution stochastic single particle trajectories in cell biology. *Annual Review of Statistics and Its Application* **4**, 189-223, doi:10.1146/annurev-statistics-060116-054204 (2017).

69 Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* **39**, 1-38 (1977).

70 Karlin, S. T., H. E. A second course in stochastic processes. *Elsevier* (1981).

71 Amitai, A., Kupka, I. & Holcman, D. Computation of the mean first-encounter time between the ends of a polymer chain. *Phys Rev Lett* **109**, 108302, doi:10.1103/PhysRevLett.109.108302 (2012).

72 Deindl, S. *et al.* ISWI remodelers slide nucleosomes with coordinated multi-base-pair entry steps and single-base-pair exit steps. *Cell* **152**, 442-452, doi:10.1016/j.cell.2012.12.040 (2013).

73 Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-260, doi:10.1038/38444 (1997).

74 Basu, S. *et al.* FRET-enhanced photostability allows improved single-molecule tracking of proteins and protein complexes in live mammalian cells. *Nature communications* **9**, 2520, doi:10.1038/s41467-018-04486-0 (2018).

75 Edelstein, A., Amodaj, N., Hoover, K., Vale, R. & Stuurman, N. Computer control of microscopes using microManager. *Curr Protoc Mol Biol* **Chapter 14**, Unit14 20, doi:10.1002/0471142727.mb1420s92 (2010).

76 Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature methods* **9**, 671-675 (2012).

77 Herbert, A. Peak Fit, a Single-Molecule Plugins. *http://www.sussex.ac.uk/gdsc/intranet/microscopy/imagej/smlm_plugins* (2014).

Acknowledgements

We thank the EU FP7 Integrated Project "4DCellFate" (277899), the Medical Research Council (MR/P019471/1) and the Wellcome Trust (206291/Z/17/Z) for programme funding. We also thank the MRC (MR/R009759/1 to BDH, and MR/M010082/1 to EDL), the Wellcome Trust (106115/Z/14/Z to IB, 210701/Z/18/Z to CS), and the Isaac Newton Trust (17.24(aa) to BDH) for project grant funding, and we thank the Wellcome Trust/MRC for core funding to the Cambridge Stem Cell Institute (including a starter grant 203151/Z/16/Z to SB). We thank Tessa Kretschmann for help in preparing the figures for publication, Luke Lavis for providing the JF549 dye, Joanna Wysocka for sending published constructs for 2D enhancer tracking, Andy Riddell for flow cytometry and Aelun Crombie for preliminary analysis of 2D enhancer tracking experiments.

Author contributions

SB, BDH, DH and EDL designed the experiments. DL, SB, WB and TJS carried out the in-nucleus Hi-C data collection and analysis. NR, WB, RR, IM and BDH carried out the ChIP-seq data collection and analysis. LM, EB, MAB and LDC carried out preliminary ChIP-seq experiments and analysis. BDH, NR, JC, RF and SB made and characterised the cell lines used for the live-cell single-molecule imaging. ARC, AP, SFL and DK designed and built the 3D DH-PSF microscopes for the live cell single molecule tracking experiments, and L-MN, SFL and DK designed and built the microscope for the 2D enhancer tracking. SB, DS and LHS carried out the live-cell 3D single-molecule imaging of NuRD complex subunits and analysed the data together with OS, PP, DK, BDH, EDL and DH. GB, KG, DS and SB carried out the live-cell 2D tracking of enhancer movement. WB developed the code to generate 3D single molecule tracks from the localisation data and dissociation time analysis, OS, PP, AJ and DH developed the code for segmentation of the 3D single-molecule trajectories. The in vitro biochemical experiments were carried out by WZ, JB, TD and TB, after purifying NuRD components and complexes expressed by AA, GC, IB and CS. SB, DH and EDL wrote the manuscript with assistance from all the other authors.

Data Availability

The Hi-C and ChIP-seq datasets reported in this study are available from the Gene Expression Omnibus (GEO) repository under accession code GSE147789. The single-molecule imaging movies are available from the authors upon request.

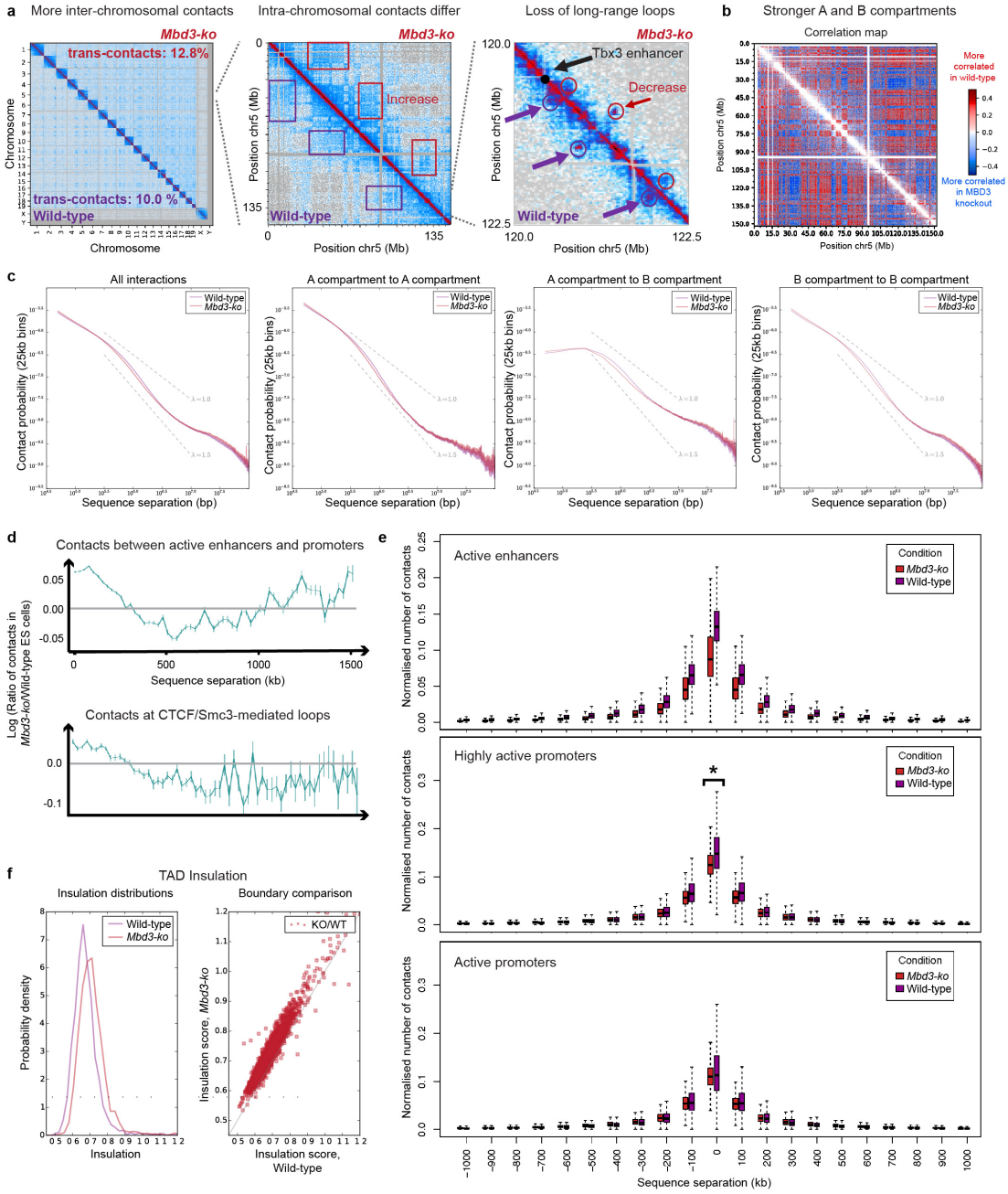
Extended Data

Extended Data Figure 1. Processing of the sequencing data for i*n-nucleus* Hi-C experiments of mouse ES cells in the presence and absence of MBD3. *In-nucleus* Hi-C was carried out and analysed as previously described[16] for wild-type E14 ES cells and *Mbd3* knockout 7g9 ES cells[29].

