Combined SPT and FCS methods highlight a mechanism of RNAP II oversampling in cell nuclei

Marie Fournier[1,6,§], Pierre Leclerc[1,6,§], Aymeric Leray[2,6], Dorian Champelovier[1,6], Florence Agbazahou[1,6], Gabriel Bidaux[3,6], Alessandro Furlan[1,4,5,6,*], Laurent Héliot[1,6 *]

1. University of Lille, CNRS, UMR 8523, PhLAM Laboratoire de Physique des Lasers, Atomes et Molécules, Lille, France
2. Laboratoire Interdisciplinaire Carnot de Bourgogne, UMR 6303 CNRS, Université de Bourgogne Franche Comte, Dijon, France
3. INSERM UMR 1060, CarMeN laboratory, Univ Lyon1, IHU OPERA, Hôpital Louis Pradel, Hospices Civils de Lyon, Lyon, France
4. Univ. Lille, CNRS, Inserm, CHU Lille, Institut de Recherche contre le Cancer de Lille, UMR9020 – UMR-S 1277 - Canther, F-59000 Lille, France
5. Unité Tumorigenèse et Résistance aux Traitements, Centre Oscar Lambret, F-59000 Lille, France
6. Groupement de Recherche ImaBio, CNRS Villeneuve d'Ascq, F−59655 France

[§]: These authors contributed equally to this work

*: Co-senior authors; for correspondence: alessandro.furlan@univ-lille.fr; laurent.heliot@univ-lille.fr

Keywords: molecular dynamics, diffusion, single-particle tracking, spectroscopy, model, transcription, biophotonics, RPB1

ABSTRACT

The orchestration of gene expression is a key question in both fundamental and applied research. Different models for transcription regulation were proposed, yet the dynamic regulation of RNA polymerase II (RNAP II) activity remains a matter of debate. To improve our knowledge of this topic, we investigated RNAP II motility in eukaryotic cells by combining Single Particle Tracking (SPT) and Fluorescence Correlation Spectroscopy (FCS) techniques, to take advantage of their different sensitivities. Thanks to calibrated samples, we first developed a benchmark for measurements and analysis to eliminate the main potential instrumental biases. We then applied this workflow to study the diffusion of RPB1, the catalytic subunit of RNA pol II. By a cross-analysis of FCS and SPT, we were able to highlight different RPB1 motility states and characterize them. We could distinguish between a very slow Brownian diffusion state, a stationary state, and two different modes of confined diffusion. Interestingly, our analysis also highlighted the oversampling by RPB1 of nuclear subdomains. We propose a novel model of highly dynamic spatio-temporal transcription regulation based on these results. Altogether, our results highlight the importance of combining microscopy approaches at different time scales to get a full insight into the real complexity of molecular diffusion distribution kinetics.

INTRODUCTION

Biological activities are classically modulated through the dynamic binding and unbinding of molecular partners (Presman et al., 2017). Depending on the process considered, molecules may exist in several distinct complexes, which differ in their motility and availability features. In the frame of gene expression modulation, diffusion and target search constitute critical features for transcription actors (Halford and Marko, 2004; Izeddin et al., 2014). It was acknowledged for a long time that RNA Pol II was recruited on the target genes by general transcription factors during pre-initiation complex (PIC) formation (Kim et al., 1997; Kornberg, 2007; Lee and Young, 2000). Indeed, RNAP II was initially observed in mammalian cells in cluster-like structures by immunostaining of the polymerase active form in fixed cells and was co-localized with nascent RNAs (Chakalova and Fraser, 2010; Jackson et al., 1993). This led to the interpretation of static transcription factories. Recent advancements in the field however suggest that the assembly of RNA Pol II is very dynamic throughout time and that several RNA pol II can be co-recruited and form transient clusters which disintegrate after a few seconds, depending on the transcriptional activity (Cho et al., 2016; Cisse et al., 2013). How these clusters match with the concepts of transcription factories or TAD (Topologically Associated Domains), and how they organize over time remain elusive. These divergent models probably highlight distinct facets of the same highly dynamic process, which we aimed to better understand with state-of-the-art microscopy tools designed for molecular dynamics studies.

Various microscopy approaches have been developed in the last decades to quantify and/or track molecular motilities. Fluorescence microscopy has notably been widely used to decipher the dynamics of molecular orchestration and is a key element to push forward the understanding of spatiotemporal molecular mechanisms. The mobility of a given molecule in cells classically encompasses several diffusion modes (Hager et al., 2009; Hansen et al., 2018), and retrieving quantitative data about the diffusion of both fast and slow molecules is critical to understanding the dynamics of molecular events, including transcription, in the nucleus. To this end, we combined FCS (Fluorescence Correlation Spectroscopy) and SPT (Single-Particle Tracking) methods to take advantage of the complementary spatiotemporal information they provide.

The calculation of molecular scattering using FCS is based on the measurement at the µs scale of the fluorescence fluctuations due to the transit of fluorescent molecules through the excitation focal volume of a laser beam (Magde et al., 1972). Its temporal sensitivity is thus excellent. Although the focal volume is generally small (about 0.1 μm^3), it corresponds at the nucleus scale to a subdomain that can encompass several distinct molecular crowding conditions or molecular interactions, and each experimental point thus already constitutes an average of several behaviors (Langowski, 2008). Moreover, such fluctuation-based techniques are blind to strong binding events in which molecules barely move (Altan-Bonnet and Altan-Bonnet, 2009; Elson, 2011), which constitutes a drawback that has to be taken into account.

In parallel, super-resolution SPT techniques have been developed thanks to the improvement of fluorophores and camera sensitivity over the last decade, and allow tracking of individual molecules in the field of view over time (Cheezum et al., 2001; Manley et al., 2008). Their temporal sampling rate (classically

10-100 ms/frame for a 10 µm² area) unfortunately restrains the capacity to measure molecular motilities with diffusion coefficients higher than 5-10 µm²/s (Hansen et al., 2018; Harwardt et al., 2018). In addition, SPT techniques require the analysis of a large number of individual traces (10^3-10^5 trajectories) that may not be equally represented as a function of their length, depending on the mobility of the molecules and the experimental set-up. These biases may result in an over-representation of slower molecules with longer trajectories and in less accurate statistics concerning faster molecules, which is often underestimated in SPT analyses (Izeddin et al., 2014).

In this study, we first evaluated the results obtained by FCS and SPT methods with reference samples consisting of beads of known diameter in water/glycerol mixtures of different viscosities, to establish a benchmark for the use of these methods to unravel RPB1 behavior in the nucleus. This allowed us to overcome the potential biases that could derive from the different timescales and different analytical protocols of the two techniques. We reasoned that it was important to deal with experimental data rather than simulations because the former describe at best the reality of experimental images and instrumental constraints and contributions.

We came up with a robust process displaying optimized conditions, which we applied to the analysis of RPB1 dynamics. Thanks to it, we show that FCS and SPT techniques are complementary and allow the identification of a wide range of POL II subpopulations with distinct diffusion features in cell nuclei.

RESULTS

We first characterized the sensitivity of both FCS and SPT techniques, thanks to a system based on the diffusion of fluorescent beads in different media of controlled viscosity. To recapitulate the apparent diffusion values observed in the literature for proteins of the transcription complex in the nucleus, i.e. between 0.1 and 15 µm²/s (Grünwald et al., 2008; Izeddin et al., 2014), we selected appropriate parameters to perform FCS and SPT experiments (see Material and Methods for more details).

Characterization of calibrated bead diffusion by ACF analysis of FCS measurements

The FCS experiments were conducted with fluorescent microspheres in a water/glycerol mix, with a proportion of glycerol ranging from 0 to 80 %. These combinations provided us with control solutions displaying theoretical diffusion coefficients ranging from 15 µm²/s to 0.2 µm²/s (Table S1).

Autocorrelation curves resulting from these experiments were then fitted with an anomalous model. While the Brownian model is clearly more suited to analyzing freely diffusing particles, we chose to fit all data sets with an anomalous model to avoid the potential bias that locking the anomaly coefficient to 1 may have brought. Moreover, the anomalous model is classically used when studying the molecular motility in living cells (Mazza et al., 2012; Stasevich et al., 2010; Wachsmuth et al., 2000), and allowed us to maximize the dispersion range of the results and to explore at the same time the goodness of this model that we would use later to investigate RNA pol II diffusion in cells. We globally obtained a good correlation between the theoretical values and our FCS experimental values (Fig.1 and Table S1).

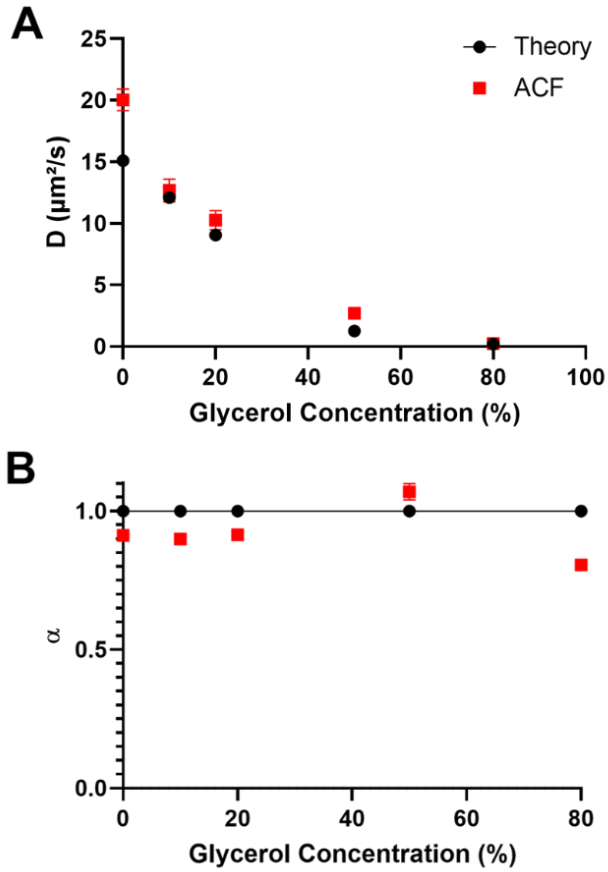


Figure 1. Characterization of diffusion by Fluorescence Correlation Spectroscopy. Fluorescent microspheres were resuspended in water/glycerol mixtures, with proportions of glycerol ranging from 0 to 80 %. The ACF curves obtained from FCS acquisitions were fit with an anomalous model to infer the diffusion coefficients (A) and anomalous coefficients (B). Experimental values (red squares) can be compared to theoretical values (black dots). Experimental results are expressed as means +/- SD from at least 20 independent measurements.

We noticed a slight overestimation of our computed results for the diffusion coefficients when compared to the theoretical values, especially for the diffusion of beads in pure water (Fig.1A). This could stem from a minimal error in the estimated value of viscosity inferred, as previously reported by (Takamura et al., 2012) and described in Material and methods.

Besides, up to a glycerol content of 50 % in water, the ACF analysis yielded an anomalous coefficient value very close to the theoretical value of 1 (between 0.96 and 1.02), as expected for Brownian diffusion (Fig.1B). Only the diffusion of beads in the solution containing 80 % glycerol was described as anomalous, with an alpha of 0.70 +/- 0.02. However, it is noteworthy that this mixture might be non-perfectly homogenous due to the high viscosity of glycerol, and that may result in anomalous diffusion, as already described in the literature (Banks et al., 2016; Banks and Fradin, 2005).

Characterization of calibrated bead diffusion by SPT analysis

To compare the measurements by FCS and SPT, we applied a similar strategy to analyze the diffusion of reference beads with SPT.

We performed our SPT acquisitions in HILO (Highly Inclined and Laminated Optical sheet) microscopy mode (Tokunaga et al., 2008), a technique classically used for nucleus imaging because it allows a selection of the plan of observation thanks to a very inclined optical sheet. We analyzed our temporal image series with the MTT (Multiple-Target Tracing) algorithm tracking software (Sergé et al., 2008) to detect and reconstruct trajectories of the particles. Since short trajectories strongly penalize the stability of the adjustment function methods (Michalet, 2010), we only selected trajectories with more than 10 detections, which are more reliable. We first wondered whether the distribution of the trajectory lengths varied as a function of the diffusion coefficients in our benchmark (between 0.1 and 4 $\mu m^2/s$) and observed that all conditions included trajectories with a mean number of detections superior or equal to 20 (Fig.S1).

In concordance with FCS analysis, the trajectories were analyzed by an anomalous model in order not to make any *a priori* on the diffusion mode. We found a computed anomalous coefficient alpha very close to 1, which is consistent with the fact that beads in solution should present a Brownian movement (Fig.2A). Once again, a high concentration of glycerol (80%) resulted in an alpha computed at 0.8, with this apparent anomaly probably associated with a certain heterogeneity of the mixture.

