Chromatin organization, transcription factor dynamics and gene expression in higher eukaryotes with superresolution imaging and single molecule tracking

THÈSE N^{O} 6361 (2015)

PRÉSENTÉE LE 12 JANVIER 2015

À LA FACULTÉ DES SCIENCES DE BASE
LABORATOIRE DE BIOPHYSIQUE EXPÉRIMENTALE
PROGRAMME DOCTORAL EN PHYSIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Aleksandr BENKE

acceptée sur proposition du jury:

Prof. M. Q. Tran, président du jury
Prof. S. Manley, directrice de thèse
Prof. J. Elf, rapporteur
Prof. B. Schuler, rapporteur
Prof. D. M. Suter, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE
Suisse
2015

Abstract

Gene expression is dynamic and heterogeneous across all cell types. It serves several fundamental functions such as adaptation for changing environment and developing tissue specific phenotype and functionality during development and in adult organism. Control of gene expression mostly occurs at the transcription stage. Two main factors that regulate transcription in high eukaryotes are chromatin organization and transcription factor (TF) protein dynamics. Until recently they were mostly studied by methods that are performed in vitro or/and require ensemble averaging. We have been contributing to a development of new single molecule imaging methods that allow quantifying chromatin organization and TF dynamics in vivo and at single cell/single molecule level. Here we present our methodological developments and apply them to study how chromatin organization and TF dynamics regulate transcription.

On the technical side of the study we, first, developed a method of imaging DNA with enhanced resolution in living cells. Super-resolution (SR) point localization based microscopy allows visualizing sub-diffraction organization of cellular structures and their dynamics in vivo. However, until recently almost exclusively proteins were imaged in SR microscopy. We found photoswitching conditions to image with stochastic optical reconstruction microscopy (STORM) DNA directly using the site-specific DNA –binding dye Picogreen (Chapter 2). We achieved a resolution of 50-70 nm which is ~5 fold better than conventional microscopy. Due to the excellent dye preservation we were able to do time lapse SR imaging. This study was a first demonstration of using a site-specific dye for STORM SR imaging in living cells.

Photoswitching principle used in SR microscopy can be applied to single molecule (SM) tracking. It provides up to 1000x increase in density of trajectories compared to classic SM tracking. Initially, photoactivatable (PA) fusion proteins were used for high density SR based tracking. However, PA proteins are relatively dim and there is a limited choice of them. We demonstrated a new approach to SR based high density tracking by using organic dyes' photoswitching to track proteins (Chapter 4). We showed that different dyes can be photoswitched in similar live-cell compatible media on membrane and in different organelles. Together with orthogonal protein labelling schemes this allows us to perform multicolour high density tracking across different compartments. This approach gives flexibility of labelling and increased track length

compared to PA-FP SR based tracking. In a separate study, using tracking we developed algorithm which can account for long-living molecules in live cell SR imaging and allows us to correct clustering artefacts (Chapter 5).

These single molecule imaging methods allow measuring parameters that were previously inaccessible. On the biological side of this study we apply them to study transcription control. First, we study how chromatin organization regulates gene transcription in a model system of HoxD gene cluster. Hox genes are responsible for body plan formation during development. Their expression is highly regulated in time and space during development. It was shown that local chromatin organization influences transcription of Hox genes. Chromatin conformation capture (4C) data suggest that within the HoxD cluster active or passive genes are preferentially contact other active or passive genes respectively and form two sub-compartments. Activation of a gene is then accompanied by its relocation from inactive to active sub-compartment. These data provided the first insight on HoxD spatial organization. However, there is a limit in interpretation of 4C results due to methods' drawbacks: 4C measures spatial proximity indirectly and requires hundreds of thousands cells. Here we use fluorescent in-situ hybridization (FISH) imaging which allows the direct measurement of chromatin spatial organization at the single cell level and combine it with STORM SR microscopy to take advantage of its sub-diffraction resolution. We study the HoxD genes cluster spatial organization in tissues with different HoxD transcription levels and test 4C based model (Chapter 3). We demonstrate that the shape of the HoxD gene cluster is different in different parts of the embryo. HoxD in the forebrain region where the cluster is known to be silent shows a round shape. In contrast, HoxD has a more elongated shape in the forelimb region where some of its genes are active. Our results support the 4C based model and show that these sub-compartments can be indeed the physical sub-clusters.

In parallel, we study how TF dynamics regulates transcription. In eukaryotes transcription initiation requires over 30 different TF proteins. Dynamic assembly of these components during initiation of the complex formation determines the transcription rate. Although key players have been identified, a quantitative model of transcription initiation describing how transcription rate is regulated by TFs in eukaryotes remains to be elucidated, partially, because suitable methods were not available. Here we use SR based single molecule (SM) tracking to directly image

single molecule dynamics of TFs. We study binding dynamics of p65 transcription factor and test different transcription regulation models (Chapter 6). P65 is a conditionally active gene specific TF whose activity dysregulation was associated with many diseases. To discriminate between different models of initiation we test p65 mutants with different binding affinity to p65-specific DNA motif. Using high density SM tracking we quantify their binding on different time scales. We see that having a wide range of affinities mutants and wild type protein bind DNA similarly at millisecond range although there is a correlation between affinity and transcription outcome. New evidence show that functional TF binding time is likely to be at longer (seconds) range. We continue these experiments in a new configuration to image this slow binding dynamics of mutants and wild-type p65 towards quantification of transcription initiation.

Keywords:

Transcription Regulation, Transcription Factor Dynamics, Chromatin Organization, Super Resolution Imaging, Single Molecule Tracking

Résumé

L'expression des gènes est dynamique et hétérogène à travers toutes les cellules. Ce processus est important car il aide à l'adaptation des différents environnements cellulaires, au développement des tissus et à leur fonctionnalité pendant le développement et dans l'organisme adulte. La régulation et l'expression des gènes sont principalement contrôlés à l'étape de transcription. Il existe deux facteurs essentiels qui contrôlent la transcription dans les eucaryotes : 1) l'organisation de la chromatine et 2) les dynamiques des protéines régulatrices, également appelées facteurs de transcriptions (TF). Jusqu'à récemment les TFs ont été étudiés avec des méthodes *in vitro* où un ensemble de TFs est étudié. Dans cette thèse, nous contribuons au développement des méthodes de microscopie à haute résolution pour la localisation de molécules individuelles. Avec l'aide de ces méthodes, nous quantifions l'organisation de la chromatine et les dynamiques des TFs dans les cellules vivantes. Nous présentons nos développements méthodologiques et nous les appliquons à l'étude de l'organisation de la chromatine et de la dynamique des TFs.

Premièrement, nous présentons une méthode d'imagerie de l'ADN dans les cellules vivantes avec une résolution améliorée. La localisation de molécules individuelles par microscopie 'super-résolution' (SR) permet la visualisation d'objets dont la taille est normalement limitée par la diffraction. Jusqu'à récemment, cette technique était exclusivement appliquée à des protéines. Nous avons trouvé des conditions qui permettent le clignotement nécessaire d'un marquage lié à l'ADN, ainsi qu'un marquage 'site specifique', PicoGreen (Chapitre 2). Nous sommes arrivés à une résolution de 50-70 nm qui représente un progrès d'un facteur cinq par rapport à la microscopie conventionnelle. Grâce à la performance du PicoGreen, nous sommes capables d'étudier les dynamiques de la chromatine avec une grande résolution. Ceci est la première analyse d'une teinture 'Site Specifique' pour la SR dans les cellules vivantes.

Le clignotement utilisé pour la SR peut être également appliqué à la technique de suivi des particules (SPT). Dans ce cas-là, une augmentation de mille fois des trajets est atteint par rapport à les techniques de SPT classiques. Initialement, seules les protéines activées par lumière, comme les protéines 'photoactivatable' (PA), étaient utilisées pour le SPT. Cependant, les protéines PA ne sont pas assez brillantes et il

n’en existe que très peu qui fonctionnent pour la SR. Nous avons montré que les expériences de SPT peuvent être effectuées en utilisant des marquages organiques avec une haute densité des trajets (Chapitre 4). Nous avons montré que les différents marquages peuvent clignoter dans des environnements divers dans la cellule ; par exemple la membrane, les mitochondries et le noyau. De plus, toutes les conditions déterminées sont compatibles avec la santé des cellules. En combinaison avec les protéines PA, nous avons donc la capacité d’étudier les trajets des différentes molécules, avec différentes couleurs sur des organelles variées. Dans une analyse séparée, nous avons développé un algorithme qui permet l’évaluation des trajets longs des molécules dans les cellules vivantes. Finalement, nous avons corrigé les artefacts pouvant se produire avec ce type de trajet (Chapitre 5).

Grâce à nos méthodes des molécules individuelles, nous avons accès à des paramètres précédemment inaccessibles. Sur le côté biologique, nous appliquons toutes ces méthodes à l’analyse du contrôle de la transcription. En premier lieu, on étudie la relation entre l’organisation de la chromatine et son rôle dans la régulation de transcription des gènes ; notre système modèle est le groupe des gènes Hox. Les gènes Hox sont responsables de la formation du corps pendant le développement embryonnaire. Leur expression est régulée dans espace et le temps pendant le développement. Il est connu que l’organisation de la chromatine a une influence sur la transcription des gènes Hox. Les données de ‘capture de la conformaiton de la chromatin’ (4C) suggèrent que les gènes actifs ou passifs ont respectivement une preference de contacts pour les autres gènes actifs ou passifs. En pratique, cette préférence est responsable de la formation des deux compartiments. Par conséquent, l’activation d’un gène est accompagnée par son mouvement d’un compartiment actif à un compartiment inactif. Ces données donnent la première idée de l’organisation des genes hox. Cependant, l’interprétation des données de 4C est limitée par la méthode. En particulier, une mesure de la proximité dans l’espace est déterminée indirectement avec un échantillon de cent à mille cellules. Dans ce travail, nous utilisons la technique d’hybridation in situ fluorecente (HISF) qui permet une mesure directe de l’organisation de la chromatine au niveau d’une seule cellule. En combinaison avec la microscopie SR, on peut ainsi visualiser la chromatine à très haute résolution. Nous analysons l’organisation groupée (cluster) des gènes Hoxd dans différents tissus qui ont différentes combinaisons et différents niveaux de

transcription des gènes Hoxd. Avec nos données, nous faisons un test direct de l'organisation spatiale proposée par les modèles de 4C. Nous montrons que la forme des groupes de gènes Hoxd est différente dans les diverses parties de l'embryon. Dans le téléencéphale (cerveau embryonnaire), une partie où tous les Hoxd sont silencieux, le cluster ont une forme compactée. Au contraire, les formes des clusters de Hoxd sont plus longes dans les membres antérieurs où les gènes sont actifs. Non seulement nos résultats valident le modèle basé sur les données de 4C , mais nous montrons également que les deux différents compartiments peuvent être physiquement distincts dans l'espace.

En parallèle, nous analysons les dynamiques des TF et comment ces protéines régulent la transcription. La transcription dans les eucaryotes a besoin d'une trentaine de protéines TF différentes. La vitesse de transcription est déterminée par l'assemblage de tous ces TFs en formant un complexe d'initiation. Les TFs principaux ont été dèjà identifiés. Cependant, un modèle quantitatif d'initiation de transcription qui décrit comment la vitesse de transcription est régulée par des TFs n'existe pas. Dans ce travail, nous analysons les dynamiques des TF en utilisant SPT en combinaison avec SR. En particulier, nous étudions les dynamiques de la liaison entre le TF p65 et ADN. Avec ces données, on teste les différents modèles qui décrit la régulation de transcription (Chapitre 7). Ce TF, p65, est actif à certaines conditions ; son dérèglement est associé à plusieurs maladies. Pour distinguer les différents modèles d'initiation, on teste les différents gènes mutants de p65 qui ont une différent affinité au motif d'ADN spécifique à p65. En utilisant le SPT à haute densité avec des marquages organiques, nous analysons la liaison dynamique de p65 sur les différentes échelles de temps. Avec les différents gènes mutants, nous ne notons pas de différences supérieures à une milliseconde dans les dynamiques; cependant, nous notons une relation entre l'affinité et la capacité de transcription. Ces nouvelles preuves montrent que les dynamiques des TFs fonctionnels se produisent dans l'intervalle de quelques secondes. Nous poursuivons ces expériences dans une nouvelle configuration avec le but de rassembler des trajets des TFs dans un échelle temporelle de quelques secondes.

Mots-clés:

Régulation de la Transcription, Dynamique des Facteurs de Transcription, Organisation de la Chromatine, Microscopie à Haute Résolution, Suivi des Molécules Individuelles

Краткий обзор

Экспрессия генов – это динамический, гетерогенный процесс, характерный для всех типов клеток. Среди ее функций адаптация к изменениям внешних условий, развитие тканеспецифического фенотипа и функциональности в процессе развития и во взрослом организме. Контроль экспрессии генов происходит в основном на уровне транскрипции. Два основных фактора регуляции транскрипции в эукариотах: организация хроматина и динамика транскрипционных факторов. До недавнего времени, они изучались с использованием методов, которые проводятся in vitro и требуют усреднения. Мы принимали участие в разработке новых методов визуализации единичных молекул, которые позволяют количественно описать организацию хроматина и динамику транскрипционных факторов in vivo и на уровне единичных клеток/ единичных молекул. В этом тезисе мы представляем разработанные нами методы и применяем их для изучения влияния организации хроматина и динамики транскрипционных факторов на регуляцию транскрипции.

В методологическом плане вначале мы разработали метод прямой визуализации ДНК со сверхвысоким разрешением в живых клетках. Микроскопия сверхвысокого разрешения позволяет визуализовать субдифракционную организацию клеточных структур и их динамику in vivo. Однако, до недавнего времени, практически только белки были объектом изучения микроскопии сверхвысокого разрешения. В данной работе мы нашли фотофизические условия для визуализации ДНК с помощью микроскопии сверхвысокого разрешения, используя ДНК- специфичный краситель Picogreen. Мы сумели достичь разрешение 50-70 нм, что является пятикратным улучшением, по сравнению с обычной оптической микроскопией. Это также было первой демонстрацией использования компартмент -специфического красителя в микроскопии сверхвысокого разрешения в живых клетках.

Фотофизический принцип, используемый в микроскопии сверхвысокого разрешения также может быть использован в трекинге единичных молекул. Он позволяет получить 1000х увеличение количества траекторий. Мы продемонстрировали новый подход к трекингу единичных молекул, основанному на принципе микроскопии сверхвысокого разрешения, используя

органические красители. Мы показали, что эти красители позволяют проводить трекинг до 3х типов белков одновременно и получить более длинные траектории по сравнению с используемыми ранее фото активируемыми белками.

Эти методы визуализации единичных молекул позволяют измерять ранее недоступные параметры. Мы используем в данной работе эти методы для изучения контроля транскрипции. Мы изучаем, как организация хроматина влияет на транскрипцию генов в модельной системе кластера генов HoxD. Эти гены ответственны за формирование морфологии тела во время развития эмбриона. Было показано, что локальная организация хроматина влияет на транскрипцию этих генов, в частности, что активные и неактивные гены кластера в основном контактируют с другими активными и неактивными генами кластера соотвественно. В данной работе мы используем метод флуоресцентной гибридизации, который позволяет производить измерения пространственной организации хроматина в единичных клетках и сочетаем его с микроскопией сверхвысокого разрешения. Мы изучаем пространственную организацию кластера генов HoxD в тканях с различной экспрессией этих генов. Мы демонстрируем, что организация кластера различна в тканях, где эти гены не активны, и в тканях, где они активны. Мы предлагаем модель организации кластера генов HoxD, в соответствии с которой активные и неактивные гены физически не контактируют друг с другом и формируют два отдельных пространственных субкомпартмента.

Ключевые слова:
Регуляция транскрипции, Динамика транскрипционных факторов, Организация хроматина, Микроскопия сверхвысокого разрешения, Трекинг единичных молекул

Table of contents

Table of figures

1 Introduction and aims of the work:

1.1 Chromatin organization and regulation of gene expression

1.1.1 Chromatin organization in higher eukaryotes

The human cell DNA polymer molecule has a size of ~3000Mb (mega base pairs). The famous double helix structure discovered by Crick, Watson, Franklin and Wilkinson gives a characteristic distance in this polymer chain: 3.4 nm per 10 base pairs [1, 2]. Extrapolating it to the size of the genome would give a distance of more than 1 m meaning that to be put into the 5x5 um^2 nucleus of a typical cell the DNA molecule should be significantly compacted.

There are several levels of DNA packaging: nucleosomal, chromatin fibers and chromosomes (Figure 1). As DNA is always found in living cells together with different proteins providing DNA organization and functions we will henceforth refer to chromatin – the complex of DNA and these proteins. The nucleosome is an octamer complex made of proteins dimers called histones. The first level of DNA compactization is achieved by wrapping DNA around nucleosomes. DNA makes a 720 degrees loop around the nucleosome and 146 bp of DNA are used for that (gene control). Histones can be post-translationally modified at specific amino acid positions. There are several types of modifications: acetylation, methylation, ubiquitination, phosphorylation and others. Usually residues facing the surface of the nucleosome are affected. This can tune specific interaction of chromatin with different nuclear regulatory proteins [3]. This lowest level of organization is called the ‘’beads-on-string’’ model.

The highest level of organization is chromosomal. The human genome is divided into 23 chromosome pairs. Chromosomes condense during cell division when they are clearly visible and can be physically separated. During interphase they occupy distinct territories inside the nucleus called chromosome territories [4]. Gene-poor chromosomes are preferentially located at the nuclear outer region while gene-rich ones are in the centre of the nucleus. So there is a correlation between the number of genes and the relative position of chromosomes.

Intermediate, so called higher-order, levels of DNA organization cannot be visualized by conventional microscopy in live cells due to the diffraction limit and are, therefore, less studied. In Chapter 2 we present a method of direct super-resolution imaging of DNA in live cells. It is naturally believed that at higher-order levels of DNA organization nucleosomes are organized in fibres of different size called chromatin fibers. In vitro reconstituted chromatin fibres of 30 nm were visualized with electron microscopy [5]. These fibres are probably cross-linked providing further condensation of chromatin to create the required packaging level.

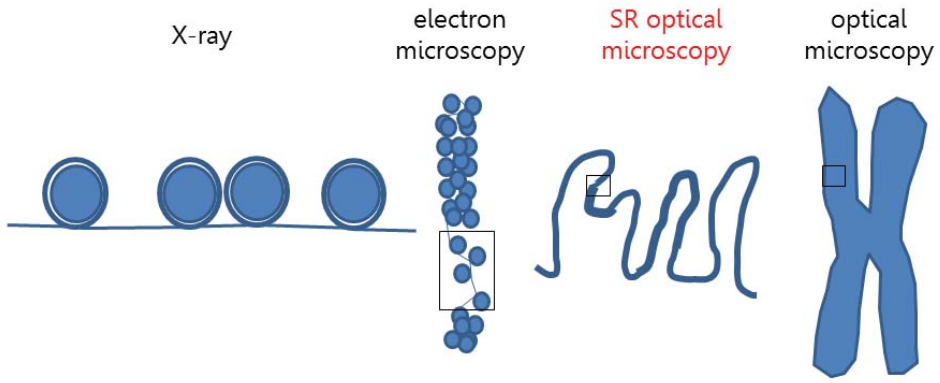


Figure 1: Levels of the DNA packaging in eukaryotic cells

DNA is wrapped around nucleosomes which probably organize themselves in fibres. These fibres are further compacted into chromosomes. Note the heterogeneity of packaging density. Methods that are used to study them are noted above.

First quantitative in vivo based data describing higher-order chromatin organization were obtained using recently developed chromatin capture techniques (3C,4C,5C and Hi-C) [6]. There, pieces of chromatin that are in close proximity in 3D are captured together by cross-linked. After that, cross-linked DNA is randomly restricted; resulting linked DNA short fragments are circularized so that circular DNAs contain joined together pieces of DNA from two separate locations in the genome which are in close proximity in 3D. Quantitative sequencing of these linked DNA fragments gives information how often these pieces of DNA are found in one vector and therefore contact probability between different sites in the genome. In 4C, for example, contact probability between a particular site in the genome and the whole genome is retrieved (one against all).

Using these methods, it was found that globally, chromatin in high eukaryotics is organized into a fractal globular structure at the megabase scale during interphase. This structure provides very high density packaging remaining knot-free and easily allows given part of chromatin to condense and decondense at any scale. In this model, the spatial 3D distance between two sites in the genome is inversely proportional to the genomic 1D distance to the power of 1.5 [7].

It was also found that the genome is divided into two compartment types with relatively high and low transcription activity at the megabase scale corresponding to open and closed chromatin domains respectively. Interestingly, the mitotic chromosome has a different and simpler structure. It consists of linearly oriented compacted chromatin loops [8]. Domain organization seems to be multilevel, as it was also discovered at the sub-megabase scale. These 500 kb sized domains were called topologically associating domains (TADs) [9, 10]. Their functional role is not well established, as there is no significant difference in transcription activity or histone modifications between domain types. One idea is that they can compartmentalize the genome properly to be able to properly establish contacts and interactions between different genes and enhancers.

1.1.2 Chromatin organization and transcription in HoxD gene cluster during development

At each time point, only a subset of genes in a cell is active. One mechanism, which determines how a cell chooses these genes can be that local chromatin organization has a permissive environment only at specific regions corresponding to potentially active genes. We study how chromatin local organization is related to gene expression on HoxD gene cluster (Figure 2 A).

Hox genes are a group of genes that is responsible for body plan formation during embryogenesis [11]. There is remarkable colinearity between the position of these genes on the chromosome and the order of their expression along the anterior-posterior axis in the developing animal [12]. Their importance and organized expression pattern make them an interesting model system. The open question is how chromatin organization influences the precise spatial and temporal expression patterns of Hox genes.

One aspect of how chromatin organization regulates transcription, which is common for all genes, is histone modifications. Histone modification patterns accompany specific transcription patterns of HoxD genes (Figure 2 C)[13]. For example in the anterior trunk where Hox genes d1-d9 are active they are marked by the presence H3K4m3 (histone H3 is tri-methylated at lysine 4), while genes d13-d10 that are silent are associated with a H3K27m3 modification [14]. Generally, in different parts of the embryo, inactive Hox genes are covered by H3K27m3 modifications, while active genes correlate strongly with the presence of an H3K4me3 modification [13].

The chromatin capture technique 4C, which measures the contact probability between different pieces of DNA is a useful tool to study chromatin organization of the HoxD cluster. First, it was demonstrated that Hox genes make long-range contacts with specific sequences at regions that are up to 1 Mb from both centromeric and telomeric sides of the cluster (Figure 2 B) [15]. These regulatory regions are important for transcription. Removal or mutation there leads to downregulation of transcription and severe phenotypic effects in mouse models. In a separate study, 4C was used to look at the organization of the HoxD cluster itself (Figure 2 D). It was shown that contact probability between genes within the cluster is far from being uniform [14]. It was found that genes that are silent (more centromeric) mostly interact with each other. Similarly, active genes (more telomeric) also preferentially contact each other. This implies that non-active and active genes are separated from each other and form two distinct domains. Within a domain, genes have similar transcription states. For Hox genes virtually moving along the anterior-posterior axis means activation of the more centromeric genes from Hox d1 to d13. It was demonstrated that this corresponds to relocation of these genes one by one from the non-active to the active domain in 4C data [14] [16]. In other words, to be activated, a gene should translocate to an active domain. Recent studies [9] show that, interestingly, HoxD cluster seem to lie at the border between two TADs. TAD organization provides the physical separation of HoxD cluster necessary for the centromeric part of the HoxD cluster associates to one TAD and the telomeric part lies in another [17] [18]. This matches well with the two-domain model organization of clusters discovered by Noordermeer and cowokers [14, 19].

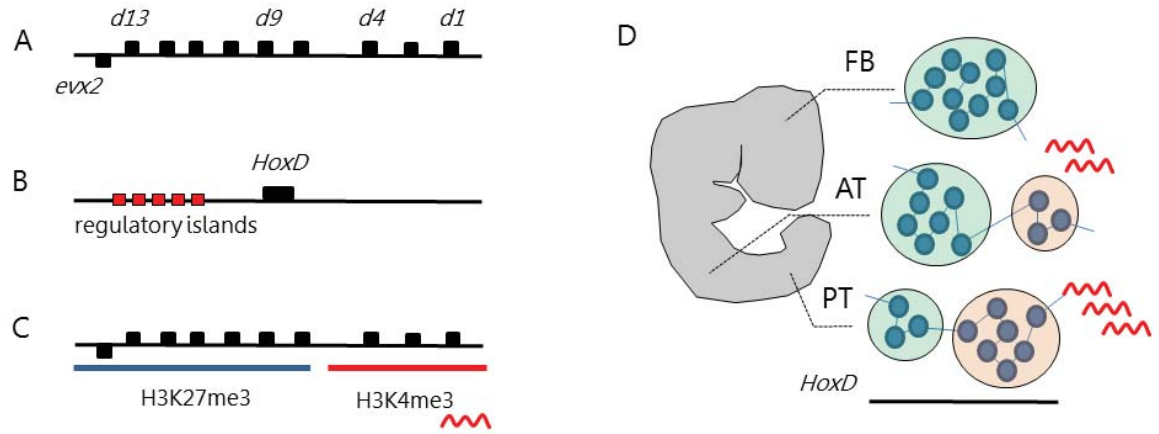


Figure 2: The HoxD gene cluster organization

(A) Ten genes make the HoxD gene cluster. (B) Centromeric and telomeric long range interactions regulate HoxD genes expression. (C) Histone modifications correlate with HoxD gene expression. (D) HoxD architecture in the developing E10.5 embryo according to 4C data from Nordeermeer [14]. Genes in forebrain (FB) are silent and make a round structure. Activation of telomeric genes in anterior (AT) and posterior trunks (PT) is accompanied by their redistribution to active sub-compartment so that active and passive genes are separated into two distinct sub-compartments according to their expression status.