***In situ* Hi-C**

Clipping reads 1	**E14 wt**		**7g9 *Mbd3-ko* 1**		**7g9 *Mbd3-ko* 2**	
Input reads	178,070,576		180,696,850		373,576,259	
Junction clipped	39,957,770	22.44%	43,222,931	23.92%	102,847,306	27.53%
Quality clipped	729,378	0.41%	394,263	0.22%	12,948	0.00%
Too short	493,359	0.28%	525,137	0.29%	1,045,193	0.02%
Mean length	112.136		111.647		131.243	0.28%
Clipping reads 2	**E14 wt**		**7g9 *Mbd3-ko* 1**		**7g9 *Mbd3-ko* 2**	
Input reads	178,070,576		180,696,850		373,576,259	
Junction clipped	39,505,735	22.19%	42,122,060	23.31%	99,762,884	26.70%
Quality clipped	220,766	0.12%	687,337	0.38%	337,648	0.09%
Too short	473,068	0.27%	508,990	0.28%	1,000,237	0.02%
Mean length	112.397		111.771		131.783	0.27%
Genome alignment 1	**E14 wt**		**7g9 *Mbd3-ko* 1**		**7g9 *Mbd3-ko* 2**	
Input reads	177,577,217		180,171,713		372,531,066	
Unique	123,875,191	69.76%	126,839,098	70.40%	260,263,747	69.86%
Ambiguous	42,780,380	24.09%	45,399,724	25.20%	79,712,271	21.40%
Unmapped	10,921,646	6.15%	7,932,891	4.40%	32,555,048	8.74%
Genome alignment 2	**E14 wt**		**7g9 *Mbd3-ko* 1**		**7g9 *Mbd3-ko* 2**	
Input reads	177,597,508		180,187,860		372,576,022	
Unique	122,922,826	69.21%	124,204,619	68.93%	254,526,379	68.32%
Ambiguous	42,562,288	23.97%	44,438,144	24.66%	78,139,428	20.97%
Unmapped	12,112,394	6.82%	11,545,097	6.41%	39,910,215	10.71%
Pairing reads	**E14 wt**		**7g9 *Mbd3-ko* 1**		**7g9 *Mbd3-ko* 2**	
End 1 alignments	219,779,807		224,941,398		451,036,668	
End 2 alignments	219,555,080		223,979,780		449,458,833	
Unpaired ends	1,182,505		1,276,263		2,463,286	
Unmapped end	21,129,477	11.93%	17,552,027	9.77%	51,933,444	13.98%
Unique	95,917,216	54.16%	97,420,589	54.22%	204,160,435	54.95%
Ambiguous	60,058,576	33.91%	64,691,425	36.01%	115,438,776	31.07%
Total contacts	177,105,269		179,664,041		371,532,655	
Filtering pairs	**E14 wt**		**7g9 *Mbd3-ko* 1**		**7g9 *Mbd3-ko* 2**	
Input pairs	95,917,216	100.00%	97,420,589	100.00%	204,160,435	100.00%
Accepted	59,308,322	61.83%	64,618,110	66.33%	136,110,000	66.67%
Near cis pairs	9,383,963	9.78%	9,794,426	10.05%	20,581,971	10.08%
Far cis pairs	43,952,899	45.82%	46,620,273	47.85%	97,592,769	47.80%
Trans pairs	5,971,460	6.23%	8,203,411	8.42%	17,935,260	8.78%
Internal re1	9,074,617	9.46%	6,342,544	6.51%	12,109,977	5.93%
Adjacent re1	4,481,520	4.67%	3,235,190	3.32%	5,948,023	2.91%
Circular re1	312,592	0.33%	316,412	0.32%	633,425	0.31%
Overhang re1	4,363,251	4.55%	2,867,348	2.94%	6,856,366	3.36%
Too close	1,047,026	1.09%	724,556	0.74%	1,176,158	0.58%
Too small	13,786	0.01%	17,874	0.02%	19,820	0.01%
Too big	17,296,356	18.03%	19,275,742	19.79%	41,270,054	20.21%
Excluded ambig group	0	0.00%	0	0.00%	0	0.00%
Internal re2	0	0.00%	0	0.00%	0	0.00%
No end re2	0	0.00%	0	0.00%	0	0.00%
Unknown contig	19,746	0.02%	22,813	0.02%	36,612	0.02%
Final output	**E14 wt**		**7g9 *Mbd3-ko* 1**		**7g9 *Mbd3-ko* 2**	
Total pairs	54,176,154		56,152,429		124,812,397	
Total contacts	54,176,154		56,152,429		124,812,397	
Ambiguous contacts	0	0.00%	0	0.00%	0	0.00%
Cis homolog	0	0.00%	0	0.00%	0	0.00%
Cis near (<10kb)	7,416,921	14.31%	8,478,383	15.10%	46,320,661	37.11%
Cis far (>10kb)	39,023,307	75.28%	40,508,436	72.14%	61,992,415	49.67%
Trans	5,475,044	10.11%	7,165,610	12.76%	16,499,321	13.22%

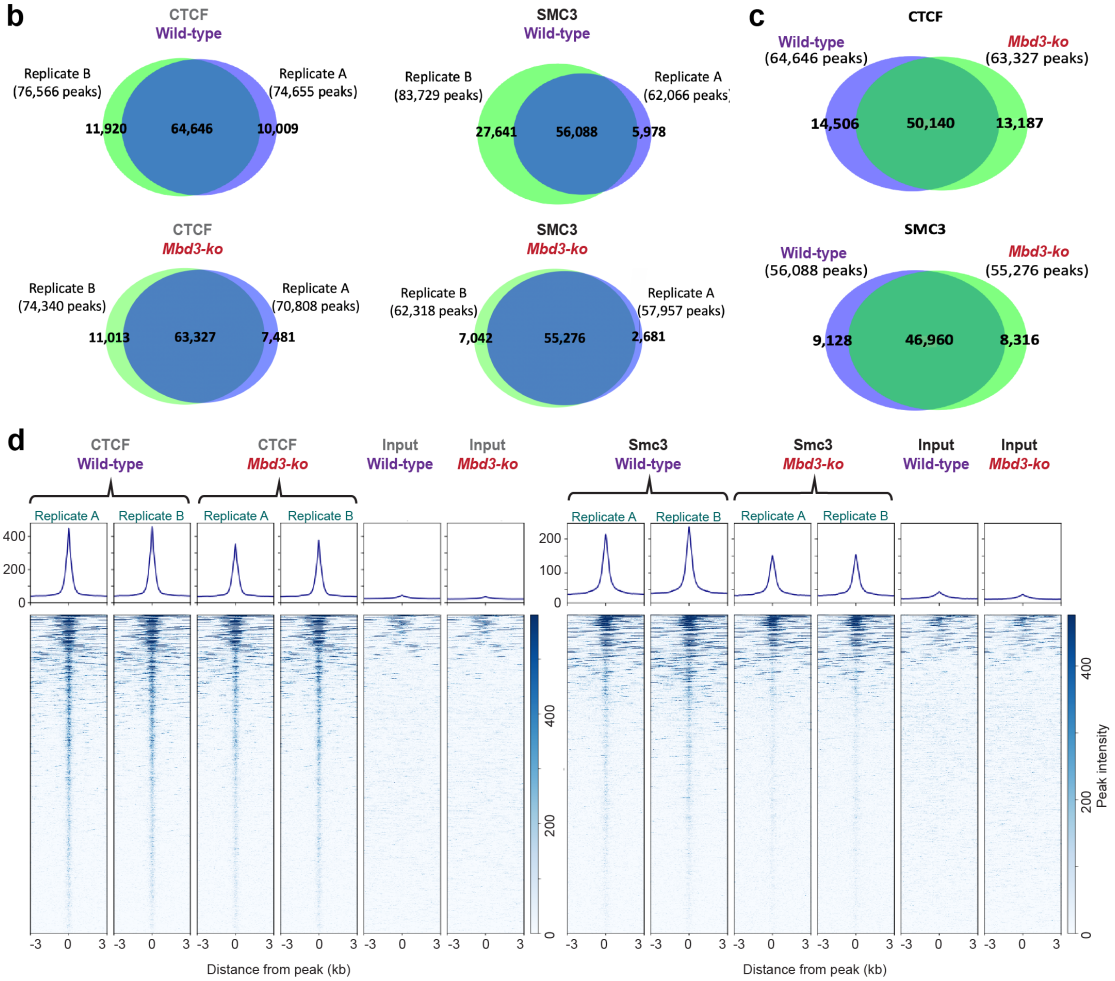
Extended Data Figure 2. MBD3-dependent holo-NuRD complex assembly increases intermediate-range chromosomal interactions. (a) Normalised 25 kb resolution contact maps from *in nucleus* Hi-C of wild-type (bottom left of diagonal) and *Mbd3* knockout ES cells (top right of diagonal); represented at the scale of inter-chromosomal contacts (left), intra-chromosomal contacts in chromosome 5 (middle) and looping interactions near the *Tbx3* gene (right). The boxes show regions where there is an increase in inter-compartmental contacts, whilst the circles show a decrease in looping interactions in the *Mbd3* knockout (red) vs wild-type ES cells (purple). (b) A differential correlation map shows more inter-compartmental interactions in wild-type ES cells and more intra-compartment interactions in *Mbd3* knockout ES cells. (c) Contact probability as a function of genomic distance in the A/B compartments, averaged across the genome for wild-type (purple) and *Mbd3* knockout ES cells (red). Removal of MBD3 results in a decrease in TAD scale (500 kb – 2 Mb) cis-contacts. (d) (Top) Using active enhancers and promoters defined in wild-type ES cells (Stevens et al, 2017)[16], a decrease was observed in interactions between active promoters and nearby enhancers. (Bottom) The normalised number of contacts in the *Mbd3* knockout vs wild-type ES cells at CTCF/cohesin-mediated loop regions defined by Pekowska et al. [55]. The data are represented as the log-ratio of normalised median counts in *Mbd3* knockout vs wild-type ES cells (error bars are the standard error of the mean). (e) Boxplot of normalised contacts in 100 kb bins centred at indicated distances either side of enhancers and promoters shows a greater loss in contacts at enhancers than promoters in *Mbd3* knockout versus wild-type ES cells. (Top) Enhancers show a significant decrease in normalised contacts for all bins in the absence of MBD3 ($p < 0.001$, t-test and Kolmogorov-Smirnov test). (Middle) Highly active promoters show significant decrease only at the central 100 kb bin ($p < 0.001$, t-test and Kolmogorov-Smirnov test), and (Bottom) active promoters show no significant change in normalised contacts for all bins ($p > 0.05$, t-test and Kolmogorov-Smirnov test). (f) Using TADs defined by Dixon et al.[10], and insulation scores derived from the contacts in our Hi-C data, a global increase in TAD insulation was observed upon knockout of MBD3.