We then implemented two methods to compute diffusion coefficients: D_α and D_{60ms}. D_α is obtained from a linear regression of the logarithmic MSD (Mean Square Displacement) of all trajectories as a function of time, alpha being the slope of the line and D_α the value at the origin. Besides, it is classically assumed that, at short time intervals, molecules behave like Brownian particles; that's why Etoc et al. proposed to calculate the diffusion coefficient D_{60ms} for each trajectory by considering the slope of the vector from the origin to its value at 60 ms (Etoc et al., 2018).

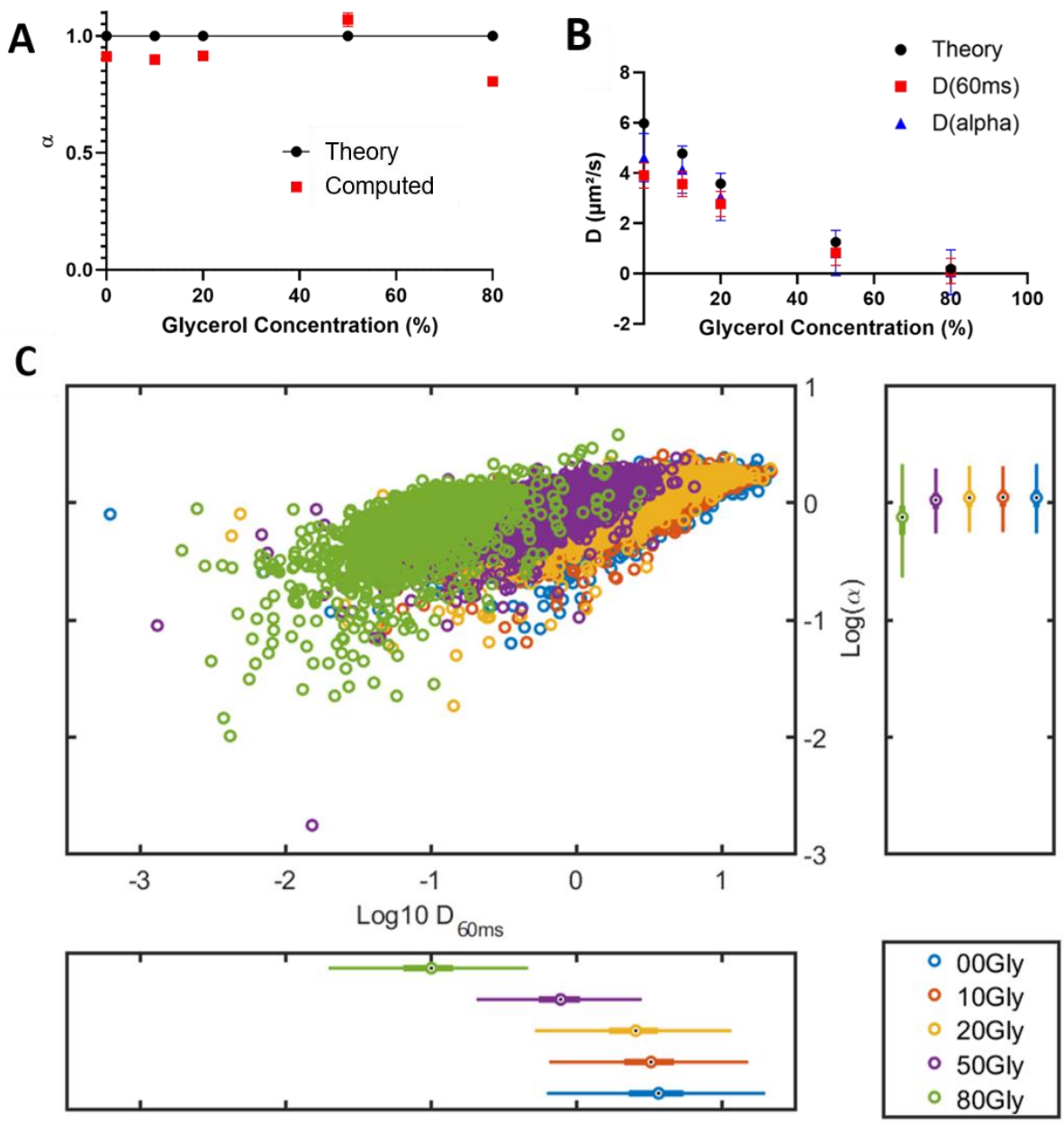


Figure 2. Characterization of diffusion by Single Particle Tracking. Fluorescent microspheres were resuspended in water/glycerol mixtures, with proportions of glycerol ranging from 0 to 80 %. SPT acquisitions were performed at a frame rate of 100 Hz. A. The anomalous coefficient α is provided as a function of the glycerol proportion. Experimental values (red squares) can be compared to theoretical values (black dots). B. The diffusion coefficients were computed by two methods: D_a (blue triangles) were obtained from the slopes of TEA-MSD (Time Ensemble Average Mean Square Displacement) expressed as log, while D_{60ms} were calculated for each trajectory during the first 60ms of its occurrence (red squares). These experimental values could be compared to theoretical values (black dots). Both methods yielded an asymptotic trend for high D values, that can be linked to physical limitations such as the framerate of the camera used to record the event. C. The distributions of D_{60ms} and α from the different experimental conditions are represented with log scales in a scatter plot. Distinct subpopulations can be distinguished (green, purple, and yellow dots) within a limited range of parameters, whereas some others overlap (yellow, red, and blue dots). Data come from at least one hundred measurements per condition.

Both methods yielded diffusion coefficients slightly smaller than the theoretical values (Fig.2B). This underestimation of diffusion coefficients with SPT was previously described (Wieser and Schütz, 2008). Interestingly, the standard deviation associated with D_{60ms} was two times lower when compared to that of D_{α}, thus showing a clear advantage of the first method over the second one.

We noticed that, for theoretical values of diffusion coefficients above 4 µm²/s, the values of D_{60ms} and D_{α} were lower than the expected theoretical values, suggesting that we had reached a plateau (Fig.2B). Thus, our SPT experiments did not allow the study of diffusions faster than 4 µm²/s, which stands in agreement with the classical measurement ranges observed in the literature with this method, due to limitations in camera frame rates, tracking algorithms and fluorophore brightness.

Consistently, the representation of measurements as a scatter plot (Fig.2C) nicely illustrated the fact that the populations that we detected at different glycerol concentrations could be easily distinguished, up to this plateau where they overlapped. The distribution of bead diffusion measurements still evidenced that this method was well suited to detect particles diffusing with coefficients over several orders of magnitude.

h-MSD discriminates quite well trajectories mixed from different subpopulations

In SPT, the individual trajectories of every single molecule display an imprint of the convolution between the instrumental function and the local environment fluctuation. To extract global rules for the diffusion of molecules, the Time Ensemble Average-Mean Square Displacement (TEA-MSD) analysis method averages a large number of trajectories, which improves the Signal-to-Noise Ratio (SNR) in an inherently noisy observation. However, this method is only valid when there is a single diffusion law (i.e. no subpopulations), otherwise, it can lead to a loss of and/or erroneous information because the result will be the weighted average of these populations. Since we aim at characterizing biological samples in which there are mixtures of subpopulations, we chose to use the histogram analysis (called H-MSD), which is based on the MSD analysis of each trajectory and allows to get a comprehensive overview of diffusion features in complex samples with a histogram representation (Etoc et al., 2018).

Under the assumption that D_{60ms} and alpha distributions follow sums of power laws, one can discriminate different subpopulations emerging from a heterogeneous sample, by using a Gaussian Mixture Model (GMM) to fit the data of log(D) and log(alpha) distributions. Using a methodology derived from unsupervised machine learning, we used the BIC (Bayesian Information Criteria) and AIC (Akaike Information Criteria) to infer the proper *k* number of populations (ranging from 1 to 10) by analyzing the point of inflection of the BIC and AIC curves as a function of *k*. Then we computed GMM with this *k* to fit our distributions.

We first tested this methodology by mixing batches of data originating from distinct bead acquisitions with various glycerol contents and assessed its ability to discriminate subpopulations and retrieve both their right proportions and accurate coefficient values. Consistently, the BIC and AIC analyses indicated that two-

population models best described our different mixtures (Fig.3A and B for a mixture of trajectories from the 0 % and 50 % glycerol conditions). We showed that the H-MSD/GMM method allowed for discriminating distinct subpopulations, with good quality in retrieving both the proper values and proportions for each subpopulation (Fig.S2 and Table S2). Logically, in these settings in which particles correspond to beads in aqueous solutions, i.e. Brownian particles, the discrimination happened in the diffusion coefficient dimension while alpha logically remained centered around 1 (Fig.3C). This method was pushed towards its limits only when mixing subpopulations that were very close to the upper limit of detection, namely when merging populations with diffusion coefficients centered around 3.8 and 4.0 µm²/s, that were considered as a single population (Table S2D).

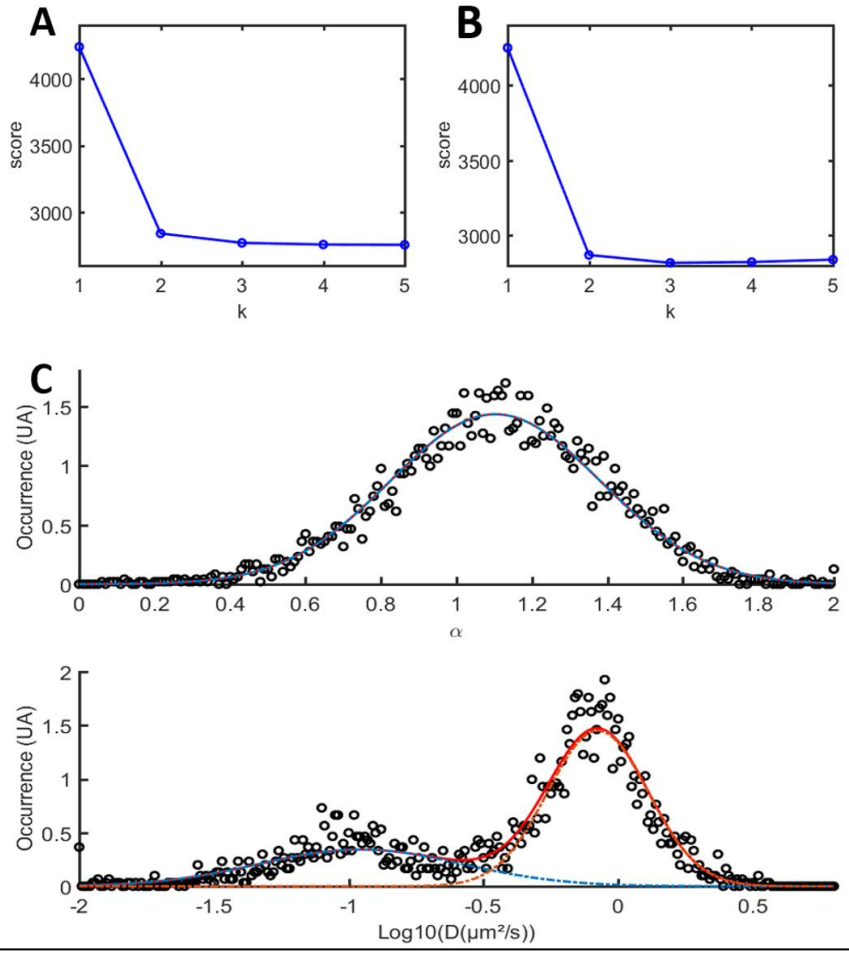


Figure 3. Discrimination of merged subpopulations thanks to h-MSD. SPT trajectories from experiments with fluorescent microspheres diffusing in 20 % and 50 % glycerol solutions respectively, were merged and submitted to the h-MSD analysis workflow to compute for each trajectory the alpha and log(D). The number of Gaussian functions to be found in our distribution was inferred from the AIC (Akaike Information Criteria) (panel A) and the BIC (Bayesian Information Criteria) (panel B) respectively. In this example, the inflection point occurs for a number of populations k=2, consistently with the mixture of two experiments. Using the results of the h-MSD analysis, we represented in panel C the distributions of alpha (upper graph) and D (lower graph). The distribution was then fitted with a GMM (Gaussian Mixture Model). The details of all combinations and the comparison of calculated values with the initial values are provided in the Supplemental Material.

FCS detects a wide array of RPB1 diffusion modes in living cells.

