Analysis of transcription factor and histone modification dynamics in the nucleus of single living cells using a novel antibody-based imaging approach

Sascha Conic

UNIVERSITÉ DE STRASBOURG

École doctorale des Sciences de la Vie et de la Santé

IGBMC – CNRS UMR 7104 – Inserm U 1258

THÈSE présentée par:

Sascha CONIC

soutenue le : **27 septembre 2018**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Spécialité : Aspects moléculaires et cellulaires de la biologie

Analysis of transcription factor and histone modification dynamics in the nucleus of single living cells using a novel antibody-based imaging approach

THÈSE dirigée par:

M. László TORA Directeur de Recherche, Université de Strasbourg

RAPPORTEURS:

M. Christof GEBHARDT Professeur, Ulm University

M. Lothar SCHERMELLEH Directeur de Recherche, University of Oxford

---

AUTRES MEMBRES DU JURY:

Mme Evi SOUTOGLOU Directeur de Recherche, Université de Strasbourg

Acknowledgements

First of all, I would like to thank Dr. Evi Soutoglou, Prof. Dr. Christof Gebhardt and Dr. Lothar Schermelleh for accepting to be members of my jury. Thank you all for taking the time to evaluate my thesis.

I also thank Laszlo for being my supervisor and for giving me the opportunity to work in such an excellent scientific environment. The last four years were very exciting but also sometimes exhausting. However, I could always come to you if I had problems or questions and you were always there for me to help me. Thanks so much for being so supportive, trustful and encouraging!

Next, I would like to thank Etienne for all his support over the last 8 years which started already during my studies at the ESBS and continued all through my PhD and therefore actually for my whole scientific career until now. I was able to start my scientific career in your lab and I would not be where I am now without your constant help and support. Thank you!

Now, I would like to thank both past and current members of the Tora lab. First, I would like to thank Tiago, Federica and Ivanka who more or less started together with me in the lab and with whom I had a lot of laughs and fun moments (even during the Halloween nights which should be scary normally). Next, I thank Didier and Stéphane for all their supportive comments through all the years and especially in the lab meetings which helped a lot to build my work. I thank Nikolaos for all the funny moments we spend together and his help if I had some questions concerning microscopy. I want to thank Matthieu and especially Eli for all their help in the lab and for their tips and tricks concerning bench work I learned from them over the years. Next, I thank Farrah, Pooja, Fang, Gizem and Vincent for the wonderful moments we shared in the lab and outside of it during all the fun stuff we made over the time. I want to thank Paul and Changwei for all the interesting discussions over the years and for being the best “module-mates” that you can have. I also thank our master students David and Emma and especially Emma for being the best and most motivated master student a supervisor can dream of. I would like to thank Kenny for all the funny discussions we had which were always “super easy, barely an inconvenience”. A very special thank you goes to one of my best friends Veronique who is always there if I

need help and for all the discussions and laughs we shared which could lighten every dark day. Finally, I would like to thank Chen-Yi, Marjorie and Sarina for all the great moments we shared over the time I had the pleasure to share the same lab with you.

I also want to thank everyone in the IGBMC who helped to improve my work and shape me as a scientist. Therefore, I want to thank all the facilities of the IGBMC especially the imaging, FACS and antibody facilities. I want to thank especially Mustapha from the antibody facility for all the nice discussions over results and of course all the antibodies I got over the years for my project. A special thank you goes to Alexia Loynton-Ferrand for all the time we spend together at the 3D SIM microscope, the resulting wonderful images we got and especially for all the discussion we had during the acquisitions about science but also other subjects of life. I also would like to thank the ERC for the financing of my work.

Moreover, I want to thank Dr. Lothar Schermelleh and Dr. Thomas Sexton for their helpful support and comments during my mid-thesis.

Next, I would like to thank everyone outside from my scientific life who helped and supported me over the years. I would like to thank all my friends and my family who were always there for me. Especially my brother Oliver and my father Sava who supported me over all the years of my study as well as my mother Eveline who always believed in me and motivated me to go on. I couldn’t have a better family! I would also like to thank the Lehn family for all their support (and cakes) and all the wonderful time we spent over the years. Last but never least, I would like to thank the love of my life, Désirée, for everything she has done for me. Her love and constant support kept me going over the years and I hope that it will go on forever. Thank you so much!

Abstract

In eukaryotic cells, gene transcription is controlled by a plethora of different proteins which preassemble in multiprotein complexes. In case of class II transcription this process is controlled by RNA Polymerase II (RNA Pol II) and several general transcription factors (GTFs). However, most of our basic knowledge about transcription and transcription regulation originate from biochemical experiments using cell extracts as well as purified proteins or from immunofluorescence (IF) experiments using fixed cells. Consequently, many efforts have been devoted recently to obtain information about the dynamic movements, assembly or nuclear localization/distribution of transcription factors involved in the subsequent steps of transcription directly from living cells.

Therefore, the visualization of protein complexes like RNA Pol II in single living cells is of high importance and can give new insights about their natural behavior. A crucial step to obtain reliable results is the fluorescent labeling of the target proteins. However, labeling of proteins for live cell studies is often performed by overexpressing fluorescently tagged proteins (FTPs) which can behave differently to their endogenous counterparts and are unable to visualize specific posttranslational modifications (PTMs). Thus, there is a demand for imaging tools which can be used to gain insights into the dynamic behavior of endogenously expressed proteins and PTMs in single living cells.

Therefore, we developed a labeling strategy, named versatile antibody-based imaging approach (VANIMA), in which fluorescently labeled antibodies are introduced into living cells to image specific endogenous proteins or PTMs. We were able to show that VANIMA can be used to study dynamical processes of fundamental biological mechanisms including factors of the transcription machinery like RNA Pol II and TAF10 as well as histone modifications in form of phosphorylated histone H2AX in living human cancer cells using conventional or super-resolution microscopy. Initial experiments also indicated that VANIMA can be combined with genetic labeling strategies to study RNA Pol II recruitment dynamics directly at a gene array. Hence, in the future VANIMA will serve as a valuable tool to uncover the dynamics of endogenous biological processes including transcription directly in single living cells.

Résumé

Introduction

Dans les cellules eucaryotiques, la transcription des gènes est contrôlée par un ensemble de protéines différentes, capables de former des complexes multi-protéiques. Dans le cas de la transcription de classe II, ce processus est contrôlé par l'ARN Polymérase II (ARN Pol II), ainsi que par des facteurs de transcription généraux (General Transcription Factors, GTFs). Les complexes multi-protéiques suivants font partie des GTFs : TFIIA, TFIIB, TFIID (contenant la « TATA-binding protein » (TBP)), TFIIE, TFIIF et TFIIH. Ces protéines jouent un rôle fondamental dans la reconnaissance du promoteur, le recrutement de l'ARN Pol II, l'interaction avec des facteurs de régulation, l'ouverture de la double hélice d'ADN ainsi que dans la reconnaissance du site d'initiation de la transcription (Transcription Start Site, TSS). La première étape dans la transcription d'un gène codant pour une protéine est la liaison des GTFs sur le promoteur du gène et le recrutement de l'ARN Pol II par les GTFs, qui résulte à la formation du complexe de pré-initiation (Preinitiation Complex, PIC). Les GTFs vont par la suite ouvrir la double hélice d'ADN et guider l'ARN Pol II au TSS. Ces évènements vont aboutir à la transition de l'ARN Pol II en phase d'élongation, pendant laquelle l'enzyme va transcrire de façon active l'ADN en ARN.

La majorité de ces connaissances concernant la transcription et sa régulation a été acquise grâce à de biologie moléculaire, de génétique et des expériences de liaison statique ou à des expériences d'Immunofluorescence (IF) sur des cellules fixées. Par conséquence, peu de choses sont connues concernant le mouvement dynamique des facteurs de transcription qui s'impliquent dans les phases de transcription qui suivent. Afin d'étudier la régulation de l'expression des gènes *in vivo*, des expériences de suivi de molécules uniques (single particle tracking) sont nécessaires, pour analyser la dynamique et la fonction de la machinerie de transcription Pol II en utilisant des complexes de facteurs de transcription, comme TFIID ou m'ARN Pol II elle-même.

De ce fait, la visualisation de complexes protéiques comme l'ARN Pol II dans des cellules vivantes est très importante et pourrait révéler des informations concernant le comportement de la protéine au sein de la cellule. Le marquage fluorescent des protéines-cibles serait dans ce cas crucial pour l'obtention de résultats fiables. Pourtant, la majorité des protéines étudiées à l'aide de la microscopie à fluorescence sont des protéines marquées avec un fluorochrome (Fluorescently-Tagged Proteins, FTPs) surexprimées, qui ne se comportent pas toujours comme les protéines endogènes. Il est connu que la fonction des facteurs de transcription impliqués dans les processus dépendants de la chromatine est étroitement liée à leurs interactions avec des diverses modifications post-traductionnelles (Post-Translational Modifications, PTMs) au sein du noyau. Ces interactions ne peuvent pas être observées sans l'utilisation de FTPs. Il serait ainsi impératif de développer des nouveaux outils afin l'étudier la dynamique des protéines endogènes au sein de cellules uniques vivantes. Par conséquence, mon projet se base sur le développement d'une nouvelle technique de marquage de protéines, nommée « approche versatile d'imagerie basée sur des anticorps » (Versatile Antibody-based Imaging approach ; VANIMA). Cette technique se base à l'électroporation afin d'introduire des anticorps marqués dans des cellules vivantes. Ces anticorps peuvent se lier spécifiquement à leur cible endogène (soit une protéine, soit une PTM), permettant de la visualiser sous un microscope à fluorescence. En utilisant cette méthode, plusieurs protéines nucléaires faisant partie de la machinerie de transcription ou encore des PTM peuvent être étudiées par des approches de microscopie « conventionnelle » ou à super-résolution. En plus, nous avons combiné VANIMA avec un marquage génétique et un système qui permet d'activer la transcription, afin d'étudier la dynamique de l'ARN Pol II endogène au sein de cellules vivantes, à un locus génétique spécifique.