1.1.3 Aim One: HoxD gene cluster spatial organization and regulation of transcription with SR based FISH

C methods in general lack single-cell resolution and provide an indirect model rather than direct observation of chromatin organization. Here in Chapter 3 we look at local organization of HoxD with super-resolution (SR) FISH imaging. FISH allows observation of gene cluster organization directly at the single cell level. We combine it with STORM SR microscopy to take advantage of its sub-diffraction enhanced resolution. We look at the shape of HoxD gene clusters and measure morphological parameters of these clusters. We correlate them with the position within the embryo and therefore with expression patterns of Hox genes.

1.2 Transcription factor dynamic and regulation of gene expression

1.2.1 Transcription initiation and transcription factors

Regulation of gene expression mostly occurs at the transcription level. This process of RNA production from the DNA template is achieved by collaborative work from many proteins. Synthesis of the RNA molecule itself is performed by the multisubunit enzyme RNA polymerase (RNAP). The rate of transcription (for RNAP polII) in eukaryotes was measured recently: it takes about 20 min to transcribe an average gene or 6.3 nucleotides per second [20]. To start, RNA synthesis polymerase should be phosphorylated at the serine 5 residue. Next, after processing 20-30 bp polymerase stalls and waits until further phosphorylation occurs. So in terms of polymerase activity three types exist: inactive genes with no polymerase, active genes with uniform polymerase distribution and stalled genes with a polymerase waiting near a transcription start site (TSS) for phosphorylation. Promoter regions around TSSs contain specific sequence elements, which allow RNAP to start transcription. Few of them, such as the TATA box are present at all promoters and are required for all genes to be transcribed. However most are specific DNA motifs that are typical for particular genes. A typical gene may contain tens of them. A combination of general and specific promoter elements enables specific gene expression.

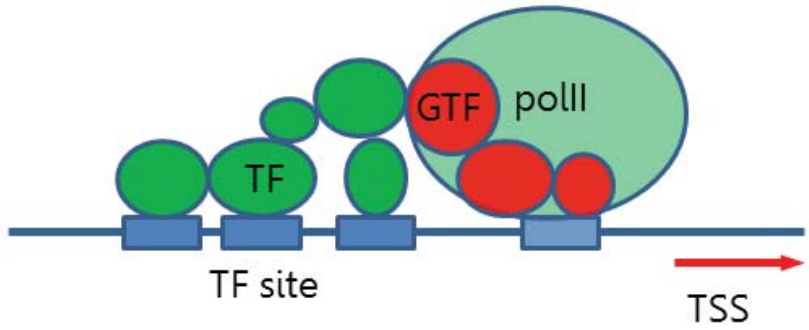


Figure 3: Transcription initiation machinery

Multiple gene specific TFs (green) and general TFs (red) bind specific sequences (dark blue) and Tata box (light blue) respectively. They recruit RNA polymerase (light green).

Initially, RNAP doesn't bind the promoter itself; there are transcription factor proteins (TFs) that recognise and bind a particular sequence in the promoter region build initiation complex and bring RNAP there (Figure 3). General TFs bind the TATA box and provide the requirements for RNAP. There are over 30 general TFs that are required for transcription initiation [21]. Specific factors bind motifs that are present at specific genes promoters. They are considered as pioneering factors as they bring general TFs to start transcription. TFs usually have a modular structure. Particularly, they have DNA binding domains and transactivation domains. The DNA binding domain is responsible for binding specific motifs of DNA. The length of the DNA motif, which is recognized by TF is 6-12 nucleotides for most TFs [22]. In prokaryotes, with their small genomes, the content of the motif is sufficient to provide single gene specificity for TFs. In contrast, in eukaryotes, the size of the motif is usually insufficient to spot a single gene. Therefore it's generally accepted that TFs should act cooperatively to provide gene specific transcription activation [23]. Indeed, whole genome ChIP shows that most eukaryotic genes have binding sites for multiple different TFs at their promoter region. In agreement with low information content, it is demonstrated that TFs have multiple binding sites across the genome with a number exceeding the known number of genes that are regulated by this TF [24, 25]. For example p65 is shown to bind at ~20000 positions while there about 500 genes regulated by it [26]. So, the binding landscape of eukaryotic TFs is rather complicated. Transactivation domains are responsible for interactions with other TFs and chromatin remodelers so they enable the assembly of multimeric complexes at promoter regions of transcribed gene.

1.2.2 Transcription factor dynamics during transcription initiation

Transcription factors are key players in transcription regulation. They bring other TFs, they modify chromatin and they bring RNAP to the promoter. One can imagine the process of transcription initiation as a process of proteins binding DNA to bring another protein or modify chromatin. The result of this will be the chromatin template and polymerase ready for transcription. Early biochemical in vitro studies show that TFs have stable, long lasting interactions with DNA. The dissociation constant of p65 was found to be 0.1-10 nM in vitro and its binding time about 30 min [27]. In addition, when particular TFs were omitted initiation complex assembly was inhibited and

transcription efficiency dramatically decreased. If a particular factor was knocked out/omitted, PIC assembly could only process until the step when this factor is brought to DNA [28]. Based on these studies it was hypothesized that there is a particular chain of events and one pathway for transcription initiation (Figure 4A). TFs binding is ordered and forms stable, long-lasting transcription initiation complexes. Each new factor stabilizes an interaction between an initiation complex and DNA and creates favourable conditions for binding of downstream factors.

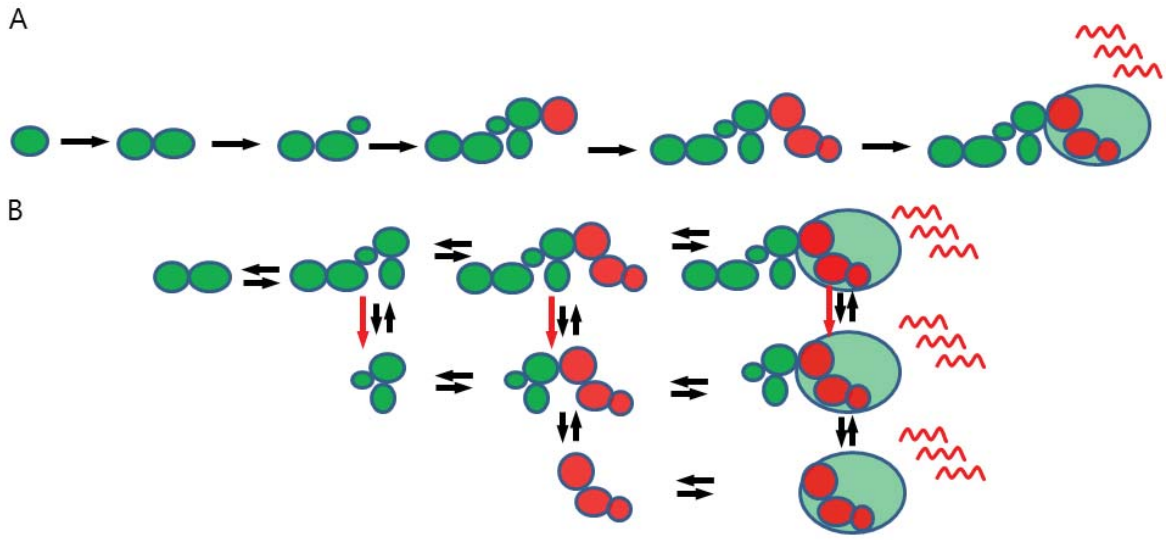


Figure 4: Different views on transcription initiation

Based on early biochemical data it was hypothesized that there is a particular chain of events and one pathway for transcription initiation. TFs binding is ordered and forms stable, long-lasting transcription initiation complexes. Live cell microscopy data show that TFs are actually very dynamic and their interactions with DNA are transient. It implies that TFs are not necessarily all at the promoter at the same time to start transcription. They can interact with chromatin briefly and unbind before transcription initiation. In this scenario there are multiple pathways to transcription initiation through a variety of intermediate states. Non-equilibrium ATP-driven processes (indicated by red arrows) are likely to regulate transcription initiation.

Live cell microscopy added a new layer of information on TF binding during transcription initiation. FRAP and single particle tracking data show that TFs are actually very dynamic in the nucleus. It was found that many TF interactions with DNA are transient [29, 30]. Their binding times are on the order of seconds rather than minutes or hours obtained in biochemical studies. It implies that, in contrast to early models where proteins bind sequentially to make a long lived stable initiation

complex, it can actually be that TFs are not necessarily all at the promoter at the same time to start transcription (Figure 4B). It can be that each of them interacts with chromatin briefly and unbinds before transcription initiation. In this scenario there are multiple pathways to transcription initiation through a variety of intermediate states and transcription is initiated, not by presence of all factors, but by their cumulative action.

1.2.3 Models of transcription factor regulation of transcription

Quantitative models of transcription predict a relationship between transcription and TF concentration. Given the relative simplicity of transcription in prokaryotes, quantitative models were initially developed in these systems. In many cases, equilibrium thermodynamics can explain the behaviour of the system. In this case binding energies and TF concentrations are the only important parameters. It gives occupancies of states and eventually predicts transcription rates [31]. Cellular processes are out of equilibrium, as there is a directionality of processes and energy consumption. Indeed in the equilibrium case, rates of direct and opposite reactions should be equal which is not true for transcription initiation. There is increasing awareness that non-equilibrium models which don't assume equal rates should be used [32-34]. Emerging live cell single molecule imaging methods allow scientists to get quantitative information for these models. Single molecule tracking directly probes TF dynamics [35-37]. It allows one to measure rates between states, for example between bound and unbound TF states directly in vivo. With such information, we can assess how TF dynamics regulates transcription outcome. Accordingly, the first steps have already been taken in this direction [32, 38-40].

1.2.4 Aim Two: P65 transcription factor dynamics and regulation of transcription with single-molecule tracking

Although the key players of transcription regulation are identified, a quantitative model of transcription initiation is still to be worked out. With the development of new methods that allow for direct measurement of TF dynamics, we can address this question. One of the most studied specific TFs is p65, which is also called NfkB and RelA [41, 42]. P65 is involved in many cellular processes, including the production of cytokines as a response to different extracellular signals. It is a conditionally active

TF, which translocates into the nucleus upon stimulation with different factors (TNFalfa, LPS, H2O2 etc.). The mechanism of activation and translocation and its target genes has been extensively characterized [43-46]. All of this makes p65 a convenient model system. Here in Chapter 6 we study p65 binding dynamics and how p65 regulates transcription of its target genes. To test transcription initiation models we modulate rates between states using p65 mutants with different binding affinity to p65-specific DNA motif. We examine how in vitro binding affinity of different p65 mutants is related to their in vivo binding dynamics. Further, we study how binding dynamics of TFs with different binding affinities regulates transcription outcome.

2 Live-cell dSTORM of cellular DNA based on direct DNA labeling

Introduction

Super-resolution (SR) fluorescence imaging [47-50] relies on specially adapted microscopes and analysis software, but equally important are the fluorescent probes used to label biological proteins and molecules of interest. In point-localization SR, the on- and off- rates of fluorophore photoswitching must be controlled to temporally separate the signals from individual molecules. The fluorescence from each molecule is then spatially localized [51], allowing their positions to be obtained with nanometric precision [52]. Molecular positions are then rendered to give a reconstructed image with spatial resolution up to several tens of nanometers [53, 54]. Although fluorescent proteins can be used as labels [55], synthetic dyes have the advantage of higher photon yields, leading to higher attainable resolution [56]. A special challenge for chemical biologists is to develop or identify synthetic dyes that are compatible with live-cell SR imaging, since most dyes are not cell membrane-permeable. Recently, the point-localization SR method of direct stochastic optical reconstruction microscopy (dSTORM) [50, 57, 58] was used to image histone H2B proteins in living cells [59, 60]. Proteins were labelled with rhodamine and oxazaline dyes using the genetically encoded chemical trimethoprim [61] or SNAP [62] tags. In a separate study, dSTORM was applied *in vitro* to image purified DNA by direct labelling using the cyanine-based YoYo-1 dye [63]. However, most DNA-associating dyes are not compatible with live cell imaging, due to their cell-impermeance and cytotoxicity. Thus, live-cell super-resolution imaging of DNA structure has never been demonstrated.

We present here the imaging of DNA in living cells with dSTORM based on direct labelling with the commercially available cyanine-based Picogreen dye (Invitrogen). Picogreen presents several advantages for live-cell imaging over other DNA-specific dyes, including minimal perturbations to DNA structure and an increase in fluorescence upon binding, resulting in low background fluorescence. We identified a live-cell imaging medium that optimizes the reversible photoswitching of the fluorophores, and used it to resolve nuclear and mitochondrial DNA structure directly.

Furthermore, due to the excellent preservation of these dyes, we were further able to perform time-lapse dSTORM imaging of directly labelled DNA.

Results

Theoretically, by choosing an appropriate reducing-oxidizing system one can achieve the controllable reversible photoswitching required for the dSTORM of nearly any cyanine derivative. Several cyanine-based dyes have already been used in dSTORM imaging [57, 58]. Typically, enzymatic oxygen scavenging buffers are used to prevent the irreversible photobleaching of dyes and mercaptoethylamine (MEA) was used as a reducing agent to induce photoswitching or photoblinking. MEA is responsible for converting the dye into long-lived dark non-fluorescent states from which single molecules can spontaneously recover to the ground state upon interaction with residual oxygen: this is the photochemical basis of dSTORM. Buffer conditions determine oxygen concentration and reduction reaction rate, and are therefore important for optimizing photoswitching rates.

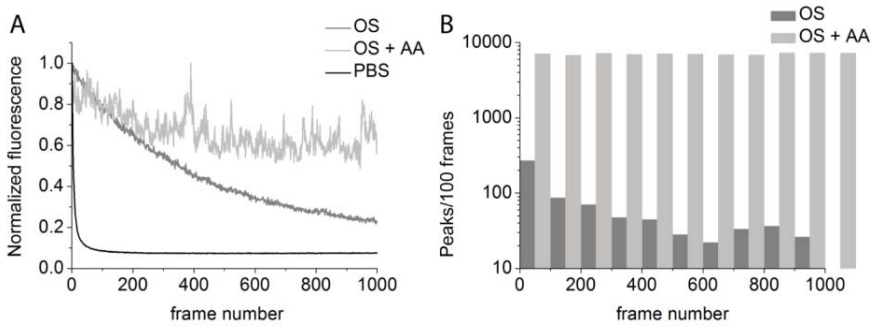


Figure 5: Picogreen photoswitching dynamics

U2OS cells labeled with Picogreen were imaged with a 488 nm laser with excitation power ~1Kw/cm^2 in the focal plane. (A) Fluorescence decay curves for a 300 nm x 300 nm square region of interest and (B) time dependence of the number of localized peaks averaged over multiple frames in different imaging buffers (note the logarithmic scale).

To perform live-cell imaging, we avoided adding the potentially toxic thiol reducing agents commonly used to induce photoblinking such as MEA [58, 64]. We tested buffer conditions that have worked for live-cell dSTORM in the past for other dyes. In phosphate-buffered saline (PBS) alone, we observed rapid photobleaching of Picogreen with no recovery (Figure 5A). When using a buffer containing an oxygen scavenging system (OS) composed of 10% glucose, 0.5 mg ml^{-1} glucose oxidase and 40 µg/ml catalase (Sigma), the fluorescence intensity over time showed a much

slower decay. Thus, enzymatic removal of oxygen indeed helps to avoid photobleaching of the dye. However, it was not sufficient to induce the single molecule blinking required in dSTORM. To improve the recovery and photoblinking of Picogreen, we applied ascorbic acid (AA) as a reducing agent and the OS (described above) to reduce irreversible photobleaching [65]. We found that 1 mM ascorbic acid combined with this oxygen scavenging system in Leibowitz media (Gibco/Invitrogen) at pH 7.2 allowed rapid and reversible photoswitching (Figure 5A). In this medium, a large number of molecules could be localized in each raw image, over thousands of images, as shown in Figure 5B.

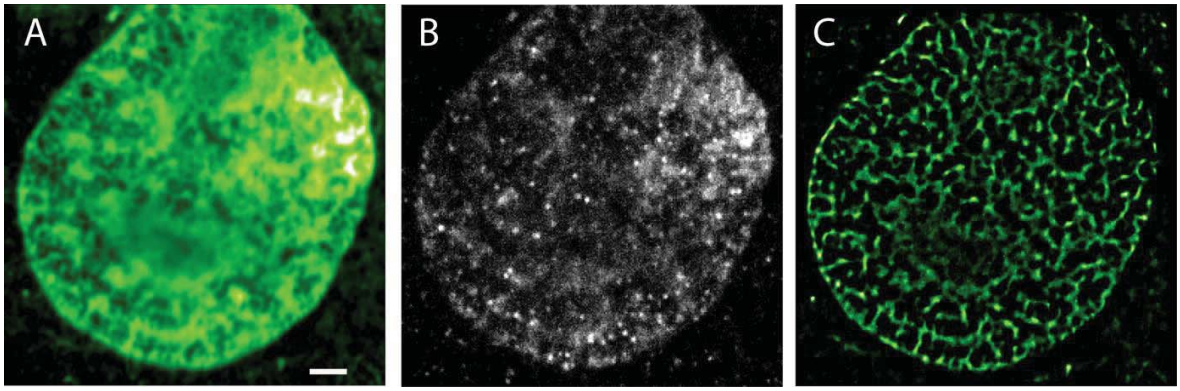


Figure 6: Live-cell dSTORM of cellular DNA based on direct DNA labelling with Picogreen
(A) Wide-field diffraction limited image, (B) single raw dSTORM frame demonstrating Picogreen single molecule photoswitching, seen as bright spots and (C) final dSTORM image reconstructed from 8000 frames. Scale bar 2.5 µm.

dSTORM of living cells requires fast acquisition. This in turn necessitates a high number of peaks per frame to satisfy the Nyquist criteria for object structure determination while maintaining nanometric molecular localization [66]. This is particularly true for DNA, with its intricate and dense structure. In our optimized medium, individual raw images show a large number of peaks, suitable for live dSTORM imaging (Figure 6B). We obtained 3500 peaks per μm^2 for a stack of 8000 raw images, acquired at 30 ms per frame - a sufficiently high density of well-localized molecules to reconstruct super-resolution structure (Figure 6C). The organization of the DNA revealed by the live dSTORM image would not be possible to deduce from the epi-fluorescence image. Due to the improved resolution and coincident axial sectioning from dSTORM, continuous fibres of DNA could be resolved.

We also observed a small pool of extra-nuclear cellular DNA representing mitochondrial nucleoids. To quantify the resolution achievable with Picogreen, we visualized these individual mitochondrial DNA nucleoids, also in living cells (Figure 7A). Nucleoids appear as bright punctate structures along elongated mitochondria, whose size is greatly reduced in dSTORM imaging (Figure 7B). The brightness, high specificity and efficient photoswitching of Picogreen enabled a resolution of 70 nm as measured by the full width half maximum (Figure 7C). This is approximately a factor of 5 better than can be obtained with conventional epi-fluorescence imaging.

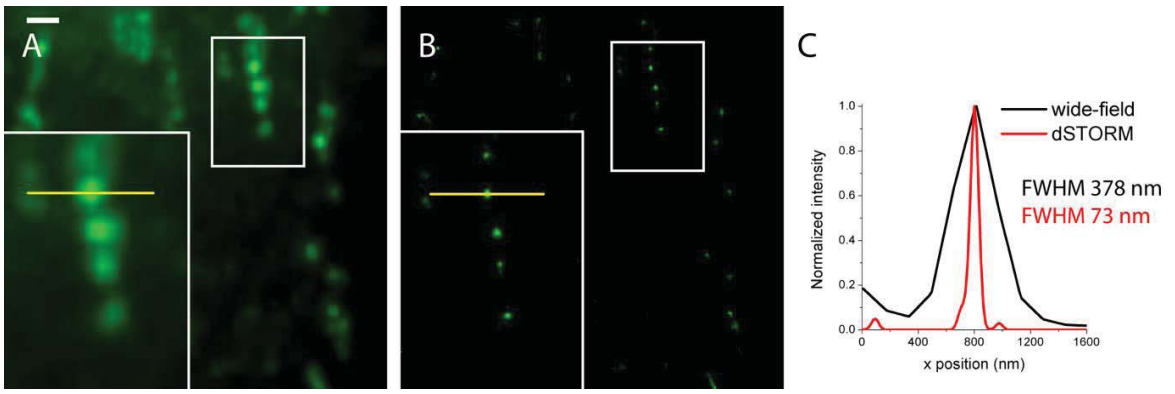


Figure 7: Mitochondrial DNA labelled with Picogreen

(A) Wide-field and (B) dSTORM images of mitochondrial DNA. Scale bar 1 µm. (C) Line intensity profiles (from lines in A and B) and calculated full width half maximums (FWHM) of corresponding structures.

A goal of live-cell imaging is to monitor changes in cellular structures over time. Many dyes do not have good enough recovery from photobleaching to offer sustained signal over many images. We tested the possibility of time-lapse dSTORM by allowing the Picogreen fluorophores 10 minutes to recover from longer-lived dark states between acquiring each stack of raw images. The individual stacks were used to construct a series of dSTORM images (Figure 8A) demonstrating that is possible to monitor the sub-diffraction limited organization of DNA in individual cells over time. The total number of molecules in each dSTORM image was quite constant, varying only by ~10-20%. The number of molecules per raw image showed more fluctuations, probably due to cellular processes (Figure 8B). The dynamics of DNA fibers can be more clearly seen by looking at the differences between consecutive dSTORM images (Figure 8C).

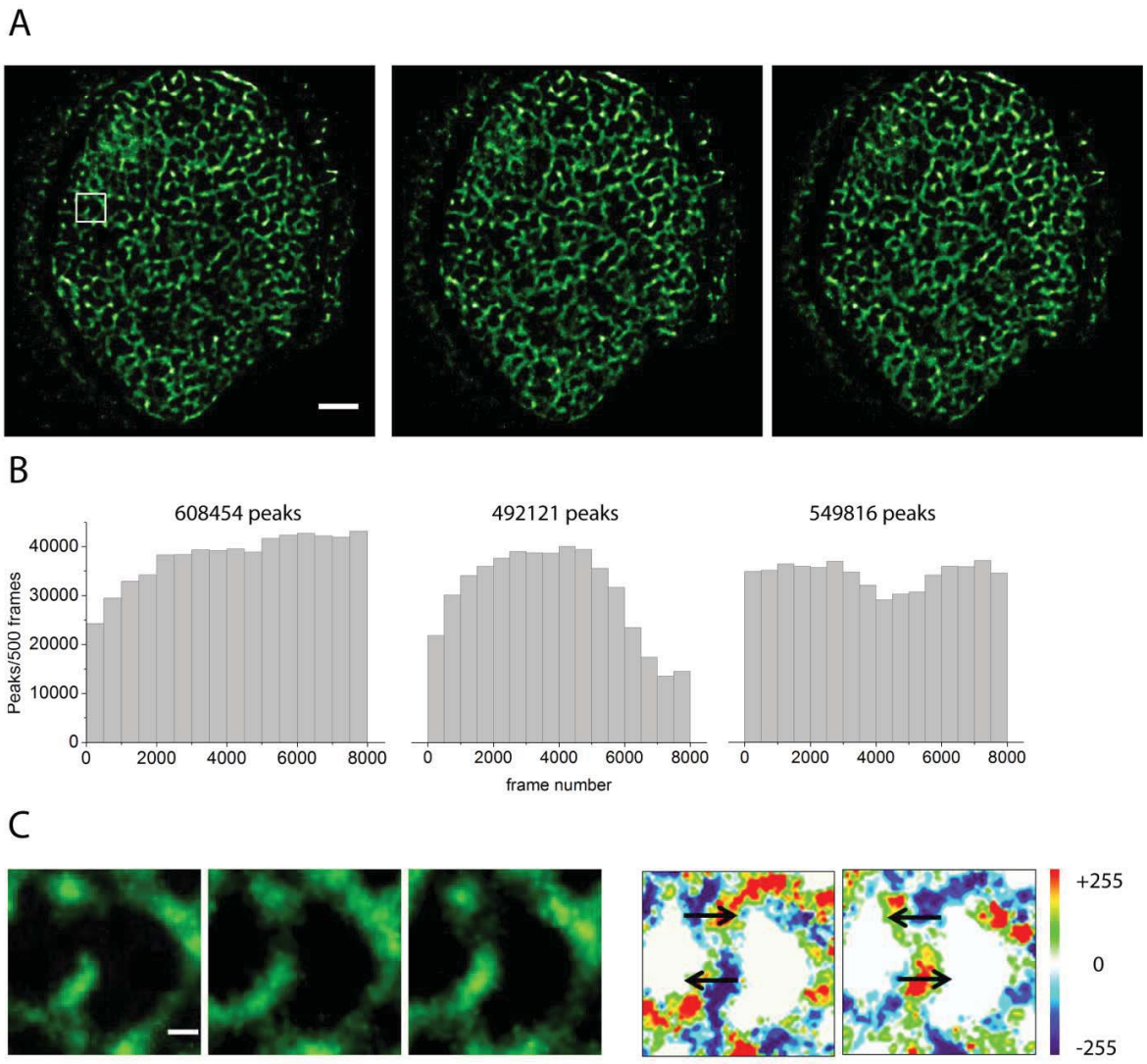


Figure 8: Time-lapse dSTORM

(A) Three consecutive dSTORM images (8000 frames each) with 10 min interim pause. Scale bar 2.5 µm. (B) Time dependence of the number of peaks for each raw image stack. The total number of localized molecules for each dSTORM image is indicated. Notably the variation in number of molecules across images is small. (C) Zoom of dSTORM images (left) show local rearrangements, highlighted by consecutive image substraction (right). Subtracted dSTORM images demonstrate chromatin fibers sub-diffractional dynamics. Arrows indicate direction of motion. Scale bar 200 nm.