Once this benchmarking of molecular dynamics analyses was validated, we applied our methods to dissect the orchestration of transcription in eukaryotic cells. For that purpose, we engineered U2OS cell lines to replace endogenous RPB1, the catalytic subunit of RNAPII, with an SYFP2 fluorescently-labeled amanitin-resistant RPB1 version (Fig.4A), following the strategy previously used by Darzacq and collaborators (Cisse et al., 2013; Darzacq et al., 2007).

A

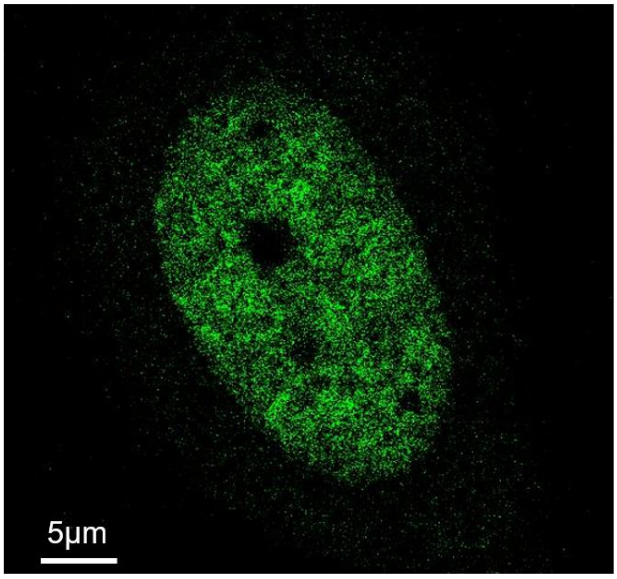
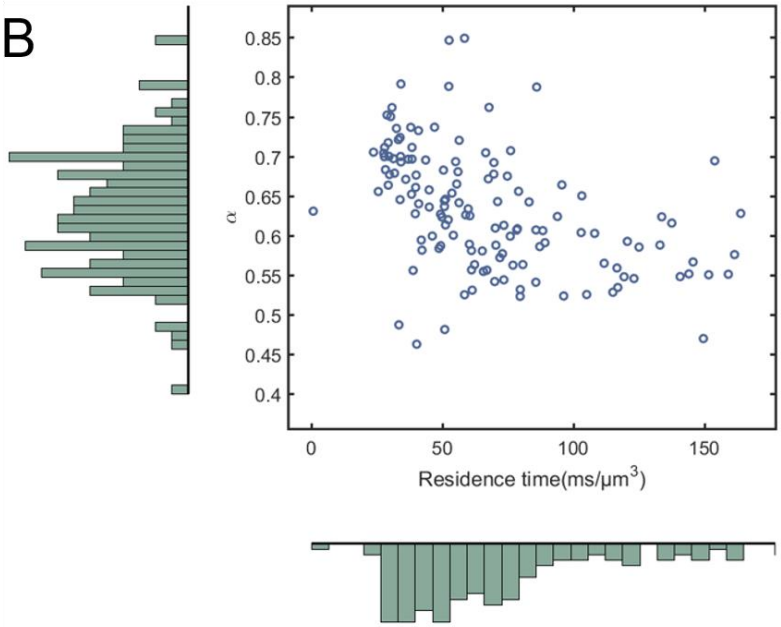


Figure 4. FCS detects a wide array of RPB1 diffusion modes in living cells.
A. Representative picture of a U20S RPB1-SYFP2 cell. B. Scatter plot of residence times versus anomalous coefficients computed from 138 FCS measurements performed in U2OS cells expressing RBP1-SYFP2 (n=6). 80% of the residence times range from 23.7 and 89.1 ms/µm^3 with an average of 61.9 ms/µm^3 (D=5.7 µm²/s$^{\alpha}$). And the anomalous coefficients α range from 0.5 to 0.8 with a mean value of 0.61.

B



The scatter plot of results from FCS experiments carried out with U20S cells stably expressing RPB1-SYFP2 indicated that residence times, as well as anomalous coefficients, were spread out, which is representative of a variety of different behaviors detected by FCS (Fig.4B). It should be noted that the anomalous dimension made a direct comparison of the coefficient values of Brownian and anomalous

diffusing molecules questionable, although the community classically acknowledges the apparent value approximation. That context drove us to propose the use of residence times to normalize the values to be compared. RPB1 diffusion was described by residence times ranging from 20 to 100 ms/µm^3, while the anomalous coefficients were mostly distributed between 0.5 and 0.8, with a mean value of 0.61, indicating a certain level of confinement. Of note, the distribution of diffusion values was imperfectly fitted by a Gaussian model, suggesting the existence of minority subpopulations and/or the presence of some outliers.

Besides, it has to be noted that the RPB1 diffusion coefficients measured by FCS (mean value of 5.74 µm^2/s) mostly exceed the SPT detection plateau (4 µm^2/s) determined during our benchmarking, suggesting that FCS and SPT methods will lead to the complementary identification of different RPB1 subpopulations.

H-MSD SPT highlights the existence of distinct RPB1 subpopulations.

We then performed SPT experiments to identify those putative additional RPB1 subpopulations not detected by FCS. SPT requires a dedicated labeling strategy to get only a limited number of molecules that can be tracked over time in the field of view. These molecules should be bright enough to be detected in a single particle tracking regime and possess a high photostability to be followed over a sufficient period (at least a few seconds, and ideally several tens of seconds). For that purpose, we took advantage of the novel Janelia fluorophores coupled to the Halo ligand JF549 (Grimm et al., 2017) and engineered a stable cell line expressing RPB1 fused to the Halo tag.

In that context, we were able to follow RPB1 molecules, whose SPT trajectories were processed as aforementioned. We observed a great variety of molecular behaviors, as illustrated by the different kinds of tracks reported in Fig.S3. The BIC and AIC analyses of these data suggested the presence of at least four distinct subpopulations (Fig.5A and B), and it is plausible that other subpopulations are encompassed within this distribution although they cannot be efficiently discriminated. Following this workflow, we generated a scatter plot of anomalous coefficients and residence times corresponding to RNAP II (RPB1-Halo/JF549) trajectories (Fig.5C). Two main populations were highlighted in this plot: a majority one, probably encompassing several subpopulations as suggested by the distorted shape of the Probability Density Function, representing 75% of the trajectories, centered around a residence time TR=266 ms/µm^3 (i.e. D=0.02 µm^2/s) and an anomaly coefficient α=0.31, and a second one, corresponding to 25 % of the trajectories, with mean TR=0.63 ms/ µm^3 (i.e. D=0.5 µm^2/s) and α=0.95.

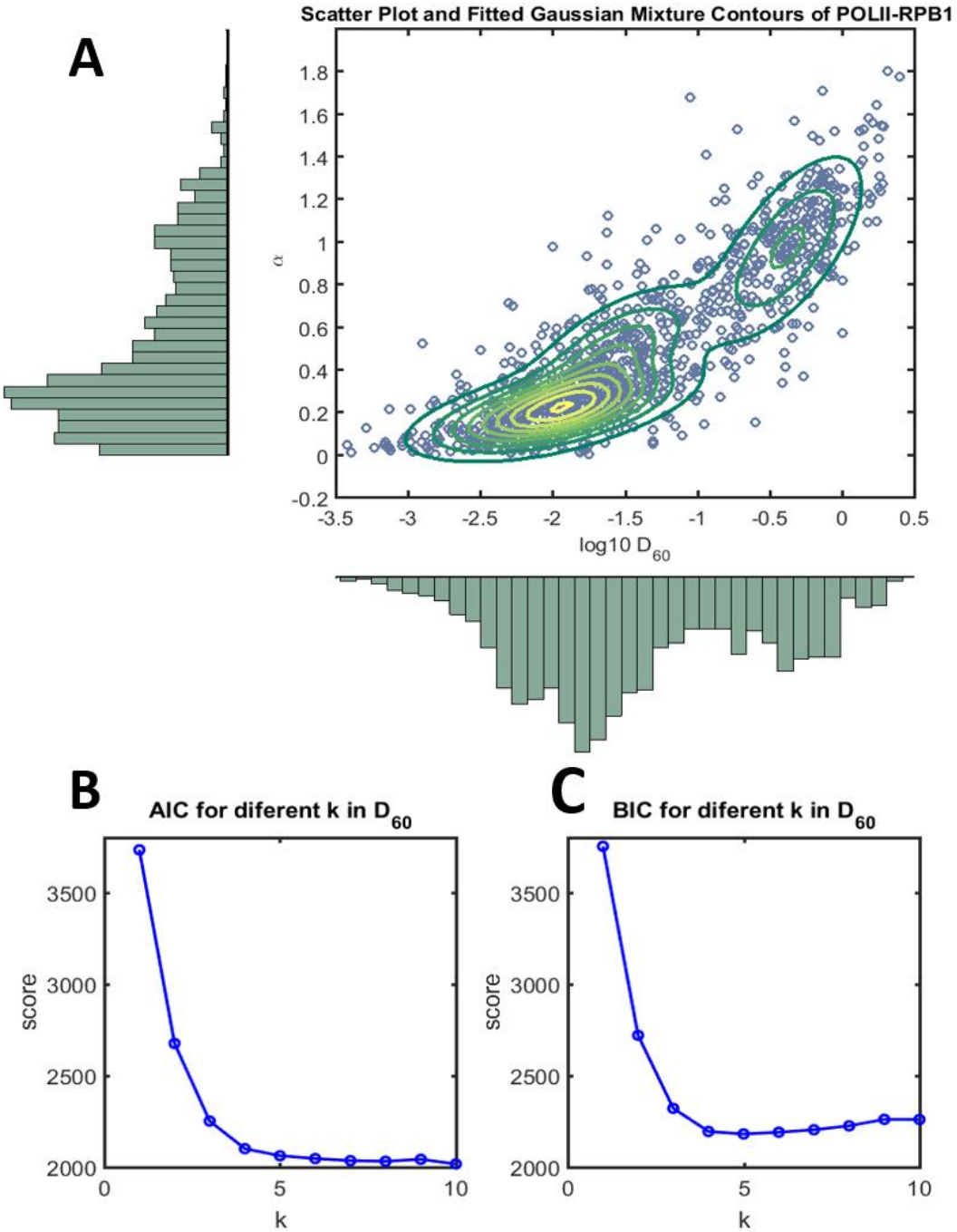


Figure 5. H-MSD SPT highlights the existence of distinct RPB1 subpopulations. A: 1051 trajectories from SPT experiments carried out in the U2OS RPB1-Halo cell line (n=10), stained with JF549 at 0.1nM, were analyzed by h-MSD. AIC (Akaike Information Criteria) (panel A) and BIC (Bayesian Information Criteria) (panel B) suggest the existence of four populations in the dataset. C. D and α.values from these experiments are represented in a scatter plot, in which concentric lines show the Probability Density Function given by the Gaussian Mixture Model with four populations. At least two distinct populations are clearly visible, a first population (75 %) centered around D_{60ms}=0.02 µm^2/s and α =0.31+/- 0.25, and a second one, corresponding to 25 % of the trajectories centered around D_{60ms}=0.5 µm^2/s and α =0.95.

223

Spatio-temporal localization of RPB1 subpopulations.

Another intriguing observation that we made resides in the fact that no static regions with increased RPB1 concentrations could be observed within our super-resolution experiments. To get enough detections to infer the spatial information, we chose to stain the cells with a JF549 concentration higher than that used for multiple particle tracking and trajectory analysis. This level of signal allowed us to image the whole nucleus and investigate the global spatio-temporal distribution of RPB1 (see supplementary movies). When checking the mean intensity projection signal along the movie, we could observe that some areas were highlighted, that could correspond to transcription areas, over a diffuse background, showing that RPB1 explores the whole nucleus (except for nucleoli (Fig.6A). Moreover, when summing all the detections occurring throughout the acquisition, we noticed that some regions were more frequently visited by RPB1 molecules, suggestive of a previously undescribed mode of dynamic spatio-temporal regulation of gene transcription by RNAP II oversampling (Fig.6B). Importantly, when averaging the detections from all frames, we obtained a quite homogenous noisy background (Fig.6C), indicating that the regions highlighted in Fig.6A and B did not result from a static phenomenon. When comparing the signals from Fig.6A and 6B (merged in Fig.6D), one can see that regions with higher intensity mainly match, still, a great proportion of the nucleoplasm was visited by RPB1 but did not lead to detections by SPT. Finally, the analysis of detections from successive batches of 50 frames indicated that spots appeared and disappeared over time, demonstrating the highly dynamic nature of these elements (Fig.6E and F).