Résultats

La première étape a été de mettre au point cette nouvelle technique de marquage, appelée VANIMA, afin de pouvoir étudier des protéines endogènes ou des PTMs dans des cellules vivantes. Pour cela, nous avons électroporé un anticorps marqué avec un fluorochrome dans des cellules U2OS vivantes. Cet anticorps est dirigé contre RPB1 (ainsi nommé anti-RPB1), la plus grande sous-unité d'ARN Pol II. L'anticorps était détectable dans le cytoplasme des cellules 6 heures après

électroporation ; 24 heures post-électroporation, l'anticorps se trouvait dans le noyau. Sachant que les anticorps sont très grands (150 kDa) pour pouvoir diffuser de façon passive dans le noyau, nos résultats indiquent que l'anticorps s'est lié sur sa cible néo-synthétisée, RPB1, et a été ensuite « piggybacked » au noyau (Figure 1).

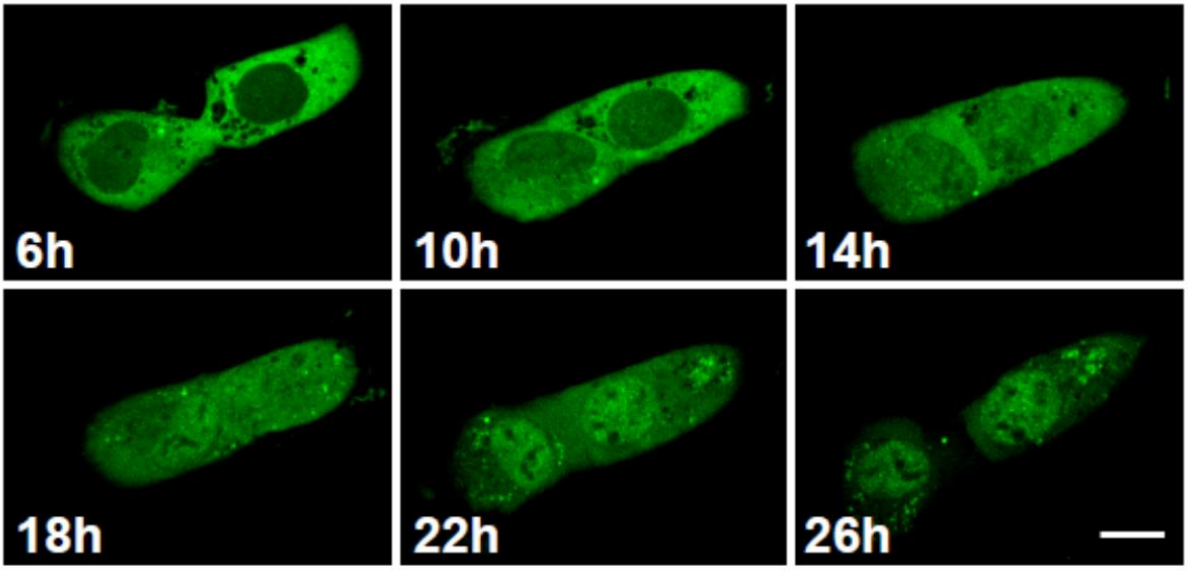


Figure 1: Transport d'anti-RPB1 mAb transduit du cytoplasme au nucleus dans des cellules U2OS. *Après la transduction d'anticorps anti-RPB1 marqué avec de l'Alexa Fluor 488, les cellules ont été visualisé en utilisant la microscopie confocal à disque rotatif après 6 h d'incubation et ensuite toutes les heures pendant une période de 20 h. Echelle : 15 µm.*

Dans le but de démontrer que l'anticorps se lie vraiment à la protéine-cible *in cellulo*, nous avons fait des expériences d'immunoprécipitation (IP). Des cellules U2OS ont été électroporées soit sans anticorps (contrôle négatif), soit avec des quantités croissantes d'anticorps anti-RPB1. 24 heures post-électroporation, les cellules ont été utilisées pour une extraction de cellules entières ; les extraits ont été purifiés en utilisant des billes recouvertes de protéine G, afin de tester si l'anticorps était toujours lié à RPB1. Nous n'avons pas détecté de protéine RPB1 dans le contrôle négatif. Dans le reste des échantillons, la quantité d'anticorps électroporé était inversement proportionnelle à la quantité d'RPB1 détectée dans les extraits cellulaires, ainsi indiquant que l'anticorps anti-RPB1 reste lié à sa cible dans la cellule (Figure 2).

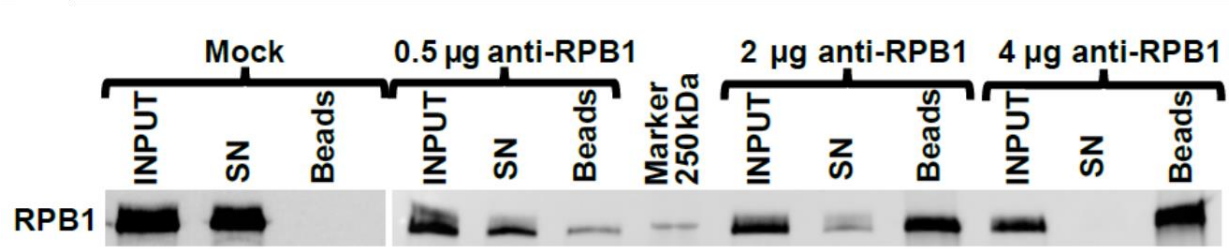


Figure 2: ***Capacité de liaison de l'anti-RPB1 mAb dans des cellules U2OS.*** *Les cellules ont été électroporer avec 0 (mock), 0.5, 2, et 4 µg de l'anti-RPB1 mAb et des extraits entiers des cellules préparer 24 h après transduction (INPUT) ont été mélanger avec des billes de protéine G. Le matériel lié et non-lié a été analyser par Western blot. Le blot montre la proportion des molécules d'ARN Pol II liés aux anticorps absorber sur les billes (beads) ou restant dans le surnageant (SN), détecter avec un anticorps secondaire.*

Afin de tester si d'autres cibles nucléaires pouvaient être marquées en utilisant VANIMA, nous avons électroporé des anticorps ciblant les facteurs de transcription TAF10 (anti-TAF10) et TBP (anti-TBP), qui font partie des sous-unités de TFIID. Tous les deux anticorps ont donné les mêmes résultats concernant le marquage des facteurs de transcription endogènes. Au contraire, dans le cas où des anticorps n'ayant pas de cible cellulaire (anti-MBP) ou ciblant une protéine cytoplasmique (anti-α-tubuline) ont été électroporés, le signal fluorescent était cytoplasmique, même après 24 heures d'incubation (Figure 3).

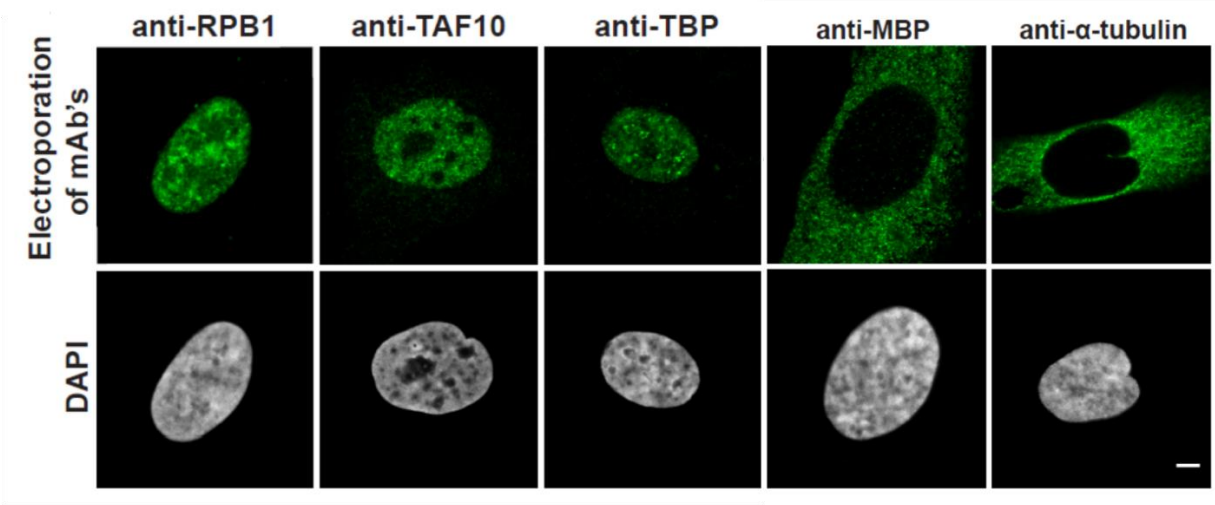


Figure 3: ***Visualisation des facteurs de transcription endogènes avec VANIMA.*** *Des mAbs marqués lient spécifiquement aux facteurs de transcription RPB1, TAF10 et TBP ainsi que des mAbs marqués de control contre MBP et α-tubulin ont été transduit dans des cellules U2OS et leur localisation dans les cellules a été monitorer par microscopie confocale 24 h après traitement. Un seul plan de z est montré par condition. Les images représentent un nucleus typique enregistré dans chaque cas après fixation des cellules et contre-coloration subséquente avec DAPI. Echelle: 5 µm.*