Commercially available, DNA-specific cyanine dyes can be used for live-cell dSTORM, and even time-lapse dSTORM imaging, giving insight into the dynamics of DNA organization at unprecedented resolution. This provides an advantageous approach to study processes related to DNA dynamic structural rearrangements such as those occurring during cell division or in response to stress. Furthermore, this

represents a complementary tool to those previously developed [60, 67], permitting combined DNA and protein super-resolution imaging. Potential applications include the study of chromatin reorganization and the dependence of DNA organization on gene expression in high transcription versus silent regions. We further suggest that other commercially available affinity dyes can be used to directly target cellular structures and organelles for super-resolution imaging.

Methods

Cell culture

U2OS cells were cultured in DMEM Glutamax-I media (Gibco/Invitrogen) in a 5% CO2 atmosphere at 37 °C supplemented with 10% FBS (Invitrogen). Cells were normally split every 2-3 days depending on their confluence.

Cell labelling

Cells were plated 24 hours before labelling on 25 mm glass coverslips (approximately 75K cell per coverslip). On the day of the experiment cells were washed in PBS (Sigma) and labelled with Picogreen (1/500 dilution from original stock of Quant-iT PicoGreen (Invitrogen)) in DMEM without phenol red (Gibco/Invitrogen) and incubated for 15 min at 37 °C. Then cells were washed again and pre-warmed imaging buffer was added (1mM ascorbic acid, 10% glucose, 0.5 mg ml^{-1} glucose oxidase, 40 µg ml^{-1} catalase in Leibowitz media, pH 7.2).

Live-cell dSTORM imaging

Wide-field and dSTORM imaging was performed on an inverted microscope (Axio Observer.D1, Zeiss) equipped with a TIRF module. A 488 nm laser (Sapphire 488-50, Coherent) used to excite Picogreen was focused on the back focal plane of the oil-immersion objective (α Plan-Apochromat, 100x, NA=1.46, Zeiss). Fluorescent light collected by this objective then was projected onto an EMCCD camera (iXon+, Andor). Additional lenses resulted in a final pixel size of 100 nm. To minimize drift, imaging was performed on a motorized stage (ProScan III, Prior), equipped with a Piezo (N-725, PI). Typically, we acquired 2000-8000 images (30 ms per frame) with a laser excitation intensity of 1-5 kW cm^{-2} for a total acquisition time of 1-4 minutes for a single dSTORM image.

Data analysis

Raw data was analyzed using Peakselector software provided by Harald Hess (previously described in [47]). Peaks were selected based on their parameters: peak width, number of photons and localization precision. Peaks with a width between 120 nm and 400 nm, 100-3000 photons and localization precision less than 50 nm were considered as good single molecule localizations. Standard grouping procedure was also applied to group molecules appearing in consecutive frames. dSTORM images were rendered as a sum of superimposed filtered peaks with a width corresponding to their empirically determined localization precision.

3 HoxD gene cluster spatial organization on single cell level with super-resolution FISH imaging

Introduction

The precise control of HoxD gene expression in time and space is crucial for correct development in high eukaryotes. Hox genes encode developmental regulator transcription factors and their major role is to establish a proper body plan along the anterior-posterior axis during development [68]. Their transcription activation pattern was extensively studied [11]. There is colinearity between the order of expression in time and space within the embryo and the order of Hox genes in a genome.

In eukaryotes, genes exist in a chromatin context that modulates environment for expression. There are indications that distinct chromatin organization provides the Hox genes transcription pattern [12, 18]. Chromatin conformation capture (4C) data, that give probability of contacts between different pieces of genome [6], suggest that HoxD genes are compartmentalized according to their activity [14-17]. Genes that are active/silent preferentially contact other active/silent genes respectively. 4C data provide a first insight on organization of the HoxD cluster but this method's properties limit result interpretation. First, chromatin conformation capture assays measure spatial proximity indirectly. Second, they usually require hundreds of thousands cells so there is a lot of averaging.

We complimented the 4C study of HoxD and tested the compartmentalization model using super-resolution (SR) based FISH [69]. FISH [70] allows us to perform studies at the single cell level and to directly observe genes fluorescently labeled within HoxD cluster. STORM SR imaging [47, 50] adds the ability to resolve sub-diffraction spatial organization of HoxD. We found that in active tissue, the HoxD cluster becomes elongated. This elongation occurs around the border between active and non-active genes' regions. Our results taken together with 4C data suggest that active and non-active regions physically separate from each other and form spatial sub-compartments.

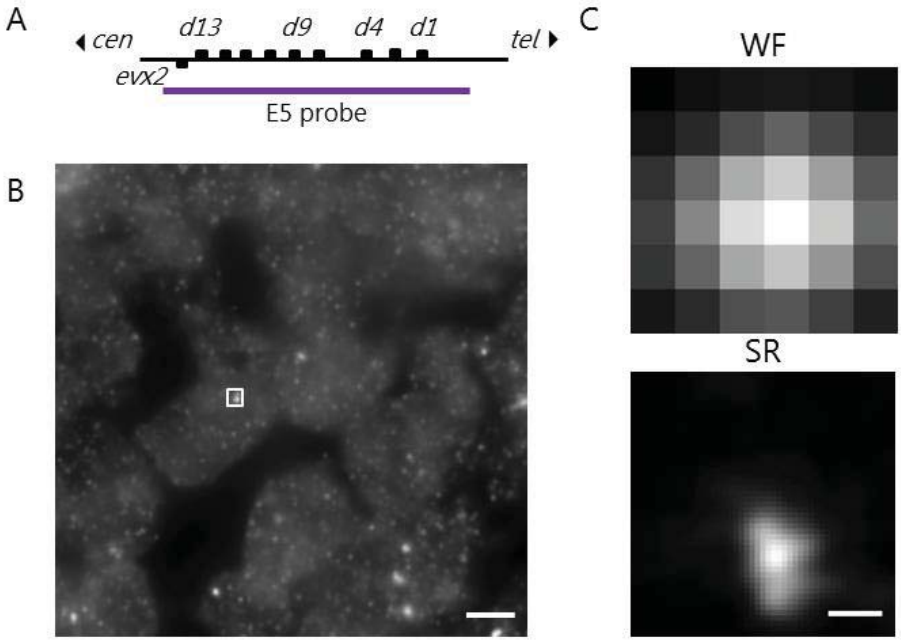


Figure 9: SR imaging of the HoxD gene cluster

(A) HoxD gene cluster and RP24-448E5 BAC probe (E5) covering the whole cluster used for labelling. (B) Image of labelled tissue (scale bar 5 µm). (C) Wide-field (top) and SR (bottom) images of a labelled single HoxD cluster (scale bar 200 nm).

Results

We performed SR FISH imaging of 10.5 day-old (E10.5) mouse embryos. Collected tissue samples were then fixed and labelled with biotinylated HoxD probe called E5 (Figure 9 A,B). The probe was stained with the dye Alexa 647. The HoxD cluster has a size of 100 kb and consists of 9 genes. They are named from hoxd1to hoxd13 with hoxd1 being the closest to the end of chromosome (most telomeric gene) and hoxd13 being the closest to the center of chromosome (most centromeric gene). E5 probe is 159 kb and covers the whole cluster. Clusters were identified as bright spots in nuclei. Normally a maximum of 3-4 bright spots were identified per cell (for details of cluster identification see Methods section). Despite our initial concern that we wouldn't obtain satisfying photoswitching in a thick tissue sample, a standard buffer containing oxygen scavengers and β-mercaptoethanol used in previous studies [71] for cell culture STORM worked well here. It induced cycling of molecules through dark states required in STORM. One can see that SR imaging significantly improves resolution (Figure 9 C) allowing the observation of nanoscale details of HoxD organization hidden normally by the diffraction barrier in conventional FISH imaging.

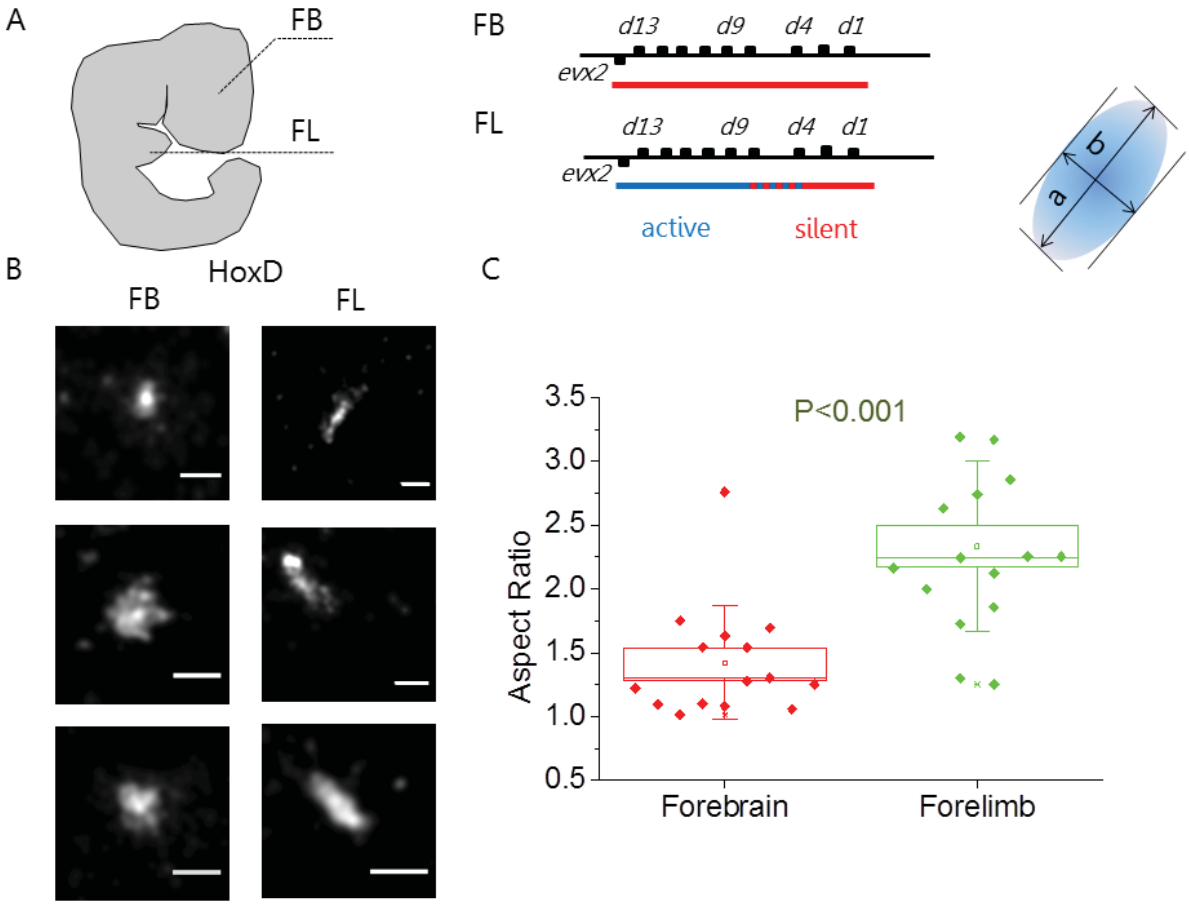


Figure 10: The HoxD gene cluster shape in FB and FL

(A) Forebrain (FB) and forelimb (FL) areas in developing E10.5 mouse embryo. Expression status of HoxD genes in corresponding tissues. (B) Typical images of HoxD cluster in FL and FB. Scale bar 400 nm. (C) The shape of the cluster obtained from SR images was fitted with an ellipse and an aspect ratio was measured in FB and FL. HoxD has significantly more elongated shape in FL compared to FB (p-values of two-sample t-test reported).

To study how the organization of chromatin template is related to transcription activity of HoxD genes, we imaged cells at specific positions in tissue with different HoxD expression patterns: forebrain (FB) and forelimb (FL) (Figure 10). FB is the region where HoxD genes are inactive. In FL (E10.5), genes from d13 to d8 are active so there is a telomeric silent part and centromeric active part [72]. We imaged multiple clusters in FL and FB tissues and quantified their shape. We noticed changes in the cluster shape between FL and FB. Clusters in FL often looked elongated compared to FB. We fit the shape with an ellipse, found its major axes and calculated the aspect ratio to quantify elongation of the cluster (Figure 10, see details in Methods section). For example, a perfectly round European football has an aspect ratio of 1

and an American football has an aspect ratio of 2. We found that in FB, where HoxD genes are non-active, the cluster has a round shape with a mean aspect ratio of 1.3. In FL the HoxD cluster is partially active. When we imaged HoxD there, we observed that these clusters have elongated shapes on average with a mean value of aspect ratio 2.3. We, therefore, noticed that the HoxD gene cluster is significantly more elongated in partially active FL tissue compared to non-active FB tissue.

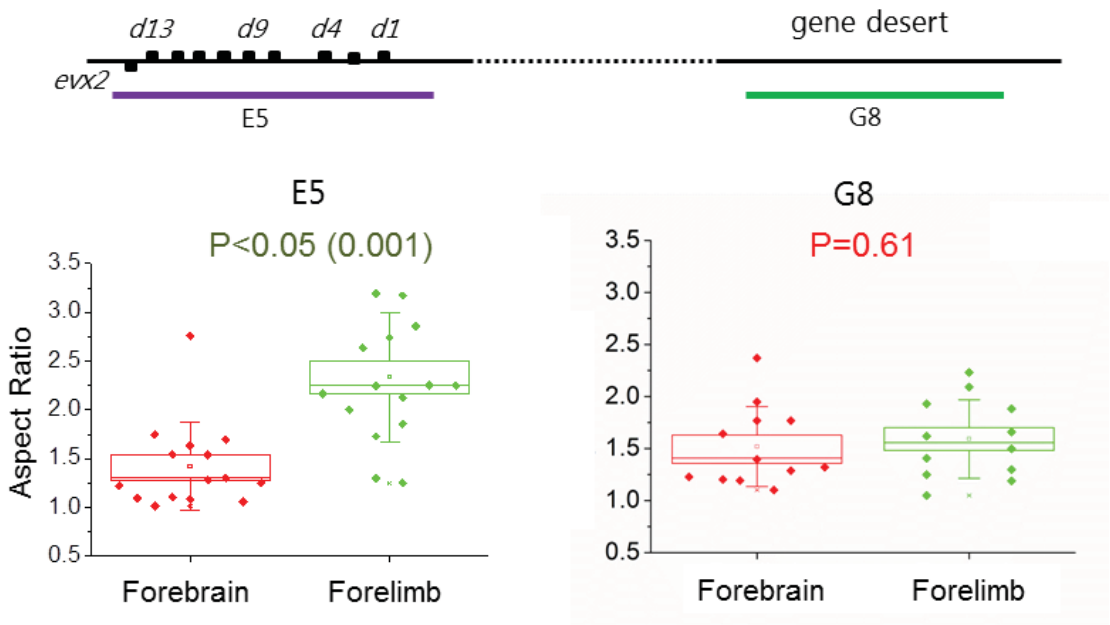


Figure 11: Specificity of HoxD gene cluster elongation

The shape of the HoxD gene cluster (labelled with E5) in FB and FL was compared to the shape of gene desert region of similar size (labelled with RP23-374G8 G8 probe). There is no significant difference in the shape of G8 region between FL and FB illustrating cluster-specific elongation of HoxD (p-values of two-sample t-test reported).

The observed shape difference between FL and FB may be due to differences in the general properties of chromatin, which are not related to HoxD cluster activation itself. In order to account for this, we run control experiment where we look at the shape of control chromatin region (labelled with G8 probe). This control region we chose has a similar size but is gene free, so there is no transcription activation in FL. We observed no significant difference in shape of the control region between FB and FL. It turned out to be quite round with a measured aspect ratio of 1.5 (Figure 11). These data imply that the difference we observed for HoxD is specific to the HoxD cluster and difference in organization can be explained by a gene specific effect rather than a general difference in chromatin properties in the FB and FL.

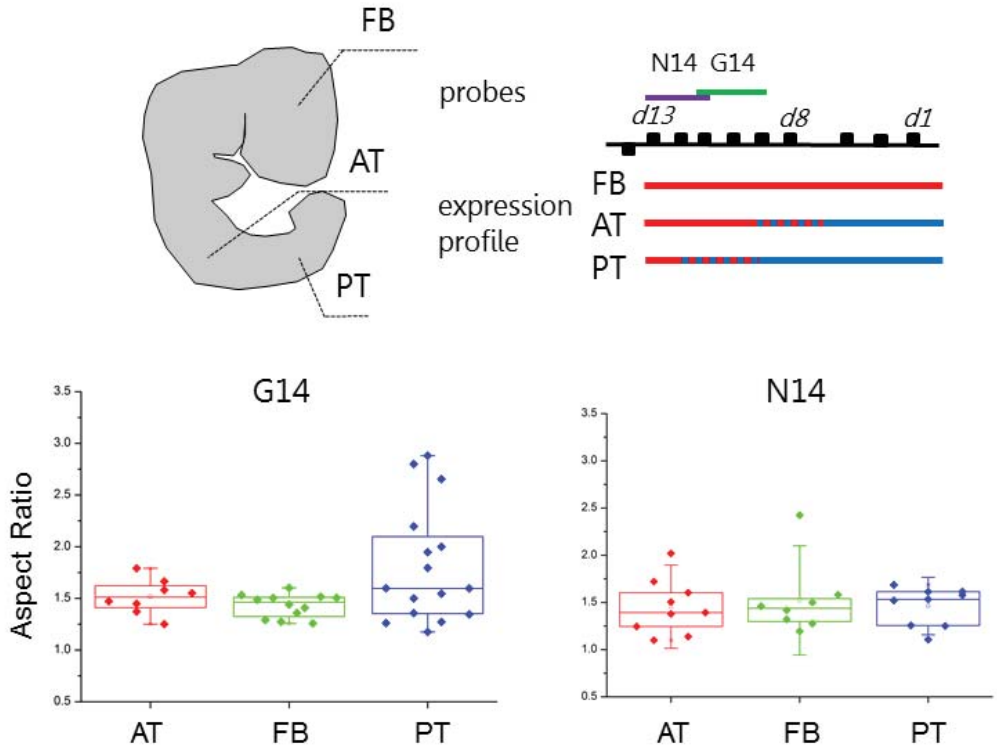


Figure 12: Sub-cluster organization of HoxD gene cluster

(Top) Studied forebrain (FB), anterior trunk (AT) and posterior trunk (PT) areas in developing E10.5 mouse embryo. Expression status of HoxD genes in corresponding tissues. Probes' position within HoxD cluster. (Bottom) The elongation of d13-d11 (N14 probe) and d11-d9 (G14 probe) regions measured as the aspect ratio in FB, AT and PT. The N14 region doesn't show elongation in studied tissues. Note that these genes are non-active there. The G14 region shows elongation in PT. Note that some of its genes become active there.

Next, we imaged HoxD sub-cluster organization using smaller probes (Figure 12). We use them to further study cluster-specific elongation observed with whole-cluster probe E5. We imaged two probes corresponding approximately to regions d13-d11 (N14 probe) and d11-d9 (G14 probe). These probes have a size of 28 kb and 40 kb respectively and were labelled with Alexa647. To study the effect of cluster organization on gene expression, d13-d11 (N14) and d11-d9 (G14) were imaged in tissues with different levels of expression of these genes. We imaged them 1) in a control forebrain tissue (FB) where all HoxD genes are non-active, 2) in the anterior trunk (AT) where active (d8/d9-d1) and non-active (d13-d10/d9) regions exist and 3) in posterior trunk where d11/d10-d1 are active and d13-d12/d11 are non-active (see Figure 12 for visual representation). Therefore genes that are covered by the G14

probe are active mostly in the posterior trunk. Genes covered by N14 are non-active in all three studied tissue types with the exception of d11 and d13 being partially active in a very posterior trunk. We quantified the shape of labelled regions. Similarly to the whole cluster shape we measured an aspect ratio (AR) of sub-cluster regions. We obtained that the d13-d11 region (N14) has a similar AR in the forebrain, anterior trunk and posterior trunk. These values demonstrate that this region doesn't show elongation in these tissue types. As we noted these genes are mostly non-active in these tissues with the exception of the very posterior trunk. We can see that the fact that d13-d11 are mostly non-active in these tissues corresponds to our measurement that region d13-d11 doesn't show elongation. Generally speaking, the non-active region stays round. G14 which represents d11-d9 region gives the following results. It is round in the forebrain and anterior trunk where these genes are not expressed. It gets significantly elongated on average in the posterior trunk where some of its genes are expressed. So we can see that the area of activation of d11-d9 genes coincides with an area where this region gets elongated. In other words, the region gets elongated when in is at the border or within an active sub-cluster region.

Discussion

In this study, we investigated how chromatin organization of the HoxD gene cluster is related to its expression profile. We labelled the HoxD cluster and imaged it with SR microscopy. Using a whole cluster probe we found that in forelimb tissue where HoxD genes are partially active the whole cluster is significantly more elongated when compared to the forebrain where these genes are non-active. We also discovered that this effect is specific to the cluster. When a region of similar size but not containing genes was labelled and imaged in the forebrain and forelimb, we saw no significant difference in shape. These experiments then show that in active tissue the HoxD cluster experiences specific elongation. We further looked at the shape of the HoxD gene cluster at the sub-cluster level. We found that the centromeric region d13-d11 which is non-active in almost the entire embryo has a similar round shape in all studied tissue types. Region d11-d9 has a similar behavior: in the forebrain and anterior trunk where it is not expressed, it is round. So we concluded that if a region is non-active it stays approximately round. We detected that region d11-d9 becomes elongated when we looked at the posterior trunk where these genes are partially

activated. So, in general, we observed that this region gets elongated when it is at the border or within active cluster.

Our results show that at different scales (whole cluster and sub-cluster), transcription activation of HoxD genes is accompanied by elongation of the cluster and active sub-region in particular. In its turn, biochemical 4C data suggest that activation is accompanied by redistribution of genes into two domains: active and passive [14] so that in 3D space active genes contact other active genes and non-active genes mostly contact other non-active genes. Taken together, we therefore propose following model (Figure 13).

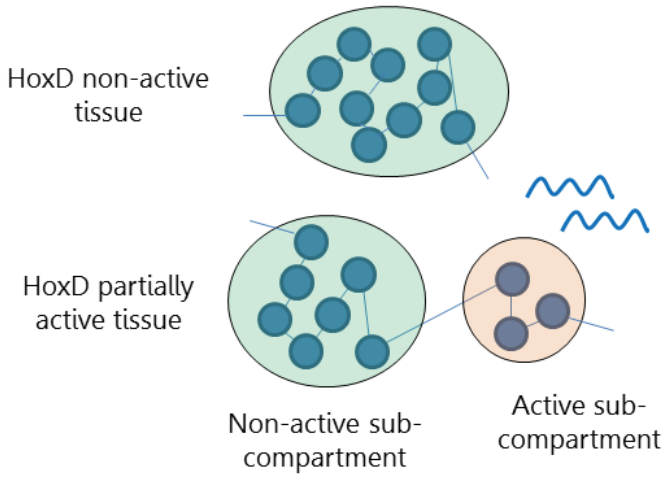


Figure 13: The model of HoxD gene cluster organization

We suggest that in terms of chromatin organization, the HoxD gene cluster is organized in spatial sub-compartments. We think that activation of certain genes leads to the redistribution of HoxD genes into two physically separated spatial sub-compartments with active and non-active genes correspondingly. When a gene gets activated, it is accompanied by its translocation to the active sub-compartment. This compartmentalization and physical separation explains preferences in contact probability observed by 4C and our observation of region-specific cluster elongation upon activation.

This spatial compartmentalization can be advantageous since it is a simple mechanism that can potentially facilitate switching of the right genes to provide the required expression pattern. Possibly, when genes are compartmentalized according

to their activity, specific enzymes which provide transcription activation can access and locate their target in a more efficient way.