a
More inter-chromosomal contacts
Intra-chromosomal contacts differ
Loss of long-range loops
Mbd3-ko
trans-contacts: 12.8%
trans-contacts: 10.0 %
Wild-type
Chromosome
Increase
Decrease
Tbx3 enhancer
Position chr5 (Mb)
b
Stronger A and B compartments
Correlation map
More correlated in wild-type
More correlated in MBD3 knockout
c
All interactions
A compartment to A compartment
A compartment to B compartment
B compartment to B compartment
Contact probability (25kb bins)
Sequence separation (bp)
d
Contacts between active enhancers and promoters
Contacts at CTCF/Smc3-mediated loops
Log (Ratio of contacts in Mbd3-ko/Wild-type ES cells)
Sequence separation (kb)
e
Active enhancers
Highly active promoters
Active promoters
Normalised number of contacts
Condition
Mbd3-ko
Wild-type
Sequence separation (kb)
f
TAD Insulation
Insulation distributions
Boundary comparison
Probability density
Insulation
KO/WT
Insulation score, Mbd3-ko
Insulation score, Wild-type

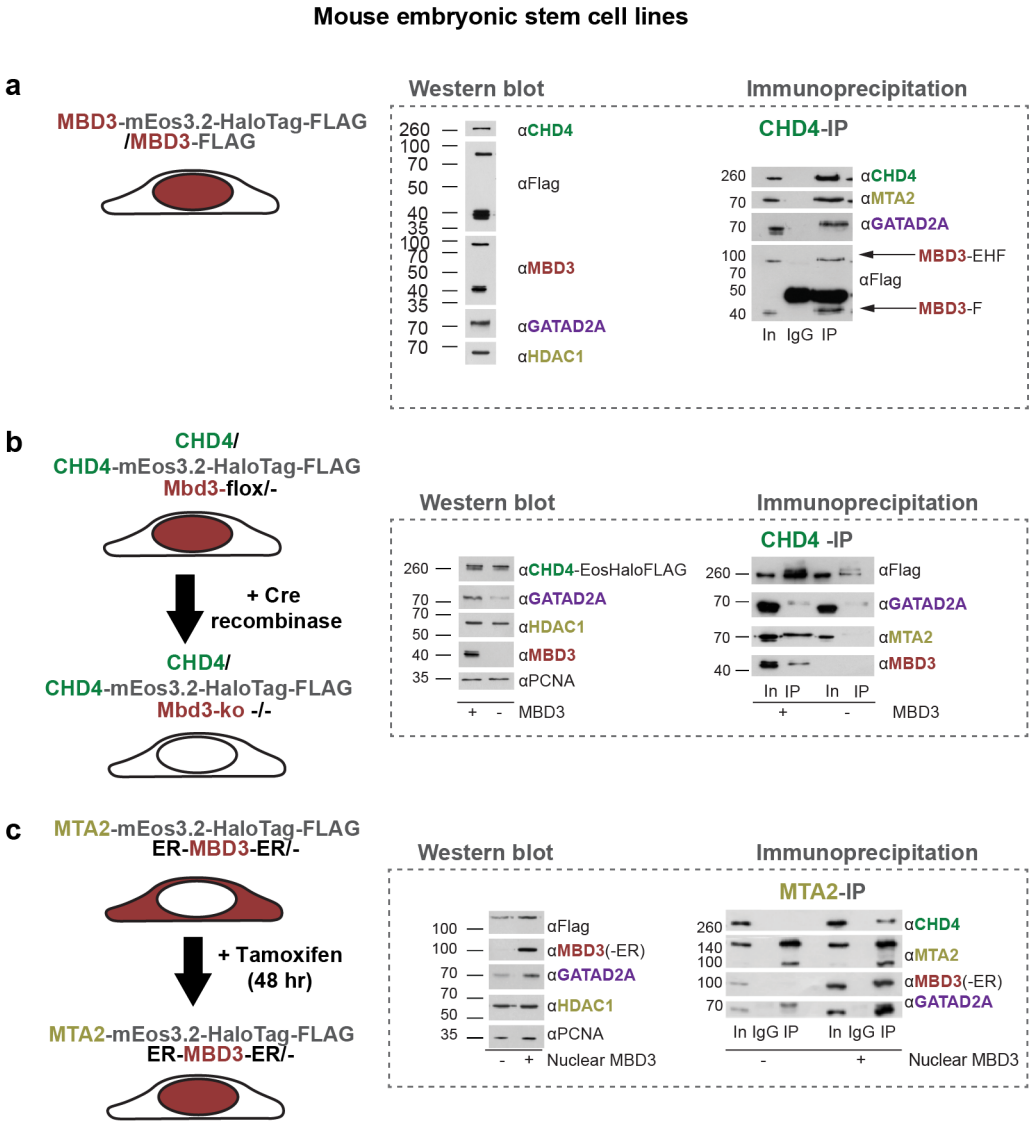
Extended Data Figure 3. ChIP-seq experiments upon removal of MBD3 reveal a loss in CTCF and SMC3 (a cohesin subunit) near MBD3-regulated genes in ES cells. (a) Sequencing statistics for ChIP-seq experiment replicates. (b) Venn diagrams showing the number of CTCF and SMC3 peaks in replicate samples for wild-type and *Mbd3* knockout ES cells. (c) Venn diagrams comparing the number of CTCF and SMC3 peaks in wild-type versus *Mbd3* knockout ES cells. (d) (Top) Average peak profile and (Bottom) heatmap of CTCF and SMC3 signal 3 kb either side of identified peaks shows a loss in cohesin signal intensity (and to a lesser extent CTCF) upon removal of MBD3.

a

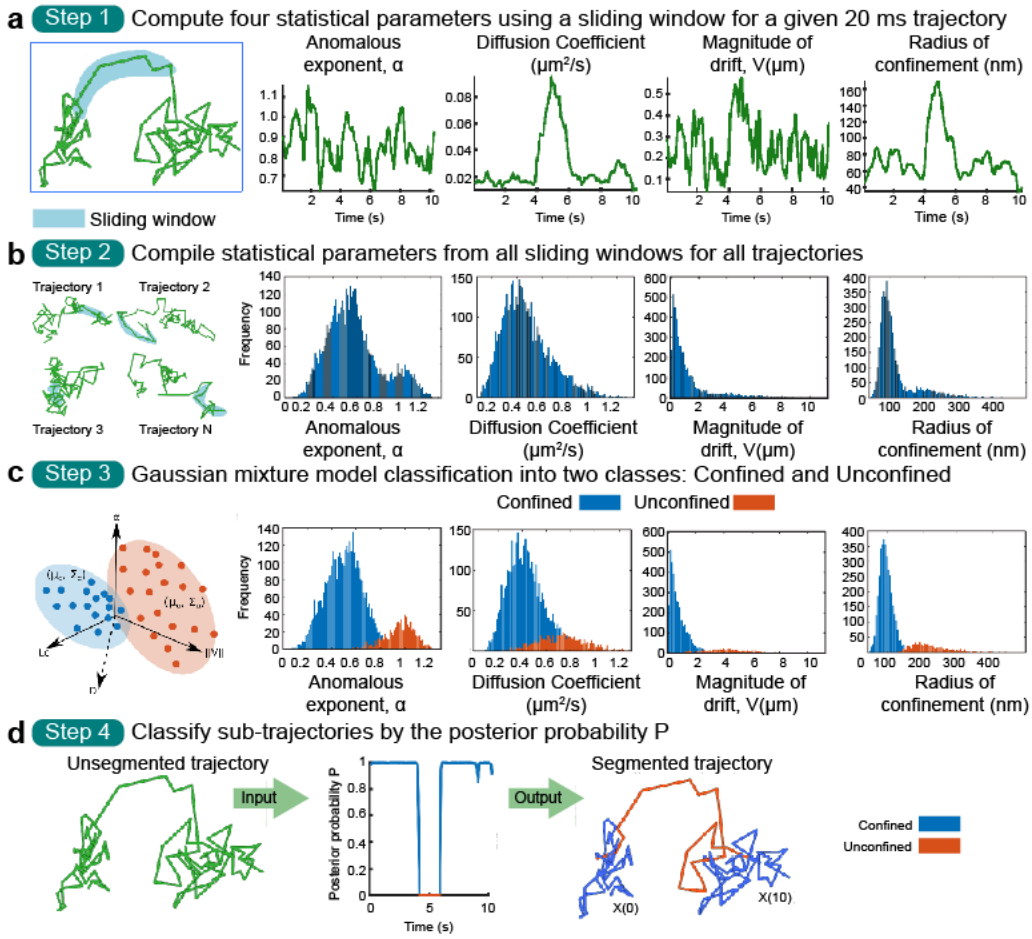
Cell line	Antibody	Replicate	Mapped reads/Total reads	Peaks
E14 Wild-type	Input	-	13,094,679/38,522,985	76,566
7g9 *Mbd3-ko*	Input	-	11,444,825/39,049,951	70,808
E14 Wild-type	CTCF	A	27,305,997/40,393,951	74,655
E14 Wild-type	CTCF	B	31,926,704/44,621,373	76,566
7g9 *Mbd3-ko*	CTCF	A	24,187,417/38,180,143	70,808
7g9 *Mbd3-ko*	CTCF	B	23,961,709/38,655,876	74,340
E14 Wild-type	SMC3	A	23,244,946/38,689,486	62,066
E14 Wild-type	SMC3	B	26,076,922/44,776,616	83,729
7g9 *Mbd3-ko*	SMC3	A	20,052,192/39,822,477	57,957
7g9 *Mbd3-ko*	SMC3	B	22,010,795/44,285,003	62,318



Extended Data Figure 4. Mouse embryonic stem cell lines expressing mEos3.2-HaloTag-FLAG tagged NuRD complex subunits. Schematic of the (a) MBD3, (b) CHD4 and (c) MTA2 cell lines generated. CHD4 was tagged as previously described[23]. MTA2 was tagged in ES cells expressing the ER-MBD3-ER (estrogen receptor-MBD3-estrogen receptor) fusion protein so that nuclear localisation of MBD3 is tamoxifen-inducible[22]. (Left) Expression of NuRD complex subunits was confirmed by western blot. Note that the stability of MTA2 and GATAD2A are both dependent upon MBD3, but that of CHD4 is not[78]. (Right) Immunoprecipitation of MTA2 and CHD4 confirm that the Eos-Halo tags do not prevent association with other NuRD components, and that NuRD complex integrity is dependent upon the presence of MBD3.