Pour démontrer l'utilité de notre approche en ce qui concerne l'imagerie, nous avons comparé les anticorps marqués électroporés avant (150 kDa) avec leurs fragments Fab correspondants (50 kDa), étant donné que les Fabs peuvent entrer de façon passive dans le noyau des cellules. Cette comparaison a démontré que les mAbs marqués et leurs fragments Fab correspondants se comportent de la même façon en ce qui concerne le marquage des facteurs de transcription endogènes. Une remarque importante était que les fragments Fab marqués, ciblant des protéines nucléaires arrivent au noyau 6h après électroporation, alors que les anticorps y sont « piggybacked » après environ 24-48 h (Figure 4).

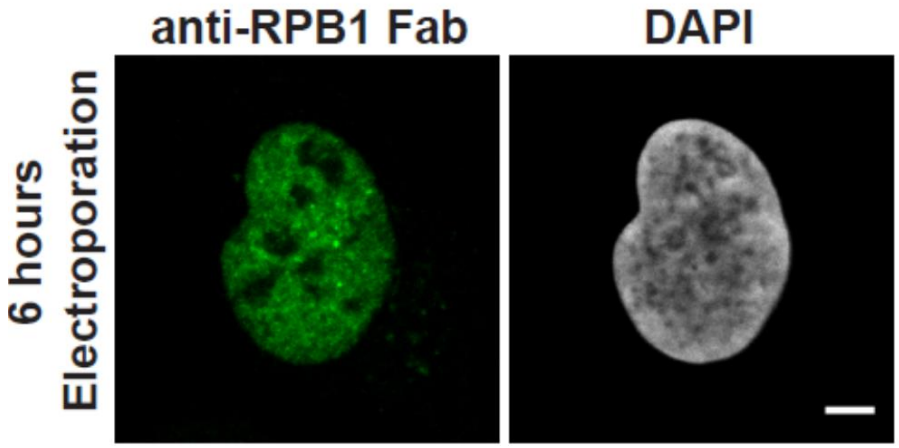


Figure 4: Visualisation du RPB1 endogènes en utilisant des fragments Fab d'anti-RPB1 marqués. *Le fragment de Fab contre RPB1 marqué avec de l'Alexa488 a été transduit dans des cellules U2OS et monitorer par microscopie confocale 6 h post-électroporation. Echelle : 5 µm.*

Nous avons également testé si les anticorps marqués reconnaissaient des PTMs associés à la chromatine. Pour ce faire, nous avons utilisé des fragments anticorps (Fab) ciblant γH2AX, un marqueur de cassures double-brin de l'ADN. Des Fabs anti-γH2AX ont été électroporés dans des cellules « contrôle », ainsi que dans des cellules où des cassures dans l'ADN étaient induites. Comme attendu, des Fab marqués à l'Alexa 488 étaient capables d'entrer dans le noyau et de se lier sur des foci de H2AX phosphorylée, au sein des cellules traitées avec de l' hydroxyurée ou du néocarzinostatine. Les Fabs électroporés peuvent donc se lier à des PTMs sur la chromatine de cellules vivantes. Afin de vérifier si les anticorps électroporés ont un impact sur la transcription, nous avons isolé des ARN néo-synthétisés à partir de cellules électroporés et nous les avons analysés par RT-qPCR. Nos résultats montrent qu'il n'y a pas d'effet détectable sur la transcription, pour tous les anticorps testés (Figure 5).

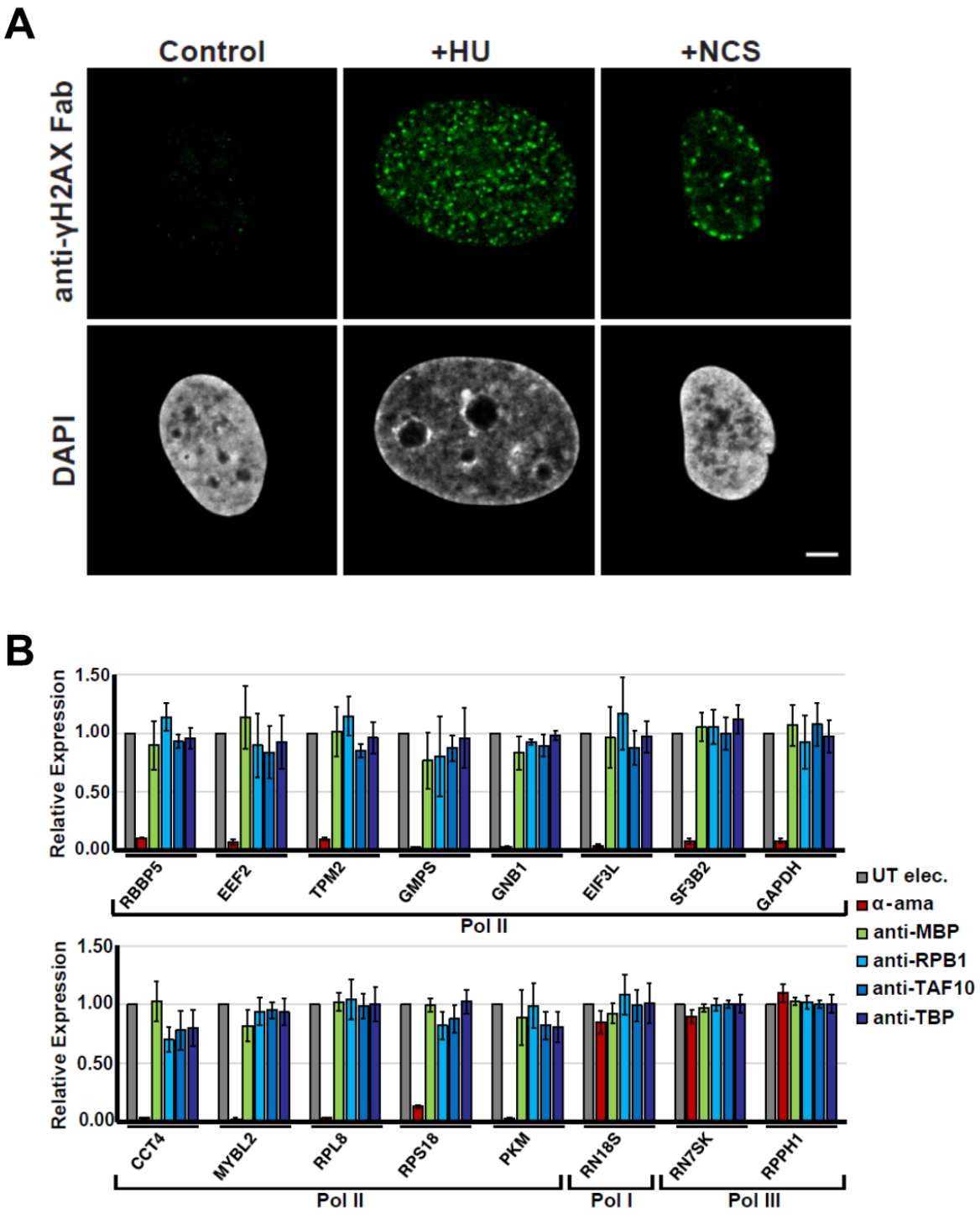


Figure 5: Légende à la page suivante →

Figure 5: ***Visualisation de l'H2AX phosphorylé endogènes et expérience de control pour tester que les mAbs n'inhibent pas la transcription des ARN pré-messager. (A)*** *Le Fab marqué contre yH2AX a été transformer dans des cellules U2OS et leur localisation a été enregistrer 24 h après électroporation en utilisant la microscopie confocale et après traitement des cellules électroporées avec soit du NCS (pour 15 min) ou soit du HU (pour 48 h). Control, cellules non traitées. Un nucleus typique est représenté pour chaque cas après fixation des cellules et contre-coloration subséquente avec DAPI. Echelle : 5 µm.* ***(B)*** *Des cellules U2OS électroporées mais sans anticorps (UT elec), électroporées et traitées avec de l'α-amanitin (α-ama), électroporées avec un anticorps de control se lient au MBP bactériale (anti-MBP), ou électroporées avec des mAbs reconnaitront spécifiquement RPB1, TAF10, ou TBP (anti-RPB1, anti-TAF10, ou anti-TBP). 24 h après électroporation, les ARN totaux ont été isolé, et l'expression des gènes de Pol I, Pol II, et Pol III a été analyser par PCR quantitative en temps réel. Des transcrits de Pol III ont été utilisé pour la normalisation. Les ARN nouveau synthétisés des gènes indiqués ont été quantifié avec des pairs d'amorce validés. L'histogramme correspond à la valeur moyenne obtenue de trois expériences indépendantes.*