Methods

Tissue Preparation

E10.5 mouse embryos were dissected in PBS and then fixed in 4% paraformaldehyde overnight. The next day the embryos were washed in PBS and then further dissected to separate different tissue types. Sectioning was performed using a Microm COOLCUT HM325 microtome to cut 5μm thick slices. Slices were then dried on 25mm round glass coverslips and stored at 4C overnight.

3D DNA-FISH

3D DNA fluorescent in-situ hybridization (FISH) was conducted as described previously [73]. Probes were prepared by nick-translation with biotin- or digoxigenin-UTP using fosmid and BAC clones obtained from BACPAC Resources Center (https://bacpac.chori.org/). 100ng (fosmid) or 200ng (BAC) of probes were used with 5μg of Cot1-DNA and 10μg of sonicated salmon sperm DNA.

SR imaging

Samples were imaged on a custom-built inverted microscope equipped with UPlanSApo 100x/1.40 oil objective (Olympus) and a piezo objective scanner (N-725, PI). Dye molecules were excited by 641nm laser (Cube 640nm-100C, Coherent) with a typical laser irradiance 10-20 kW cm^{-2}. Emitted light was directed through the system of mirrors (Chroma) and Optosplit device (Cairn-research) to EMCCD camera (iXon+, Andor Technology). Samples were imaged in GLOX (0.5 $mg.ml^{-1}$ glucose oxidase, 40 mg ml^{-1} catalase and 10% glucose (all Sigma)) plus 143 μM β-mercaptoethanol (Sigma) buffer with the pH adjusted to 8-8.5 using HEPES (Sigma). First, GLOX buffer was prepared, after 30 min of buffer incubation β-mercaptoethanol was added and pH adjusted. Typically 10000-15000 snapshot images with a pixel size of 100nm with exposure time 0.03s were acquired to create one SR image.

Data analysis

Octane software was used to reconstruct SR images [74]. Peak detection threshold was set individually for each sample set due to variation in image quality. Localizations were grouped as belonging to the same molecule if the distance

between them in space was less than 100 nm and in time equal or less than 2 frames. SR images were created using pixel size of 10 nm; there each localized molecule was represented as a spot with a given localization precision. Initially clusters were identified using standard deviation (STD) images. In STD image each pixel value is a standard deviation of intensity values in this pixel across image stack acquired to reconstruct a single SR image. Only spots with a signal 10x over other objects within nucleus in STD images were considered as labeled clusters; this enabled us to avoid analyzing non-specific staining. The precise shape of clusters was determined by thresholding based on Otsu's algorithm integrated into Fiji software (http://fiji.sc/Fiji). The algorithm sets a threshold by minimizing a weighted sum of variances of the background class and the image class pixels. The shape of the thresholded image then was used to determine the aspect ratio of clusters. The aspect ratio was calculated as a ratio between major and minor axes of ellipse fitted to the shape of the cluster. To evaluate difference between samples statistically we used two-sample t-test for means.

4 Multicolor single molecule tracking of stochastically active synthetic dyes

Introduction

Dynamic measurements of single molecules in living cells offer a window into nano-scale processes obscured by conventional imaging methods. Single-molecule tracking in living cells has been improved by exploiting temporal separation of signals to allow up to thousands of molecules to be tracked in a single cell in the time span of minutes [75]. This has been implemented using two different approaches to fluorescently label proteins of interest : either via protein fusions to photoswitchable or photoactivatable fluorescent proteins [75] or more recently via ligand binding [76]. In the case of protein fusions, there are few constraints on the nature of the protein of interest; since dyes are genetically encoded proteins they can be studied in virtually any cellular compartment. However, the relatively low photon yields of fluorescent proteins limit the time scales over which a given molecule can be tracked, and the accuracy with which it can be localized. In the case of ligand binding, ligands are conjugated to synthetic dyes that have high photostability and long lifetimes, but limit the possible proteins studied to extracellular receptors with efficiently targeted ligands. Moreover, the pool of unactivated receptors without a ligand bound is invisible to this method. Here, we combine the advantages of protein target flexibility and dye properties by genetically encoded tags to target synthetic dyes as a means of probing the dynamics of membrane and intracellular proteins.

Genetically encoded protein tags enable rapid and specific protein labelling via the formation of a stable, covalent bond between the protein and the label [77]. Thus, they offer an interesting alternative to other forms of labelling by fluorescent proteins or ligands. Unlike antibody labelling, they are also compatible with live-cell imaging. Self-labelling proteins such as the SNAP-tag [62] and Halo-tag [78], which are smaller than fluorescent proteins, are fused to the protein of interest and react directly with the labelling compound, benzylguanine (BG) or Halo. These two tags have been used in live cell imaging, including super-resolution imaging [59, 67, 79, 80]. Enzyme-mediated tags are even smaller than self-labelling proteins and have higher specificity, but require the addition of an enzyme to covalently link the label to the tag and therefore can only be used extracellularly [77]. We combine both self-labelling

proteins and enzyme-mediated tags for targeting synthetic dyes, which we then use to track single molecules with multiple colours and in different cellular compartments. To track thousands of molecules in a single cell, we take advantage of the photophysical processes that allow fluorophores to cycle through reversible dark states. These processes allow us to control the number of molecules that are in a fluorescent state at any given time. This is the same mechanism for temporal separation of fluorescent signals exploited to create super-resolution images [47, 48, 50].

Results

Temporal separation of single molecule signals can be obtained by a number of different mechanisms, including chemical control of the rate of recovery from a dark state [57]. A combination of b-mercaptoethanol (BME) and oxygen scavengers was originally used to control the dark states of Cy5 dyes and enable point-localization of single molecules. Since then, other thiols such as cysteamine (MEA) have been used in combination with oxygen scavengers to induce the blinking of several different dyes [81, 82]. Although live-cell imaging is possible for a limited time in such a buffer, we restricted our screen to combinations of less cytotoxic components previously used for super-resolution imaging, including the reducing agents glutathione [58, 83, 84] and ascorbic acid [65, 85]. We tested buffer conditions using cells transiently transfected with the b2 adrenergic receptor coupled to the SNAP-tag, SNAP-b2AR, and labelled with Alexa 488. Cells were exposed to moderate levels of excitation light to maintain cell viability while effectively putting most of the fluorophores in the dark state, and a stack of images was collected. These images were analyzed to localize single molecules by Gaussian fitting. At the same time, we could extract their photon counts in each frame, a measure of their brightness (Figure 14B). The molecular positions were then subjected to an automated tracking algorithm [75, 86] to reconnect molecular trajectories as a function of time (Figure 14A): molecular localizations were connected as belonging to one molecular trajectory if they were present in consecutive frames and within a distance calculated based on the expected diffusion coefficient of the protein. To ensure correct trajectory reconnection, the density of fluorescent molecules per frame was kept low. Track densities, as defined by the number of tracks per area lasting more than 8 frames, were compared for different buffers (Figure 14C). We found that by both quantitative

measures (brightness and track density), a combination of Leibovitz medium, glutathione and a low-concentration enzymatic oxygen scavenger system composed of glucose oxidase and catalase was optimal for Alexa 488. Under these conditions, on average more than 400 photons were detected per localization, and more than 16 trajectories per square micron were reconstructed over a series of 5000 image frames. Surprisingly, Leibovitz medium alone yields a high number of tracks for Alexa 488 without requiring a complex mixture of buffers (Figure 14C); this is a promising result since Leibovitz is frequently used as a medium for imaging live cells in culture over a span of hours in the absence of CO_2. Therefore, we subsequently used the Leibovitz buffer as a starting point for testing other dyes for multicolor tracking applications.

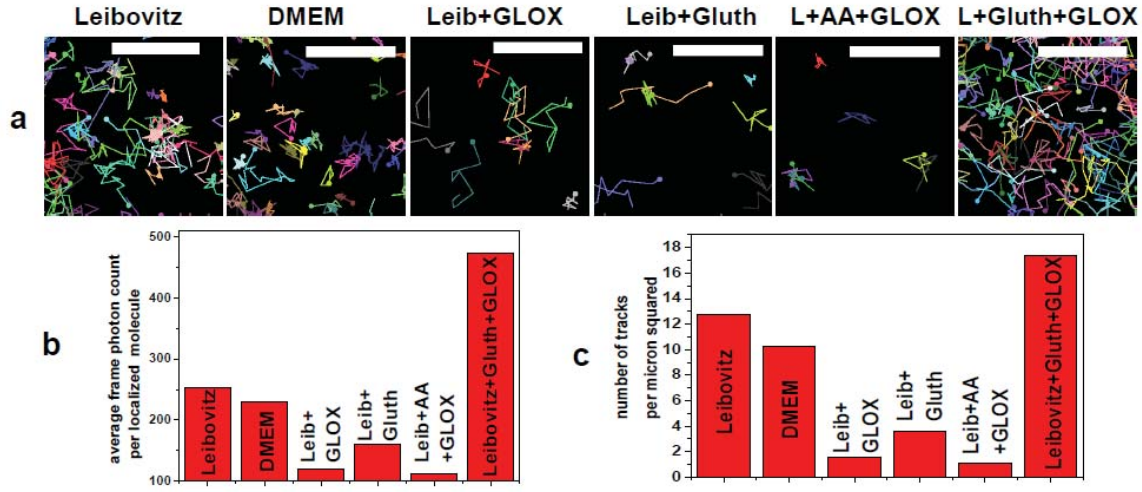


Figure 14: Optimizing buffer conditions for tracking of Alexa 488

(a) Example of typical SNAP-b2AR tracks obtained in the different buffer conditions tested on a 2x2 micron area (scale bar = 1 µm). The starting point of each track is indicated by a dot. (b) Average number of detected photon per track point (threshold set at 100 photons) and (c) average track density per micron squared (minimum track size: 8 points) measured in 5000 frames for the different buffer conditions.

The primary obstacle to make several fluorophores blink simultaneously is to identifying a buffer that is appropriate for all the dyes used. In super-resolution imaging, this balance is necessary to achieve both high localization accuracy and a high density of molecules, as required by the Nyquist criteria [66]. In the case of tracking, live imaging further restricts the possible buffer conditions. On the other hand, the Nyquist criteria no longer applies in single particle tracking, since the goal is rather to obtain a statistically significant number of molecular trajectories and does

not require complete sample coverage. As a consequence, it is acceptable to use a buffer that yields a lower density of molecules than required for super-resolution imaging. Conversely, the constraints on the on-times of dyes are more stringent in single molecule tracking, as one needs sufficiently long on-times relative to the acquisition rate to obtain long trajectories of single molecules. Optimum conditions are therefore different for imaging and tracking.

We tested buffers and dyes for multicolor live-cell imaging using cells transfected with SNAP- b2AR, and labelled with either Alexa 488, TMR*, Dy-547, Alexa 647 or Dy-647. In screening a range of buffers, we found that Leibovitz medium allowed us to track dyes with emission in the green, red and far-red parts of the visible spectrum (Figure 15). Interestingly, Alexa 647, which is one of the best dyes for super-

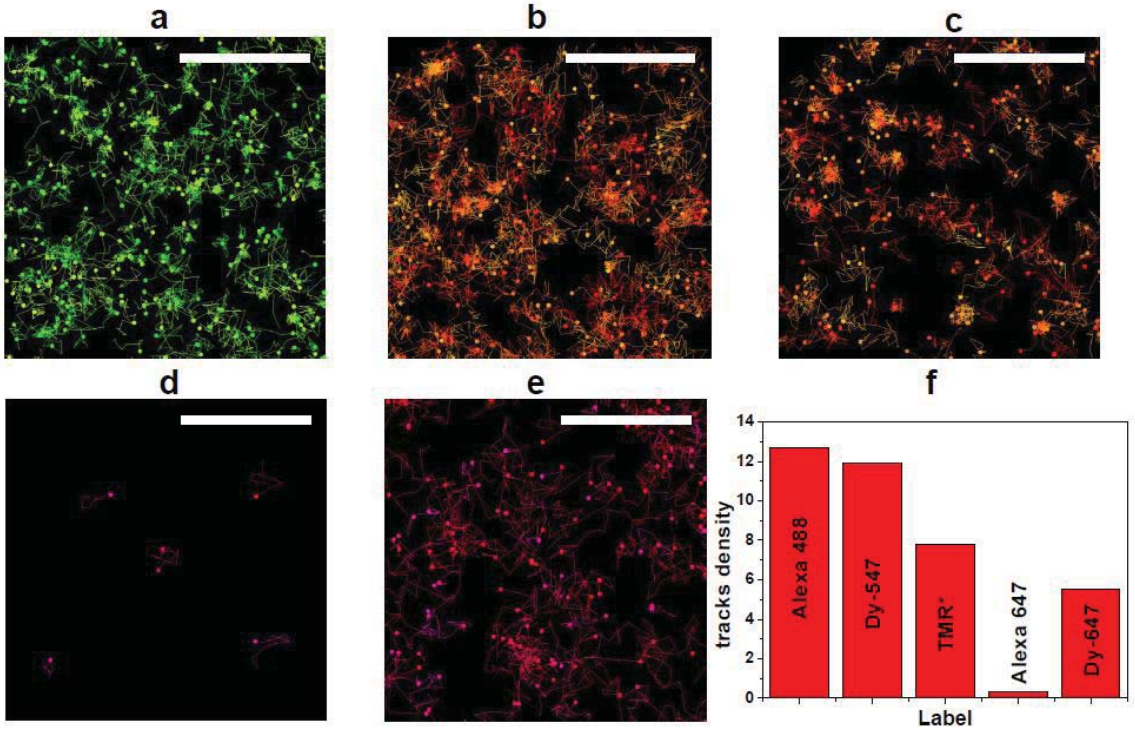


Figure 15: Tracking in Leibovitz media with different dyes

Tracks of b2AR receptors fused to a SNAP-tag and labelled with (a) Alexa 488 (b) Dy-547 (c) TMR* (d) Alexa 647 (e) Dy-647. The starting point of each track is indicated by a dot. (f) Track density obtained for 5000 frames in Leibovitz media for these different dyes (minimum track size: 8 points). Scale bar=2µm.

resolution imaging [82], turns out to be a poor dye for tracking performed in Leibovitz. This underscores the difficulty in predicting which dyes will perform well in a live cell setting based on in vitro or fixed cell data.

To enable multicolor single-molecule tracking, it is necessary to target dyes with well-separated spectra to different proteins using orthogonal labeling schemes. Thus, we tested the SNAP-tag, Halo-tag and ACP-tag fused respectively to b2AR, transferrin receptors (TfR), or a glycosylphosphatidylinositol anchor sequence (GPI). There was no detectable difference when the same dye was targeted to a different tag on the same protein, indicating that the tag itself does not play an important role in dye performance. We note that the combination of Alexa 488 (Halo), TMR* (SNAP) and Dy-647 (ACP) represents three orthogonal tags that were fused to spectrally separated dyes, which all performed well in Leibovitz buffer alone (Appendix Figure 2).

Once compatible blinking buffers and orthogonal labeling schemes have been established, it becomes possible to perform multicolor tracking. We co-transfected cells with the plasma-membrane receptor proteins SNAP- b2AR and Halo-TfR. Cells were co-labeled with Dy-547 (SNAP) and Alexa 488 (Halo) and imaged sequentially. Overlaid maps of trajectories from each species reveal distinct zones of occupation for each protein (Figure 16C), suggesting spatial decorrelation. A closer analysis demonstrates that the distribution of mobilities as parametrized by the diffusion coefficient is similar for each receptor, with values in agreement with those measured using other methods [87, 88]..

Imaging buffer conditions are crucial for controlling dye blinking. Traditionally used buffers usually contain BME or MEA which are cytotoxic, so it is challenging but important to find conditions which would be good both in terms of dye blinking and cell viability for live cell imaging and tracking. We tested the different imaging buffers used in this work to quantify their cytotoxicity by performing the MTT survival assay [89]. Cells were exposed to a buffer for approximately 2-3 hours and the MTT test was run 24 hours later. Our results indicate that cells incubated in all imaging buffers remain viable at least for 24 hours after incubation and the number of cells in the system was not significantly decreased compared to control experiments (Appendix Figure 3). From this, we conclude that the buffer conditions used here are live-cell compatible.

An even greater challenge arises in performing single-particle-tracking on proteins located intracellularly. Because the local environment in a living cell is determined largely by cellular homeostasis, it is difficult and probably not desirable to strongly perturb these conditions. Thus, creating the appropriate environment for dyes to blink

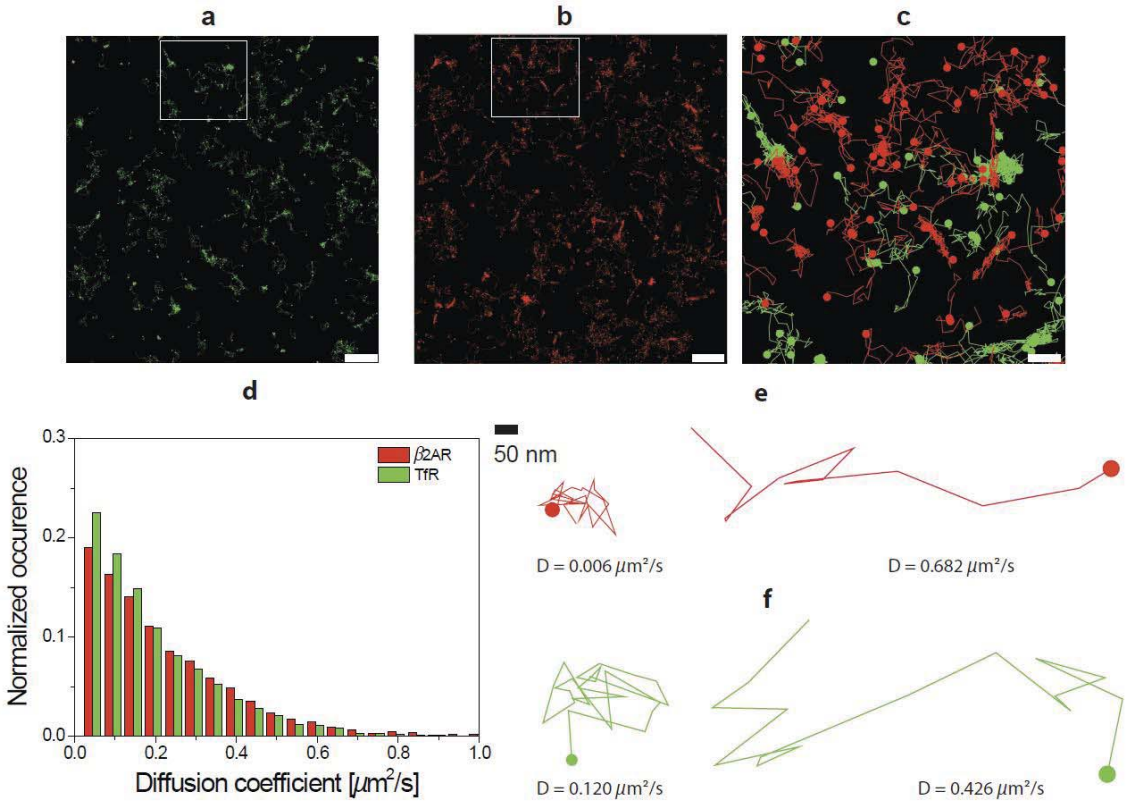


Figure 16: Two-color tracking of membrane receptors

(a) Tracks of TfR-Halo-Alexa 488 (b) tracks of SNAP-b2AR-Dy-547 (scale bars: = 2 µm). The starting point of each track is indicated by a dot. (c) Zoom of the overlaid tracks in the region of interest delimited in (a) and (b) (scale bar = 500 nm). (d) Distribution of the diffusion coefficient of the mobile fractions (D > 0.02 mm2/s) of the two receptors (TFR:4770 tracks, b2AR: 2120 tracks). Example of individual tracks of b2AR (e) and TfR (f) and associated diffusion coefficient.

depends heavily on naturally occurring cellular processes. Moreover, relatively few dyes are cell permeable, further limiting the possibilities for tracking on intracellular targets: of the handful of commercially available cell-permeable dyes that can be targeted using the SNAP-, CLIP-, or Halo-tags, we focused on three, Oregon Green, Dy-505, and TMR*.

Mitochondria are enclosed by two membranes, with a proton gradient across the inner membrane that drives the synthesis of cellular ATP via the citric acid cycle.

They represent a unique chemical environment, which could impact the blinking of dyes within the compartment. It was recently demonstrated that the spontaneous blinking of TMR* in mitochondria allows tracking of Halo-tagged proteins in the inner or outer mitochondrial membrane [80]. We transfected cells with the inner mitochondrial membrane protein Cox8A-SNAP, and found that Cox8A-SNAP-TMR* labelled mitochondria specifically (Figure 17A), and that under optimized imaging conditions (Supplementary Methods) it blinked well and yielded a high density of molecules that could be rendered as a super-resolution image (Figure 17B). A large number of tracks were reconstructed from the molecular localizations, and allowed us to create a spatial map of the determined mean diffusion coefficient. This map reflects the limited spatial variations of the diffusion coefficient, except at the edge of mitochondria where the reduced diffusion could be due to the fact that we measure 2D diffusion in a 3D sample (Figure 17CD).

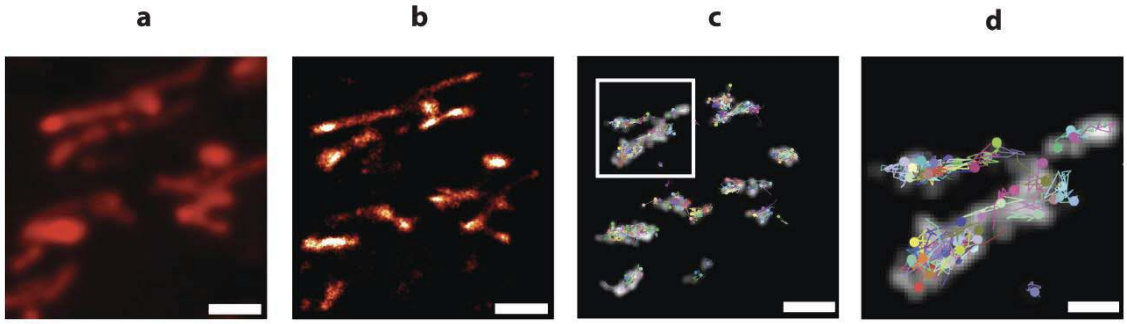


Figure 17: Intracellular tracking in mitochondria with TMR*

(a) Wide-field image of Cox8A-SNAP stained with TMR*. (b) Reconstructed super-resolution image. (c) Overlay of tracks from re-connected single molecule localizations and the measured map of diffusion coefficients color-coded in gray (see Supplementary Methods). Higher intensities correspond to higher diffusion coefficients. The starting point of each track is indicated by a dot. (scale bars = 2 µm). (d) Zoom of the region of interest delimited in (c) (scale bar = 700 nm).

The nucleus represents another cellular compartment with unique chemical characteristics. We tested the possibility of tracking dyes fused to proteins in the nucleus. Cells were transfected with the histone protein H2B-SNAP and labelled with Oregon Green, Dy-505, or TMR* which were previously used to demonstrate live cell super-resolution imaging of H2B [59, 60, 90]. We found that all three dyes could be made to blink, and we used them to both create live-cell super-resolution images (Figure 18A) and track the position of H2B proteins (Figure 18BC). We found the protein to be essentially immobile within the nucleus, confirming predictions based on

fluorescence recovery after photobleaching experiments, which showed no recovery over a timescale of hours [90]. Although all three dyes could be tracked at high density, the number of molecules localized over time was the most stable for TMR*, indicating its superior performance. Intriguingly, when these same dyes were targeted to cytoskeletal proteins, they did not undergo the sustained blinking necessary for high-density tracking in the imaging conditions used here. Such a striking difference in dye properties between the cytosol and the nucleus can be attributed at least in part to the differences in their chemical environments. Cellular compartments may contain different levels of natural reducing agents as well as molecular oxygen. This can have a strong impact on dye photoswitching behaviour, and could explain the observed differences in blinking.

Discussion

Here we have demonstrated that synthetic dyes in different regions of the spectrum can be made to blink in a single buffer, chosen for its compatibility with single molecule tracking, and long-term imaging in live cells. Targeted to specific proteins with orthogonal tags, these dyes can then be combined for multicolor imaging. With the dyes, tags and buffers described here, up to three different species of proteins can be tracked on the surface of a single cell. Intracellular imaging is also possible with this method, as we show with mitochondria and the nucleus, illustrating the flexibility of this approach to high-density molecular tracking. We also demonstrated the possibility of two-colour single-particle-tracking in the nucleus. We expect that an increasing number of dyes will be designed and synthesized to be cell-permeable and compatible with live-cell imaging, further extending the range of possible protein targets. The future of this method is to combine the best features of existing techniques in high-density single molecule tracking, to allow the maximum flexibility in choice of proteins and colours.