Extended Data Figure 5. Segmentation of single-molecule trajectories of JF_{549}-HaloTag-tagged NuRD complex subunits. (a) (Left) A single molecule trajectory, where we show an example sliding window (blue). (Right) Four biophysical parameters are extracted from sliding windows within this trajectory: the anomalous exponent α, the effective diffusion coefficient D, the norm‖V‖ of the mean velocity, and the length of confinement Lc, were all estimated from a sliding window of size w=10 Delta t (see Supplementary Methods). (b) (Left) Several trajectories with example sliding windows (blue). (Right) Histograms of the biophysical parameters extracted in (a) from individual sub-trajectories are computed over the ensemble of all the recorded trajectories. (c) (Left) Scheme of the Gaussian mixture model (GMM) used to separate the histograms of the four parameters (four-dimensional feature space) into confined (C) and unconfined (U) populations (for 20 ms trajectories). The mean vectors μC, μU and covariance matrices ΣC, ΣU are estimated from the maximum likelihood. (Right) Separation of histograms in (b) into confined (blue) and unconfined (orange) populations. (d) Classification procedure applied to each trajectory, resulting in confined (blue) and unconfined (orange) sub-trajectories. The posterior probability P of the GMM (panel b) is computed on the four parameters for each point [$Xi(k\Delta t) \in C$ with $P(k\Delta t) > 1 - P(k\Delta t)$ (blue); otherwise $X(k\Delta t) \in U$ (orange)] (see Supplementary Methods. The result is a segmented trajectory where each time point is assigned as confined (C) or unconfined (U).


a
Step 1
Compute four statistical parameters using a sliding window for a given 20 ms trajectory
Anomalous exponent, α
Diffusion Coefficient (μm²/s)
Magnitude of drift, V(μm)
Radius of confinement (nm)
Sliding window
Time (s)
b
Step 2
Compute statistical parameters from all sliding windows for all trajectories
Trajectory 1
Trajectory 2
Trajectory 3
Trajectory N
Frequency
Anomalous exponent, α
Diffusion Coefficient (μm²/s)
Magnitude of drift, V(μm)
Radius of confinement (nm)
c
Step 3
Gaussian mixture model classification into two classes: Confined and Unconfined
Confined
Unconfined
Frequency
Anomalous exponent, α
Diffusion Coefficient (μm²/s)
Magnitude of drift, V(μm)
Radius of confinement (nm)
d
Step 4
Classify sub-trajectories by the posterior probability P
Unsegmented trajectory
Input
Posterior probability P
Time (s)
Output
Segmented trajectory
X(0)
X(10)
Confined
Unconfined

Extended Data Figure 6. Measuring the accuracy of the classification algorithm using synthetic trajectories. a. Parameter estimation from trajectories with Δt = 20ms. **b.** Parameter estimation from trajectories with Δt = 500 ms. **c.** Parameter estimation from trajectories with Δt = 500 ms after applying the four-parameter classification algorithm. In a, b and c, the blue boxes indicate the preferred model based on BIC analysis. **d.** We generated synthetic trajectories with known classes (confined- *C* and unconfined *U*) to form a ground-truth set on which the classification algorithm could be tested. The switching behaviour is described in Eq. 26 of the Online Methods, where the class (C, U) is determined at each time point using a Markov chain (left) with switching rates λ, μ for the transition between C and U and vice-verse, respectively (right). **e.** After the ground truth set of trajectories was generated, we applied the classification procedure as described in the Online Methods, and computed the accuracy measure *M* (Eq. 33) defined as the fraction of correctly assigned classes by the segmentation algorithm according to the known classes in the ground-truth set. We computed *M* for trajectories generated by Eq. 28 with diffusion coefficients $D_c \in [0.02, 0.1]$ $\mu m^2/s$ for class C and $D_U \in [0.1, 1]$ $\mu m^2/s$ (Left) for U. We find that $M > 0.76$ for all tested values of D_C, D_U, where a well separated D_C, D_U resulted in $M > 0.9$. The accuracy *M* for switching rates λ, $\mu \in [0, 1]$ and $D_C = 0.008$, $D_U = 0.01 \mu m^2/s$, *M*, resulted in $M > 0.85$ for all tested λ, μ (Right).

a

Exps.	Single Diffusion State		Two Diffusion States				Three Diffusion States							
	D ($\mu m^2/s$)	BIC	K_c	D_1 ($\mu m^2/s$)	D_2 ($\mu m^2/s$)	BIC	K_c	K_{u1}	D_1 ($\mu m^2/s$)	D_2 ($\mu m^2/s$)	D_3 ($\mu m^2/s$)	BIC	N disps.	N exps.
CHD4	0.166	-1719	0.33	0.057	0.280	-2886	0.14	0.62	0.034	0.158	0.574	-3262	206894	5
CHD4-GATA	0.140	-1691	0.31	0.046	0.229	-2881	0.13	0.56	0.025	0.121	0.398	-3197	131287	8
MBD3	0.127	-1696	0.30	0.041	0.203	-2895	0.14	0.60	0.025	0.117	0.389	-3330	285096	7

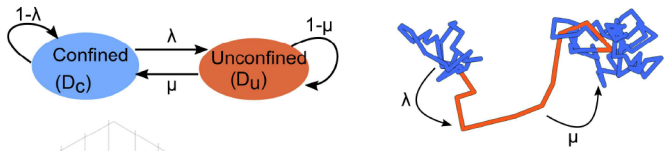
b

Exps.	Single Diffusion State		Two Diffusion States				Three Diffusion States							
	D ($\mu m^2/s$)	BIC	K_c	D_1 ($\mu m^2/s$)	D_2 ($\mu m^2/s$)	BIC	K_c	K_{u1}	D_1 ($\mu m^2/s$)	D_2 ($\mu m^2/s$)	D_3 ($\mu m^2/s$)	BIC	N disps.	N exps.
CHD4	0.005	-1868	0.55	0.003	0.010	-3134	0.18	0.65	0.002	0.005	0.017	-3244	118169	9
CHD4-MBD3	0.004	-1852	0.52	0.002	0.008	-3141	0.21	0.63	0.002	0.004	0.015	-3219	63164	4
CHD4-GATA	0.004	-1865	0.44	0.002	0.006	-2975	0.26	0.65	0.002	0.004	0.014	-2831	18189	3
MBD3	0.003	-1679	0.48	0.001	0.006	-2933	0.19	0.58	0.001	0.003	0.011	-3111	105984	8
MTA2	0.004	-1792	0.49	0.002	0.007	-3006	0.14	0.66	0.001	0.004	0.013	-3192	27551	3

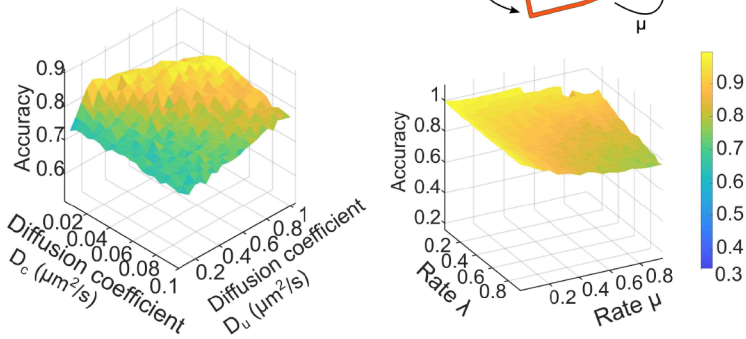
c

Experiments	Decomposition (all trajectories)				Decomposition (unconfined trajectories)					
	K_c	D_c ($\mu m^2/s$)	D_u ($\mu m^2/s$)	BIC	K_{u1}	D_{u1} ($\mu m^2/s$)	D_{u2} ($\mu m^2/s$)	BIC	N disps.	N exps.
CHD4	0.69	0.003	0.007	-2078	0.16	0.004	0.012	-2151	118169	9
CHD4-MBD3	0.71	0.003	0.006	-2085	0.12	0.004	0.010	-2149	63164	4
CHD4-GATA	0.72	0.003	0.005	-1959	0.07	0.002	0.006	-2014	18189	3
CHD4-DRB	0.72	0.002	0.004	-1963	0.10	0.002	0.006	-2055	26272	3
MBD3	0.68	0.002	0.005	-1896	0.16	0.003	0.010	-1933	105984	8
MBD3-FK228	0.75	0.003	0.006	-2070	0.12	0.004	0.011	-2091	29757	3
MTA2	0.76	0.003	0.005	-1933	0.05	0.002	0.007	-2002	27551	3