Ensuite, nous avons voulu obtenir des images des facteurs de transcription endogènes en haute résolution ; pour cela, nous avons utilisé de la microscopie à super-résolution. En observant RPB1 et TAF10 en utilisant VANIMA et de la microscopie à super-résolution 3D-SIM nous a permis de détecter des spots bien définis, correspondant à l'ARN Pol II ou à TFIID dans le noyau des cellules. Ces images ont été utilisées pour quantifier le nombre et la taille des foci observés, afin de mesurer les changements dynamiques de la distribution d'ARN Pol II au sein du noyau. De façon intéressante, la distribution du volume des foci RPB1 (ARN Pol II) changeait après traitement au flavopiridol (flavo), qui est un inhibiteur de la phase d'élongation. Le nombre foci RPB1 de grande taille, ayant un volume supérieur à 10^{-2} µm^3 diminue drastiquement après traitement au flavo. Au contraire, la distribution de la taille des foci TAF10 (TFIID) n'est pas impactée. Ceci indique qu'il existe des assemblages multimoléculaires (« clusters ») d'ARN Pol II de grande taille, qui se dissocient après traitement au flavo parce que les molécules d'ARN Pol II se dissocient de la chromatine et deviennent plus mobiles. En utilisant de la microscopie à super-résolution 3D-SIM sur cellules vivantes, nous avons pu observer que les grands assemblages d'ARN Pol II sont dynamiques et qu'ils s'associent et se dissocient constamment (Figure 6).

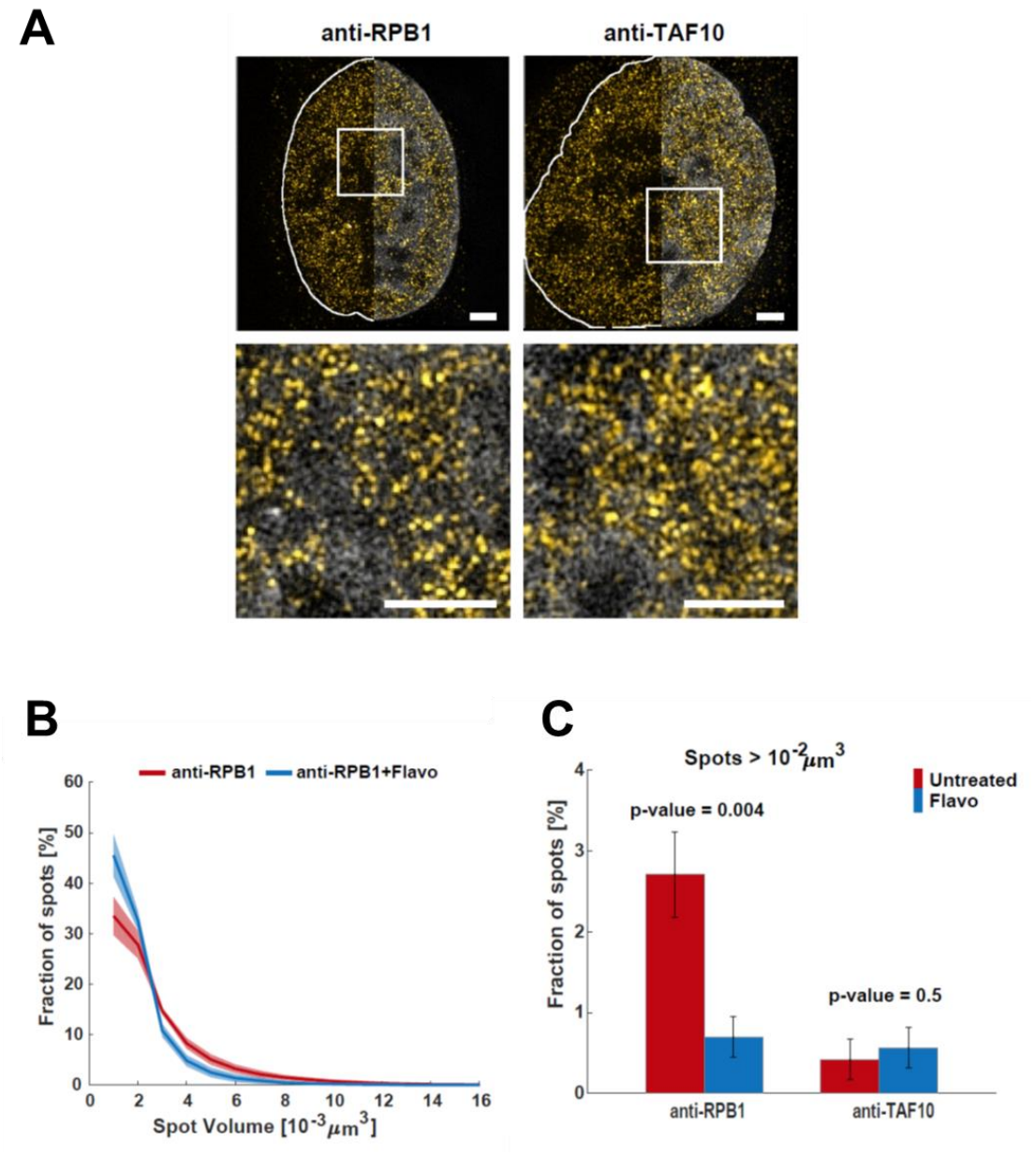


Figure 6: Légende à la page suivante →

Figure 6: ***Visualisation et quantification de la distribution des facteurs de transcription en utilisant VANIMA et la microscopie à super-résolution. (A)*** *Les mAbs marqués se lient aux facteurs de transcription RPB1 et TAF10 (jaune) ont été transformé dans des cellules U2OS, et leur localisation dans les cellules a été monitorer 24 h après transformation en utilisant 3D-SIM. Les images montrent un nucleus typique enregistré dans chaque cas après fixation et traitement avec DAPI (gris). La projection d'intensité maximale de Z de cinq coups montre les mAbs marqués avec (partie droite) ou sans (partie gauche) contre-coloration avec DAPI (gris). La ligne blanche solide décrie le contour du nucleus. En bas : Magnification des régions d'intérêt blanche, sous l'image correspondant. Echelle : 2 µm.* ***(B)*** *Des cellules U2OS ont été transformé avec de l'anti-RPB1 mAb marqués avec de l'Alexa Fluor 488 et après avoir été traité avec Flavo (2 µM) pour 1 h ou pas (untreated). 24 h après traitement des cellules ont été fixé et analyser par 3D-SIM. Le nombre de points individuels et leurs volumes dans des nuclei individuelles ont été quantifié en utilisant les logiciels Fiji/ImageJ et Matlab. Le graphique montre le pourcentage des points avec un volume donné dans des cellules non traitées (rouge) et traitées avec du Flavo (bleu) acquis de 10 cellules individuelles pour chaque condition.* ***(C)*** *Les volumes des points ont été extraire et le pourcentage des points de RPB1 et TAF10 avec un volume >10^{-2} μm^3 dans des cellules non traitées (rouge) et traitées avec du Flavo (bleu) sont montrés. Les indices d'erreur représentent le SE pour 10 cellules individuelles de chaque condition.*

Par la suite, nous avons combiné VANIMA avec une lignée cellulaire U2OS, dans laquelle une séquence répétée d'ADN (gene array) a été intégrée de façon stable dans le génome. Ce gene array, contenant l'opérateur de l'opéron lactose, permet de marquer le locus génétique en utilisant une protéine LacI marquée avec un fluorochrome. En plus, le gene array contient des Tet Response Element (TRE), permettant d'induire la transcription des gènes de l'array en utilisant un activateur fluorescent. En électroporant également l'anticorps anti-RPB1 dans cette lignée cellulaire, c'est possible d'étudier la dynamique de l'ARN Pol II dans des cellules vivantes sur un locus spécifique après induction de la transcription. Nous avons commencé à calculer la dynamique du recrutement de l'ARN Pol II sur le gene array en observant des cellules vivantes sous le microscope confocal. De façon intéressante, l'accumulation de l'ARN Pol II et de l'activateur sur l'array commencent en même temps, indiquant que le recrutement de l'ARN Pol II commence très rapidement après la liaison de l'activateur sur le gene array (Figure 7).