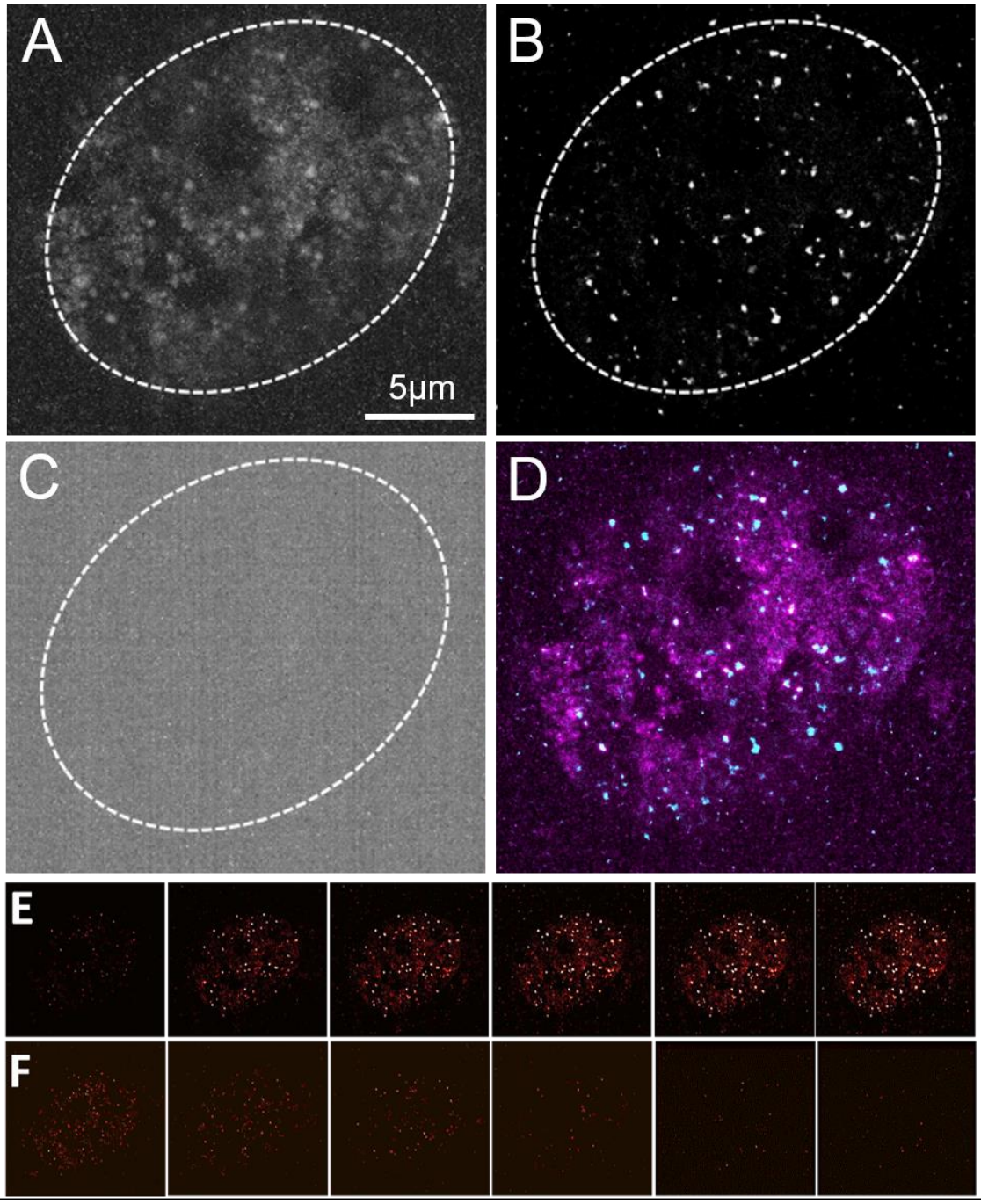


Figure 6. Pictures of the nucleus of an RPB1-Halo expressing U2OS cell stained with JF549. These images were acquired as a separate experiment within a 10s movie, after staining with JF549 at 1 nM. Panel A displays the mean intensity projection throughout the movie. Panel B represents the projection of all detected localizations using MTT in the whole film. Picture C shows the mean projection of detected localizations for the whole movie. Picture D merges the information from pictures A (in purple) and B (in cyan). E and F represent a series of 50 frames, displaying projections of MTT detections summed over time (E) or considered independently (F). The movie is also provided in supplementary data (Movies S1, S2, and S3), with raw data (S1), ongoing tracking superimposed (S2), or detections summed over time (S3).

DISCUSSION

Quantifying the mobility of molecules in living cells is a very important issue for understanding the regulation of physiological mechanisms. The FCS and SPT techniques are highly suitable for such studies. However, because they strongly differ in acquisition methods, analysis, and experimental parameters, they are not easily comparable. In addition, the photo-physical properties of fluorophores (i.e. blinking, lifetime, etc.) affect the two techniques in different ways. For these two techniques to produce complementary information, it is necessary to properly characterize their windows of sensitivity and identify their limitations.

FCS is based on the fluctuations of the averaged fluorescence signal emitted by several molecules in the focal volume. It is particularly suited for the quantification of fast diffusion events but is less sensitive to slower molecular motilities with respect to SPT. In addition, low-frequency events are classically missed by the ACF analysis. In contrast, SPT is based on the study of single-molecule trajectories and therefore allows the study of diffusion heterogeneities on a molecule-by-molecule basis, and potentially gives access to the monitoring of rare molecular events. A relative consensus is emerging around the tracking methods, but the methods of trajectory analysis and the models used for interpretation are still a subject of discussion (Etoc et al., 2018; Hansen et al., 2018). The main limitation of SPT remains its low sensitivity to fast displacements,

because a large number of photons must be detected to allow accurate localization of single molecules, which requires integration times in the millisecond (ms) scale, while FCS is using sampling in the microsecond scale (µs). However, labeling with bright fluorophores presently allows obtaining good frame rates and track particles for tens of time points (Grimm et al., 2015). The optical settings and configurations used in FCS or SPT are also quite different, with each instrument response functions (IRF) that must be considered to define their influence.

We, therefore, reasoned that a measurement-driven comparison between the two techniques was instrumental. For that purpose, we first characterized the influence of the setups on the diffusion measurements of calibrated beads. Whereas the use of simple models such as beads is widespread in FCS characterization, few works have dealt with this concept in SPT. We selected this solution, over *in silico* simulations of random behavior generation, because we reasoned that our approach would display the great advantage of taking into account instrumental functions and real experimental noise. Thus, we could compare the performances of both techniques with relatively simple experimental settings. We show that SPT gives reliable results for diffusion coefficients below 4 $\mu m^2/s$, closer to the theoretical values when compared to FCS, while this later allows measurements up to 14 $\mu m^2/s$ at least.

To better control the measurement and analysis parameters of heterogeneous samples, we developed a benchmark based on a mixture of tracks from bead populations with different diffusion coefficients. Interestingly, our SPT analysis workflow, relying on the result distribution fitting by a Gaussian Mixture Model (GMM), discriminated quite well between the two initial subpopulations and obtained values very close to the expected ones (i.e., measured in the initial mono-populations). Moreover, the proportions

of populations given by the analysis were also close to those used to make the mixtures, with very low variability. It has to be noticed that some mixtures displayed slightly stronger deviations from the input values, probably induced by the measurement conditions, which confirms the interest of our approach to validate the accuracy level of the analysis process, in comparison with the *in-silico* modeling approach. Taken together, we can consider that the proportions are reliable with a relative uncertainty as low as 10 %.

Once our workflow coupling SPT and FCS was validated, we applied it to the elucidation of the RNA polymerase II choreography in cell nuclei. Labeling of the RPB1 subunit of polymerase II (RNAP II) with fluorescent proteins, or Halo tags and adequate ligands, allowed us to measure RPB1 molecular dynamics with the microscopy approaches previously described. The analysis of a thousand RNAP II SPT trajectories with H-MSD showed that 20 % of RPB1 molecules appeared immobile for more than 60ms. Logically, FCS, that analyses fluctuations, was blind to this fixed fraction. Reciprocally, we established that most RPB1 diffusion measurements by FCS exceeded the values caught by SPT in our benchmark, even though such molecules probably contribute to the background noise of SPT acquisitions. Among the diffusive RPB1 population, the SPT/H-MSD unsupervised clustering allowed us to discriminate at least two main populations. The fastest population observed by SPT/h-MSD, representing approximately 25 % of RNAP II mobile particles analyzed, displayed a mean diffusion coefficient of 0.5$\mu m^2/s$ (Tr=0.63 $ms/\mu m^3$), and an anomalous coefficient very close to 1 (mean α=0.95), which reflects fairly well a free Brownian diffusion. This suggests that free RNAP II explores the nucleus in a stochastic and rather slow manner, that may derive from a certain level of viscosity in the regions scanned by RNAP II, although this “viscosity” would not be sufficient to trigger an anomalous behavior. This value is in the same order of magnitude as those found for the free diffusion of transcription factors (Jankevics et al., 2005; Mazza et al., 2012), but it is an order of magnitude lower than the measurements of GFP multimer diffusion in the nucleus (Bancaud et al., 2009). This could indicate that free RPB1 and transcription factors have a slowed down Brownian diffusion in the nucleoplasm.

The second mobile population that we identified with SPT/h-MSD had a mean diffusion coefficient of 0.02 $\mu m^2/s$ (Tr=266 $ms/\mu m^3$), and a mean anomalous coefficient of 0.31, and represented 75 % of trajectories analyzed by h-MSD. This RNAP II subpopulation thus displayed a highly constrained diffusion and a very slow motion, even if it was distinct from bound immobile RNAP II previously described (Tr≥600 $ms/\mu m^3$). When we analyzed RPB1 diffusion thanks to FCS, we obtained mean diffusion parameters very close to those that we previously observed by FCS for CYCT1 (Bidaux et al., 2018; Furlan et al., 2019), which suggests that RPB1 measurements may correspond to RNAP II molecules involved in the transcription pause release process. FCS, which is more sensitive to high-frequency fast diffusions, hardly detected the subpopulations analyzed in SPT, with most molecules displaying residence times 2 to 4 orders of magnitude higher than those identified by FCS, thus highlighting the complementarity of the two approaches.

The FCS data histogram distribution was imperfectly fit by Gaussian functions, and displayed an array of relatively fast-diffusing molecules with a continuous range of diffusion parameters. We assume that within this range, there are no strict categories of RPB1 behaviors but rather transitions between states of interactions progressively evolving. Such behaviors may refer to the recent concepts that multivalent and dynamic interactions can drive phase separation (Lu et al., 2018) and that the function of large complexes could be maintained while molecules are constantly exchanged (Darzacq and Tjian, 2022). We found it interesting to notice within the scatter plot of parameters that confinement was not necessarily linked to low diffusion coefficient values.

FCS also informed us of the molecular concentration, and indicated to us the presence in average of 2.6 +/-1.4 RPB1 molecules per focal volume, i.e. approximately 14 molecules per μm^3. With respect to the previous estimation of 80,200 RNAP II molecules per nucleus in U2OS cells, i.e. 37 molecules per μm^3 (Zhao et al., 2014), this would mean that FCS detected almost one-third of RPB1 molecules. Although we must be cautious with this extrapolation of data, this would mean that almost two-thirds of the remaining RNAP II molecules, 23 out of 37 molecules per femtoliter, could be observed by SPT. Within this category, 20% of SPT trajectories (i.e. 13 % of total RNAP II) were immobile (7-8 molecules/μm^3), which stands in agreement with the estimation of RNAP II being bound on chromatin (Cisse et al., 2013). A similar proportion of RNAP II molecules detected by SPT could be considered free, according to their anomalous coefficient α=0.95, in agreement again with the literature, including recent FRAP measurements concluding on a minority fraction of freely diffusing polymerases (Steurer et al., 2018). Taken together, these data point out a global view of a majority of RNAP II molecules that are strongly confined, and potentially involved in regulatory steps. Interestingly, the proportion of RPB1 molecules detected by FCS and involved in a confined (mean α=0.61) yet fast sampling corresponds to that of genes estimated as under the control of the transcription pause (Zhou et al., 2012).