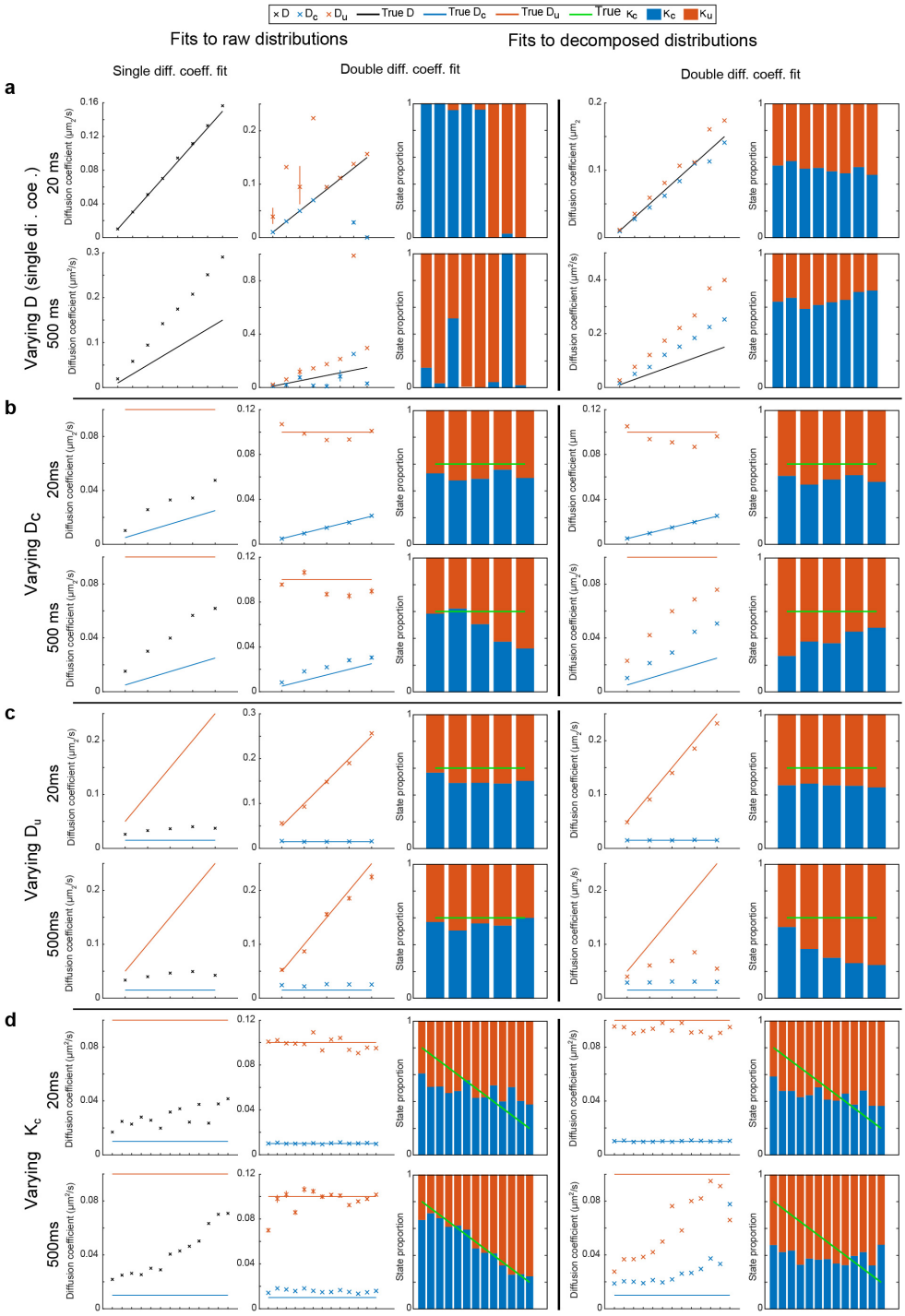
d



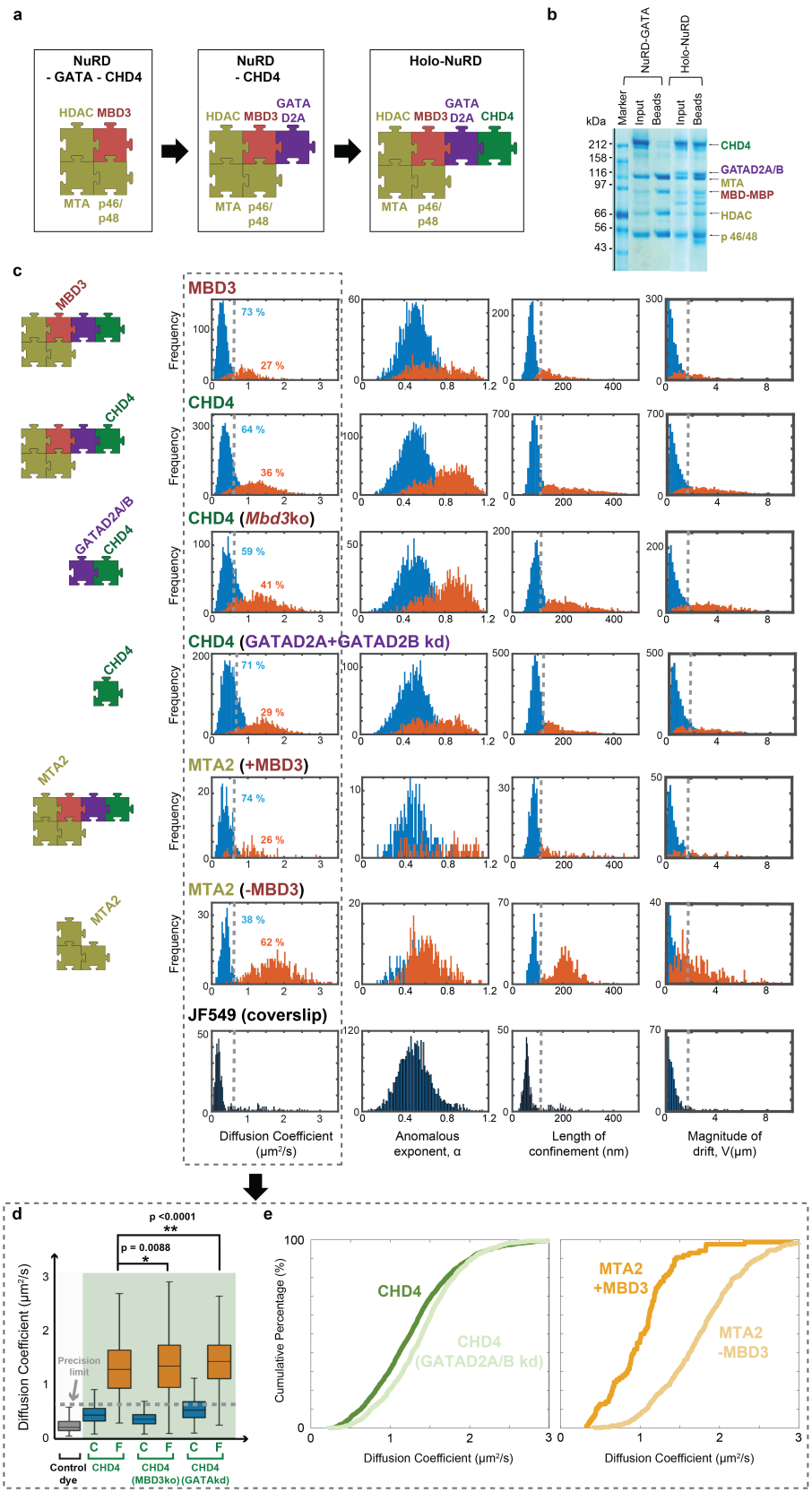
e



Extended Data Figure 7. Parameter estimation from displacement histograms. Displacement distributions (left column) determined when fitting the confined and unconfined distributions obtained using the four parameter classification algorithm (right column). We fitted either one (left) or two (right) diffusion coefficients. **a**. Effect of varying a single diffusion coefficient: for D = 0.01,0.03,0.05,0.07,0.09,0.11,0.13,0.15 μm^2/s and Δt = 20 (top) or 500 ms (bottom). **b**. Effect of varying D_c for two state diffusion when D = 0.005,0.01,0.015,0.02,0.025 μm^2/s with D_u = 0.1μm^2/s, $\kappa_{c\to u}$ = 0.4, $\kappa_{u\to c}$ = 0.6 for Δt = 20 (top) or 500 ms (bottom). **c**. Effect of varying D_u for two state diffusion when D = 0.05,0.1,0.15,0.2,0.25 μm^2/s and with D_c = 0.015μm^2/s, $\kappa_{c\to u}$ = 0.4, $\kappa_{u\to c}$ = 0.6 for Δt = 20 (top) or 500 ms (bottom). **d** Diffusion coefficients and proportions when varying the transition coefficients [$\kappa_{c\to u}$,$\kappa_{u\to c}$] = [0.2,0.8], [0.75,0.25], [0.7,0.3], [0.65,0.35], [0.6,0.4], [0.55,0.45], [0.5,0.5], [0.45,0.55], [0.4,0.6], [0.35,0.65], [0.3,0.7], [0.25,0.75] ,[0.8,0.2] with D_c = 0.01μm^2/s, D_u = 0.1μm^2/s for Δt = 20 (top) or 500 ms (bottom).