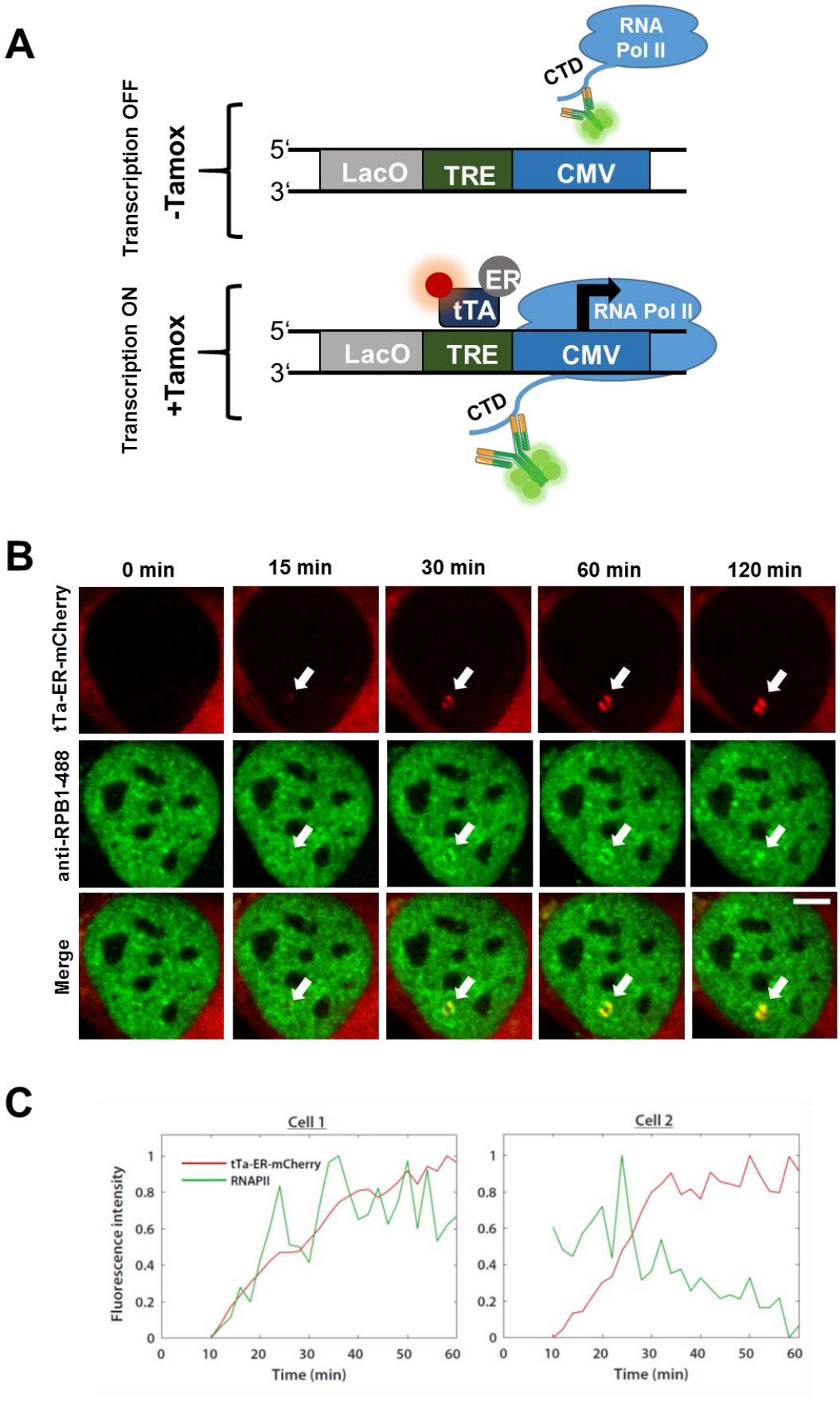


Figure 7: Légende à la page suivante →

Figure 7: ***Combinaison de VANIMA avec une « array » inductible de gêne pour mesurer le recrutement de l'ARN Pol II endogène dans des cellules vivantes. (A)*** *Représentation schématique du array de gêne comprenant l'opérateur de Lac (LacO), l'élément réceptive au Tet (TRE) et un promoteur CMV ainsi que les éléments se lient au array de gêne avant et après l'induction de transcription avec tamoxifen (+tamox) ou pas (-tamox). L'activateur fusionné avec du mCherry (tTa-ER) va se lier et marquer l'ADN seulement en présence de tamoxifen. Des cellules ont été electroporer avec des anti-RPB1 mAbs marqués avec un fluorophore pour marquer l'ARN Pol II qui va se lier après l'activation de la transcription.* ***(B)*** *Des cellules U2OS 2-6-3 ont été électroporer avec du plasmide codant pour le tTa-ER activateur marqué avec du mCherry et du mAbs marqué avec de l'Alexa488. Après 24 h, les cellules ont été traité avec tamoxifen et visualiser en utilisant la vidéomicroscopie confocal pour 3 h avec une image tous les 15 min. Les têtes de flèche blanche indiquent des accumulations de l'activateur ou de l'ARN Pol II à l'array de gêne. Echelle : 5 µm.* ***(C)*** *Les fluctuations de fluorescence ont été mesuré au point du array de gêne du tTa-ER-mCherry activateur et de l'anti-RPB1-Alexa488 (ARNPII) pour 1 h post-induction avec tamoxifen. L'intensité de fluorescence maximale a été mise à 1. L'activateur transcriptionnelle et les profils d'intensité de l'ARNPII sont représentés en rouge et verte, respectivement. Deux cellules représentatives sont montrées pour indiquer la variabilité élevée observé entre cellules.*

Conclusion

Pour conclure, nous avons développé une stratégie simple pour visualiser des antigènes-cibles sous leur forme native dans des cellules uniques vivantes, qui n'est pas toxique pour les cellules traitées. Le marquage de protéines endogènes en utilisant VANIMA correspond aux vrais complexes antigène-anticorps qui se forment dans la cellule après électroporation de l'anticorps. Cette approche peut être utilisée pour détecter une grande variété de facteurs et de PTMs en faisant de l'imagerie de super-résolution sur des cellules vivantes et uniques. De plus, en utilisant VANIMA, des processus biologiques fondamentaux et dynamiques peuvent être visualisés dans des cellules non fixées à haute résolution.

Nos résultats suggèrent que des larges foci détectés pourraient contenir plusieurs assemblages multimoléculaires d'ARN Pol II ou même des « trains » d'ARN Pol II, possiblement organisés en domaines topologiques (TADs) ou/et sous forme d'autres régions de contrôle. Le fait que la taille des foci d'ARN Pol II native détectés avec VANIMA diminuait après inhibition de la transcription est en accord avec d'autres

études, démontrant que l'ARN Pol II quitte la chromatine et devient mobile. Nous avons aussi montré que les grands foci d'ARN Pol II se forment constamment et se dissocient ou se réassocient de façon dynamique.

En plus, en couplant VANIMA avec du marquage génétique, de la transcription inductible et de l'imagerie confocale sur cellules vivantes, nous pourrions observer le recrutement de l'ARN Pol II endogène sur le gene array après activation de la transcription et le comparer à la liaison de l'activateur, ainsi découvrant plus concernant la dynamique de l'ARN Pol II *in vivo*. En utilisant VANIMA couplée à l'imagerie 3D-SIM sur cellules vivantes et/ou au marquage génétique décrit, ce sera possible de caractériser et disséquer la fonction et la dynamique de la transcription ARN Pol II-dépendante dans des cellules uniques vivantes.

Table of Contents

Table of Figures

Table of Tables

Abbreviation List

AcCoA	acetyl-coenzyme A
Anch3	specific chromosome partition sequence
ANCHOR	DNA sequence able to bind bacterial partition proteins
ATAC	Ada2a-containing complex
ATM	ataxia-telangiectasia mutated
ATP	adenosine triphosphate
ATR	ataxia telangiectasia-mutated and Rad3-related
ATXN7	ataxin 7
AQR	RNA helicase aquarius
bp	base pair
BRE	TFIIB recognition elements
BTAF1	B-TFIID TATA-Box Binding Protein Associated Factor 1
Cas9	CRISPR-associated 9
CCNT1	cyclin T1 gene
CDK	cyclin dependent kinase
cDNA	complementary DNA
CFP	cyan fluorescent protein
CHD family	chromodomain-helicase-DNA binding
ChIP	chromatin immunoprecipitation
CLMS	cross-linking/mass spectrometry
CMV	cytomegalovirus
CPF	cleavage and polyadenylation factor
CpG-islands	genomic region with a high GC percentage
CPSF	cleavage and polyadenylation specificity factor
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CstF	cleavage stimulatory factor
CTD	C-terminal domain
Ctk1	cyclin-dependent serine/threonine kinase 1

CUTs	Cryptic Unstable Transcripts
DAPI	4′,6-Diamidin-2-phenylindol
DBD	DNA-binding domain
DDR	DNA damage response
DEAD/H superfamily	DEAD box helicases
DNA	deoxyribonucleic acid
DNA Pol α	DNA polymerase α
DNA-PKcs	DNA-dependent protein kinase catalytic subunit
Dox	doxycycline
DPE	downstream core promoter element
DSBs	double-strand breaks
DSIF	DRB sensitivity-inducing factor
Dub module	deubiquitination module
EM	electron microscopy
ENY2	enhancer of yellow 2 homolog
ER	estrogen receptor
ES	embryonic stem
ESCs	embryonic stem cells
Fabs	antibody fragments
FabLEM	Fab-based live endogenous modification labeling
FACT	facilitates chromatin transcription
FCP1	TFIIF-associating CTD phosphatase 1
FCS	fluorescence correlation spectroscopy
FISH	fluorescence in situ hybridization
FPALM	fluorescence photoactivation localization microscopy
FPs	fluorescent fusion proteins
fps	frames per second
FRAP	fluorescence recovery after photobleaching
FRET	Förster resonance energy transfer
FTPs	fluorescently tagged proteins