About SPT analysis, Cissé et al previously used an original method to measure the dwell times of individual RNAP II molecules in foci, and revealed dynamic RNAP II foci, with an average lifetime of 5s that was increased when inhibiting promoter escape, indicating that such structures probably refer to pre-assembly steps (Cho et al., 2016; Cisse et al., 2013). The RNAP II subpopulation that we identified as diffusing with a mean anomalous coefficient α=0.31 could correspond to such steps of RNAP II trapped for transcription initiation. Interestingly, another spatiotemporal clustering analysis described foci predominantly occupied by a single RNAP II at the same time point (Zhao et al., 2014), and this was also recently confirmed with an investigation of RNA nanodomains that evidenced small RNAP II complexes mainly displaying only one RNAP II molecule (Castells-Garcia et al., 2022). The absence of regions of high RNAP II concentration in our observations is thus completely consistent with these highlights. Still, we observed in our SPT movies subdomains oversampled by RNAPII molecules, corresponding, probably, to confinement domains with relatively fast-moving molecules. We can also assume that the FCS measurements displaying sub-diffusive behavior (mean α=0.61) are consistent with a model of RNAP II relatively confined to regions of potential transcription by transient associations, similar to balls between bumpers in a pinball machine. It is interesting

to note that we previously highlighted that CYCT1, a cofactor of the transcriptional pause release, displays very similar diffusion properties, in subdomains whose size is in the same order of magnitude as that of TAD (Topologically Associated Domains), as described in the literature (Bidaux et al., 2018; Pombo and Dillon, 2015). Moreover, CYCT1 was also described to oversample the nuclear environment (Izeddin et al., 2014). Such a link between confinement, topological frameworks and recruitment/regulation steps in transcription is emerging in the field (Cardozo Gizzi, 2021) and starts to be described at the molecular level (Lee et al., 2022).

Taken together, our results thus reinforce the arguments in favor of the "spatio-temporal foci" concept. The proper execution of RNAP II activity would rely on the increased probability of RPB1 to visit subdomains (oversampling), without necessarily increasing its concentration in a specific site. Interestingly, the study of Herpes Simplex Virus replication recently evidenced the possibility for RNAPII to visit in a repeated manner nearby DNA binding sites without clear boundaries between territories, which may be a common process (McSwiggen et al., 2019). In that frame, highly dynamic interactions would ensure the transcription complex activity is maintained despite the exchange of molecules (Darzacq and Tjian, 2022).

Altogether, our combined experiments suggest an RNAP II dynamic model with at least four kinetic subpopulations: a) chromatin-bound RNAP II, b) RNAP II slowly-diffusing in a Brownian manner, and low levels of confined RNAP II, c) fast-diffusing oversampling RNAP II molecules, and d) highly confined RNAP II. Even if we limited our SPT analyses of diffusing particles to two main subpopulations, the Bayesian information criterion (BIC) and Akaike information criterion (AIC) inflection points strongly suggest the existence of other subpopulations. Future work will be aimed at elucidating whether we can identify and investigate these minority subpopulations to refine our knowledge of the transcription process.

At this stage, two hypotheses can be envisioned about RNAP II mobilities and oversampling: on one hand, a model with three transition states for each molecule, corresponding to successive stages of regulation (increased confinement), typically in domains such as TAD; and on the other hand, a model with two co-existing distinct mechanisms corresponding to the regulation of different genes with two distinct mechanisms of confinement (e.g. transcription pause release vs others) (Fig.7). This study paves the way for future work that will be aimed at investigating which of these two hypotheses correspond to the reality of RNAP II recruitment to active genes, and at identifying the molecular mechanisms responsible for RNAP II confinements. In that frame, the coupling of FCS and SPT acquisitions on the same instrument would represent an interesting instrumental development to simultaneously acquire molecules by both methods in a spatio-temporal manner, and we are moving forward in that direction.

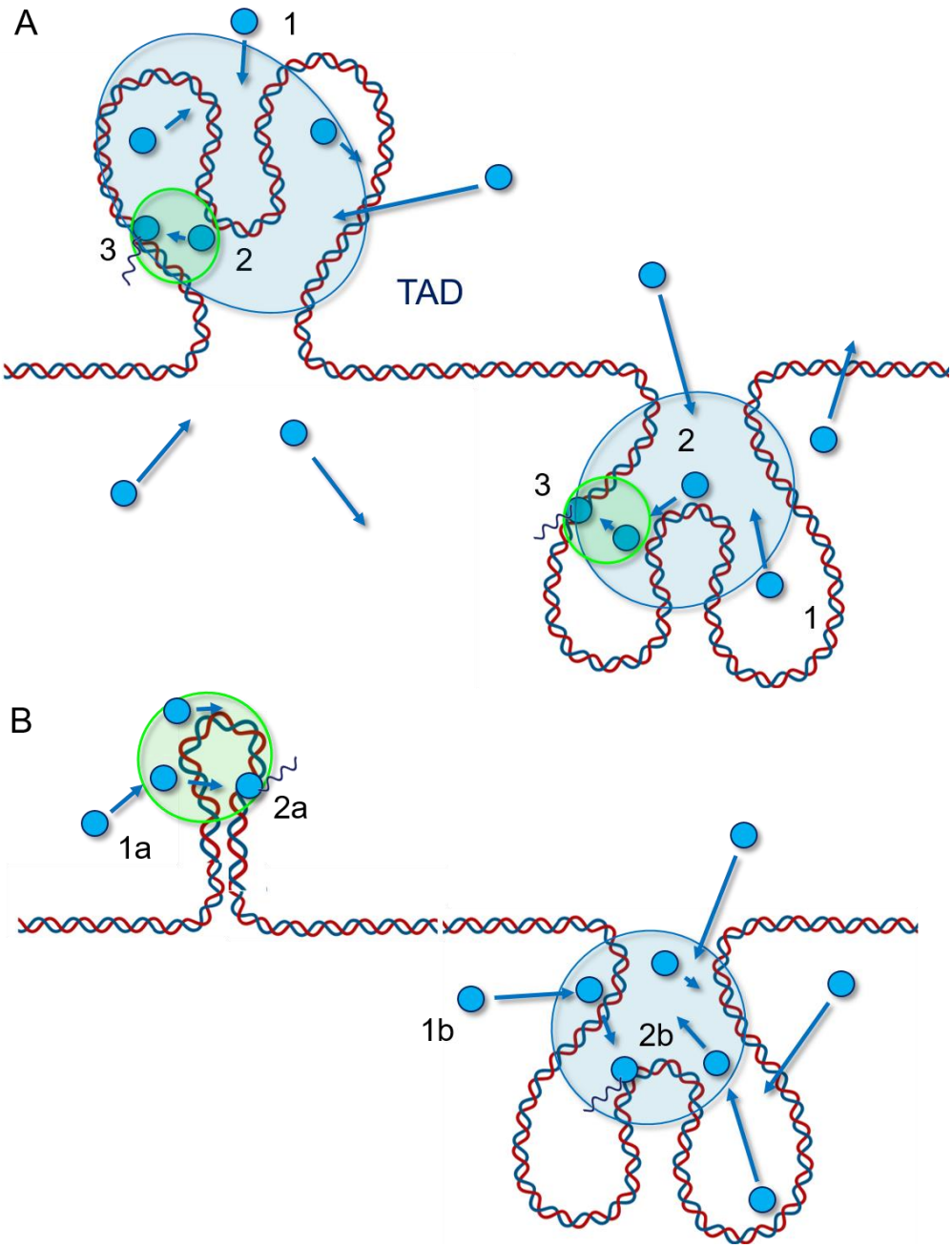


Figure 7. Models for RNAP II recruitment. This scheme displays the two hypotheses proposed to explain RNAP II mobilities and oversampling. The blue disks and arrows symbolize RNAP II molecules and their displacement vectors, respectively. A. Model with successive steps: 1) A free RNAP II (α=0.95) enters a large domain (blue), potentially a TAD, where it is confined (α=0.61); 2) it is then trapped in a subdomain (green) of higher confinement (α= 0.31), possibly through interactions with transcription factors, and binds to chromatin, 3) before being stabilized and starting transcription (fixed RNAP II).

B. Model with distinct confinement domains: 1a) and b) A free RNAPII (α=0.95) can enter two types of confined domains, leading to fast oversampling or not; 2a) and b) RNAPII is then recruited and binds to chromatin (fixed RNAPII). This second model helps to account for the existence of different mechanisms of RNAPII recruitment and in particular the existence of the proximal transcriptional pause for 40% of genes.

382

MATERIAL AND METHODS

Reagents and sample preparations

Standard samples preparation

Water/Glycerol solutions were prepared considering weight percentage with the use of a precision balance. Solutions were vortexed for 30 sec before being stored at 37°C. A few minutes before acquisitions, to avoid agglomerations of beads, we diluted our fluorescent beads stocks in two distinct manners:

- 100 nm diameter beads (Molecular Probes, Fluorescent microspheres (505/515), 2 % solid) diluted 10^4 in solutions for SPT acquisition
- 40 nm diameter beads (Molecular Probes, Fluorescent microspheres (505/515), 2 % solid) diluted 10^2 in solutions for FCS acquisition

To compare the SPT and FCS workflows, we applied a similar strategy. However, since the experimental focal volume in FCS was in the order of 200 nm diameter in the x-y axis, the use of 100nm-diameter beads resulted in huge bursts of fluorescence during signal acquisition, which were misinterpreted during the Autocorrelation Function (ACF) process. We therefore chose smaller fluorescent beads (40nm-diameter) for FCS experiments, which allowed us to measure more reliably fluorescence fluctuations and infer their diffusion coefficients.

Bead solutions were finally deposited in 35-mm glass-bottom dishes (Ibidi, Clinisciences, France) and kept at 37° in a thermostatic chamber during acquisitions.

The theoretical values of beads' diffusion were calculated according to the Stokes-Einstein formula:

$$D = \frac{kT}{6\pi r\mu}$$

With *k* Boltzmann constant, T temperature in Kelvin, r the hydrodynamical radius of the bead and μ the viscosity.

Viscosity values of different (w:g) solutions were calculated following Takamura et al. (Takamura et al., 2012) and are indexed in the following table.

% glycerol	Calculated D value (μm²/s) for 40 nm beads	Calculated D value (μm²/s) for 100 nm beads
0	15.1	5,98
10	12.1	4,78
20	9.05	3,58
50	3.24	1,25
80	0.4	0,19

Biological resources

Cell line culture

Human U2OS osteosarcoma cells (HTB-96 accession number from the American Type Culture Collection, Manassas, VA) were tested for mycoplasma contamination before establishing the RPB1 cell lines. Cells were cultured in DMEM medium (Gibco Laboratories, Gaithersburg, MD) supplemented with 10% FCS (V/V) and penicillin/streptomycin (100 mg/mL). Cells were grown in a humidified incubator at 37°C / 5 % CO2. Cells were passaged every 2–4 days before reaching confluency.

We generated plasmids encoding for human RPB1 (N792D) α-amanitin-resistant mutated fused to an N-terminal Halo tag, by molecular cloning.

We stably transfected U2OS cells with this plasmid using Fugene 6 following the manufacturer's instruction (Promega). α-Amanitin (SIGMA #A2263) was used to select transfected cells, at a concentration of 2 µg/mL and was used thereafter in permanence in the culture of the cells at a concentration of 1µg/mL to avoid endogenous RPB1 re-expression as described in (Boehning et al. 2018). For live-cell imaging, cells were plated on 35-mm glass-bottom dishes (Ibidi, CliniSciences, Nanterre, France), filled with L-15 medium without Phenol Red (Life Technologies, Carlsbad, CA), and incubated at 37°C in a microscope thermostatic chamber (OKO, Italia) for imaging.

Labeling protocol

We labeled cells just before imaging. Following HaloTag technology protocol (Promega), cells were incubated for 30 min in DMEM 10%FCS supplemented with 0.1 nM JF 549 for SPT. Then, cells were rinsed 3 times with DMEM + 10% SVF and were re-incubated for 30 min. Finally, the cell medium was replaced by L-15 medium (Life Technologies, CA, USA) before imaging, to avoid DMEM auto-fluorescence.

SPT set up

SPT measurements were performed with an inverted microscope (Ti-Eclipse, Nikon Instrument, Japan) adapted with an x100 NA=1,49 TIRF water immersion objective and an appropriate dichroic 405/488/561/635 nm (ref Di03-R405/488/561/635 SEMROCK, USA). The excitation laser (488nm Argon LASER, Melles-Griot, USA and 561nm, OXXIUS, France) was injected into the TIRF Nikon module into a mono-mode fiber and focused on the backplane of the objective. Illumination was performed in HILO mode (Tokunaga et al., 2008) to improve contrast in images. Perfect Focus System (PFS, Nikon) was set to avoid drift on Z-axis (defocusing) on the objective, relative to the coverslip. Experiments were acquired under continuous illumination (488 nm 20 kW/cm² on the sample, or 561 nm 20 kW/cm² on the sample). Emitted fluorescence was harvested on an sCMOS camera (PCO EDGE 4.2, PCO, Germany) with a pixel size of 6,5 µm. The acquisition format was 100x100 px, and the acquisition rate was set at 100 Hz (10 ms per frame). All devices were driven by NIS software (Nikon Instruments, Japan). A thermostatic chamber ('on-stage-incubator', OKOLAB, Italy) was used to keep samples at 37 Celsius degrees.