× D × Dc × Du True D True Dc True Du True Kc Kc Ku
Fits to raw distributions
Fits to decomposed distributions
Single diff. coeff. fit
Double diff. coeff. fit
Double diff. coeff. fit
a
Varying D (single di . coe .)
20 ms
500 ms
b
Varying Dc
20ms
500 ms
c
Varying Du
20ms
500ms
d
Varying Kc
20ms
500ms
Diffusion coefficient (μm²/s)
State proportion

Extended Data Figure 8. *In vitro* and live cell single-molecule imaging experiments delineate holo-NuRD complex assembly. (a) Schematic of holo-NuRD complex assembly with GATAD2A linking MBD3 to the CHD4 remodeller. (b) Pull-down experiments of MBP-tagged MBD-like with and without GATAD2A confirm that GATAD2A is required for CHD4 to interact with the deacetylase sub-complex. (c) Distribution of the four biophysical parameters described in **Figure 1** for 20ms exposure tracking of MBD3 and CHD4 in wild-type ES cells, as well as CHD4 in the absence of either MBD3 or GATAD2A/B. The data for MTA2 in the presence and absence of nuclear localised MBD3 are also shown. The grey dotted line indicates the upper bound of the precision limit calculated at the 95 % confidence interval for an immobilised JF_{549} dye control sample. (d) Box plot of diffusion coefficients extracted from chromatin bound (C) and freely diffusing (F) CHD4 molecules in wild-type, *Mbd3* knockout and GATAD2A/B knock-down ES cells (*$p < 0.01$, **$p < 0.001$, Kolmogorov-Smirnov test). The grey dotted line indicates the upper bound of the precision limit calculated at the 95 % confidence interval for an immobilised JF_{549} dye control sample. (e) Cumulative distribution functions showing a higher diffusion coefficient for freely diffusing unconfined CHD4 upon removal of GATAD2A/B, and for freely diffusing MTA2 molecules upon removal of MBD3 from the nucleus (see **Extended Data Figure 4**). (f) Table showing the number of chromatin bound and freely diffusing sub-trajectories and total number of trajectories analysed (many trajectories discarded as they were either too short for analysis or had a low probability of being classified as confined or unconfined).

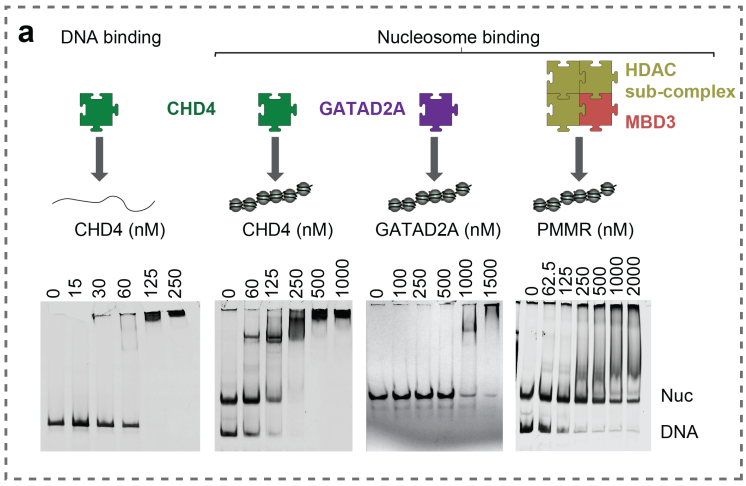


f

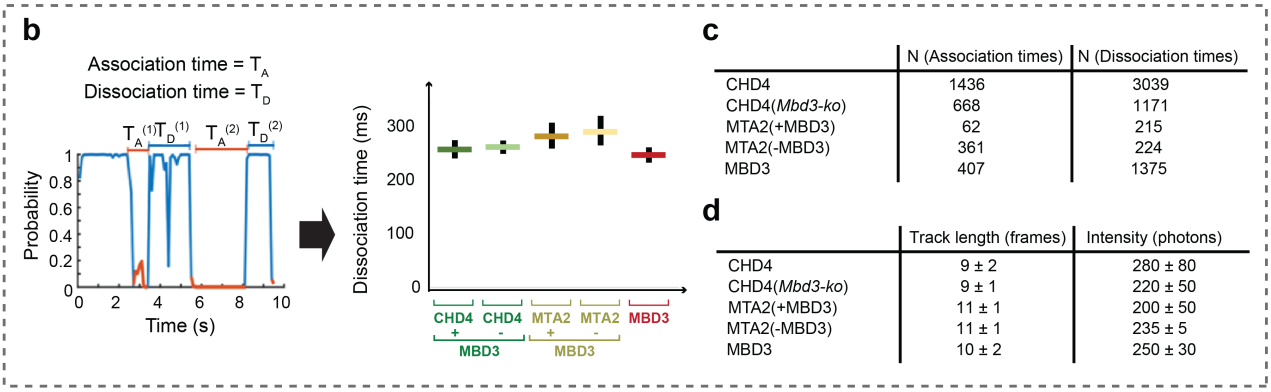
	Chromatin bound sub-trajectories	Freely diffusing sub-trajectories	Total number of trajectories analysed
CHD4	3576	1981	125521
CHD4(*Mbd3*ko)	1372	965	51215
CHD4(GATAD2A+B kd)	2760	1111	86372
MTA2(+MBD3)	250	86	8669
MTA2(-MBD3)	250	402	15130
MBD3	1617	607	48646

Extended Data Figure 9. The NuRD complex interacts with both DNA and nucleosomes, but only forms tight interactions through the CHD4 remodeller. (a) *In vitro* electrophoretic mobility shift assays confirm that CHD4 binds to both DNA and nucleosome core particles (NCPs) to form large complexes that only just enter the gel. GATAD2A alone shows low affinity binding to NCPs whilst the deacetylase complex interacts with DNA, but does not bind stably to NCPs. (b) (Left) Confinement probability allows collection of the association T_A or dissociation T_D times – defined respectively as the time a trajectory spends between periods of confined or unconfined motion. (Right) Dissociation times calculated using transitioning trajectories as periods of confined motion between two periods of unconfined motion (see also **Figure 2**). Error bars show 95 % confidence intervals. (c) Table with the number of single molecule tracks that could be used to determine the association and dissociation times. (d) Table with mean track length in frames and mean signal intensity per localised molecule in photons detected. (e) (Top) Example images demonstrating how long 500 ms exposures motion blur freely diffusing molecules, but allow detection and tracking of those that are chromatin-bound. Images of single chromatin-bound CHD4 molecules during time-lapse imaging with various dark times. Error bars show 95 % confidence intervals, ns = not significant, $p > 0.05$, 2-way ANOVA. (Bottom) Exponential fitting of time-lapse residence time histograms to extract the photobleaching rate k_b and the effective dissociation rate k_{eff}. Dissociation times ($t_{off} = 1/k_{off}$) of CHD4 were then compared to those of MBD3 and CHD4 in the absence of MBD3.

In vitro Electrophoretic mobility shift assays



20 ms exposure data



c

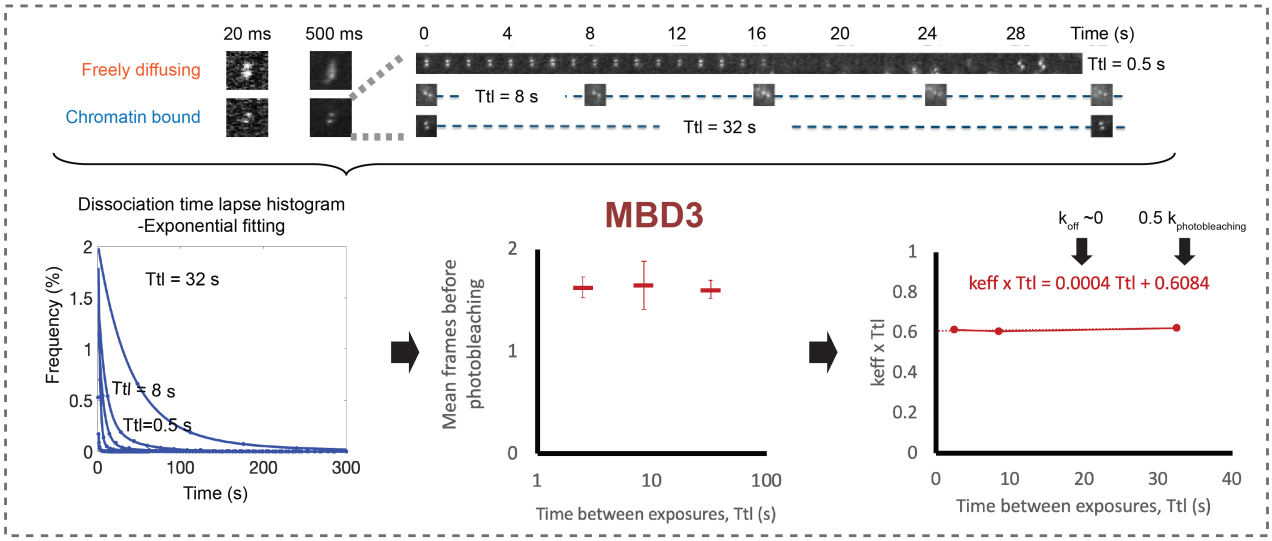
	N (Association times)	N (Dissociation times)
CHD4	1436	3039
CHD4(*Mbd3-ko*)	668	1171
MTA2(+MBD3)	62	215
MTA2(-MBD3)	361	224
MBD3	407	1375

d

	Track length (frames)	Intensity (photons)
CHD4	9 ± 2	280 ± 80
CHD4(*Mbd3-ko*)	9 ± 1	220 ± 50
MTA2(+MBD3)	11 ± 1	200 ± 50
MTA2(-MBD3)	11 ± 1	235 ± 5
MBD3	10 ± 2	250 ± 30