GCN5	general control nonderepressable 5
GFP	green fluorescent protein
gRNA	guide RNAs
GTFs	general transcription factors
h	hour
HATs	histone acetyltransferases
HDACs	histone deacetylases
HeLa	human adenocarcinoma cell line
Hi-C	chromosome conformation capture
His-tag	histidine-tag
HKMTs	histone lysine methyltransferases
hPAF1	human RNA polymerase II associated factor 1
HRD	histidine-rich domain
HRMTs	histone arginine methyltransferases
hSAGA	human SAGA
H2AX	histone H2AX
H3K4	histone 3 lysine 4
H3K3me1	modification of histone H3 by mono-methylation of lysine 3
H3K4me2/3	modification of histone H3 by di/tri-methylation of lysine 4
H3K9	histone 3 lysine 9
H3K9ac	modification of histone H3 by acetylation of lysine 9
H3K27	histone 3 lysine 27
H3K27ac	modification of histone H3 by acetylation of lysine 27
H3K27me1	modification of histone H3 by mono-methylation of lysine 27
H3K27me3	modification of histone H3 by tri-methylation of lysine 27
H3K36	histone 3 lysine 36
H3K36me2/3	modification of histone H3 by di/tri-methylation of lysine 36
H3K79	histone 3 lysine 79
H3K79me2/3	modification of histone H3 by di/tri-methylation of lysine 79
IF	immunofluorescence

Imp	importins
INO80 family	family of chromatin remodeler enzymes
Inr/INR	initiator
ISWI family	imitation SWI family
IUPAC	International Union of Pure and Applied Chemistry
K_{off}	dissociation rates
K_{on}	association rate
kb	kilobase
KDa	kilo-dalton
LacO	lac-operon
LacI protein	lac repressor protein
LCDs	low-complexity domains
Mbp	mega base pairs
Mcm2-7	minichromosome maintenance 2-7
MCP	MS2 coat protein, MS2-coating protein
MDa	mega-dalton
min	minutes
Mot1	TATA-binding protein-associated factor in yeast
mRNA	messenger RNA
MS2	sequence from bacteriophage MS2
MTE	motif ten element
MudPIT	Multidimensional Protein Identification Technology
MYST HAT	MYST family of histone acetyltransferases
NC2	negative cofactor 2
NELF	negative elongation factor
NF-Y	nuclear transcription factor Y
NLS	nuclear localization signal
nM	nanomolar
nm	nanometer
Oct4	octamer-binding transcription factor 4

ORF	open reading frame
PALM	photoactivated localization microscopy
ParB	partition protein B
PCAF	p300/CBP-associated factor
PCNA	proliferating cell nuclear antigen
PIC	preinitiation complex
PIKKs	phosphatidylinositol-3-OH-kinase-like family of protein kinases
PHD fingers	plant homeodomain zinc fingers
PP2A	phosphatases 2A
PP2Cγ	phosphatases 2Cγ
PP4	phosphatase 4
PRC2	polycomb repressive complex 2
PSF	point spread function
P-TEFb	positive transcription elongation factor b
PTMs	posttranslational modifications
PYP	photoactive yellow protein
p8/34/44/52/53/55/62	protein 8/34/44/52/53/55/62
p300/CBP	p300/CREB binding protein
Rap1	Ras-related protein 1
RECQL5	RecQ like helicase 5
R-loops	RNA-loops
rRNA	ribosomal RNA
RNA	ribonucleic acid
RNAi	RNA interference
RNA Pol	RNA polymerases
RNAP IIA	RNA polymerase IIA
RNAP IIO	RNA polymerase IIO
RNF20/40	ring finger protein 20/40
RPB1-12	RNA Pol II subunit B1-12
Rtr1	Regulator of transcription 1

rtTA	reverse Tetracycline-controlled transactivator
Rtt109	Regulator of Ty1 transposition protein 109
SAGA complex	Spt-Ada-Gcn5 acetyltransferase complex
scFv	single-chain variable fragment antibody
Ser2/5/7	serine 2/5/7
SETX	Probable helicase senataxin
Set2	SET domain-containing protein 2
SGF29	SAGA complex-associated factor 29
siRNA	short interfering RNA
SKL	peroxisome targeting signal
SMLM	single molecule localization microscopy
snRNAs	small nuclear RNAs
Sox2	transcription factor SOX-2
Sp1	transcription factor Sp1
SPT	single particle tracking
SPT5	transcription elongation factor SPT5
SUPT3	Transcription initiation protein SPT3 homolog
SUPT7L	STAGA complex 65 subunit gamma
SUPT20	Transcription initiation protein SPT20 homolog
Ssu72	RNA polymerase II subunit A C-terminal domain phosphatase SSU72
STED	stimulated emission depletion microscopy
STORM	stochastic optical reconstruction microscopy
SV40	Simian-Virus-40
SWI/SNF	SWItch/Sucrose Non-Fermentable
S-139	serine 139
TADA1/2B/3	transcriptional adapter 1/2B/3
TAF (1-13)	TBP-associated factors (1-13)
TANDs	TAF1 N-terminal domains
TATA box	promoter element with the sequence TATAWAAR
TBP	TATA Binding Protein

TC-NER	transcription coupled nucleotide excision repair
TFs	transcription factors
TFIIA/B/D/E/F/H/S	transcription factor A/B/D/E/F/H/S
TRCs	transcription-replication-conflicts
tRNA	transfer RNAs
TRRAP	Transformation/transcription domain-associated protein
TSS	transcription start site
tTA	tetracycline transactivator
UbcH6	Ubiquitin-conjugating enzyme E2 E1
USP22	Ubiquitin carboxyl-terminal hydrolase 22
U2OS	human osteosarcoma cell line
VANIMA	versatile antibody-based imaging approach
VHHs	single domain antibodies
Xist	x-inactive specific transcript
XPB	Xeroderma pigmentosum group B-complementing protein
XPD	Xeroderma pigmentosum group D-complementing protein
XPF	Xeroderma pigmentosum group F-complementing protein
XPG	Xeroderma pigmentosum group G-complementing protein
Xrn2	5'-3' exoribonuclease 2
3D-SIM	3D Structured illumination microscopy
4-OHT	4-hydroxytamoxifen
µm	micrometer

INTRODUCTION

Introduction

The possibility of a given organism to adapt to different environmental stimuli or to be able to maintain the cellular identity is dependent on transcriptional regulation. In the center of the transcription process are the RNA polymerases (RNA Pol) which are multisubunit complexes with the ability to read and convert the genetic information stored in the DNA into RNA. They can be found in all species, however, their composition and number varies across evolution. Bacteria, for example, have only one RNA Pol whereas eukaryotic cells harbor three different types of polymerases within their nucleus.

These three polymerases in eukaryotic cells are RNA Pol I, II and III which are all responsible for the transcription of nuclear genes. However, each polymerase is responsible for a specific subset of genes which can be classified as well as class I, II or III genes. RNA Pol I transcription accounts for up to 60% of the transcriptional activity in the cell and is responsible for the transcription of ribosomal RNA (rRNA). In contrast, transfer RNAs (tRNA), 5S rRNAs and some other untranslated small RNAs are synthesized by RNA Pol III. Taken together, around 80% of all genes in dividing cells are transcribed by the multiprotein complexes RNA Pol I and III. Lastly, RNA Pol II is a multisubunit complex consisting of 12 subunits that is transcribing mainly protein coding genes to produce messenger RNA (mRNA) but also some other classes of RNA like small nuclear RNAs (snRNA).

There are existing several differences between prokaryotic and eukaryotic transcription. Firstly, transcription in prokaryotes is coupled directly with translation of the mRNA into the protein whereas these two processes are separated in eukaryotes in two different cellular compartments (nucleus and cytoplasm). Another very important difference is the fact that prokaryotic RNA polymerase can bind directly to the transcription loci without any help of other factors whereas the eukaryotic polymerases need additional protein complexes, so called general transcription factors (GTFs), to be able to recruit the polymerase to the chromatin. All these complexes are recruited to a specific gene sequence upstream of the transcription start site (TSS) which is called the promoter region to form the preinitiation complex (PIC). Once the PIC is formed, transcription can be initiated and RNA Pol II is released to produce mRNA during transcription elongation.

Moreover, another difference is the tight packaging of the genome in eukaryotic nuclei to form the chromatin. Therefore, chromatin and not only DNA is the target of any DNA-related processes. The basic structural unit of the chromatin is the nucleosome core which consists of a histone octamer and 147 bp of DNA wrapped around it. The histone octamer is composed of two copies of the histones H2A and H2B (forming two H2A/H2B dimers) and two copies of H3 and H4 (forming one H3/H4 tetramer). However, even if the extensive packaging of the chromatin can act as a barrier for any DNA-related processes, there are factors which are able to “read” the chromatin and induce processes like transcription. Therefore, activators which are also known as gene specific transcription factors can bind to specific sequences on the chromatin to induce the recruitment of co-activator complexes. These co-activator complexes can harbor different functions to enable transcription by remodeling of the chromatin (by sliding or evicting nucleosomes) or by modifying histones (like acetylation or methylation). On the other side, there are also co-repressors which can interact with repressor transcription factors to inhibit the transcription of specific genes. Thus, the interplay of gene specific activators and co-activators leads to GTF binding at the promoter, PIC assembly and transcription initiation.