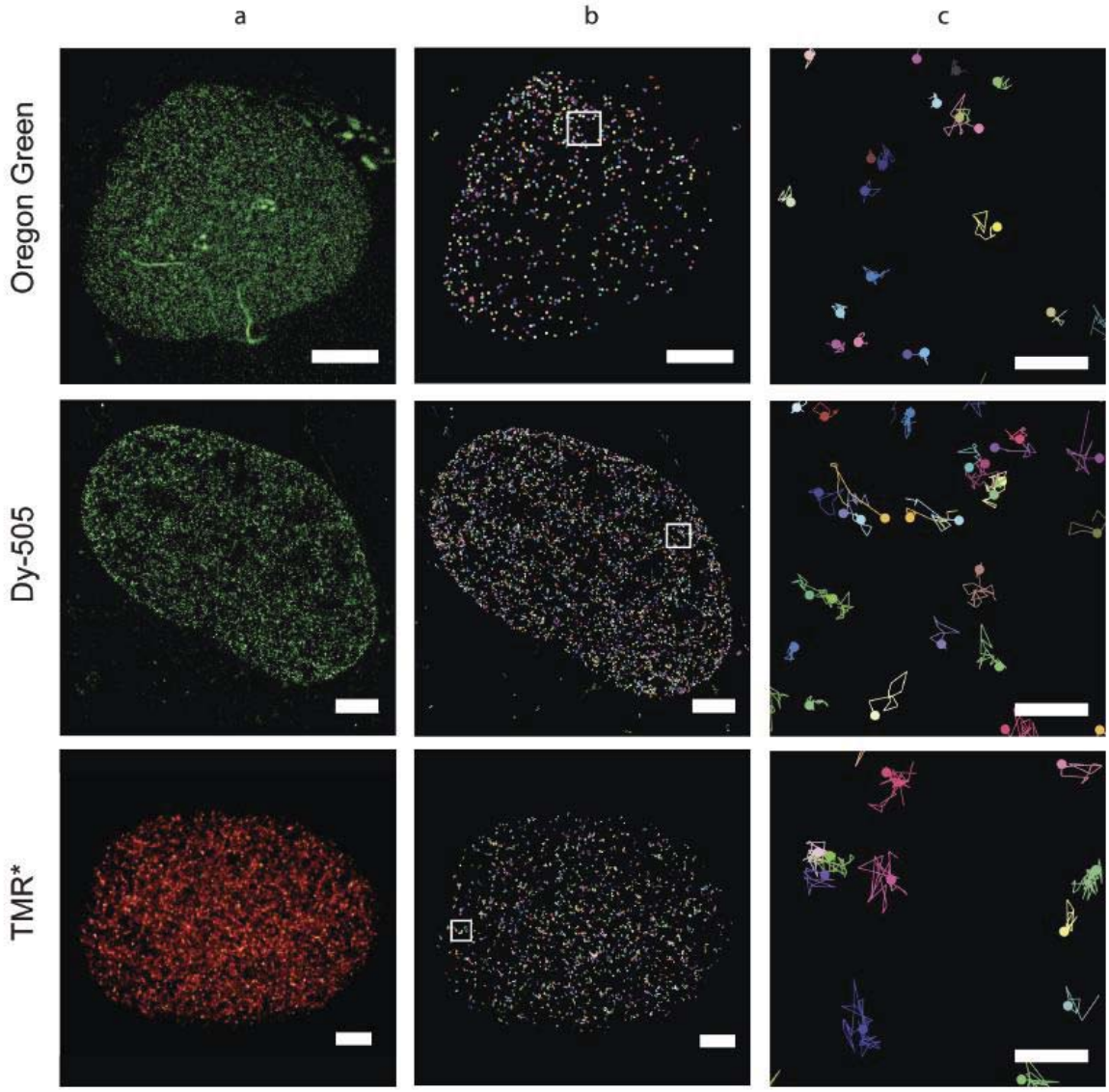


Figure 18: Intracellular tracking in nucleus with cell permeable dyes

(a) Reconstructed super-resolution image of H2B-SNAP stained with different dyes (scale bars = 2 mm). Overlay of tracks from re-connected single molecule localizations (b) for the whole nucleus (scale bars = 2 μm) and (c) for zoom of the boxed region (scale bars = 100 nm). The starting point of each track is indicated by a dot. Different rows correspond to cells labelled with Oregon Green, Dy-505 and TMR* (from top to bottom).

Methods

Microscopy

Imaging was performed on a Zeiss Axio Observer D1 inverted microscope using a 100x 1.46 objective (Zeiss) and continuous illumination from 3 different lasers: 488 nm (Sapphire, Coherent), 561nm (Crystal Laser) and 640nm wavelength (Cube 640nm-100C, Coherent), with a power at the sample of 0:1-1kW/cm^2 depending on

the dyes used. Fluorescence emission was detected with an electron-multiplying CCD camera (iXon+, Andor Technology) with a pixel size of 160nm. The filters used were a 4 color dichroic 89100bs (Chroma) with three different emission filters: ET525/50 (Chroma) for Alexa 488/Dy-488/Oregon Green/Dy-505, ET605/70 for TMR* and Dy-547, and a ET700/75 (Chroma) for Alexa 647 and Dy-647. To increase stability, the microscope is equipped with a piezo objective scanner (N725, PI) and a linearly encoded sample stage (PRIOR).

Sample preparation for tracking

U2OS cells were cultured in DMEM (Gibco) with GlutaMAX-1 (Invitrogen) supplemented with 10% FBS (fetal bovine serum, Gibco) in a humidified 5% CO_2 incubator at 37C. Cells were plated 48 hours before imaging on 25mm round size one coverslips (Mentzell) cleaned in pure Ethanol (Sigma) and flamed. Imaging was performed 24 hours after transient transfection (2mg DNA+3mL Fugene6 - Roche) and were stained (see below for details) just before imaging. Imaging was performed in Lebovitz medium (Gibco) (except for figure1, see below) at room temperature (24C) by recording 5.000 frames at 30 frame per second.

Buffer composition (Figure 14)

-"DMEM" is phenol-free DMEM (Gibco)

-"Leib+GLOX" is Leibovitz with 0.5 $mg.ml^{-1}$ glucose oxidase, 40 $mg\ ml^{-1}$ catalase and 10% glucose (all Sigma)

-"Leib+Gluth" is Leibovitz with 100mM Glutathione (LuBio Science)

-"L+AA+GLOX" is Leibovitz with 1mM Ascorbic Acid (Sigma)

-"L+Gluth+GLOX" is Leibovitz with glutathione 50mM, 0.25 $mg.ml^{-1}$ glucose oxidase, 20 $mgml^{-1}$ catalase and 5% glucose.

All buffers pH 7.2-7.6.

Sample staining for tracking

SNAP: Cells were stained for 30-60min with 250-500nM BG-fluorescent substrate (NEB or Covalys) in DMEM and washed with PBS. Cells were then incubated for 30 min in phenolfree DMEM without BG substrate and washed again.

Halo: Cells were labelled for 30-60min with 1-5mM Halo-tag ligand (Promega) in DMEM and washed with PBS. Cells were then incubated for 30 min in phenol-free DMEM without a substrate and washed again.

ACP: Cells were stained for 30-60min with 1-5mM CoA-fluorescent substrate (NEB) in phenol-free DMEM containing ACP synthase and $MgCl_2$ and washed once.

Data analysis

The data is processed in two steps similarly to sptPALM: First molecules are localized in each frame using a previously described algorithm written in IDL (Research Systems, Inc., Boulder, CO) (Peakselector, courtesy of Harald Hess). The localized peaks were then tracked using a tracking program written in Matlab (The Mathworks, Inc., Natick, MA). Peaks are identified as belonging to the same trajectory when they are present in consecutive frames, and are within a user-defined distance r. All subsequent analysis is done with custom-designed Matlab routines: Diffusion coefficients are calculated by measuring the mean squared displacements for all trajectories containing at least 10 steps, and then performing a linear fit using the first 6 time lags. Tracks are rendered by choosing a random color, displaying the first point as a dot and linking the successive steps with a line. The diffusion map is reconstructed by gray-scaling each point of a given trajectory according to the measured, weighted diffusion constant onto a 50 nm grid. For better visibility a 3x3 window around each trajectory point is gray-scaled. Higher gray values correspond to higher diffusion coefficients. The mean value is displayed if trajectories cross in a given pixel. The image is rendered median-filtered (3x3 window) and Gaussian-blurred (s = 5 pixels). TetraSpeck beads (Invitrogen) imaged in the green and red channel were used for linear chromatic shift correction.

Cell viability assay

Cells were plated in 96-well plate and incubated overnight in DMEM GlutaMAX-I with + 10% FBS. Next day cells were washed and incubated in different buffers for 2 hours mimicking imaging conditions. Then they were washed again and cultured for further 24 hours in standard medium. We performed MTT assay based on detection of reductase enzymes activity to assess the cytotoxic effect of the imaging buffer. These enzymes reduce MTT in living cells to formazan which is detected with

fluorescence plate reader. Total fluorescence signal is then proportional to number of living cells.

5 Analysis for live cell super-resolution imaging adapted for photoswitching properties and molecular dynamics

Introduction

Fluorescence imaging with photoactivated localization microscopy (PALM) [47, 48] or stochastic optical reconstruction microscopy (STORM) [50]can be used to produce super-resolution (SR) images. A key part of these and related methods is to induce the temporal separation of fluorescence from molecules in close proximity by stochastic photoswitching. Thus, in each imaging frame, fluorescent spots corresponding to the images of single molecules can be observed. These frames are then analysed to localize molecular positions, which are combined and rendered to form a reconstructed SR image. In fixed samples, a combination of localization precision and the density of localized points primarily determine image quality. Thus, to sample biological structures well, several thousands of frames must typically be acquired [66].

Live cell SR imaging aims to visualize dynamic objects, so it additionally requires rapid image acquisition and rapid photoswitching. If the frame rate is too slow, molecule motion can give rise to an underestimation of the size of concave structures, or fill in the contours of convex structures [91, 92]. Photoswitching should also be rapid to provide sufficient sampling of dynamic structures. For live-cell structural imaging, ideally, switching and imaging rates should be commensurate. Indeed, if the cycling rate between bright and dark states is much faster than the imaging rate too many molecules may be on simultaneously and the signal-to-noise ratio degraded, while if the on time is much longer than the exposure, the same molecule will be imaged over several frames.

There are two main classes of fluorophores used in live cell SR imaging: synthetic dyes [57, 60, 67] and photoactivatable fluorescent proteins (PA-FPs) [66, 93]. These fluorophores demonstrate significantly different photophysical properties. PA-FPs allow tight control of photoswitching rates to maintain the desired density of molecules per frame, but have relatively low photon yields. Synthetic dyes have high photon yields, but at the cost of reduced control over photoswitching rates in live cell SR imaging, since few photoswiching buffers have been identified that are not toxic for cells. For both classes of fluorophores, the presence of long-lived fluorescent

states spanning multiple acquisition frames requires a peak grouping procedure, as previously described for fixed cells [75, 94-96]. However, in living cells one should adapt the analysis to account for real molecular motions and take into account poor control of photoswitching for synthetic dyes. We show that the distinct photophysical properties of PA-FPs and synthetic dyes can impact the manifestation of cellular structures and dynamics in live SR imaging, and further outline a procedure designed to identify and address potential artifacts.

Results

To investigate how fluorophore properties can impact live-cell SR imaging, we imaged the histone protein H2B (Appendix Figure 4) – part of the nucleosome complex, which packages nuclear DNA into sub-diffraction limited structures. Cells

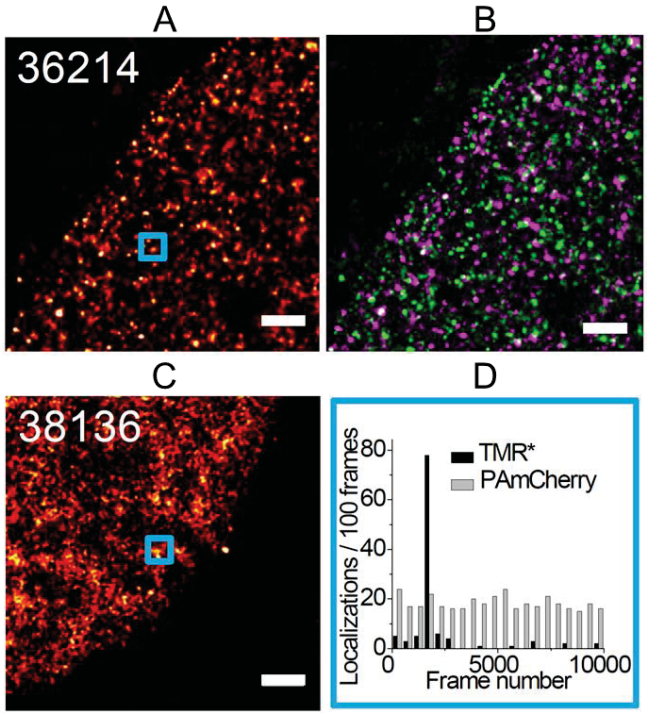


Figure 19: SR imaging of H2B in live U2OS cells

(A) SR image of live U2OS cell expressing H2B-SNAP labelled with TMR* (scale bar 1 µm). The number of peaks localized in the image is shown in white. 10,000 raw images were acquired at 30 frames per second. (B) Time-lapse image of H2B-TMR* shows dynamics of raw localizations. The first half of all localizations is displayed in magenta and the second half in green. (C) SR image of live U2OS cell expressing H2B-PAmCherry (scale bar 1 µm). The number of peaks localized in the image is shown in white. 10,000 raw images were acquired at 30 frames per second. (D) Distribution of localizations per 100 frames over the acquisition period from boxed regions in A and C.

expressing H2B-SNAP labeled with widely used in live cell SR imaging synthetic dye BG-TMR* (H2B-TMR*) revealed spatially separated, well-resolved clusters ~50-100 nm in size (Figure 19 A). These clusters appear highly dynamic when rendered in a color-coded time-lapse sequence where chronologically first half of localized points in the image is displayed in magenta and second half in green (Figure 19 B). Indeed, the majority of the clusters are single-colored and therefore present only for part of the acquisition time which implies that they are dynamic. This was surprising and required further examination to find a more plausible explanation, since we didn't expect to observe histone dynamics for several reasons. Firstly, this rapid turnover of structures contradicts the slow histone dynamics previously measured with FRAP and FCS [97]. Furthermore, when we perform SR imaging of H2B-TMR* in fixed cells, clusters still appear dynamic (Appendix Figure 5) which is an obvious artifact. Notably, these dynamic clusters were previously reported for H2B-eDHFR labeled with TMP-ATTO655 imaged in live cells with STORM imaging [60]. In contrast to H2B-TMR*, images of the PA-FP fusion H2B-PAmCherry revealed a static meshwork of fibers (Figure 19). This organization is consistent with STORM SR imaging of DNA itself which we previously reported [85], our stimulated emission depletion (STED) images of DNA (Appendix Figure 6) and the fibrous high-order structure of packaged chromatin [98].

To understand the discrepancy in observed H2B organization in living cells, we examined the switching behavior of TMR and PAmCherry at the single-molecule level. We chose sub-regions of equal size (~0.1um^2) in Figure 19 A and C, and plotted the time course of the number of localized single molecules per frame (Figure 19 D). While H2B-PAmCherry localizations were distributed throughout the acquisition period, the majority of H2B-TMR* localizations occurred in a narrow time window. This potentially indicated single molecules with long fluorescent on-times spanning multiple acquisition frames.

To identify molecules with long on-times we performed single-particle tracking [75, 99] (Figure 20 A,B). Tracking algorithms can be used to connect spatially and temporally correlated signals in an automated fashion. In this case, tracking revealed that there are many H2B-SNAP TMR* trajectories that are longer than several frames (Figure 20 C). In particular, more than 5% of tracked molecules have trajectories longer than 30 frames. As a consequence, in Figure 19 A many single molecules will

be localized several times and therefore represented by several spots, which manifest as clusters that appear and disappear as molecules are switched between long-lived on and off states. Thus, these photophysical properties give rise to false structures (clustering) and false dynamics (appearance and disappearance of structures). Conversely, there are few tracks exceeding several frames in the case of H2B-PAmCherry (Figure 20 C), and therefore they have a negligible effect on the structure.

To correct for molecules that are on for several frames we grouped localizations in time and space as belonging to one molecule. A careful choice of parameters is important, since unlike in fixed cells, some local motion of the living system in

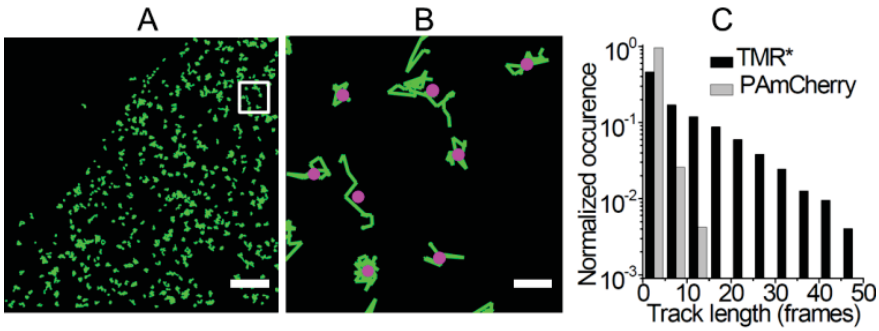


Figure 20: Single molecule tracking of H2B-TMR*

(A),(B) Successive locations along the trajectory are linked with a green line. Only tracks longer than 10 frames are displayed. Magenta dots in B (zoom of boxed region in A indicate positions of molecules after raw localizations were grouped as described in the text. Scale bars 1 μm and 100 nm respectively. (C) Distribution of track lengths for H2B-TMR* and H2B-PAmCherry.

addition to drift from the imaging setup is to be expected. Therefore, an empirical determination of intrinsic motion is a logical part of the grouping procedure for live-cell SR. We used the distribution of single step molecular displacements from the tracking to define a radius within which to group these localizations (see Appendix Figure 7 and Supplementary Note). Grouping of localizations naturally decreased number of localizations per reconstructed PALM/STORM image. Due to longer on-states we observed larger drop in the relative number of localizations of H2B-TMR* reflected by the ratio of grouped versus frame localizations represented in Figure 21 A. In the case of PAmCherry one molecule is counted on average approximately 1.5 times, while in the case of TMR* the average is ~3 times. When we rendered localizations belonging to one molecule as a single peak, H2B-TMR* dynamic

clustering artifacts disappeared (Figure 21B and Appendix Figure 8). Meanwhile, the H2B-PAmCherry structure didn't change significantly when grouped localizations were used for rendering (Figure 21C).

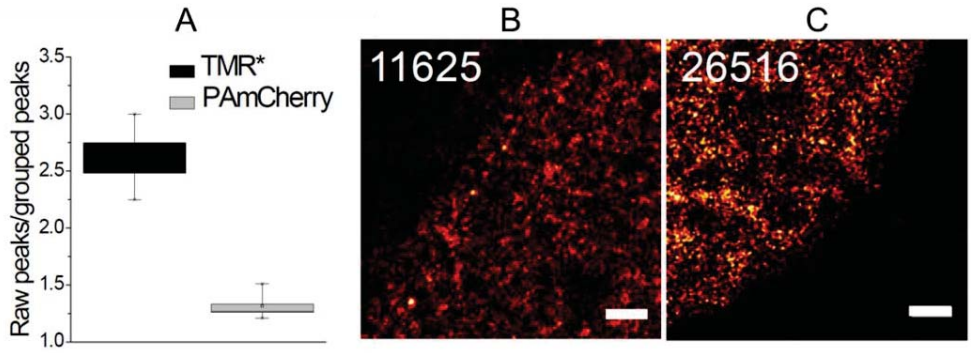


Figure 21: SR images of H2B in live U2OS cells corrected by groping procedure

(A) Ratio of raw peaks to grouped peaks for TMR* and PAmCherry images (n = 5 cells; whiskers range: min-max). (B),(C) SR images of live U2OS cells expressing H2B-SNAP labeled with TMR* or expressing H2B-PAmCherry respectively(scale bars 1 µm). Here, grouped peaks were used for rendering. The number of peaks is shown in white.

Our results demonstrate that the choice of fluorophores can impact the structure and dynamics observed in live-cell SR images. We identified a primary reason for these differences: single molecule photoswitching behaviour of fluorophores. Specifically, long-lived fluorescence states of individual molecules can lead to artifacts due to multiple localizations of single molecules. Such artifacts should be avoided, or corrected using appropriate grouping procedure. When imaging fixed samples, one can sometimes decrease the on- times of synthetic dyes by optimizing the imaging buffer or/and increasing laser excitation power. However, in live-cell SR imaging the use of buffers containing toxic reducing agents and high laser powers compromises cell viability. We therefore implemented a peak grouping procedure that uses the measured molecule motions to define a grouping radius. This procedure is intuitive and easy, and it inherently avoids the problems of an arbitrary grouping radius (Appendix Figure 9): overgrouping that can arise from erroneous linkage between distinct molecules if the grouping radius is too large, and undergrouping that fails to eliminate false dynamics if the grouping radius is too small.

Discussion

In conclusion, the ability to control PA-FP photoswitching makes their usage more straight-forward in live-cell SR imaging compared to synthetic dyes. However,

synthetic dyes offer a distinct advantage due to their superior brightness, which permits molecules to be localized with higher precision, and allows for longer trajectories when used in single molecule tracking applications. As we have shown, they are also suitable for live SR microscopy if corrected using a procedure such as the one defined here.

Methods

Cell culture and labeling

U2OS cells were cultured in GlutaMAX-I medium (Invitrogen) supplemented with 10% FBS (Gibco). Cells were trypsinized (Sigma), plated into a new flask and supplied with fresh medium every 2-3 days. For imaging they were placed on 25 mm size 1 coverslips (Mentzell) and transfected using FuGene6 transfection reagent (Roche) on the next day. ~24 hours after transfection, cells were imaged (H2B-PAmCherry) or stained with SNAP-tag substrate BG-TMR* and imaged (H2B-TMR*). For staining, cells were incubated with 500 nM BG-TMR* (NEB) for 30 minutes in PBS, washed and further incubated for another 30 minutes. Imaging was performed in Leibovitz medium (Gibco) at room temperature (~24C).

Live-cell imaging

Images were acquired on an IX71 inverted microscope (Olympus) using a 100x 1.4NA oil objective (UPLANSAPO, Olympus). Light from 405 nm (Cube, Coherent) and 561 nm lasers (Sapphire, Coherent) was reflected by a 4-color dichroic (89100BS, Chroma) and focused into the sample with an excitation intensity < 0.05 kW cm^{-2} and ~5 kW cm^{-2} respectively. Emitted fluorescence was collected through an emission filter (HQ605/70, Chroma) and detected by an EMCCD camera (iXon+, Andor Tech) with a resulting pixel size of 160 nm. Samples were mounted on a xyz-piezo stage to minimize drift (PINANO, PhysikInstrument). Typically, 10,000 raw images were acquired at 30 frames per second.

Data analysis

Fitting: Extraction of single molecule localizations for SR images was done using a software written in IDL (Research Systems, Inc.) (Peakselector, courtesy of Harald Hess). Peaks of fluorescent intensity were detected in each frame and then were fitted with 2D-Gaussian functions to determine their coordinates and width. Only peaks for which more than 100 photons were detected, and localized with a precision

better than 50 nm were considered as single molecule localizations and used for further analysis. SR images were created by superposition of all such localizations, each represented by a Gaussian spot with full width at half maximum (FWHM) 30 nm using PALM-siever software (https://code.google.com/p/palm-siever/).

Single Particle Tracking was performed in Matlab (The MathWorks Inc., version 7.10.0, Natick, MA, 2010). In this analysis peaks localized in Peakselector were characterized as belonging to the same single molecule trajectory if they were present in consecutive frames and were within a search radius. Tracks were rendered as a line joining the successive localizations.

Grouping: This procedure was done using Peakselector software. Localizations were considered as belonging to one molecule if the time gap between consecutive localizations was equal or less than 2 frames and the distance between the average position of the group and the new localization was less than a grouping radius *d. d* was derived from single molecule tracking where we first determined the distribution of step sizes between frames (Appendix Figure 7B) and then used the 95% cut-off of the distribution. See Supplementary Note for further details.

6 P65 live cell transcription factor dynamics with single molecule tracking

Introduction

Transcription factor (TF) dynamics at the promoter region plays major role in regulation of transcription. TFs modify chromatin before transcription starts, bring other TFs to a promoter region and bring polymerase to start transcription [100]. However, until recently TF dynamics could only be measured indirectly using ensemble averaging techniques and showcased, to some extent, inconsistent results.

Recently developed in vivo SR based single molecule (SM) tracking [35-37, 75, 99] allows the direct observation of TF dynamics in vivo and also allows one to measure kinetic parameters, such as TF-promoter binding and unbinding rates [32, 35-37]. We use SM tracking to study how dynamics of the human p65 transcription factor regulates transcription initiation. To distinguish between different models of transcription initiation, we modulate TF-DNA binding energy by mutating the p65 DNA binding domain and measure TF p65 binding dynamics along with the transcription outcome. This work is ongoing; preliminary results are presented in this chapter.