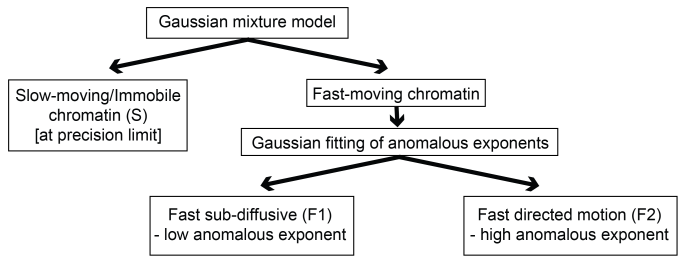
e

500 ms exposure data --- Dissociation rate from time lapse imaging

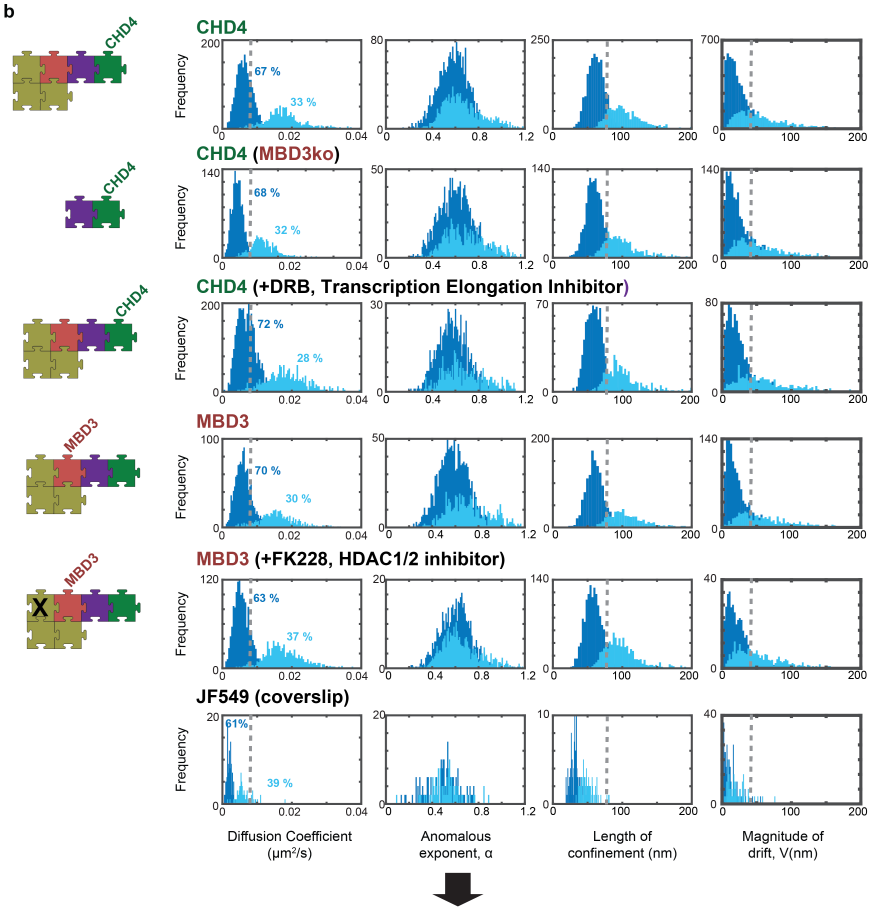


Extended Data Figure 10. **Live cell single-molecule imaging experiments to study chromatin-bound NuRD complex**. (a) Schematic of the approach used for analysing the 500 ms exposure trajectories of chromatin-bound NuRD complex subunits. First, a Gaussian mixture model is used to distinguish slow-moving/immobile and fast-moving chromatin bound sub-trajectories based on the 4 biophysical parameters (see also **Figure 1** and **Extended Data Figure 5**). Then, Gaussian fitting of the anomalous exponent distribution of the fast-moving chromatin bound sub-trajectories revealed two fast-moving populations (F1 and F2) characterised by distinct anomalous exponents. (b) Distribution of the four biophysical parameters for 500 ms exposure tracking of: (i) chromatin bound CHD4 in wild-type ES cells, in the absence of MBD3, and in the presence of DRB (an inhibitor of transcriptional elongation); (ii) chromatin bound MBD3 in wild-type ES cells, and in the presence of the HDAC1/2-specific inhibitor FK228; (iii) JF549 dye bound at the coverslip. The grey dotted line indicates the upper bound of the precision limit calculated at the 95 % confidence interval for an immobilised JF_{549} dye control sample. (c) (Left) Fitting of 1, 2 or 3 Gaussians to the anomalous exponent distributions for chromatin bound CHD4 in wild-type ES cells – the R^2 values above indicate the goodness of fit. (Right) The Bayesian information criterion (BIC) was calculated for all the datasets shown in (b) to determine that 2 Gaussians are the best minimal model to account for the data – have the lowest BIC value (light blue box). (d) Table summarising the changes in anomalous exponent of the slow and fast chromatin bound NuRD complex subunits in the presence and absence of MBD3, or in the presence of specific inhibitors. Errors given are for 95 % confidence intervals.

a



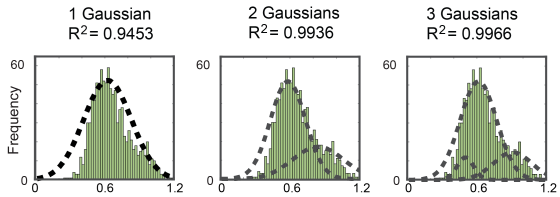
b



c

Gaussian fitting of fast-diffusing anomalous exponents

Anomalous exponents of CHD4



Bayesian information criterion (BIC) from Gaussian fitting

	No. of Gaussians (n)		
	n = 1	n = 2	n = 3
CHD4	-721	-811	-795
CHD4-MBD3	-436	-471	-457
CHD4+DRB	-199	-219	-204
MBD3	-313	-342	-324
MBD3+FK228	-643	-730	-710
Coverslip	-82	-78	-69

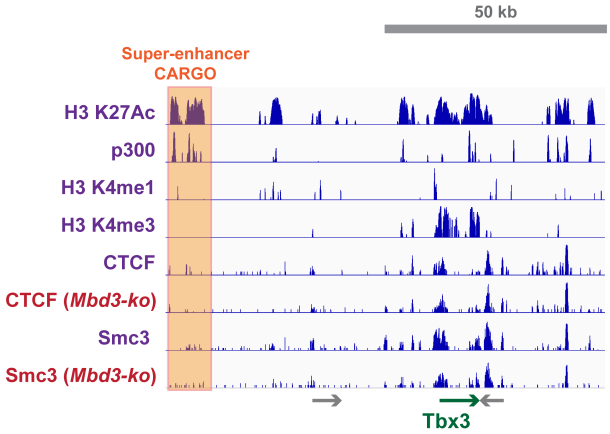
d

Anomalous exponents of and percentage of molecules in S, F1 and F2 states

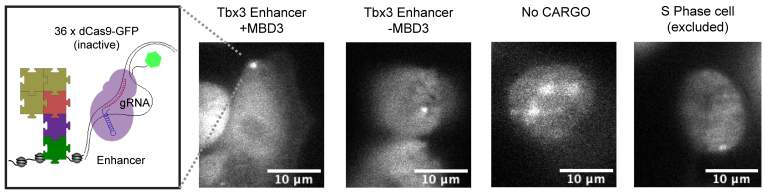
	Slow confined (S)	Fast confined (F1)	Fast confined (F2)	N
CHD4	0.593 ± 0.007 (67 %)	0.601 ± 0.014 (25.8 ± 1.1 %)	0.89 ± 0.02 (7.4 ± 0.6 %)	15528
CHD4-MBD3	0.598 ± 0.006 (68 %)	0.559 ± 0.019 (14 ± 6 %)	0.77 ± 0.12 (18 ± 5 %)	9270
CHD4+DRB	0.598 ± 0.006 (72 %)	0.556 ± 0.015 (19 ± 3 %)	0.90 ± 0.04 (8.3 ± 1.1 %)	5768
MBD3	0.580 ± 0.004 (70 %)	0.650 ± 0.016 (22 ± 2 %)	0.97 ± 0.03 (8.0 ± 1.5 %)	7973
MBD3+FK228	0.580 ± 0.004 (63%)	0.572 ± 0.010 (28 ± 5 %)	0.82 ± 0.06 (8.8 ± 1.8 %)	10614

Extended Data Figure 11. 2D dCas9-GFP tracking of the *Tbx3* enhancer. (a) Genomic location of the *Tbx3* super-enhancer (orange box) to which dCas9-GFP was targeted using CARGO vectors[4]. The corresponding ChIP-seq profiles indicate the location of active enhancers (determined from the profiles for H3K27ac, p300 and H3K4me1), active promoters (H3K4me3), and CTCF/Cohesin (SMC3) in wild-type ES cells and in the absence of MBD3. (b) Representative images of 36 gRNAs targeted to the *Tbx3* enhancer in the presence or absence of MBD3 with a negative control expressing no gRNAs. (Right) Example of cell excluded from analysis with doublets indicating cell is in S phase. (c) Distribution of the four biophysical parameters extracted from sliding windows within the 2D single-molecule trajectories of bound dCas9-GFP at the *Tbx3* enhancer in wild-type ES cells imaged using 500 ms exposures – (top) before and (middle) after classification based on the anomalous exponent α, the effective diffusion coefficient D, the length of confinement Lc, and the drift magnitude, norm‖V‖ of the mean velocity. (Bottom) Distribution of the four biophysical parameters after classification for *Tbx3* enhancer in *Mbd3-ko* cells. (d) (Left) Fitting of 1, 2 or 3 Gaussians to the anomalous exponent distributions of the *Tbx3* enhancer in either wild-type ES (Top) or *Mbd3-ko* cells (Bottom) – the R^2 values above indicate the goodness of fit. (Right) The Bayesian information criterion (BIC) was calculated for both the datasets shown in (c) to determine which number of Gaussians best modelled the data – that with the lowest BIC value (light blue box). (e) Table showing the Gaussian fitted anomalous exponent values for slow- and the fast-moving chromatin bound dCas9-GFP at the *Tbx3* enhancer in the presence and absence of MBD3.