FCS set up

FCS measurements were performed on a Nikon A1 confocal microscope (Nikon Instrument, Tokyo, Japan) with a 40x objective NA=1,25 (Nikon Instrument, Tokyo, Japan). Excitation Argon LASER 488 nm (Coherent, USA) was used to scan samples in confocal illumination. To spectrally select the excitation from the emitted light, we used a bandpass dichroic mirror (405/488 nm). Fluorescence photons were collected on a photodiode and images were reconstructed with NIS software (Nikon Instruments, Tokyo, Japan).

To perform FCS, the CLSM was adapted with a TCSPC module (PicoQuant, Germany), and a pulsed laser at 488 nm (LDH-D-C 488, PicoQuant), a TSCPC unit, and a single molecule counting detector (SPCM CD 3516H, Excelitas Technologies Corp., USA). To drive this photon counting system, we used SymphoTime software (PicoQuant, Germany) in FCS mode. A high pass filter (BLP-488R, SEMROCK, USA) was set before the detector.

SPT data analysis:

Tracks were reconstructed (detection and tracking) with MTT algorithm tracking software (Sergé et al., 2008) running on MATLAB (MathWorks), which is well adapted for a high concentration of particles (up to 10 for a 100x100px field of view). For each condition, we had at least 1000 trajectories lasting more than 10 detections (smaller trajectories were not included in the analysis). Tracks were then analyzed with custom code kindly provided by Laura Caccianini (Coppey/Hajj lab, Institut Curie, France). H-MSD (histogram Mean Square Displacement) technique is based on the Time-Average MSD calculation $TA-MSD(\tau) = \langle r^2(\tau)\rangle_T = \langle [x(t+\tau)-x(t)]^2 + [y(t+\tau)-y(t)]^2\rangle_T \propto 4D\tau^{\alpha}$ where Δt , with n the number of frames considered for the sliding average (0<n<L_{traj}, $L_{traj}$the length of the trajectory), Δt the lag time between two frames (10 ms), D the diffusion coefficient (µm²/s), and α is the anomaly coefficient.

With H-MSD each track was treated individually as described in (Etoc et al., 2018; Normanno et al., 2012). The TA-MSD curves were fitted with a linear fit until $\tau = 60\,ms$ to obtain the instantaneous diffusion coefficient (D_{60ms}) following the previous works of Michalet et al. describing a procedure to analyze the individual MSD curves (Michalet, 2010). In accordance with this work, the first point MSD (1Δt) is not considered in the fit because it is the same order of magnitude as the static localization error (50 nm, calculated on fixed RPB1-Halo in U2OS cells.

Anomaly coefficient alpha was calculated by fitting the log-log representation of the TA-MSD with a power-law function TA-MSD(t)∝ $4Dt^{\alpha}$. For individual tracks, it is common that the second half of the MSD curve is very noisy because of the low statistics for each MSD(nΔt) value. We decided to consider the first third of the curve for the linear fit because more than the second half of the MSD curve is noisy and cannot be interpreted. The slope of the regression line gave alpha. Another D could also be computed from the intercept of the slope, D_{α} , and was compared to D_{60ms} in the results.

A Gaussian mixture model (GMM) was applied to fit the data. GMM was used as a clustering algorithm where the number k of populations is chosen from an analysis of the Bayesian Information Criterion and the Akaike Information Criterion. This population number was chosen as small as possible, following those criteria. By evaluating the number of discernible populations with a statistical criterion we tried to minimize any bias that can occur when using such a supervised machine learning algorithm and create a robust and reproducible analysis pipeline.

To be able to compare the diffusion coefficients calculated by SPT and FCS, we converted the SPT diffusion coefficient D_{60ms} into residence time per cubic micrometer (ms/ µm^3). We assumed that the diffusion measurements of the molecules were isotropic in the nucleus and that we could therefore extrapolate the exploration of the surface (µm^2) to that of a volume (µm^3).

For the two sub-populations studies, we mixed batches of trajectories coming from different conditions. The different conditions and the different fractions for each mix are indexed in the following tables.

Condition A		Total number of trajectories	Nb of D1 trajectories	% D1	Nb of D2 trajectories	% D2
D1=4 µm²/s (w :g =100 :0) & D2=0.1 µm²/s (w :g=20 :80)	m1	2013	845	42	1168	58
	m2	3118	2339	75	780	25
	m3	3369	842	25	2527	75

Condition B		Total number of trajectories	Nb of D1 trajectories	% D1	Nb of D2 trajectories	% D2
D1=0.9 µm²/s (w :g =50 :50) & D2=0.1 µm²/s (w :g=20 :80)	m1	2027	932	46	1095	54
	m2	3207	1860	58	1347	42
	m3	3633	908	25	2725	75

Condition C		Total number of trajectories	Nb of D1 trajectories	% D1	Nb of D2 trajectories	% D2
D1=4.0 µm²/s (w :g =100 :0) & D2=0.9 µm²/s (w :g=50 :50)	m1	4470	894	20	3576	80
	m2	4791	2443	51	2348	49
	m3	5119	3993	78	1126	22

Condition D		Total number of trajectories	Nb of D1 trajectories	% D1	Nb of D2 trajectories	% D2
D1=4.0 µm²/s (w :g =100 :0) & D2=3.8 µm²/s (w :g=90 :10)	m1	5293	688	13	4605	87
	m2	5310	2443	46	2867	54
	m3	5398	4049	75	1350	25

FCS analysis:

FCS setup was calibrated with ATTO 488 diluted at 10 nM in Milli-Q water to define the focal volume by best practice. In the water at 36°, the temperature of our incubator, the Atto 488 diffusion coefficient is 536 µm²/s (source: PicoQuant Application Note Kapusta and GmbH, n.d.).

The number N of molecules in this observation volume V was inferred from the autocorrelation function with

$$N = \frac{1}{G(0)}$$

And the observation volume determined with:

$$V_{eff} = \frac{1}{G(0).N_A.c}$$

N_A is the Avogadro number and c is the concentration of the calibration solution.

Once calibration was done, we had a good estimate of the focal volume of the microscope and could proceed to the analysis of fluorescence fluctuations. Measurements displaying a linear regression of the fluorescent signal characteristic of photobleaching were excluded. For data analysis, we use a home-written MATLAB code named EASY FCS (Furlan et al., 2019) which can fit the autocorrelation function (ACF).

In the case of normal diffusion the ACFs were adjusted with a Gaussian function (assuming the beam as a Gaussian profile) :

$$G(\tau) = \frac{1}{N}\left(\frac{1}{1+\left(\frac{\tau}{\tau_D}\right)}\right)\left(\frac{1}{\sqrt{1+\frac{w_0^2}{z_0^2}\left(\frac{\tau}{\tau_D}\right)}}\right)$$

Where *N* is the number of molecules and τ_D is the characteristic residence time.

For anomalous diffusion, the mean square displacement is no longer linearly dependent on time, but it evolves proportionally with time to the power of α. In this case, the autocorrelation function becomes:

$$G(\tau) = \frac{1}{N}\left(\frac{1}{1+\left(\frac{\tau}{\tau_D}\right)^{\alpha}}\right)\left(\frac{1}{\sqrt{1+\frac{w_0^2}{z_0^2}\left(\frac{\tau}{\tau_D}\right)^{\alpha}}}\right) \tag{4}$$

The fitting of the ACFs was done with the Trust-Region-Reflective Least Squares Algorithm. To evaluate the goodness of the fit, we minimized the sum of the squared error.

Since the diffusion coefficient (D) obtained by anomalous model fit is given with time dimensions dependent on exponent α, it is recommended to give the results in residence time (Bidaux et al., 2018; Mazza et al., 2012; Stasevich et al., 2010) to compare data. To be able to compare FCS and SPT data from different experiments with different focal volumes, it is just necessary to normalize the Residence Time (Tr) by μm^3.

AUTHOR CONTRIBUTIONS

MF, PL, GB, AF, and LH designed research. MF, PL, DC, and FA performed experiments. MF, PL, AL, GB, AF, and LH analyzed and interpreted the data. MF, PL, AF, and LH wrote the manuscript.

DATA AND MATERIAL AVAILABILITY

Data and material are available on request to laurent.heliot@univ-lille.fr

SUPPLEMENTARY DATA

Supplementary Movies are available at:
https://nextcloud.univ-lille.fr/index.php/s/KjXnoYcG26AnFXL
https://nextcloud.univ-lille.fr/index.php/s/HQW5rXBSg78JHMw
https://nextcloud.univ-lille.fr/index.php/s/eFTsxyXaPnApE9C

ACKNOWLEDGEMENTS

We thank all members of our team and GDR Imabio-CNRS, particularly Hugues Berry, for their help and fruitful discussion/comments. We are much grateful to Raoul Torero-Ibad for his great work and a precious help in providing us with a functional lab environment, and for his enlightened opinion on this study. We are much grateful to Jonathan Grimm and Luke Lavis (Janelia Farm, VA, USA) for kindly providing us with JF dyes to perform our experiments.

FUNDING

This work was supported by CNRS and Ministerial Funding; Agence Nationale de la Recherche (Dynam-ic-12-BSV5-0018-02 and ABC4M- ANR-20-CE45-0023); the LABEX CEMPI (ANR-11-LABX-0007), I-SITE ULNE (ANR-16-IDEX-0004) as well as by the Ministry of Higher Education and Research, Hauts de France council and European Regional Development Fund (ERDF) through the Contrat de Projets Etat-Region (CPER) Photonics for Society (P4S), Region «Hauts de France» I-PRIMER, CNRS and GDR-ImaBio Interdisciplinary Master's Scholarship and Nikon partnership agreement.

Funding for open access charge: ANR-20-CE45-0023

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

REFERENCES

Altan-Bonnet, N., Altan-Bonnet, G., 2009. Fluorescence Correlation Spectroscopy in Living Cells: A Practical Approach. Curr. Protoc. Cell Biol. 45, 4.24.1-4.24.14. https://doi.org/10.1002/0471143030.cb0424s45

Bancaud, A., Huet, S., Daigle, N., Mozziconacci, J., Beaudouin, J., Ellenberg, J., 2009. Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. EMBO J. 28, 3785–3798. https://doi.org/10.1038/emboj.2009.340

Banks, D.S., Fradin, C., 2005. Anomalous Diffusion of Proteins Due to Molecular Crowding. Biophys. J. 89, 2960–2971. https://doi.org/10.1529/biophysj.104.051078

Banks, D.S., Tressler, C., Peters, R.D., Höfling, F., Fradin, C., 2016. Characterizing anomalous diffusion in crowded polymer solutions and gels over five decades in time with variable-lengthscale fluorescence correlation spectroscopy. Soft Matter 12, 4190–4203. https://doi.org/10.1039/C5SM01213A

Bidaux, G., Le Nézet, C., Pisfil, M.G., Henry, M., Furlan, A., Bensaude, O., Vandenbunder, B., Héliot, L., 2018. FRET Image Correlation Spectroscopy Reveals RNAPII-Independent P-TEFb Recruitment on Chromatin. Biophys. J. 114, 522–533. https://doi.org/10.1016/j.bpj.2017.11.3783

Cardozo Gizzi, A.M., 2021. A Shift in Paradigms: Spatial Genomics Approaches to Reveal Single-Cell Principles of Genome Organization. Front. Genet. 12, 780822. https://doi.org/10.3389/fgene.2021.780822

Castells-Garcia, A., Ed-daoui, I., González-Almela, E., Vicario, C., Ottestrom, J., Lakadamyali, M., Neguembor, M.V., Cosma, M.P., 2022. Super resolution microscopy reveals how elongating RNA polymerase II and nascent RNA interact with nucleosome clutches. Nucleic Acids Res. 50, 175–190. https://doi.org/10.1093/nar/gkab1215

Chakalova, L., Fraser, P., 2010. Organization of Transcription. Cold Spring Harb. Perspect. Biol. 2, a000729. https://doi.org/10.1101/cshperspect.a000729

Cheezum, M.K., Walker, W.F., Guilford, W.H., 2001. Quantitative comparison of algorithms for tracking single fluorescent particles. Biophys. J. 81, 2378–2388. https://doi.org/10.1016/S0006-3495(01)75884-5