Results

P65 (also called Nf-kB and RelA) is a conditionally active gene-specific transcription factor. It is an important TF that controls hundreds of genes related to different cellular processes, such as cellular growth, apoptosis and the immune response [41]. Dysregulation of p65 activity is involved in various diseases such as different types of cancer and inflammatory diseases. To perform in vivo SM tracking of p65 we fused p65 with different PA-FPs and peptide tag Halo and tested their behavior. Normally, p65 gets activated and translocated into the nucleus by many different stimuli such as TNFa, BMP, EGF, LPS and others [45, 101, 102]. We demonstrated that p65-Halo is present in cytosol and upon stimulation with TNFa translocates into the nucleus (Appendix Figure 10). We also showed that p65-Halo activates p65 dependent genes by RT-PCR (Appendix Figure 11). These experiments functionally validated our p65-Halo construct.

Figure 22: P65 domain organization and mutations used in this work

Four p65 domains and their functions; amino acids mutated in this work (highlighted); relative affinities of the mutants (K_d wt = 1).

We modulated TF binding affinity to DNA by introducing point mutations in p65 (Figure 22). The X-ray structure of this protein provided insight into which amino acids are responsible for the p65-DNA interaction [103]. P65 has a domain structure with an N-terminal DNA-binding domain (DBD) and a C-terminal with two transactivation (TA) domains. It was found that five amino acids Arg33, Arg35, Tyr36, Glu39 and Arg187 of region 30-50 provide interaction with a p65-specific DNA motif. In addition, amino acids in regions 120-140 and 200-240 interact with the sugar-phosphate backbone of DNA (among them Lys122, Lys123, Gln200, Lys221 etc.). Many of these amino acid sites are subject to posttranslational modifications [104]. Several amino acids that participate in p65 binding to DNA were picked for mutation. We have chosen to mutate amino acids both in the non-specific DNA binding domain and in the specific DNA binding domain to have mutants with a wide range of affinities to DNA. Particularly, we have made mutants K122A/K123A, K122R/K123R, K218A, K218R, K221A, K221R and Y36A/E39D. Dissociation constants of all but the last mutant were measured previously in vitro on naked DNA using an ELISA assay [26]. They range from 0.1 to 10 of wild-type K_d.

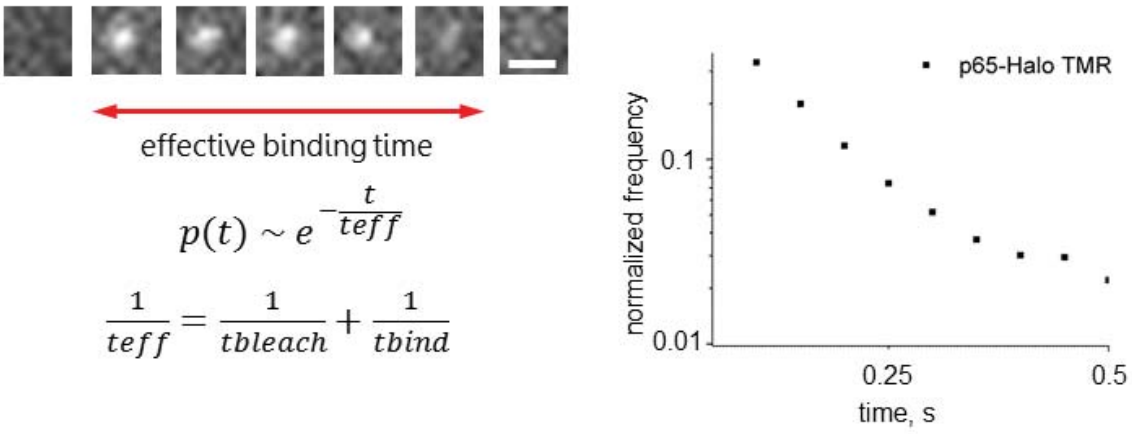


Figure 23: P65 binding time measurement

An example of a single molecule binding event (frame rate 20 Hz, scale bar 500 nm); typical distribution of p65 binding time and equations used for fitting: effective rate is a combination of bleaching rate and unbinding rate.

We imaged TF binding dynamics using SM tracking. Here we looked at dissociation off- rates of p65 wild-type protein and its mutants. To image them we tagged p65 with Halo tag and labelled it with the TMR dye. TMR can be effectively photoswitched in a nuclear environment to obtain many trajectories required for an 'off- rate' measurement [99]. Labeled protein was then tracked inside the nucleus. Lower laser power, as compared to conventional SR experiments, was used to increase fluorophore photostability and focus on non-moving, bound to DNA p65 molecules. We measured the p65 binding time by calculating how much time the immobile p65-Halo TMR is visible. A typical distribution of the effective binding times is presented in Figure 23. We fit this data with an exponential law as TF unbinding is a Poisson process. The decay rate in exponential fitting is the measured dissociation rate. In fact, bleaching of the dye is another reason for the molecule to disappear. So, the rate we measure is a combination of p65 unbinding and dye bleaching. To account for bleaching we imaged also H2B histone. This is an immobile protein which is always bound to DNA so that its measured rate is actually the bleach rate of our dye. We use this bleaching rate to correct the measured effective on- time and to get real binding time of p65 (using the equation from Figure 23). Reliably, due to fitting uncertainty, binding times that are shorter than bleaching times could only be measured. A long bleaching time increases the range of possible measured binding

times. One can notice that we register deviation from single exponential fit for p65. This is likely the result of heterogeneity of p65 binding.

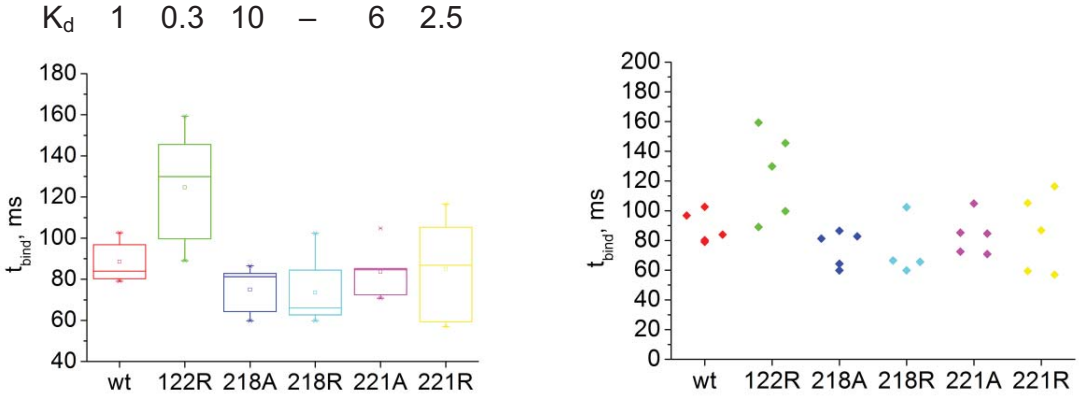


Figure 24: P65 mutants short term binding dynamics

Short term binding times of different p65 mutants in 3t3 p65-/- cells (rendered using boxes (left) and single cell data points (right); K_d values for each mutant shown). Mutant 122R has significantly longer short term binding compared to wild-type (two-sample t-test).

The TF search for a binding site is a complex, multimodal process [35, 105-107]. TFs exhibits free diffusion, sliding along DNA and binding to a given DNA motif. There are several types of TF-DNA binding. They can bind nonspecific sites and specific sites which are discovered biochemically as their primary targets. The chromatin context also plays a role. Given the variability of sequences that TF can bind, there is a spectrum of binding affinities for TFs. Therefore it is reasonable to think that TF binding occurs on different time scales. It was reported that TFs can bind from 10 ms to several hours. We first looked at short TF binding (Figure 24). Using 50 ms exposure time we compared binding time distributions of different mutants. We found that binding time distributions are very similar for these mutants. Binding rates between 50-150 ms were registered. For lower energy mutants we were not able to see significant difference even if there was a 10 times difference in K_d. We found that the mutant that has a higher energy of binding (K122R) has a longer binding time and this difference is significant. Similarly, this mutant has the highest transcription outcome. We also looked at cell-to-cell variation of p65 dynamics. Interestingly, the spread of results for different cells is quite small.

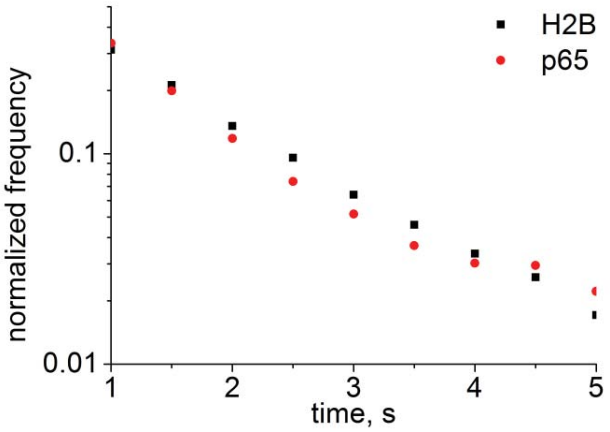


Figure 25: P65 slow dynamics

P65 long effective binding time distribution was measured. H2B distribution illustrates fluorophore's photobleaching. Both distributions have similar decay parameters (t_{p65} 1.1s, t_{H2B} 1.3s).

There is growing evidence that specific binding in eukaryotes can be on the seconds range. For example, TF Sox2 specific binding time was measured to be 12 s [39]. In another work, the dynamics of glucocorticoid receptor (GR) was studied. They obtained value of 1.45 s [37]. We are therefore now performing SM tracking with long exposure times to register long binding events (Figure 25). We first imaged H2B and p65 wt. We observed that p65 indeed has a long binding mode (t_{off} 8.5s). However one can see that H2B curve is close to p65 one and the bleaching time is quite short itself (t_{bleach} 1.3s). The challenge here is to obtain a long bleaching time to reliably measure off rates for p65. Currently, we are performing set of experiments to improve photobleaching value. Single molecules would have longer bleaching times if the laser power irradiance is decreased. Unfortunately, it cannot be infinitely decreased because of the background noise which starts dominating over the signal from single molecules at very low laser power. Therefore, our strategy is to eliminate background noise. We optimize our imaging protocol for that. First, we noticed that even with very low concentration of dye we can still efficiently label cells. Using a lower amount of dye potentially decreases non-specific signals so we plan to use 1 nM instead of 500 nM Halo-TMR for labelling. We will also use a more thorough washing procedure which will potentially eliminate off-target staining. Finally, we are going to use highly inclined illumination which allows avoiding excitation of out-of-focus molecules.

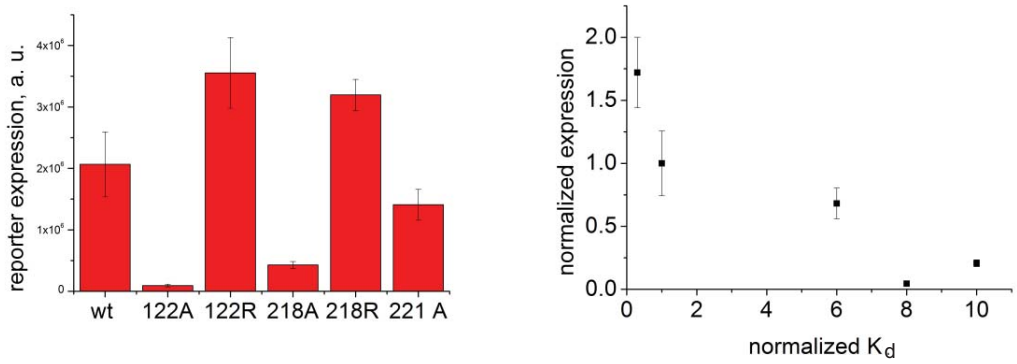


Figure 26: P65-induced transcription

3t3 p65-/- cells were transfected with a reporter plasmid containing p65-dependent promoter. Reporter levels for different mutants were measured 24h after p65 induction (Left). Correlation between normalized expression of reporter and normalized K_d of mutants (Right). Error bars mark standard deviation.

The result of TF binding is transcription initiation. We looked at the ability of different p65 mutants to induce transcription (Figure 26). The in vitro binding energy is directly related to the probability of the DNA motif being occupied by the p65 TF. This would theoretically correlate with transcription from a p65-dependent promoter as when p65 is bound it recruits RNAP. However, as there are many cofactors, enzymes and the chromatin template, combined with the non-equilibrium nature of transcription, the initiation relationship between K_d and transcription becomes less obvious in vivo. We compared transcription outcomes of different mutants using the luciferase reporter system. Here, we transfected cells with a plasmid containing an artificial p65 responsive promoter and reporter sequence. It contains four p65 binding sites derived from the natural IkB promoter. When p65 activates transcription, this leads to synthesis of the secreted luciferase substrate which can be detected in the supernatant. Without p65, transcription from this promoter is negligible. We found that mutations change transcription from this promoter. Particularly, we found that mutants that bind DNA stronger than wild-type give higher transcription and accordingly lower binding mutants result in a lower transcription outcome. In general, to our surprise, the quantitative correlation between K_d and transcription was present in vivo.

Discussion

So far, we have obtained preliminary results on TF dynamics and transcription of p65 mutants with different binding affinities. We saw that p65 binds chromatin at different

time scales. At short millisecond time scales we saw that mutants with different binding energies have similar binding times. While mutants with lower affinity to the DNA motif in vitro than wild type p65 have statistically the same binding time in vivo, the mutant with stronger binding affinity than wild type p65 shows longer binding. We observed a positive correlation between binding affinity and transcription outcome of different mutants. As we discussed above, the short binding we observed is probably not directly related to transcription initiation. Now we are acquiring data in a new setup configuration to study slow binding dynamics of p65. For the reasons described above we expect that these data will provide insight into the mechanism of transcription initiation.

Methods

Cell culture

HeLa and 3t3 p65-/- (the gift from Covert lab [44]) cells were used in this work. HeLa cells were cultured in DMEM media (Gibco) supplemented with 10% FBS (Gibco). 3t3 p65-/- cells were cultured in DMEM media (Gibco) supplemented with Glutamax and 10% FBS. For SM tracking experiments, cells were either transfected or electroporated with p65-Halo plasmid (Promega). Cells were transfected using Fugene 6 (Promega) transfection reagent (ratio 3:1 used) and imaged 24 hours later. In some experiments, instead, cells were electroporated using Amaxa electroporation kit (Amaxa) according to manufacturing instructions and imaged 24 hours later.

Luciferase assay

Cells were co-transfected with p65-Halo plasmid and a luciferase reporter plasmid pMetLuc2 (Clontech) containing a p65-dependent promoter placed in front of the luciferase reporter. 24 hours later, p65 translocation was stimulated with TNFa (Sigma). 8-24 hours after p65 activation supernatants were collected and expression of the reporter was measured using chemiluminometer (Tuscon). To account for variability in expression between different p65 mutants we measured their expression levels using flow cytometry (Accuri, BD Biosciences). There, we quantify the percentage of cells actually expressing p65-Halo. This value was used to normalize the transcription level from the p65-dependent promoter for WT and different mutants.

SM tracking

Before imaging, cells were labeled with Halo-TMR fluorophore. They were incubated 30 minutes with 500 nM Halo-TMR (Promega), washed and incubated 30 minutes further. After that, cells were incubated with 20nM TNFa (Sigma) for 20min to induce p65 translocation from cytoplasm to nucleus. Samples were imaged on a custom-built inverted microscope equipped with UPlanSApo 100x/1.40 oil objective (Olympus) and a piezo objective scanner (N-725, PI). The TMR dye was excited with a 561nm laser (Sapphire, Coherent) and the emitted light was directed through the system of mirrors (Chroma) and Optosplit device (Cairn-research) to EMCCD camera (iXon+, Andor Technology). Samples were imaged in Leibovitz CO_2-independent media (Sigma) at room temperature.

Data analysis

Molecular localizations in each frame were extracted using Octane software [74]. Peak detection threshold was set individually for each sample set due to variation in image background level. To find the time the molecule stays bound to DNA, we reconnected localizations belonging to the same bound molecule. We considered that localizations belong to the same bound molecule if the spatial distance between them was less than 160 nm to avoid registration of moving non-bound molecules. Maximum temporal distance between localizations belonging to the same molecule was set to 2 frames to account for missed localizations and fluorophores' blinking. For each molecule trajectory, the length was assumed to be equal to the time the molecule stays bound. The distribution of trajectory lengths was fit with an exponential decay (first five time points) to obtain characteristic binding times. As discussed in the main text, correction for photobleaching of the dye was performed using trajectory length data from histone H2B labelled with TMR.

7 Conclusions and perspectives

7.1 Conclusions

This work can be virtually separated into two parts: optimization and developing of super-resolution (SR) imaging techniques and their application to studies of gene expression, chromatin organization and transcription factor (TF) dynamics. The principle of super-resolution point localization microscopy (PALM/STORM) was demonstrated and its potential to study different biological phenomena was shown now almost a decade ago. However, there are still a lot of efforts to make the method broadly applicable, versatile, reliable and user-friendly. So the first part of this work was to extend the range of applications and improve performance of SR related techniques.

Site-specific dyes are a valid alternative to protein fusions as probes for microscopy. They have a simple labelling protocol and don't require transfection and protein overexpression. They also can be used as an orthogonal secondary colour to highlight the molecule of interest's localization. We developed a method of direct SR imaging of DNA in living cells. We showed that the DNA specific dye Picogreen can be efficiently photoswiched and imaged with PALM/STORM in live cells over long times. The protocol we developed gave 5-7 x improvement in resolution and allowed doing time lapse SR imaging of chromatin. It is the first demonstration that site-specific dyes which label different organelles could be used in live cell SR imaging.

Single molecule (SM) tracking performance can be boosted by using photoswitching principles of STORM/PALM. Fluorophore photoswitching can provide up to 1000x improvement in trajectory density which allows creating maps of protein dynamics with subcellular resolution. We developed a high-density SM tracking approach based on photoswitching of organic dyes. We showed that many different organic dyes covering the spectrum can be photoswitched in a live-cell compatible way and therefore also tracked. We demonstrated that using different labelling schemes and different dyes, three colour simultaneous tracking can be achieved. Notably, we found several dyes that can be used for intracellular SM tracking in mitochondria and nuclei. We also demonstrated that, as expected, dyes provide longer trajectories

compared to photoactivatable proteins; this is another major type of phoswitchable fluorophore.

In a separate study we showed that the photophysical properties of dyes can affect the appearance of a structure and lead to artefacts. We show, in particular, that molecules with long on times can introduce artificial clustering. We demonstrate a simple grouping procedure based on SM tracking to account for molecules' long on times. This procedure allows correction for these artefacts.

Improved SM tracking and protocols for SR imaging allowed us to study phenomena and measure parameters inaccessible before. We used them to study how chromatin organization and transcription factor dynamics regulate transcription. This is ongoing work. First, we are studying how organization of the HoxD gene cluster in developing mice is related to HoxD gene activity there. We found that the HoxD cluster in the forebrain where Hox genes are silent has the most round shape. In the forelimb, where some of them are active, we see HoxD-specific elongation of the cluster. This observation supports the idea that activation of HoxD genes is accompanied by their redistribution to separate active sub-compartments.

Second, we are studying how binding of TF p65 regulates transcription initiation. To find a quantitative model of initiation we vary the affinity of TF to its DNA motif. We show that TF binds on multiple time scales from miliseconds to seconds. We saw that at short time scale mutants with a wide range of binding affinity (Kd = 0.1-10 Kd wild-type) show similar binding times in vivo. We found a correlation between transcription and in vitro binding affinity of different mutants. We are currently focusing our study on slow TF binding dynamics to address the question how transcription initiation is regulated by TF dynamics. We present our plans for these projects in Perspectives section.

7.2 Perspectives

7.2.1 HoxD chromatin organization

To clarify the sub-compartmentalization model of the HoxD gene cluster, we designed several probes approximately 5 times smaller than used before (E5) corresponding to different sub-regions of the HoxD cluster. Smaller probes allow the quantification of smaller features of sub-cluster organization. In terms of chromatin

organization we can imagine several scenarios (Figure 27A). The observed elongation of the cluster in the FL can be the result of simply more open chromatin at HoxD gene clusters specifically there. Then, one can imagine that chromatin at active domain gets less condensed and this is what lengthens the cluster. Finally, it can be separation of two domains keeping similar levels of condensation. We are now measuring shape parameters of the regions corresponding to these small probes in different tissues. Looking at these shape parameters of different small probes X1, X2 and X3 will potentially allow discriminating between these scenarios (Figure 27B).

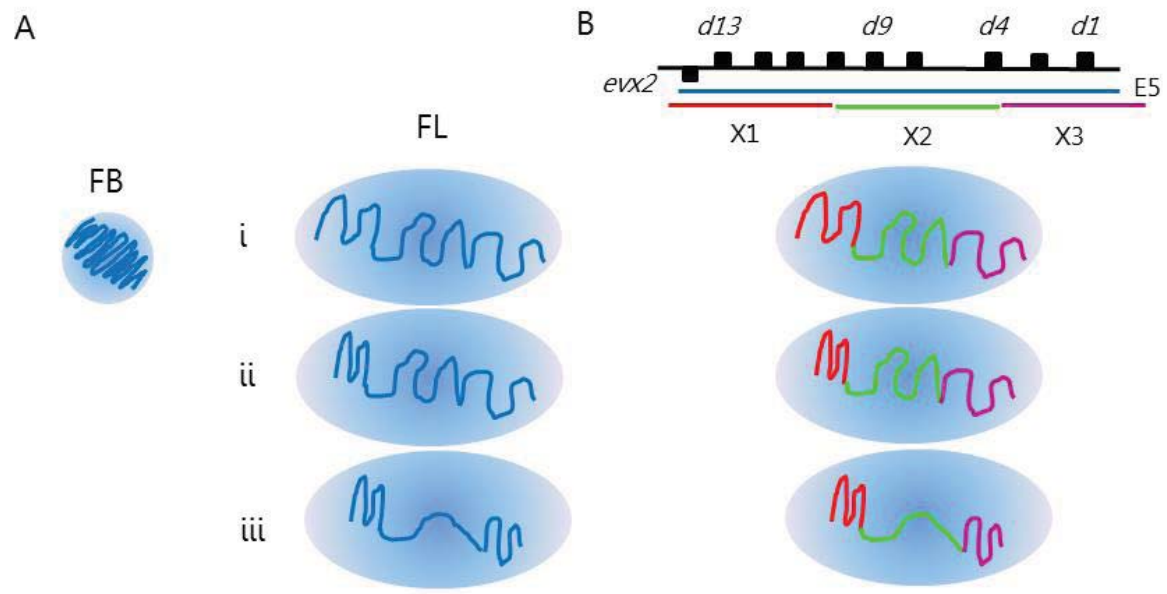


Figure 27: Several possible scenarios of the HoxD cluster elongation
See text for the details.

HoxD gene activation is known to be strongly influenced by long range interactions. Centromeric and telomeric regulatory islands up to 1Mb away from cluster make specific contacts with these clusters [18]. Removal or mutation of them leads to downregulation of transcription and severe phenotypic effects in mouse models. We would like to study how they regulate transcription by looking at how they affect chromatin organization of HoxD gene cluster. We are currently imaging samples where mutations in regulatory islands were introduced. Furthermore, we are testing if this leads to changes in HoxD organization.

7.2.2 Transcription factor dynamics and initiation of transcription

We discussed above that, technically, the limiting factor in studying TF binding dynamics is the photostability of fluorophores. To extend the range of measurable binding times we will label TF molecule with multiple fluorophores. We plan to label TF targeting nano-bodies with multiple dye molecules and directly deliver them to the nucleus using recently developed protein delivery systems [108]. This will allow us to potentially increase bleaching time proportionally to number of fluorophores per nano-body.

In this work we use p65 mutants with different affinities to the p65-specific DNA motif in order to test models of transcription initiation. When we change binding affinity of p65, the way TF binding times change gives an idea about the mechanism of assembly of the pre-initiation complex. Accordingly, the way transcription outcomes change indicates a mechanism of transcription initiation. We discussed above that it is reasonable to consider that these processes are out of equilibrium. We know that the p65 binding time in vivo is shorter than the binding time obtained in early in vitro studies. We think that this is a reflection of ATP-induced non-equilibrium TF dynamics. We would like to test if there is an active removal of the TF from the promoter region during this initiation process. This would explain a discrepancy between in vitro and in vivo TF dynamics measurements. In addition, this will be a direct demonstration that an ATP-dependent process shapes transcription initiation in eukaryotes. If TFs are actively removed from the promoter region, we expect a specific shape of the binding time-binding affinity curve for our mutants (Figure 28). At some point, the curve should reach a plateau when stronger affinity doesn't increase binding time since the TF is actively displaced from the promoter. In other words, above a certain threshold in binding time there is no difference in TF binding time at the promoter for mutants having an affinity higher than this threshold. This binding time would then determine a characteristic time for TF active displacement. We think that ATP-dependent chromatin remodelers, particularly, Brg1 are responsible for that [109, 110]. So to test this idea further we will target them with siRNA. If there is ATP induced eviction of p65, the downregulation of Brg1 will stabilize the p65-DNA interaction in vivo.