In fact, the mechanism of transcription is highly conserved from yeast to mammals and many important studies about transcription have been performed in *S. cerevisiae*. However, this introduction will mainly focus on the processes within metazoan cells with specific mentioning of yeast studies when necessary. In the first chapter, we will explore the regulation of RNA Pol II transcription and how different factors like activators, co-activators or GTFs are influencing global RNA Pol II transcription. Secondly, as my work was mainly focusing on the development of a new labeling technique to label endogenous transcription factors in living cells to study the distribution and dynamics of transcription, I will give an overview of the research that was performed concerning transcription imaging as well as discuss the advantages and disadvantages of different labeling and super-resolution microscopy techniques in the second chapter.

1. Transcription by RNA polymerase II

1.1 Core promoter architecture

The core promoter, also often contemplated as “the gateway to transcription” (Danino et al., 2015), is defined to be the DNA region where GTFs and RNA Pol II are recruited to form the PIC. This assembly of the PIC will eventually lead to successful transcription initiation and afterwards elongation to produce messenger RNA (mRNA) with the help of RNA Pol II. For long time, it was assumed that the core promoter is a generic DNA sequence that follows the same universal mechanisms but nowadays it is known that structure and function of the core promoter is much more divers (Juven-Gershon and Kadonaga, 2009).

1.1.1 Core promoter types

Two types of transcription initiation patterns were characterized: focused and dispersed (Smale and Kadonaga, 2003). The difference between these two types is that focused transcription initiation starts from a single nucleotide whereas dispersed initiation includes multiple weak start sites over a region of 50 to 100 nucleotides. In simpler organisms, focused transcription initiation occurs to be the predominant form of transcription. However, in vertebrates only around 30% of all genes follow the focused transcription initiation. In addition, it appears that regulated genes inherit focused promoters, whereas constitutive genes use the dispersed transcription mode. This regulation is consistent with the idea that it is easier to regulate a gene with a single focused promoter than one with multiple start sites. Specific core promoter elements can be found in both types of promoter. Focused promoters have several different core promoter elements like the TATA box, Initiator or MTE and DPE sequences (see also section 1.1.2) whereas dispersed promoters are mainly located in CpG rich islands and contain Sp1 and NF-Y sites (Bajic et al., 2006; Kadonaga, 2012).

But this classification is already challenged by recent studies which showed that some human promoters contain both TATA boxes and CpG islands (Carninci et al., 2006). These mixed mode promoters combine the abilities of both types by having

multiple dispersed but also one particular strong start site (Stamatoyannopoulos, 2010; Kadonaga, 2012). That is why a new sub-classification was presented which divides the promoter in three major types termed Type I, Type II and Type III (Lenhard et al., 2012). Type I promoters were found to be responsible for tissue specific transcription in adult tissues and contain TATA boxes and focused transcription start sites (TSSs) but lack CpG islands. In contrast, Type II promoters contain CpG islands and dispersed promoters but lack TATA boxes and are associated with broad expression of constitutive genes. Finally, Type III promoters are responsible mainly for developmentally regulated genes and contain large CpG islands. Another level of complexity concerning transcription initiation is the phenomenon of bidirectional promoters (Ame et al., 2001). Bidirectional transcription is mainly defined by the head-to-head transcription in both sense and anti-sense orientation within a region of less than 1 kb (Adachi and Lieber, 2002). It was shown that around 10-22% of the genes in mammals perform bidirectional transcription (Orekhova and Rubtsov, 2013). Thus, bidirectional promoters might have evolved to facilitate regulation of two different genes at the same time and maybe they consist of two separate core promoters that are dependent on each other.

1.1.2 Core promoter elements

There are a high diversity of sequence motifs or core promoter elements that exist in core promoters and it is possible that many more elements remain to be discovered. Even if 70% of all vertebrate promoters are dispersed promoters, most of the studies have been carried out on focused promoters. That is why I will focus for this introduction mainly on the core promoter elements present in focused Type I promoter (see Figure 8).

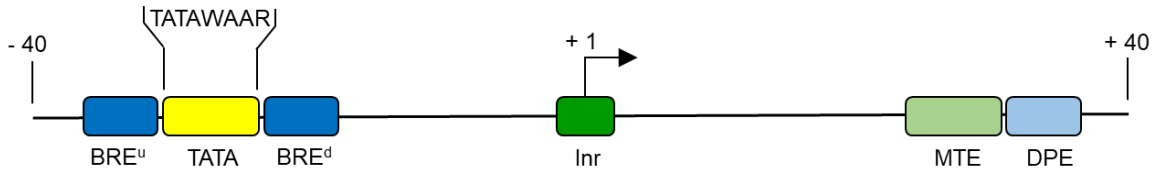


Figure 8: Scheme of core promoter elements. *It displays the location of the BRE^{U} and BRE^{d} relative to the TATA box, MTE, DPE and the Inr. Adapted from (Juven-Gershon and Kadonaga, 2009).*

One of the most prevalent core promoter elements of focused promoters is the initiator (Inr) (Ohler and Niemann, 2001; FitzGerald et al., 2006; Gershenzon et al., 2006). Early studies already assumed the existence of a specific promoter sequence around the TSS and it was later defined as the Inr (Corden et al., 1980; Smale and Baltimore, 1989). The human Inr consensus is $YYA_{+1}NWYY$ (IUPAC nomenclature) where the A nucleotide indicates the +1 TSS whether or not transcription starts at this site (Javahery et al., 1994). Nevertheless, this definition is useful as other core promoter motifs (like MTE or DPE) have a strict spacing dependence from the Inr sequence (Burke and Kadonaga, 1997). Even if many factors were found to interact with the Inr sequence, the binding of TFIID correlates best with the activity of the Inr.

The first core promoter motif that was discovered was the TATA box. The consensus of the TATA box in metazoan cells is TATAWAAR, where the upstream T is positioned at -31 or -30 to the location of the A+1 position of the Inr sequence (Carninci et al., 2006; Ponjavic et al., 2006). The TATA box is bound by the TATA Binding Protein (TBP), a subunit of TFIID, and the sequence is conserved from archaebacteria to human. However, even if the TATA box is a well-known core promoter motif, it is present in only 8-30% of all metazoan core promoters, whereas the other part is known as TATA-less promoters (Kim et al., 2005; Carninci et al., 2006). The TFIIB recognition elements (BRE) are bound by the basal transcription factor TFIIB and are positioned upstream (BRE^u) and downstream (BRE^d) of the TATA box with the consensus of SSRCGCC and RTDKKKK respectively (Lagrange et al., 1998; Deng and Roberts, 2005). Both elements are conserved from archaea to humans, operate in conjunction with the TATA box and can have an influence on transcription by either increasing or decreasing basal transcription level depending on the cellular context (Lagrange et al., 1998; Evans et al., 2001; Deng and Roberts, 2005).

In addition to the aforementioned upstream elements, there are also some promoter motifs downstream of the TSS. The downstream core promoter element (DPE) is located downstream of the Inr sequence, +28 to +33 relative to the A+1 and was originally identified as a TFIID recognition site (Burke and Kadonaga, 1996). Even if there are core promoters that can contain a TATA box, Inr, and DPE motifs, most DPE-dependent promoters only contain DPE and Inr sequences. TFIID binds cooperatively to the Inr and DPE which means that the spacing between these elements is crucial for the transcriptional activity of DPE dependent promoters (Kutach

and Kadonaga, 2000). The motif ten element (MTE) corresponded initially to an overrepresented sequence (called motif 10) which was then identified as a functional core promoter element in *Drosophila* core promoter region (Ohler et al., 2002; Lim et al., 2004). It is located upstream of the DPE at +18 to +27 relative to the A+1 of the Inr and happens to be conserved from *Drosophila* to human. Even if there is synergy between the MTE and the DPE or TATA box, it mainly acts independently to those sequence motifs but cooperatively to the Inr motif. As for DPE motifs, the MTE serves as a recognition site for TFIID, bound by its subunits TAF6 and TAF9, and is also enriched in TATA-less promoter (Burke and Kadonaga, 1997; Theisen et al., 2010).

Taken together, all the identified core promoter elements are highly important for transcription regulation, but the composition of the core promoter can vary as there are no universal core promoter elements. Thus, many studies try to examine different compositions of core promoter elements to characterize their effects on transcriptional output (Juven-Gershon et al., 2006).