Cho, W.-K., Jayanth, N., English, B.P., Inoue, T., Andrews, J.O., Conway, W., Grimm, J.B., Spille, J.-H., Lavis, L.D., Lionnet, T., Cisse, I.I., 2016. RNA Polymerase II cluster dynamics predict mRNA output in living cells. eLife 5. https://doi.org/10.7554/eLife.13617

Cisse, I.I., Izeddin, I., Causse, S.Z., Boudarene, L., Senecal, A., Muresan, L., Dugast-Darzacq, C., Hajj, B., Dahan, M., Darzacq, X., 2013. Real-Time Dynamics of RNA Polymerase II Clustering in Live Human Cells. Science 341, 664–667. https://doi.org/10.1126/science.1239053

Darzacq, X., Shav-Tal, Y., de Turris, V., Brody, Y., Shenoy, S.M., Phair, R.D., Singer, R.H., 2007. In vivo dynamics of RNA polymerase II transcription. Nat. Struct. Mol. Biol. 14, 796–806. https://doi.org/10.1038/nsmb1280

Darzacq, X., Tjian, R., 2022. Weak multivalent biomolecular interactions: a strength versus numbers tug of war with implications for phase partitioning. RNA N. Y. N 28, 48–51. https://doi.org/10.1261/rna.079004.121

Elson, E.L., 2011. Fluorescence correlation spectroscopy: past, present, future. Biophys. J. 101, 2855–2870. https://doi.org/10.1016/j.bpj.2011.11.012

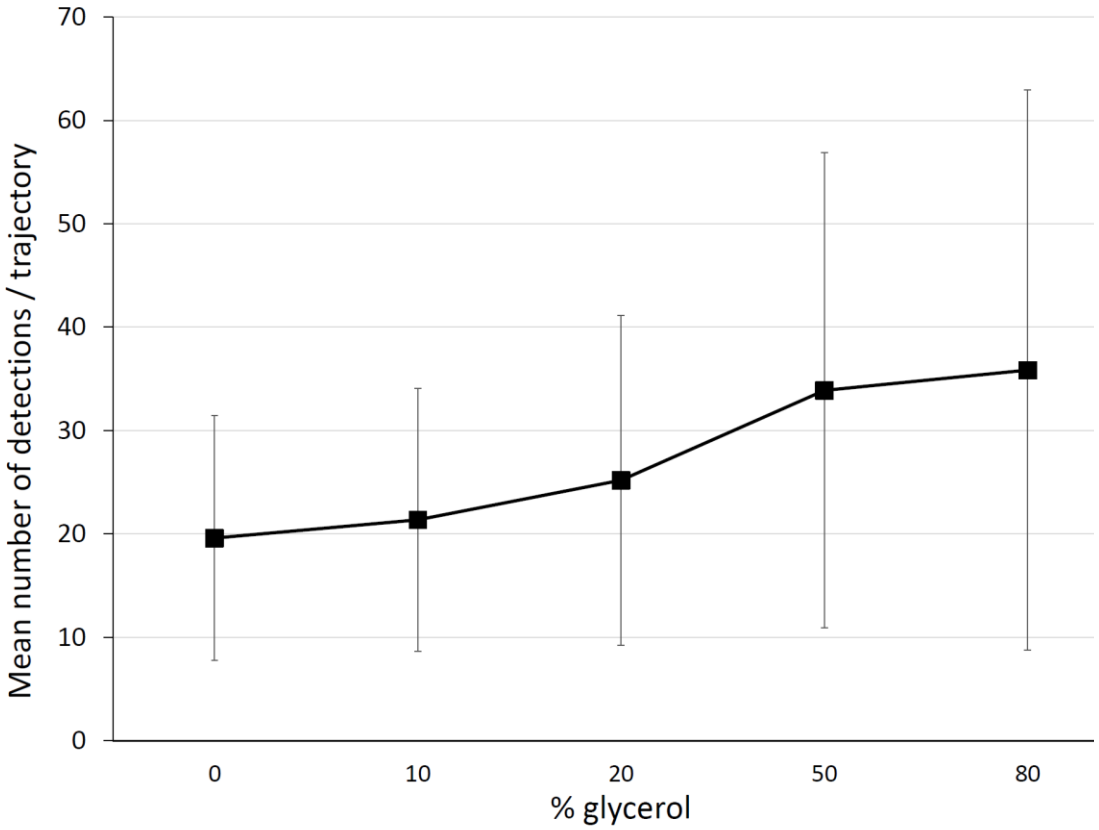
Etoc, F., Balloul, E., Vicario, C., Normanno, D., Liße, D., Sittner, A., Piehler, J., Dahan, M., Coppey, M., 2018. Non-specific interactions govern cytosolic diffusion of nanosized objects in mammalian cells. Nat. Mater. 17, 740–746. https://doi.org/10.1038/s41563-018-0120-7

Furlan, A., Gonzalez-Pisfil, M., Leray, A., Champelovier, D., Henry, M., Le Nézet, C., Bensaude, O., Lefranc, M., Wohland, T., Vandenbunder, B., Bidaux, G., Héliot, L., 2019. HEXIM1 Diffusion in the Nucleus Is Regulated by Its Interactions with Both 7SK and P-TEFb. Biophys. J. 117, 1615–1625. https://doi.org/10.1016/j.bpj.2019.09.019

591 Grimm, J.B., English, B.P., Chen, J., Slaughter, J.P., Zhang, Z., Revyakin, A., Patel, R., Macklin, J.J.,
592 Normanno, D., Singer, R.H., Lionnet, T., Lavis, L.D., 2015. A general method to improve fluorophores
593 for live-cell and single-molecule microscopy. Nat. Methods 12, 244–250, 3 p following 250.
594 https://doi.org/10.1038/nmeth.3256
595 Grimm, J.B., Muthusamy, A.K., Liang, Y., Brown, T.A., Lemon, W.C., Patel, R., Lu, R., Macklin, J.J., Keller,
596 P.J., Ji, N., Lavis, L.D., 2017. A general method to fine-tune fluorophores for live-cell and in vivo
597 imaging. Nat. Methods 14, 987–994. https://doi.org/10.1038/nmeth.4403
598 Grünwald, D., Martin, R.M., Buschmann, V., Bazett-Jones, D.P., Leonhardt, H., Kubitscheck, U., Cardoso,
599 M.C., 2008. Probing Intranuclear Environments at the Single-Molecule Level. Biophys. J. 94, 2847–
600 2858. https://doi.org/10.1529/biophysj.107.115014
601 Hager, G.L., McNally, J.G., Misteli, T., 2009. Transcription Dynamics. Mol. Cell 35, 741–753.
602 https://doi.org/10.1016/j.molcel.2009.09.005
603 Halford, S.E., Marko, J.F., 2004. How do site-specific DNA-binding proteins find their targets? Nucleic
604 Acids Res. 32, 3040–3052. https://doi.org/10.1093/nar/gkh624
605 Hansen, A.S., Woringer, M., Grimm, J.B., Lavis, L.D., Tjian, R., Darzacq, X., 2018. Robust model-based
606 analysis of single-particle tracking experiments with Spot-On [WWW Document]. eLife.
607 https://doi.org/10.7554/eLife.33125
608 Harwardt, M.-L.I.E., Dietz, M.S., Heilemann, M., Wohland, T., 2018. SPT and Imaging FCS Provide
609 Complementary Information on the Dynamics of Plasma Membrane Molecules. Biophys. J. 114, 2432–
610 2443. https://doi.org/10.1016/j.bpj.2018.03.013
611 Izeddin, I., Récamier, V., Bosanac, L., Cissé, I.I., Boudarene, L., Dugast-Darzacq, C., Proux, F., Bénichou,
612 O., Voituriez, R., Bensaude, O., Dahan, M., Darzacq, X., 2014. Single-molecule tracking in live cells
613 reveals distinct target-search strategies of transcription factors in the nucleus. eLife 3.
614 https://doi.org/10.7554/eLife.02230
615 Jackson, D.A., Hassan, A.B., Errington, R.J., Cook, P.R., 1993. Visualization of focal sites of transcription
616 within human nuclei. EMBO J. 12, 1059–1065.
617 Jankevics, H., Prummer, M., Izewska, P., Pick, H., Leufgen, K., Vogel, H., 2005. Diffusion-Time Distribution
618 Analysis Reveals Characteristic Ligand-Dependent Interaction Patterns of Nuclear Receptors in Living
619 Cells. Biochemistry 44, 11676–11683. https://doi.org/10.1021/bi050744v
620 Kapusta, P., GmbH, P., n.d. Absolute Diffusion Coefficients: Compilation of Reference Data for FCS
621 Calibration 2.
622 Kim, T.-K., Lagrange, T., Wang, Y.-H., Griffith, J.D., Reinberg, D., Ebright, R.H., 1997. Trajectory of DNA in
623 the RNA polymerase II transcription preinitiation complex. Proc. Natl. Acad. Sci. 94, 12268–12273.
624 https://doi.org/10.1073/pnas.94.23.12268
625 Kornberg, R.D., 2007. The molecular basis of eukaryotic transcription. Proc. Natl. Acad. Sci. 104, 12955–
626 12961. https://doi.org/10.1073/pnas.0704138104
627 Langowski, J., 2008. Protein-protein interactions determined by fluorescence correlation spectroscopy.
628 Methods Cell Biol. 85, 471–484. https://doi.org/10.1016/S0091-679X(08)85020-0
629 Lee, R., Kang, M.-K., Kim, Y.-J., Yang, B., Shim, H., Kim, S., Kim, K., Yang, C.M., Min, B., Jung, W.-J., Lee, E.-
630 C., Joo, J.-S., Park, G., Cho, W.-K., Kim, H.-P., 2022. CTCF-mediated chromatin looping provides a
631 topological framework for the formation of phase-separated transcriptional condensates. Nucleic
632 Acids Res. 50, 207–226. https://doi.org/10.1093/nar/gkab1242
633 Lee, T.I., Young, R.A., 2000. Transcription of eukaryotic protein-coding genes. Annu. Rev. Genet. 34, 77–
634 137. https://doi.org/10.1146/annurev.genet.34.1.77
635 Lu, H., Yu, D., Hansen, A.S., Ganguly, S., Liu, R., Heckert, A., Darzacq, X., Zhou, Q., 2018. Phase-separation
636 mechanism for C-terminal hyperphosphorylation of RNA polymerase II. Nature 558, 318–323.
637 https://doi.org/10.1038/s41586-018-0174-3