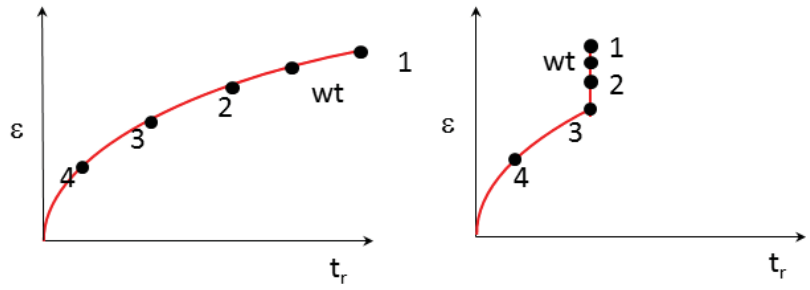


Figure 28: Effect of active ATP dependent TF eviction (hypothesis)

Relationship between binding time t_r and binding energy e without (left) and with (right) active removal of TF. Numbers represent mutants with different binding energy.

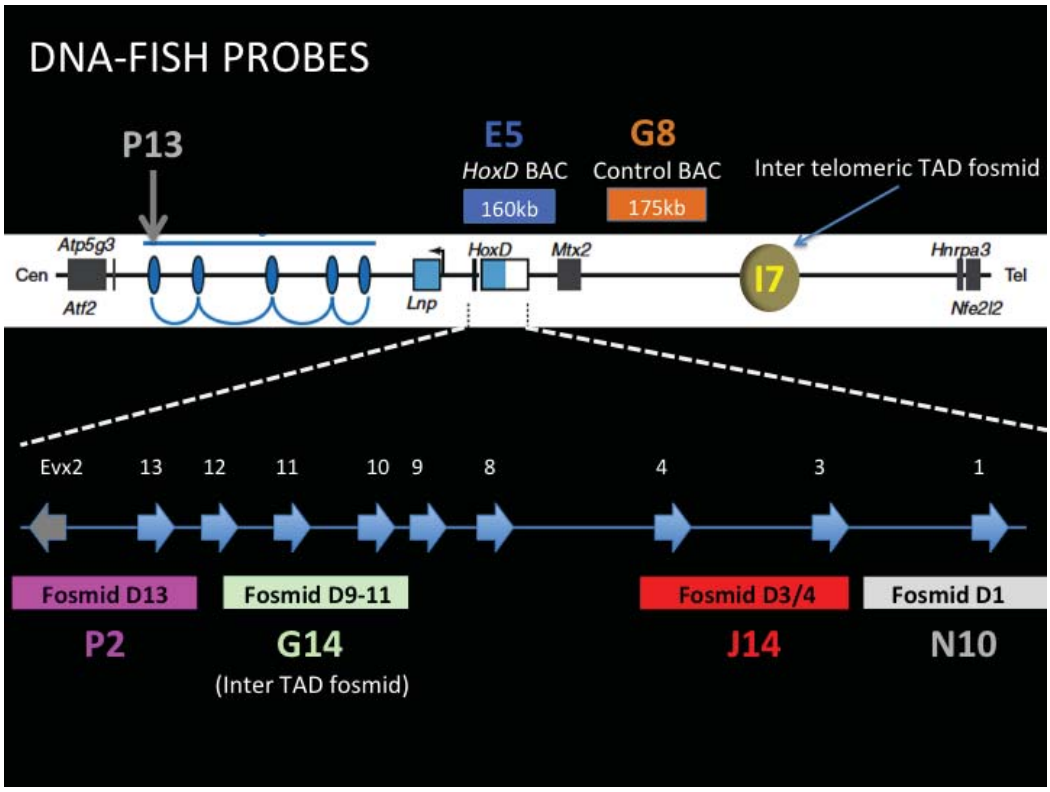
We noted already that TFs bind to a wide variety of sequences in eukaryotes so one can expect a range of binding times. The only ones that related to transcription are binding events at functional transcription sites. We would like to measure p65 binding dynamics at particular DNA motifs in vivo to distinguish the difference in binding to non-specific sites and specific, functional sites. To do that one would need to find a way to distinguish these specific events in living cells. We plan to insert an "imaging unit" into the genome with a p65-binding region derived from the Ik-B promoter, a minimal promoter containing the TATA- box and binding site for LacO-GFP for visualization of this unit. To enhance the probability of binding and of visualization of these sites we plan to introduce a cassette with multiple copies of this imaging unit in the genome. We expect that binding dynamics measured at this region would represent specific functional binding of p65. This would decouple non-specific and specific binding which would allow us to refine binding time distributions obtained without site specific visualization and potentially shed light on the mechanism of binding site search.

To directly relate transcription initiation and binding of TF, the ultimate experiment would be to simultaneously image TF binding dynamics and transcription initiation. Tools for the TF binding dynamics imaging were presented throughout this manuscript. There are also ones for transcription visualization in living cells. MS2 and PP7 systems allow for imaging of nascent RNA production in living cells [20, 111, 112]. They seem to be sensitive enough, as it is potentially possible to measure dynamics of single transcription event. Merging these two approaches would create a

valuable tool to study the dynamics of transcription initiation with single TF and single transcription event levels.

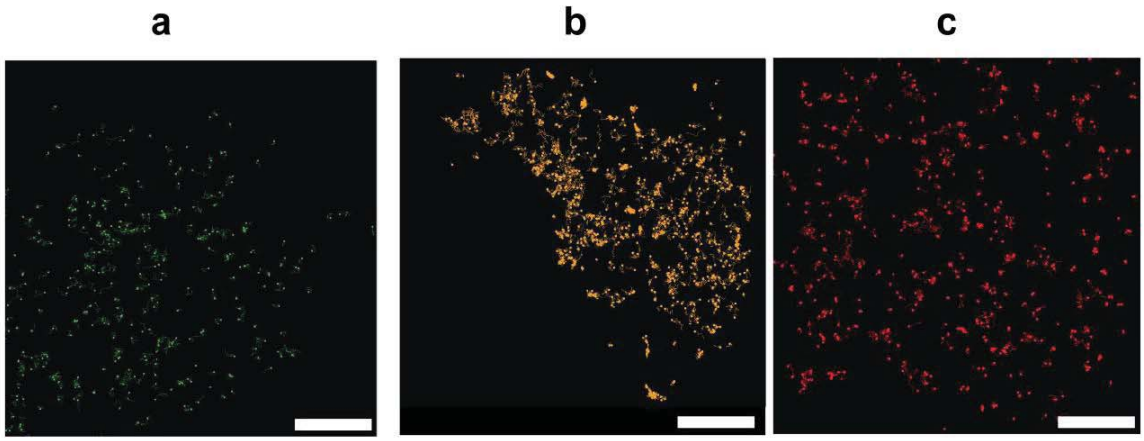
8 Appendix

8.1 HoxD gene cluster spatial organization on single cell level with super-resolution FISH imaging



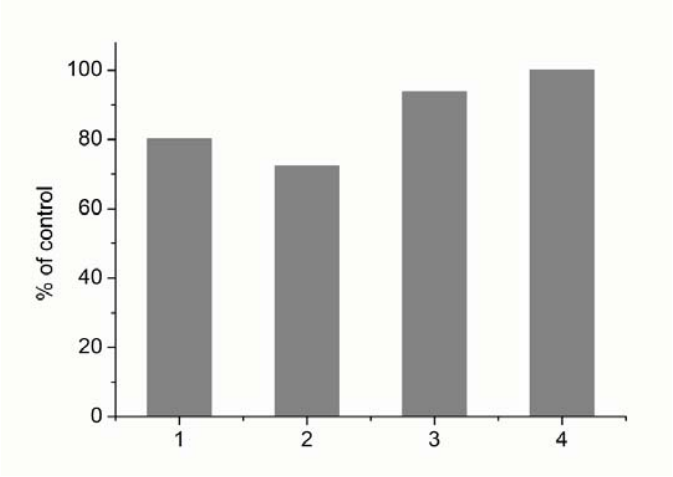
Appendix Figure 1: Map the HoxD probes

8.2 Multicolor single molecule tracking of stochastically active synthetic dyes



Appendix Figure 2: Three colour orthogonal labelling

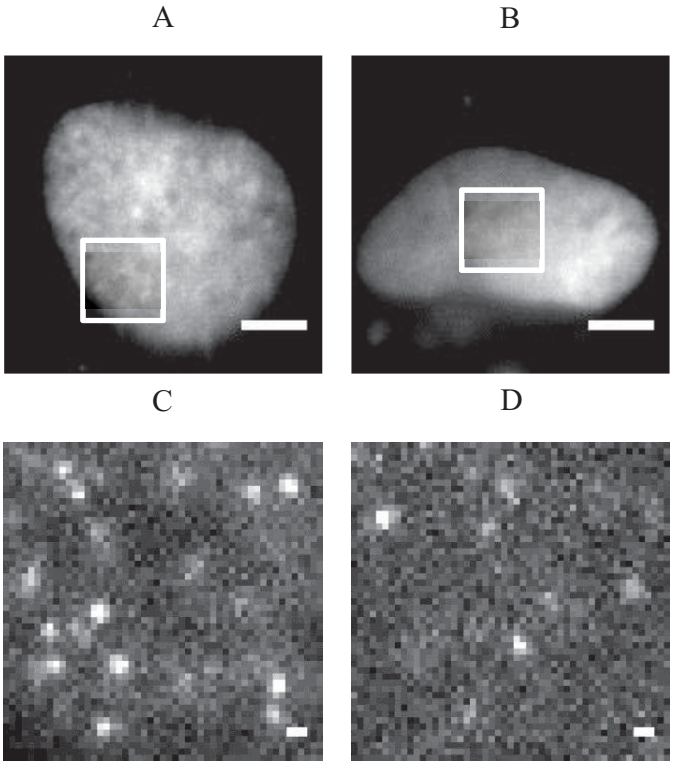
SM tracking performed on cells tagged with either (a) SNAP-b2AR Alexa 488, (b) Halo-TfR TMR* and (c) GPI-ACP Dy-647. Scale bar = 2 µm.



Appendix Figure 3: Cell viability assay

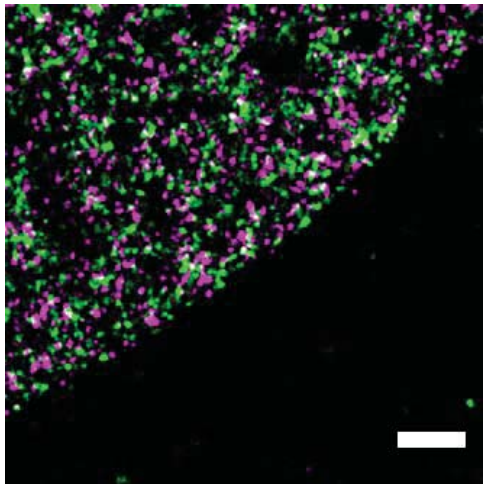
MTT test on cells exposed to following buffers: 1-Leibovitz+GLOX+1mM AA, 2-Leibovitz+GLOX+Gluthatione, 3-Leibovitz, 4-DMEM (control).

8.3 Analysis for live cell super-resolution imaging adapted for photoswitching properties and molecular dynamics



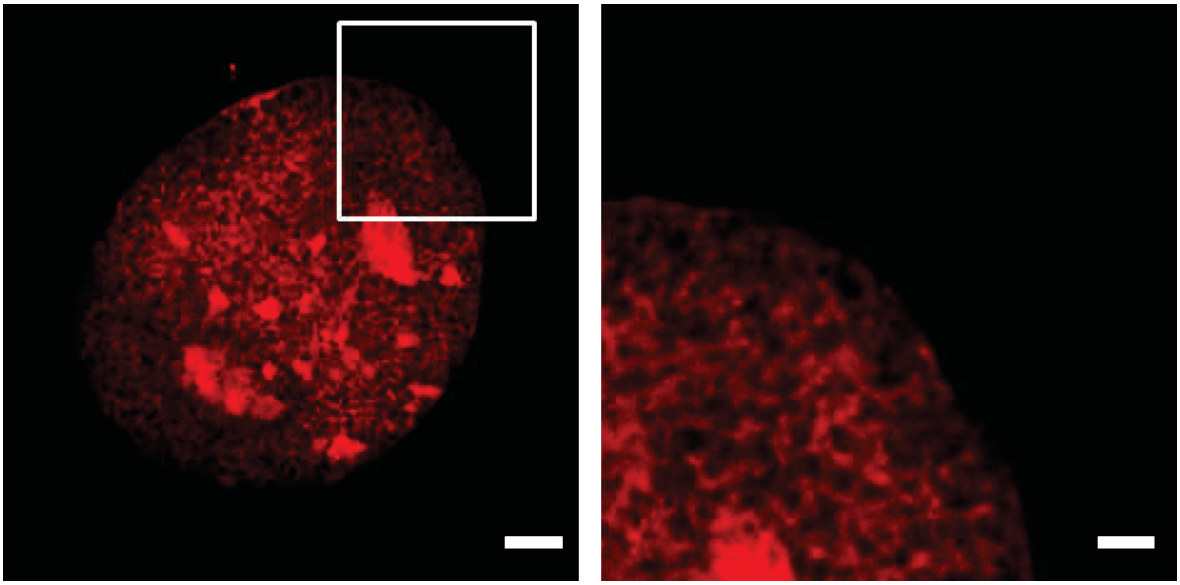
Appendix Figure 4: Wide-field and SR imaging of H2b histone

(Top)Wide-field images of U2OS cells expressing A) H2B-SNAP labeled with BG-TMR* and B) expressing H2B-PAmCherry. Scale bars 5 µm. (Bottom) Live-cell SR imaging of U2OS cells. Raw C) H2B-TMR* and D) H2B-PAmCherry single-frame images of boxed regions demonstrate fluorophores' single-molecule photoswitching. Scale bars 500 nm.



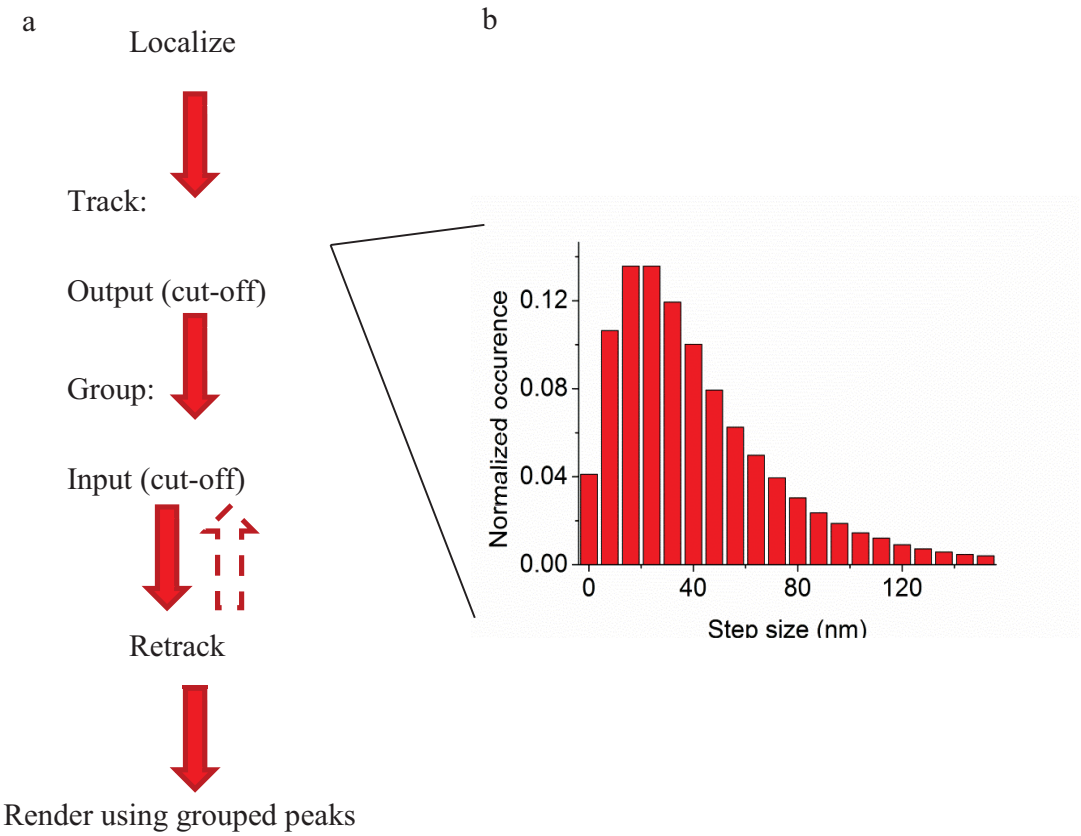
Appendix Figure 5: Time-lapse SR image of fixed cells

Time-lapse SR image of fixed U2OS cells expressing H2B-SNAP labeled with TMR* (Scale bar 1µm). Imaging was performed in PBS media containing 10% glucose, 0.5 $mgmL^{-1}$ glucose oxidase, 40 $mgmL^{-1}$ catalase and 143mM b-mercaptoethanol. The first half of all localizations are displayed in magenta and the second half in green.



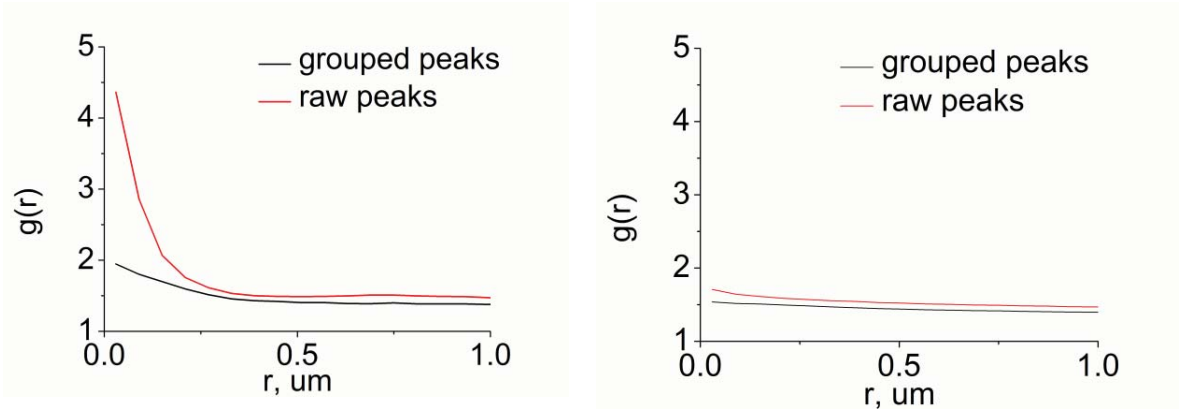
Appendix Figure 6: STED image of cell labelled with Picogreen

Stimulated emission depletion super-resolution image of U20S cell labeled with Picogreen. Scale bars are 3 µm and 1 µm.



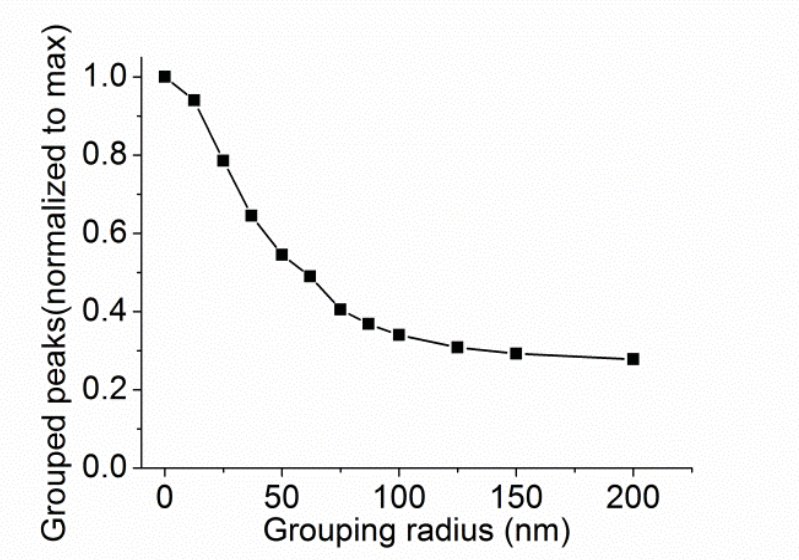
Appendix Figure 7: Grouping procedure details

(a) Schematic of the general grouping procedure scheme described in **Supplementary Note.** (b) Distribution of step sizes from single particle tracking performed on the raw H2B-TMR* data.



Appendix Figure 8: Pair correlation analysis

Pair correlation analysis is performed to compare the observed structure before and after peak grouping. g(r) function plot for H2B -TMR* image (left) and for H2B-PAmCherry (right). g(r) curve for H2B-TMR is much less steep when grouped peaks are analysed compared to raw peaks curve. Together with rendered image (Figure 21B) it shows that clustering is significantly decreased when grouping is applied. g(r) plot for H2B-PAmCherry shows no significant difference between grouped and raw peaks curves. It demonstrates together with rendered image (Figure 21C) that H2B-PAmCherry structure doesn't change strongly when grouped localizations are used for rendering.



Appendix Figure 9: Dependence of number of peaks on grouping radius for H2B-TMR*

Number of peaks after grouping procedure was calculated for different values of the grouping radius. Normalization was done to the number of peaks identified with no grouping done. The value first decreases rapidly, then begins to plateau between 75-100 nm. This corresponds well to the grouping radius determined from the tracking procedure (85 nm).

8.3.1 Supplementary Note

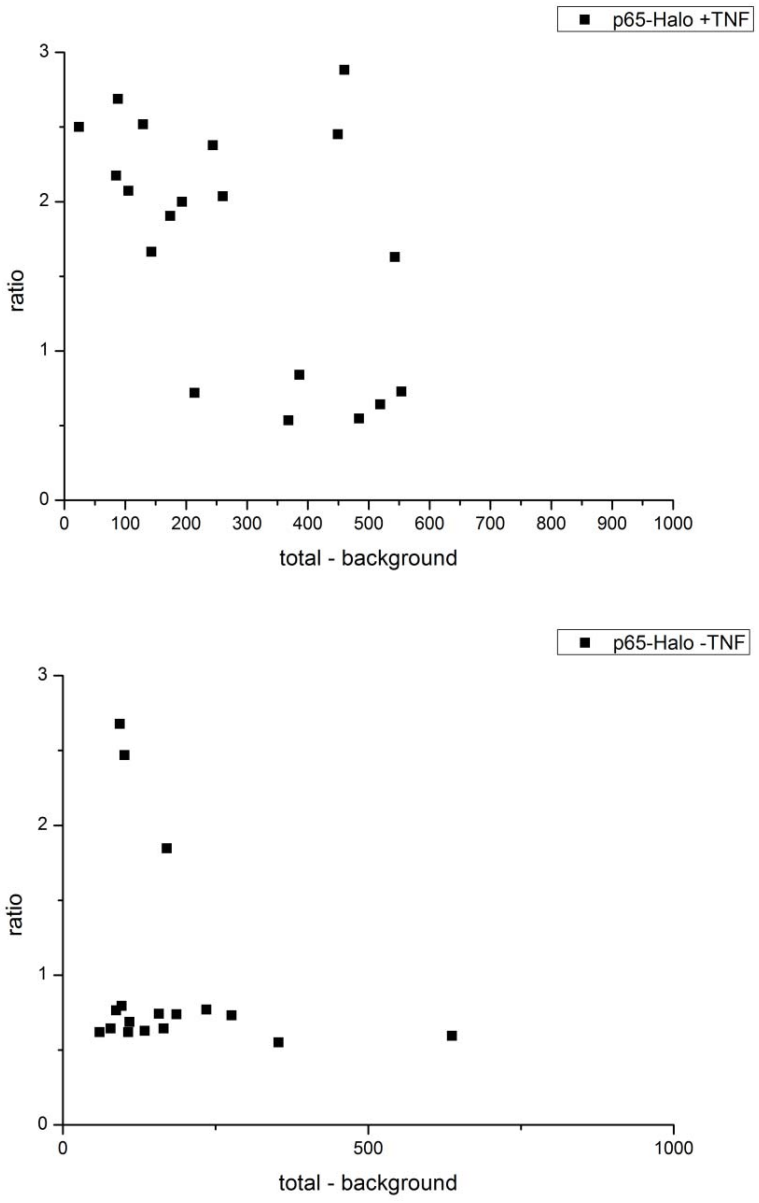
Grouping procedure details

To empirically determine the grouping radius we first had to choose a search radius for tracking. If this radius is too large, it can create artificial connections between different molecules, but if it is too small, it can miss connections. In general, it should be smaller than the minimum single frame average intermolecular distance, which in our SR images was 5 pixels. We used a search radius of 160nm (~1 pixel), significantly smaller than the intermolecular distance. Tracking with this search radius, we obtain a non-truncated distribution of step sizes between frames with a mean value far below the maximum search radius (Appendix Figure 7B), which validates our choice. The distribution of step sizes determines the grouping radius: we used a 95% cut-off (85 nm), which was sufficient to remove the great majority of the tracks (Figure 20). This procedure potentially can be applied to any set of SR data to avoid clustering artefacts from long-lived molecules (Appendix Figure 7A).