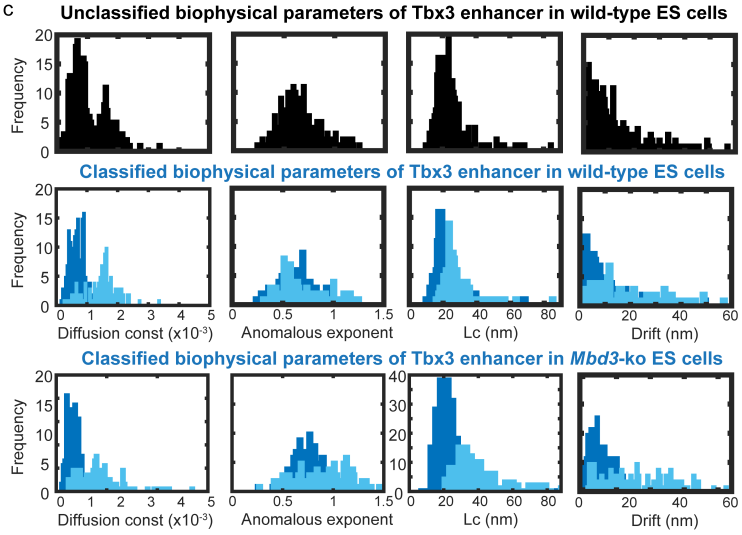
a



b

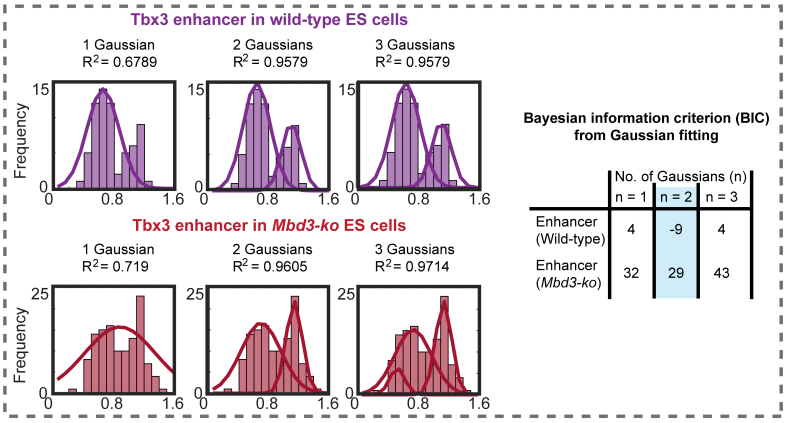


c



d

Gaussian fitting of fast-diffusing anomalous exponents



Bayesian information criterion (BIC) from Gaussian fitting

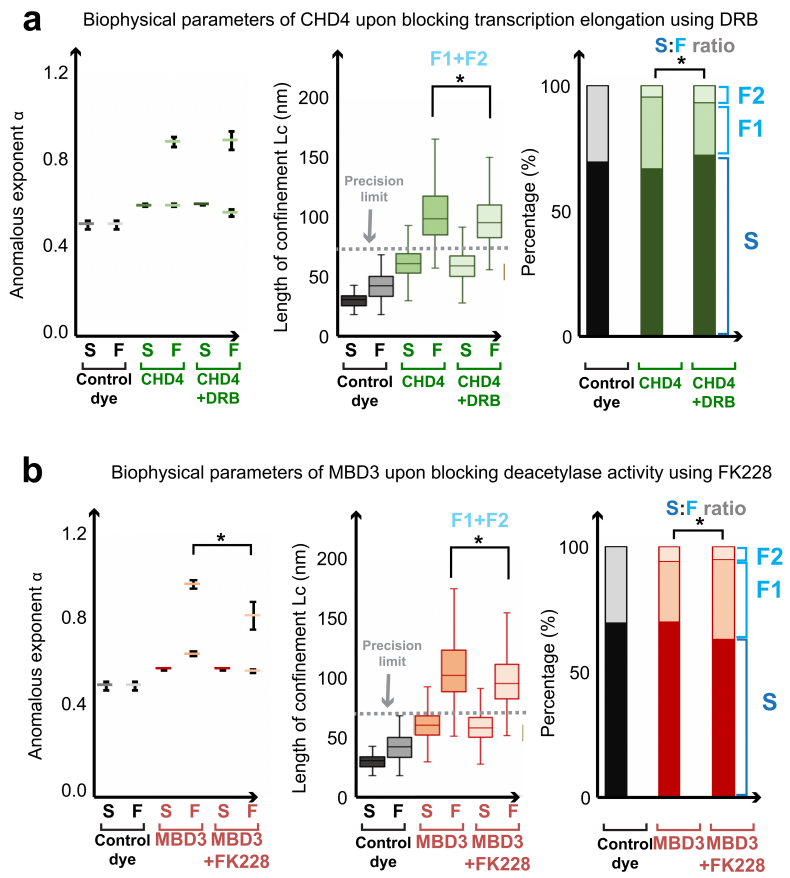
	No. of Gaussians (n)		
	n = 1	n = 2	n = 3
Enhancer (Wild-type)	4	-9	4
Enhancer (*Mbd3-ko*)	32	29	43

e

Anomalous exponents of and percentage of molecules in S, F1 and F2 states

	Slow confined (S)	Fast confined (F1)	Fast confined (F2)	N
Tbx3 Enhancer Wild-type	0.675 ± 0.008 (71.73%)	0.635 ± 0.001 (18.47%)	1.071 ± 0.001 (9.80%)	237
Tbx3 Enhancer *Mbd3-ko*	0.754 ± 0.004 (57.49%)	0.621 ± 0.007 (21.39%)	1.075 ± 0.007 (21.12%)	287

Extended Data Figure 12. **Live cell single-molecule imaging experiments of chromatin-bound NuRD in the presence of inhibitors**. (a) Biophysical parameters of chromatin-bound CHD4 in wild-type ES cells, both before and two hours after adding the transcription elongation inhibitor DRB. (Left) The anomalous exponents (error bars show 95 % confidence intervals from Gaussian fitting) (Middle) Boxplot of the localisation lengths (*p < 0.01, **p < 0.001, Kolmogorov-Smirnov test). (Right) Percentage of molecules in the slow or fast chromatin bound states (from Gaussian fitting, *p < 0.01, 2-way ANOVA). (b) Biophysical parameters of chromatin-bound MBD3 in wild-type ES cells, both before and 2 hours after adding the HDAC1/2-specific inhibitor FK228. (Left) The anomalous exponents (error bars show 95 % confidence intervals from Gaussian fitting, *p < 0.01, 2-way ANOVA). (Middle) Boxplot of localisation lengths (*p < 0.01, **p < 0.001, Kolmogorov-Smirnov test). (Right) Percentage of molecules in the slow or fast chromatin bound states (from Gaussian fitting, *p < 0.01, 2-way ANOVA). In (a) and (b), the grey dotted line indicates the upper bound of the precision limit calculated at the 95 % confidence interval for an immobilised JF_{549} dye control sample.



References (Extended Data)

4 Gu, B. *et al.* Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements. *Science* **359**, 1050-1055, doi:10.1126/science.aao3136 (2018).

10 Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380, doi:10.1038/nature11082 (2012).

16 Stevens, T. J. *et al.* 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* **544**, 59-64, doi:10.1038/nature21429 (2017).

22 Bornelov, S. *et al.* The Nucleosome Remodeling and Deacetylation Complex Modulates Chromatin Structure at Sites of Active Transcription to Fine-Tune Gene Expression. *Mol Cell* **71**, 56-72 e54, doi:10.1016/j.molcel.2018.06.003 (2018).

23 Zhang, W. *et al.* The Nucleosome Remodeling and Deacetylase Complex NuRD Is Built from Preformed Catalytically Active Sub-modules. *J Mol Biol* **428**, 2931-2942, doi:10.1016/j.jmb.2016.04.025 (2016).

29 Kaji, K. *et al.* The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. *Nat Cell Biol* **8**, 285-292, doi:10.1038/ncb1372 (2006).

55 Pekowska, A. *et al.* Gain of CTCF-Anchored Chromatin Loops Marks the Exit from Naive Pluripotency. *Cell Syst* **7**, 482-495 e410, doi:10.1016/j.cels.2018.09.003 (2018).

78 Burgold, T. *et al.* The Nucleosome Remodelling and Deacetylation complex suppresses transcriptional noise during lineage commitment. *EMBO J* **38**, doi:10.15252/embj.2018100788 (2019).