638 Magde, D., Elson, E., Webb, W.W., 1972. Thermodynamic Fluctuations in a Reacting System---
639 Measurement by Fluorescence Correlation Spectroscopy. Phys. Rev. Lett. 29, 705–708.
640 https://doi.org/10.1103/PhysRevLett.29.705
641 Manley, S., Gillette, J.M., Patterson, G.H., Shroff, H., Hess, H.F., Betzig, E., Lippincott-Schwartz, J., 2008.
642 High-density mapping of single-molecule trajectories with photoactivated localization microscopy.
643 Nat. Methods 5, 155–157. https://doi.org/10.1038/nmeth.1176
644 Mazza, D., Abernathy, A., Golob, N., Morisaki, T., McNally, J.G., 2012. A benchmark for chromatin binding
645 measurements in live cells. Nucleic Acids Res. 40, e119–e119. https://doi.org/10.1093/nar/gks701
646 McSwiggen, D.T., Hansen, A.S., Teves, S.S., Marie-Nelly, H., Hao, Y., Heckert, A.B., Umemoto, K.K., Dugast-
647 Darzacq, C., Tjian, R., Darzacq, X., 2019. Evidence for DNA-mediated nuclear compartmentalization
648 distinct from phase separation. eLife 8. https://doi.org/10.7554/eLife.47098
649 Michalet, X., 2010. Mean square displacement analysis of single-particle trajectories with localization
650 error: Brownian motion in an isotropic medium. Phys. Rev. E 82.
651 https://doi.org/10.1103/PhysRevE.82.041914
652 Normanno, D., Dahan, M., Darzacq, X., 2012. Intra-nuclear mobility and target search mechanisms of
653 transcription factors: A single-molecule perspective on gene expression. Biochim. Biophys. Acta BBA
654 - Gene Regul. Mech. 1819, 482–493. https://doi.org/10.1016/j.bbagrm.2012.02.001
655 Pombo, A., Dillon, N., 2015. Three-dimensional genome architecture: players and mechanisms. Nat. Rev.
656 Mol. Cell Biol. 16, 245–257. https://doi.org/10.1038/nrm3965
657 Presman, D.M., Ball, D.A., Paakinaho, V., Grimm, J.B., Lavis, L.D., Karpova, T.S., Hager, G.L., 2017.
658 Quantifying transcription factor binding dynamics at the single-molecule level in live cells. Methods
659 123, 76–88. https://doi.org/10.1016/j.ymeth.2017.03.014
660 Sergé, A., Bertaux, N., Rigneault, H., Marguet, D., 2008. Dynamic multiple-target tracing to probe
661 spatiotemporal cartography of cell membranes. Nat. Methods 5, 687–694.
662 https://doi.org/10.1038/nmeth.1233
663 Stasevich, T.J., Mueller, F., Michelman-Ribeiro, A., Rosales, T., Knutson, J.R., McNally, J.G., 2010. Cross-
664 validating FRAP and FCS to quantify the impact of photobleaching on in vivo binding estimates.
665 Biophys. J. 99, 3093–3101. https://doi.org/10.1016/j.bpj.2010.08.059
666 Steurer, B., Janssens, R.C., Geverts, B., Geijer, M.E., Wienholz, F., Theil, A.F., Chang, J., Dealy, S., Pothof,
667 J., van Cappellen, W.A., Houtsmuller, A.B., Marteijn, J.A., 2018. Live-cell analysis of endogenous GFP-
668 RPB1 uncovers rapid turnover of initiating and promoter-paused RNA Polymerase II. Proc. Natl. Acad.
669 Sci. 115, E4368–E4376. https://doi.org/10.1073/pnas.1717920115
670 Takamura, K., Fischer, H., Morrow, N.R., 2012. Physical properties of aqueous glycerol solutions. J. Pet.
671 Sci. Eng. 98–99, 50–60. https://doi.org/10.1016/j.petrol.2012.09.003
672 Tokunaga, M., Imamoto, N., Sakata-Sogawa, K., 2008. Highly inclined thin illumination enables clear
673 single-molecule imaging in cells. Nat. Methods 5, 159–161. https://doi.org/10.1038/nmeth1171
674 Wachsmuth, M., Waldeck, W., Langowski, J., 2000. Anomalous diffusion of fluorescent probes inside
675 living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. J. Mol. Biol.
676 298, 677–689. https://doi.org/10.1006/jmbi.2000.3692
677 Wieser, S., Schütz, G.J., 2008. Tracking single molecules in the live cell plasma membrane—Do's and
678 Don't's. Methods 46, 131–140. https://doi.org/10.1016/j.ymeth.2008.06.010
679 Zhao, Z.W., Roy, R., Gebhardt, J.C.M., Suter, D.M., Chapman, A.R., Xie, X.S., 2014. Spatial organization of
680 RNA polymerase II inside a mammalian cell nucleus revealed by reflected light-sheet superresolution
681 microscopy. Proc. Natl. Acad. Sci. 111, 681–686.
682 Zhou, Q., Li, T., Price, D.H., 2012. RNA polymerase II elongation control. Annu. Rev. Biochem. 81, 119–
683 143. https://doi.org/10.1146/annurev-biochem-052610-095910
684

SUPPLEMENTAL MATERIAL



Supplemental Figure 1. Mean number +/- standard deviation of detections obtained with Single Particle Tracking of fluorescent beads diffusing in aqueous solutions with increasing glycerol proportions.

A

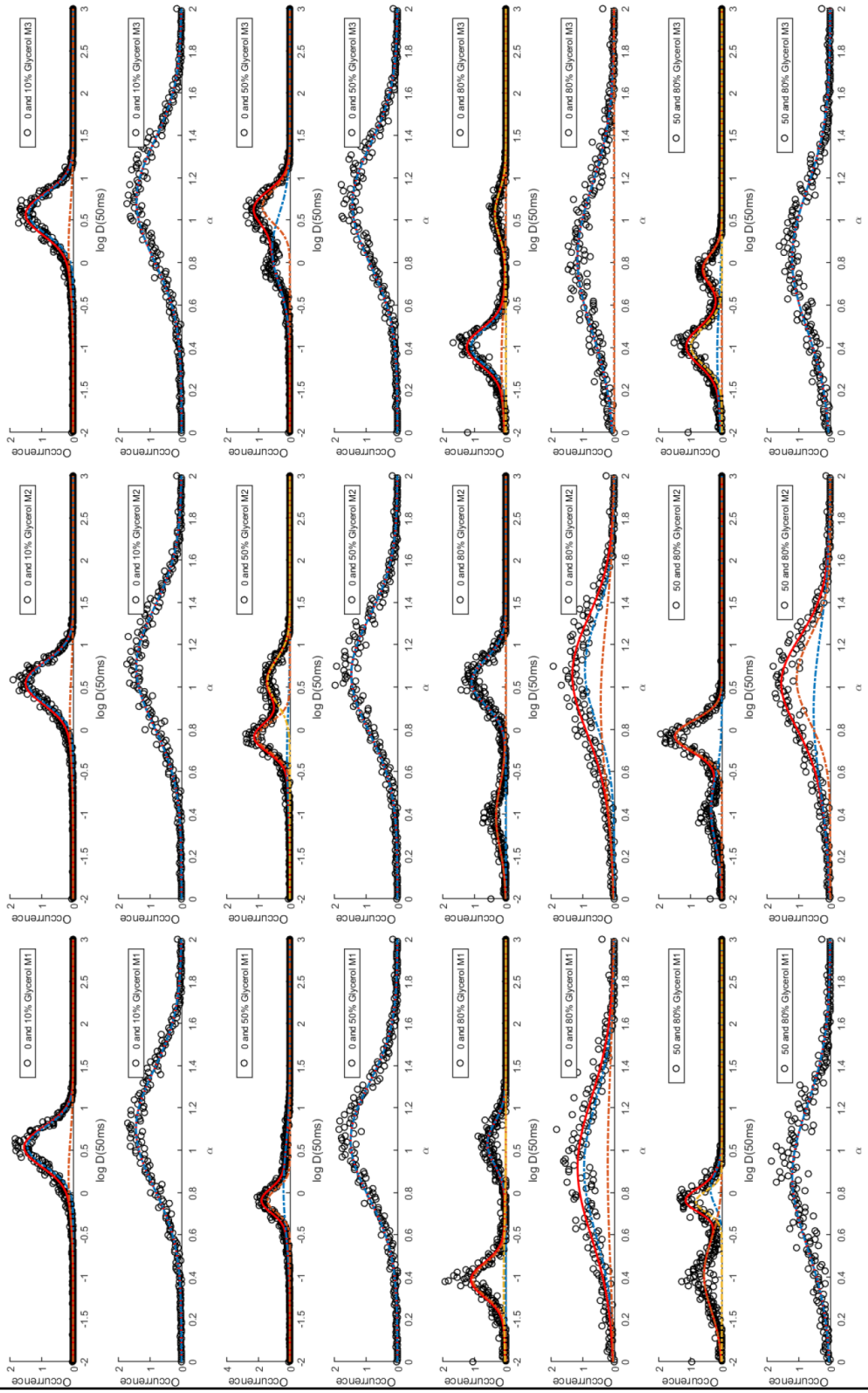
Anomaly coefficient	Experimental values		Theory
% glycerol	mean α	SD	α
0	0.96	0.01	1
10	0.94	0.02	1
20	0.97	0.02	1
50	1.02	0.02	1
80	0.7	0.02	1

B

Diffusion coefficient	Experimental values		Theory
% glycerol	mean D_α (µm²/s)	SD	D (µm²/s)
0	15.5	1.7	15.1
10	12.1	1.8	12.1
20	9.2	1.5	9.0
50	4.5	0.8	3.2
80	1.1	0.2	0.4

686

Supplemental Table 1. Comparison of experimental vs theoretical values of anomaly coefficients (A) and diffusion coefficients (B) computed from the ACF of FCS measurements performed with fluorescent beads diffusing in aqueous solutions with



687 **Supplemental Figure 2**. Representation of subpopulations identified through the GMM (Gaussian Mixture Model) after mixing batches of trajectories obtained in different conditions, as detailed within the graphics legend.

A

		Input	Calculated by H-MSD analysis
Mix 1 (m1)	D1 (µm²/s)	4.0	3.76 (SD=0.65)
	% pop 1	42%	40%
	D2 (µm²/s)	0.1	0.09 (SD=0.02)
	% pop 2	58%	60%
Mix 2 (m2)	D1 (µm²/s)	4.0	3.68 (SD=0.61)
	% pop 1	75%	73%
	D2 (µm²/s)	0.1	0.09 (SD=0.06)
	% pop 2	25%	27%
Mix 3 (m3)	D1 (µm²/s)	4.0	3.76 (SD=0.63)
	% pop 1	25%	24%
	D2 (µm²/s)	0.1	0.09 (SD=0.04)
	% pop 2	75%	76%

B

		Input	Calculated by H-MSD analysis
Mix 1 (m1)	D1 (µm²/s)	0.9	0.88 (SD=0.07)
	% pop 1	46%	43%
	D2(µm²/s)	0.1	0.10 (SD=0.05)
	% pop 2	54%	57%
Mix 2 (m2)	D1 (µm²/s)	0.9	0.84 (SD=0.06)
	% pop 1	58%	66%
	D2 (µm²/s)	0.1	0.11 (SD=0.04)
	% pop 2	42%	34%
Mix 3 (m3)	D1(µm²/s)	0.9	0.90 (SD=0.07)
	% pop 1	25%	23%
	D2 (µm²/s)	0.1	0.10 (SD=0.04)
	% pop 2	75%	77%

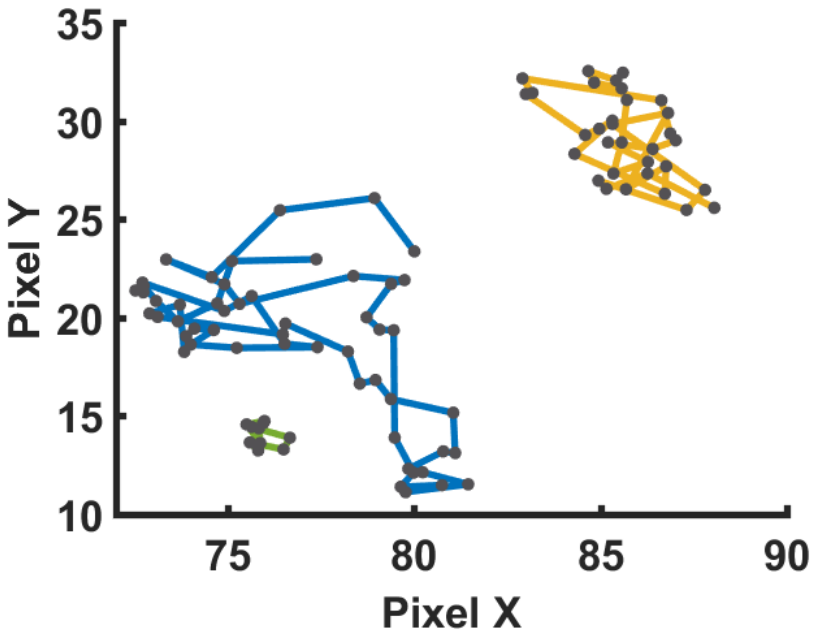
C

		Input	Calculated by H-MSD analysis
Mix 1 (m1)	D1 (µm²/s)	4.0	3.95 (SD=0.93)
	% pop 1	20%	15%
	D2 (µm²/s)	0.9	0.81 (SD=0.04)
	% pop 2	80%	85%
Mix 2 (m2)	D1(µm²/s)	4.0	4.01 (SD=0.45)
	% pop 1	51%	41%
	D2(µm²/s)	0.9	0.82 (SD=0.74)
	% pop 2	49%	59%
Mix 3 (m3)	D1(µm²/s)	4.0	4.10 (SD=0.45)
	% pop 1	78%	67%
	D2 (µm²/s)	0.9	0.87 (SD=0.42)
	% pop 2	22%	33%

D

		Input	Calculated by H-MSD analysis
Mix 1 (m1)	D1 (µm²/s)	4.0	1 population detected D=3.5µm²/s (SD=0.7)
	% pop 1	13%	
	D2 (µm²/s)	3.8	
	% pop 2	87%	
Mix 2 (m2)	D1 (µm²/s)	4.0	1 population detected D=3.6µm²/s (SD=0.9)
	% pop 1	46%	
	D2 (µm²/s)	3.8	
	% pop 2	54%	
Mix 3 (m3)	D1(µm²/s)	4.0	1 population detected D=3.8µm²/s (SD=0.8)
	% pop 1	75%	
	D2 (µm²/s)	3.8	
	% pop 2	25%	

Supplemental Table 2. Comparison of input values vs h-MSD computed values after mixing batches of trajectories obtained in different conditions.



Supplemental Figure 3. Illustrative examples of tracked RPB1 molecules