Filtering of grouped molecules

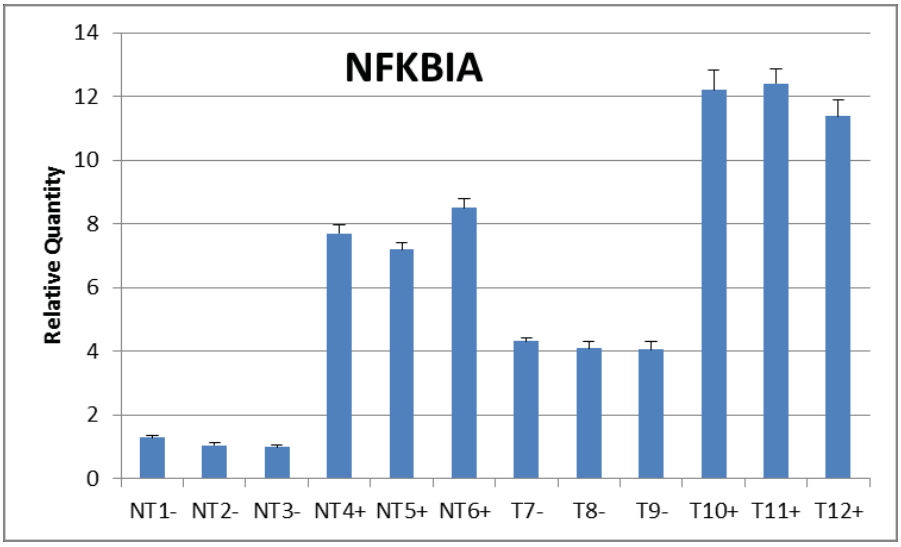
Another question is whether to retain these grouped molecules or discard them. A rule of thumb can be that if the total displacement of a given molecule is smaller than desired resolution, then one can include it in the rendered SR image. It also should be mentioned that grouping isn't always useful. Apart from studying the structure created by a protein of interest, proteins can also be used as a probe of membrane morphology (endoplasmic reticulum, mitochondria, etc.). In this case their motion is advantageous since it allows one to obtain more peaks, sampling the structure more fully with fewer molecules, if these molecules are on in several frames.

8.4 P65 live cell transcription factor dynamics with single molecule tracking



Appendix Figure 10: Validation of the plasmid – test of p65 translocation

Cells expressing p65-Halo were activated with TNFa. Translocation efficiency was measured as a ratio between nuclear and cytoplasmic fluorescence. Cells were considered as activated if after 20 min activation the ratio was above 1. (top) Cells activated with TNFa: 65% of cells show translocation. (bottom) Nonactivated cells: 15% of cells show translocation. Each spot on the graph represents one cell.



Appendix Figure 11: Validation of the plasmid - RT-PCR

Activation of p65-dependent gene NFKBIA was probed with RT-PCR in HeLa cells transfected with p65-Halo plasmid. Relative quantity of NFKBIA mRNA is reported. NT – nontransfected cells (control), T – transfected with p65-Halo plasmid cells. -/+ - TNFa/+TNFa

References

1. Watson, J.D. and F.H.C. Crick, *Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid.* Nature, 1953. **171**(4356): p. 737-738.
2. Wilkins, M.H.F., A.R. Stokes, and H.R. Wilson, *Molecular structure of nucleic acids: Molecular structure of deoxypentose nucleic acids.* Nature, 1953. **171**(4356): p. 738-740.
3. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.
4. Cremer, T. and M. Cremer, *Chromosome territories.* Cold Spring Harb Perspect Biol, 2010. **2**(3): p. a003889.
5. Song, F., et al., *Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units.* Science, 2014. **344**(6182): p. 376-80.
6. de Wit, E. and W. de Laat, *A decade of 3C technologies: insights into nuclear organization.* Genes Dev, 2012. **26**(1): p. 11-24.
7. Mirny, L.A., *The fractal globule as a model of chromatin architecture in the cell.* Chromosome Res, 2011. **19**(1): p. 37-51.
8. Naumova, N., et al., *Organization of the mitotic chromosome.* Science, 2013. **342**(6161): p. 948-953.
9. Dixon, J.R., et al., *Topological domains in mammalian genomes identified by analysis of chromatin interactions.* Nature, 2012. **485**(7398): p. 376-80.
10. Nora, E.P., et al., *Spatial partitioning of the regulatory landscape of the X-inactivation centre.* Nature, 2012. **485**(7398): p. 381-385.
11. Mallo, M. and C.R. Alonso, *The regulation of Hox gene expression during animal development.* Development, 2013. **140**(19): p. 3951-63.
12. Montavon, T. and D. Duboule, *Landscapes and archipelagos: spatial organization of gene regulation in vertebrates.* Trends Cell Biol, 2012. **22**(7): p. 347-54.
13. Soshnikova, N. and D. Duboule, *Epigenetic temporal control of mouse Hox genes in vivo.* Science, 2009. **324**(5932): p. 1320-3.
14. Noordermeer, D., et al., *The dynamic architecture of Hox gene clusters.* Science, 2011. **334**(6053): p. 222-5.

15. Montavon, T., et al., *A regulatory archipelago controls Hox genes transcription in digits.* Cell, 2011. **147**(5): p. 1132-45.
16. Noordermeer, D., et al., *Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci.* eLife, 2014. **2014**(3).
17. Andrey, G., et al., *A switch between topological domains underlies HoxD genes collinearity in mouse limbs.* Science, 2013. **340**(6137).
18. de Laat, W. and D. Duboule, *Topology of mammalian developmental enhancers and their regulatory landscapes.* Nature, 2013. **502**(7472): p. 499-506.
19. Tschopp, P., A.J. Christen, and D. Duboule, *Bimodal control of Hoxd gene transcription in the spinal cord defines two regulatory subclusters.* Development, 2012. **139**(5): p. 929-939.
20. Darzacq, X., et al., *In vivo dynamics of RNA polymerase II transcription.* Nat Struct Mol Biol, 2007. **14**(9): p. 796-806.
21. Murakami, K., et al., *Formation and fate of a complete 31-protein RNA polymerase II transcription preinitiation complex.* J Biol Chem, 2013. **288**(9): p. 6325-32.
22. Consortium, E.P., *An integrated encyclopedia of DNA elements in the human genome.* Nature, 2012. **489**(7414): p. 57-74.
23. Wunderlich, Z. and L.A. Mirny, *Different gene regulation strategies revealed by analysis of binding motifs.* Trends Genet, 2009. **25**(10): p. 434-40.
24. Neph, S., et al., *An expansive human regulatory lexicon encoded in transcription factor footprints.* Nature, 2012. **489**(7414): p. 83-90.
25. Gerstein, M.B., et al., *Architecture of the human regulatory network derived from ENCODE data.* Nature, 2012. **489**(7414): p. 91-100.
26. Schaaf, M.J., et al., *The relationship between intranuclear mobility of the NF-kappaB subunit p65 and its DNA binding affinity.* J Biol Chem, 2006. **281**(31): p. 22409-20.
27. Zabel, U., R. Schreck, and P.A. Baeuerle, *DNA binding of purified transcription factor NF-κB: Affinity, specificity, Zn2+ dependence, and differential half-site recognition.* Journal of Biological Chemistry, 1991. **266**(1): p. 252-260.

28. Fujita, T., et al., *Independent modes of transcriptional activation by the p50 and p65 subunits of NF-kappa B.* Genes & Development, 1992. **6**(5): p. 775-787.
29. Phair, R.D., S.A. Gorski, and T. Misteli, *Measurement of Dynamic Protein Binding to Chromatin In Vivo, Using Photobleaching Microscopy*, in *Methods in Enzymology*2004. p. 393-414.
30. Hager, G.L., J.G. McNally, and T. Misteli, *Transcription dynamics.* Mol Cell, 2009. **35**(6): p. 741-53.
31. Garcia, H.G., et al., *Operator sequence alters gene expression independently of transcription factor occupancy in bacteria.* Cell Rep, 2012. **2**(1): p. 150-61.
32. Hammar, P., et al., *Direct measurement of transcription factor dissociation excludes a simple operator occupancy model for gene regulation.* Nature Genetics, 2014. **46**(4): p. 405-408.
33. De Los Rios, P. and A. Barducci, *Hsp70 chaperones are non-equilibrium machines that achieve ultra-affinity by energy consumption.* eLife, 2014. **2014**(3).
34. Coulon, A., et al., *Eukaryotic transcriptional dynamics: from single molecules to cell populations.* Nat Rev Genet, 2013. **14**(8): p. 572-84.
35. Elf, J., G.W. Li, and X.S. Xie, *Probing transcription factor dynamics at the single-molecule level in a living cell.* Science, 2007. **316**(5828): p. 1191-4.
36. Mazza, D., et al., *A benchmark for chromatin binding measurements in live cells.* Nucleic Acids Res, 2012. **40**(15): p. e119.
37. Gebhardt, J.C., et al., *Single-molecule imaging of transcription factor binding to DNA in live mammalian cells.* Nat Methods, 2013. **10**(5): p. 421-6.
38. Persson, F., et al., *Extracting intracellular diffusive states and transition rates from single-molecule tracking data.* Nat Methods, 2013. **10**(3): p. 265-9.
39. Chen, J., et al., *Single-molecule dynamics of enhanceosome assembly in embryonic stem cells.* Cell, 2014. **156**(6): p. 1274-1285.
40. Morisaki, T., et al., *Single-molecule analysis of transcription factor binding at transcription sites in live cells.* Nature Communications, 2014. **5**.
41. Hayden, M.S. and S. Ghosh, *NF-κB, the first quarter-century: Remarkable progress and outstanding questions.* Genes and Development, 2012. **26**(3): p. 203-234.

42. Lawrence, T., *The nuclear factor NF-kappaB pathway in inflammation.* Cold Spring Harbor perspectives in biology, 2009. **1**(6).
43. Nelson, D.E., et al., *Oscillations in NF-kappaB signaling control the dynamics of gene expression.* Science, 2004. **306**(5696): p. 704-8.
44. Tay, S., et al., *Single-cell NF-kappaB dynamics reveal digital activation and analogue information processing.* Nature, 2010. **466**(7303): p. 267-71.
45. Nelson, G., et al., *Multi-parameter analysis of the kinetics of NF-κB signalling and transcription in single living cells.* Journal of Cell Science, 2002. **115**(6): p. 1137-1148.
46. Martone, R., et al., *Distribution of NF-kappaB-binding sites across human chromosome 22.* Proc Natl Acad Sci U S A, 2003. **100**(21): p. 12247-52.
47. Betzig, E., et al., *Imaging intracellular fluorescent proteins at nanometer resolution.* Science, 2006. **313**(5793): p. 1642-5.
48. Hess, S.T., T.P. Girirajan, and M.D. Mason, *Ultra-high resolution imaging by fluorescence photoactivation localization microscopy.* Biophys J, 2006. **91**(11): p. 4258-72.
49. Hofmann, M., et al., *Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(49): p. 17565-17569.
50. Rust, M.J., M. Bates, and X. Zhuang, *Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM).* Nature Methods, 2006. **3**(10): p. 793-795.
51. Thompson, R.E., D.R. Larson, and W.W. Webb, *Precise nanometer localization analysis for individual fluorescent probes.* Biophysical Journal, 2002. **82**(5): p. 2775-2783.
52. Yildiz, A., et al., *Myosin V walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization.* Science, 2003. **300**(5628): p. 2061-2065.
53. Huang, B., et al., *Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy.* Science, 2008. **319**(5864): p. 810-3.
54. Shtengel, G., et al., *Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(9): p. 3125-3130.

55. Patterson, G., et al., *Superresolution imaging using single-molecule localization.* Annu Rev Phys Chem, 2010. **61**: p. 345-67.
56. Fernandez-Suarez, M. and A.Y. Ting, *Fluorescent probes for super-resolution imaging in living cells.* Nat Rev Mol Cell Biol, 2008. **9**(12): p. 929-43.
57. Heilemann, M., et al., *Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes.* Angew Chem Int Ed Engl, 2008. **47**(33): p. 6172-6.
58. van de Linde, S., et al., *Photoswitching microscopy with standard fluorophores.* Applied Physics B: Lasers and Optics, 2008. **93**(4): p. 725-731.
59. Klein, T., et al., *Live-cell dSTORM with SNAP-tag fusion proteins.* Nat Methods, 2011. **8**(1): p. 7-9.
60. Wombacher, R., et al., *Live-cell super-resolution imaging with trimethoprim conjugates.* Nature Methods, 2010. **7**(9): p. 717-719.
61. Miller, L.W., et al., *In vivo protein labeling with trimethoprim conjugates: A flexible chemical tag.* Nature Methods, 2005. **2**(4): p. 255-257.
62. Keppler, A., et al., *A general method for the covalent labeling of fusion proteins with small molecules in vivo.* Nature Biotechnology, 2003. **21**(1): p. 86-89.
63. Flors, C., C.N. Ravarani, and D.T. Dryden, *Super-resolution imaging of DNA labelled with intercalating dyes.* Chemphyschem, 2009. **10**(13): p. 2201-4.
64. Vogelsang, J., et al., *Controlling the fluorescence of ordinary oxazine dyes for single-molecule switching and superresolution microscopy.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(20): p. 8107-8112.
65. Steinhauer, C., et al., *Superresolution microscopy on the basis of engineered dark states.* Journal of the American Chemical Society, 2008. **130**(50): p. 16840-16841.
66. Shroff, H., et al., *Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics.* Nature Methods, 2008. **5**(5): p. 417-423.
67. Jones, S.A., et al., *Fast, three-dimensional super-resolution imaging of live cells.* Nat Methods, 2011. **8**(6): p. 499-508.
68. Mallo, M., D.M. Wellik, and J. Deschamps, *Hox genes and regional patterning of the vertebrate body plan.* Dev Biol, 2010. **344**(1): p. 7-15.

69. Van de Corput, M.P.C., et al., *Super-resolution imaging reveals three-dimensional folding dynamics of the β-globin locus upon gene activation.* Journal of Cell Science, 2012. **125**(19): p. 4630-4639.
70. Cremer, M., et al., *Multicolor 3D fluorescence in situ hybridization for imaging interphase chromosomes*, in *Methods in Molecular Biology*2008. p. 205-239.
71. van de Linde, S., et al., *Direct stochastic optical reconstruction microscopy with standard fluorescent probes.* Nat Protoc, 2011. **6**(7): p. 991-1009.
72. Montavon, T., et al., *Modeling Hox gene regulation in digits: Reverse collinearity and the molecular origin of thumbness.* Genes and Development, 2008. **22**(3): p. 346-359.
73. Morey, C., et al., *Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation.* Development, 2007. **134**(5): p. 909-19.
74. Birant, D. and A. Kut, *ST-DBSCAN: An algorithm for clustering spatial-temporal data.* Data and Knowledge Engineering, 2007. **60**(1): p. 208-221.
75. Manley, S., et al., *High-density mapping of single-molecule trajectories with photoactivated localization microscopy.* Nat Methods, 2008. **5**(2): p. 155-7.
76. Giannone, G., et al., *Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density.* Biophysical Journal, 2010. **99**(4): p. 1303-1310.
77. Hinner, M.J. and K. Johnsson, *How to obtain labeled proteins and what to do with them.* Current Opinion in Biotechnology, 2010. **21**(6): p. 766-776.
78. Los, G.V., et al., *HaloTag: A novel protein labeling technology for cell imaging and protein analysis.* ACS Chemical Biology, 2008. **3**(6): p. 373-382.
79. Lee, H.L.D., et al., *Superresolution imaging of targeted proteins in fixed and living cells using photoactivatable organic fluorophores.* Journal of the American Chemical Society, 2010. **132**(43): p. 15099-15101.
80. Appelhans, T., et al., *Nanoscale organization of mitochondrial microcompartments revealed by combining tracking and localization microscopy.* Nano Letters, 2012. **12**(2): p. 610-616.
81. van de Linde, S., M. Heilemann, and M. Sauer, *Live-cell super-resolution imaging with synthetic fluorophores.* Annu Rev Phys Chem, 2012. **63**: p. 519-40.

82. Dempsey, G.T., et al., *Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging.* Nature Methods, 2011. **8**(12): p. 1027-1040.
83. Rasnik, I., S.A. McKinney, and T. Ha, *Nonblinking and long-lasting single-molecule fluorescence imaging.* Nature Methods, 2006. **3**(11): p. 891-893.
84. Cordes, T., J. Vogelsang, and P. Tinnefeld, *On the mechanism of trolox as antiblinking and antibleaching reagent.* Journal of the American Chemical Society, 2009. **131**(14): p. 5018-5019.
85. Benke, A. and S. Manley, *Live-cell dSTORM of cellular DNA based on direct DNA labeling.* Chembiochem, 2012. **13**(2): p. 298-301.
86. Crocker, J.C. and D.G. Grier, *Methods of digital video microscopy for colloidal studies.* Journal of Colloid and Interface Science, 1996. **179**(1): p. 298-310.
87. Kaya, A., et al., *Long and short distance movements of β2-adrenoceptor in cell membrane assessed by photoconvertible fluorescent protein dendra2-β2-adrenoceptor fusion.* Biochimica et Biophysica Acta - Molecular Cell Research, 2011. **1813**(8): p. 1511-1524.
88. Hegener, O., et al., *Dynamics of β2-adrenergic receptor-ligand complexes on living cells.* Biochemistry, 2004. **43**(20): p. 6190-6199.
89. Mosmann, T., *Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays.* Journal of Immunological Methods, 1983. **65**(1-2): p. 55-63.
90. Kimura, H. and P.R. Cook, *Kinetics of core histones in living human cells: Little exchange of H3 and H4 and some rapid exchange of H2B.* Journal of Cell Biology, 2001. **153**(7): p. 1341-1353.
91. Frost, N.A., H.E. Lu, and T.A. Blanpied, *Optimization of cell morphology measurement via single-molecule tracking PALM.* PLoS ONE, 2012. **7**(5).
92. Shim, S.H., et al., *Super-resolution fluorescence imaging of organelles in live cells with photoswitchable membrane probes.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(35): p. 13978-13983.
93. Biteen, J.S., et al., *Super-resolution imaging in live Caulobacter crescentus cells using photoswitchable EYFP.* Nature Methods, 2008. **5**(11): p. 947-949.

94. Bates, M., B. Huang, and X. Zhuang, *Super-resolution microscopy by nanoscale localization of photo-switchable fluorescent probes.* Curr Opin Chem Biol, 2008. **12**(5): p. 505-14.
95. Annibale, P., et al., *Identification of clustering artifacts in photoactivated localization microscopy.* Nat Methods, 2011. **8**(7): p. 527-8.
96. Coltharp, C., R.P. Kessler, and J. Xiao, *Accurate Construction of Photoactivated Localization Microscopy (PALM) Images for Quantitative Measurements.* PLoS ONE, 2012. **7**(12).
97. Van Royen, M.E., et al., *Nuclear proteins: Finding and binding target sites in chromatin.* Chromosome Research, 2011. **19**(1): p. 83-98.
98. Fudenberg, G. and L.A. Mirny, *Higher-order chromatin structure: bridging physics and biology.* Curr Opin Genet Dev, 2012. **22**(2): p. 115-24.
99. Benke, A., et al., *Multicolor single molecule tracking of stochastically active synthetic dyes.* Nano Letters, 2012. **12**(5): p. 2619-2624.
100. Latchman, D.S., *Eukaryotic Transcription Factors*. Eukaryotic Transcription Factors. 2007.
101. Lee, T.K., et al., *A noisy paracrine signal determines the cellular NF-kappaB response to lipopolysaccharide.* Sci Signal, 2009. **2**(93): p. ra65.
102. Cheong, R., et al., *Information transduction capacity of noisy biochemical signaling networks.* Science, 2011. **334**(6054): p. 354-8.
103. Chen, F.E. and G. Ghosh, *Regulation of DNA binding by Rel/NF-κB transcription factors: Structural views.* Oncogene, 1999. **18**(49): p. 6845-6852.
104. Perkins, N.D., *Post-translational modifications regulating the activity and function of the nuclear factor kappa B pathway.* Oncogene, 2006. **25**(51): p. 6717-30.
105. Tafvizi, A., L.A. Mirny, and A.M. van Oijen, *Dancing on DNA: kinetic aspects of search processes on DNA.* Chemphyschem, 2011. **12**(8): p. 1481-9.
106. Halford, S.E. and J.F. Marko, *How do site-specific DNA-binding proteins find their targets?* Nucleic Acids Research, 2004. **32**(10): p. 3040-3052.
107. Slutsky, M. and L.A. Mirny, *Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential.* Biophys J, 2004. **87**(6): p. 4021-35.

108. Morris, M.C., et al., *A peptide carrier for the delivery of biologically active proteins into mammalian cells.* Nature Biotechnology, 2001. **19**(12): p. 1173-1176.
109. Bultman, S.J., T.C. Gebuhr, and T. Magnuson, *A Brg1 mutation that uncouples ATPase activity from chromatin remodeling reveals an essential role for SWI/SNF-related complexes in beta-globin expression and erythroid development.* Genes Dev, 2005. **19**(23): p. 2849-61.
110. Sudarsanam, P. and F. Winston, *The Swi/Snf family: Nucleosome-remodeling complexes and transcriptional control.* Trends in Genetics, 2000. **16**(8): p. 345-351.
111. Darzacq, X., et al., *Imaging transcription in living cells.* Annu Rev Biophys, 2009. **38**: p. 173-96.
112. Larson, D.R., et al., *Real-time observation of transcription initiation and elongation on an endogenous yeast gene.* Science, 2011. **332**(6028): p. 475-8.

Acknowledgements

First of all I would like to express gratitude to my scientific advisor Prof Suliana Manley who gave me an opportunity to work on interesting and challenging projects in her lab. I'm very thankful for her guidance and advice in all aspects of scientific process which allowed me to obtain very valuable experience and skills for my future career.

I would like to thank members of my thesis committee for their time and useful comments to make this thesis better: Prof Minh Quang Tran, Prof Johan Elf, Prof Ben Schuler and Prof David Suter.

I'm very grateful to my colleagues who were contributing to the research described in this thesis:

Pierre Fabre is a co-leader of the HoxD project (Chapter 3) and a biology expert who performed all biological manipulations for these experiments and introduced me into the world of developmental biology.

Nicolas Olivier and Julia G unzenhaueser and I designed and performed experiments described in Chapter 4. We were the first generation of LEB squad and worked together a lot on development of single molecule methods in the lab.

Andrea Callegari and I have been working together on the project described in Chapter 6. We had a lot of very interesting discussions and ideas bridging biology and physics created together.

Lina Carlini and I had numerous scientific discussions and work done together to improve the technical side of our studies. We also wrote one paper together.

I would like to acknowledge other lab members for our interesting and very useful discussions: Thomas Pengo, Seamus Holden, Kyle Douglass and Niklas Berliner.

I'm also very grateful to the people I met here in Lausanne and who become my close friends: Lina Carlini, Andrea Callegari, Sergey Nazarov, Thomas Pengo. Lada Sycheva, Zoltan Spiro, Vladimir Dorodnitsyn and Lilia Ibragimova You, guys, enriched my life and were very supportive. My life inside and outside the lab became much more memorable thanks to you.

Конечно же я выражаю огромную благодарность моей семье. Ваша поддержка и вера всегда помогают мне идти вперед.

CV

Alexander BENKE

aleksandr.benke@epfl.ch

Education

2004-2010 Master Degree, Department of Molecular and Biological Physics, Moscow Institute of Physics and Technology, Moscow, Russia.

2010-present time post-graduate student, Department of Basic Sciences, EPFL, Lausanne, Switzerland

Research Experience

2007-2010 Master of Science Degree Research Program in the Laboratory of Lipids Modulators of Immune Response, Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow.

Summer 2009 participant of Summer Research Programme 2009, EPFL, Lausanne, Switzerland

2010-present time research-assistant, Laboratory of Experimental Biophysics, EPFL, Lausanne, Switzerland

Supervision

Master project: Xavier Meylan. STORM imaging of DNA binding dyes

Teaching

EPFL General Physics I

EPFL Biophysics II

Peer-reviewed publications

1. Pavillon N, **Benke A**, Boss D, Moratal C, Kühn J, Jourdain P, Depeursinge C, Magistretti PJ, Marquet P. Cell morphology and intracellular ionic homeostasis explored with a multimodal approach combining epifluorescence and digital holographic microscopy. J Biophotonics. 2010 Jul; 3(7):432-6.
2. **Benke A**, Manley S. Live-cell dSTORM of cellular DNA based on direct DNA labeling. Chembiochem. 2012 Jan 23; 13(2):298-301. Featured very important paper and cover.

3. **Benke A**, Olivier N, Gunzenhäuser J, Manley S. Multicolor single molecule tracking of stochastically active synthetic dyes. Nano Lett. 2012 May 9; 12(5):2619-24
4. Carlini L, **Benke A**, Lukinavicius G, Reymond L, Manley S. Reduced dyes enhance single molecule localization density for live superresolution imaging. Chemphyschem. 2014

Oral presentations

1. Annual Meeting of American Biophysical Society 2012: Direct live-cell super-resolution imaging of cellular DNA.
2. EMBO 10th Transcription and Chromatin Meeting 2012: Chromatin higher-order structure and dynamics with super-resolution microscopy and single-molecule-tracking.
3. Gordon Research Seminar and Research Conference 2013: Super-resolution microscopy reveals sub-diffraction higher-order chromatin organization and dynamics in living cells.

Awards

EPFL Winter Biophysical School 2012 - best poster award: A. Benke, and S. Manley. Live-cell super-resolution imaging of cellular DNA based on single-molecule localization.

Notes