1.2 Enhancer sequences and gene specific transcription factors

However, regulation of gene expression doesn't start only with regulatory events at promoter regions, but from inputs involving upstream activating or repression sequences. These distal cis-regulatory elements or enhancer sequences can be located long distances away ,as long as one megabase or beyond, from the TSS and can affect the transcription cycle at various steps with the help of specific transcription factors, co-activators or repressors (Jin et al., 2013; Hu and Tee, 2017). Enhancers promote gene transcription by establishing enhancer-promoter interactions through DNA-looping and therefore achieving contacts between co-activators and chromatin remodeler complexes with the core promoter. The first discovery of enhancers was within the SV40 genome which showed to have several elements to increase the expression of the rabbit β-globin gene in an orientation-, position- and distance-independent manner (Banerji et al., 1981; Benoist and Chambon, 1981). They can exist in three different states, where every state shows distinct histone modifications. The active state shows typically modification of histone H3 by methylation of lysine 3 (H3K3me1) and acetylation of lysine 27 (H3K27ac), whereas silent enhancers are enriched of repressive marks, namely histone H3 lysine 27 trimethylation (H3K27me3)

(Ernst and Kellis, 2010; Tee and Reinberg, 2014). The third class of enhancer states are termed as “poised” enhancers and harbor both active (H3K27me1) and repressive (H3K27me3) modifications. These enhancers are associated with developmental genes which are lowly expressed in embryonic stem cells (ESCs) but expression increases as soon as differentiation signals are present (Bernstein et al., 2006; Barski et al., 2007; Rada-Iglesias et al., 2012).

Enhancer sequences are generally bound by specific transcription factors (TFs) which function as activators for transcription. The structure of TFs is of modular nature with an effector domain which determines if the TF functions as an activator or repressor, a nuclear localization signal (NLS), a regulatory domain and a DNA-binding domain (DBD). There are different features that contribute to the binding of TFs to the DNA using its DBD. The nucleotide sequence is one of these features as it was shown that many TFs bind to specific DNA sequences. This mechanism is called “base readout” (Rohs et al., 2010). Another form of DNA recognition is the binding to specific structural features like DNA-bending or unwinding which is known as “shape readout” (Stella et al., 2010). However, it is important to mention that these two mechanisms of DNA recognition are not mutually exclusive and that it is more likely that an interplay of both readout mechanisms will lead to the binding of their cognate binding sites even if this can vary between different families of TFs (Kitayner et al., 2010; Zhang et al., 2014). The chromatin landscape plays also an important role for TF recognition as many of them show binding preferences to specific histone modifications even if it is still unclear whether a specific chromatin state contributes to TF binding or vice versa (Ernst and Kellis, 2013) (see also section 1.3.2.1). According to a recent nomenclature, TFs were divided into three different categories: pioneers, settlers and migrants (Slattery et al., 2014). “Pioneer TFs” have the ability to bind even inaccessible DNA regions to promote accessibility for other TFs and co-factors and are often a starting point for transcription initiation (Magnani et al., 2011; Zaret and Carroll, 2011). In contrast, “settler TFs” can only bind to their specific DNA-binding motif in accessible regions and cannot bind in inaccessible regions. The last category are “migrant TFs” which as settler TFs bind only to specific motifs in accessible regions but there only to a small subset of their binding sites and in a much more selective fashion (see Figure 9). Therefore, they need interactions with other co-factors to bind efficiently to their target site (Sherwood et al., 2014). Settler and migrant TFs show that DNA accessibility

is an important factor for selective TFs binding, with pioneer factors being a significant exception.

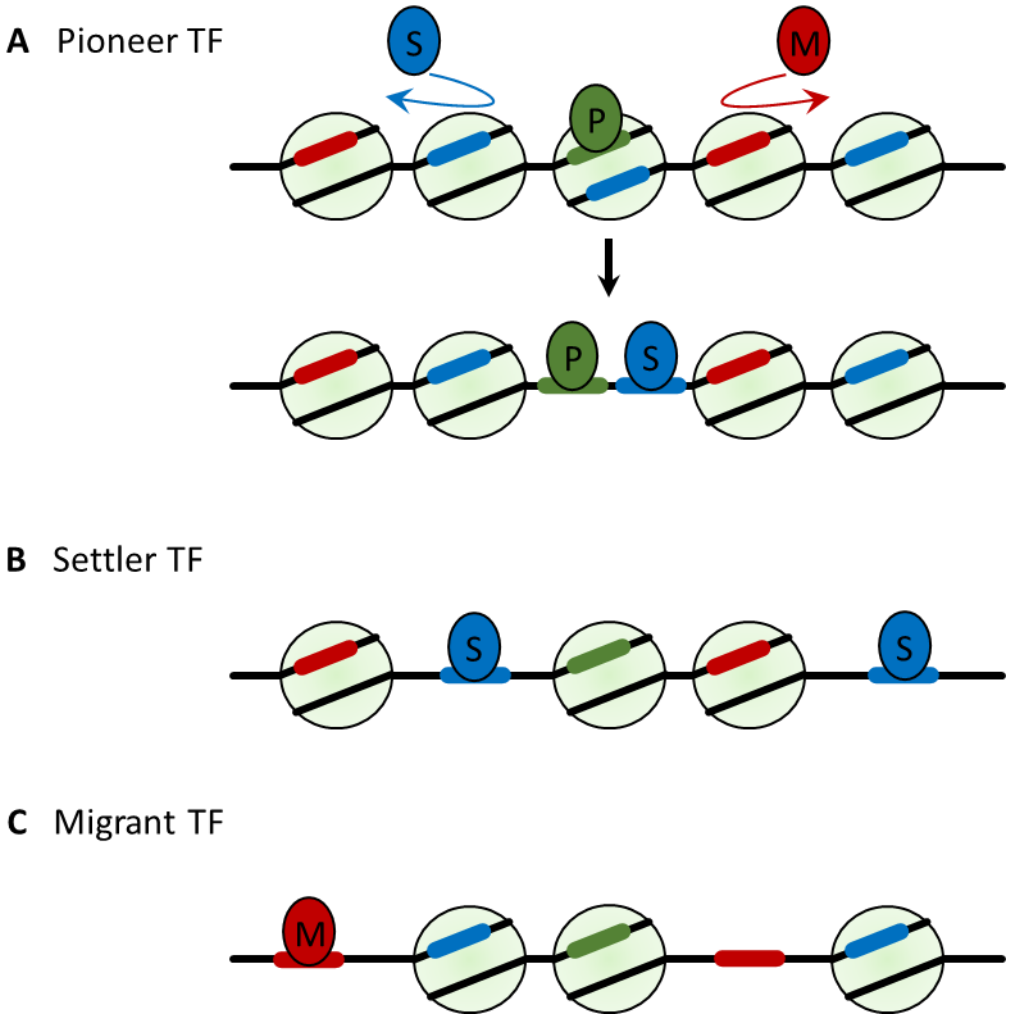


Figure 9: DNA-binding strategies of TFs. A *Pioneer TFs (P; green) can bind to nucleosome associated DNA sites to create an open chromatin environment that is needed for the binding of other nonpioneer factors.* ***B*** *Settler TFs (S; blue) can bind to all of their specific DNA binding sites.* ***C*** *Migrant TFs (M; red) bind only to a subset of their target sites on the DNA. Adapted from (Slattery et al., 2014).*

Often, enhancer need to assemble multiple TF inputs to precisely promote gene expression and there are two models which could explain how this is accomplished: the enhanceosome model and the billboard model. The enhanceosome model depends on the cooperative assembly of different TFs at the enhancer to be able to

recruit co-factors. This assembly is highly dependent on specific protein-protein interactions and a very precise pattern of TF-DNA binding sites. However, these precise DNA elements are not quite common and so the enhanceosome model may be only used to amplify the signal at enhancers or to prevent unspecific TF synergy at enhancers (Thanos and Maniatis, 1995). In contrast, the billboard model hypothesizes that even if individual TF binding sites are important for the activity of the enhancer, a cooperative assembly of the TFs at the enhancer is not important. In this case, the TFs would work very flexible in a combinatorial fashion to promote gene expression (Kulkarni and Arnosti, 2003).

1.3 Transcriptional co-activators

As the name already implies, co-activator are principally recruited by enhancer bound activators (activating TFs). Generally, co-activators are large multisubunit protein complexes that promote transcription through direct contact with general transcription factors (GTFs) (see section 1.4.3) or by their chromatin modifying activities. Many co-activators even harbor several activities within one complex and can target chromatin through posttranslational modification of histones or by remodeling of the nucleosome landscape. Therefore, these actions will lead to the exposure of the core promoter and in the initiation of transcription. In general, co-activators can be defined by their function: chromatin modifiers, chromatin remodelers and adaptors (like Mediator complex). Interestingly, most co-activators are known to regulate the expression of specific genes as they are recruited by specific activators. Nevertheless, many co-activator complexes, like the SAGA complex, harbor several chromatin interaction domains which can indicate that they can be recruited through different mechanisms and therefore their effect on transcription is potentially broader than expected. The most conserved and studied co-activators which will be discussed in the following sections are Mediator, the SAGA complex and TFIID, even if the last one mentioned is mainly known as a GTF and therefore doesn't follow the classical function of a co-activator. Thus, the question for this chapter is what mechanisms and co-activators are existing to modify the chromatin environment and to enable the initiation of transcription?

1.3.1 The Mediator complex

Mediator is a co-activator complex which was first identified in Saccharomyces cerevisiae and was found to show the ability to support the activation of RNA Pol II transcription by interacting with TFs and the PIC (Kim et al., 1994; Myers and Kornberg, 2000). The yeast Mediator complex has 25 subunits which can be divided into different modules named “head”, “middle” and “tail” modules as well as a “kinase” module that is only present in subset of Mediator complexes in the cell (Conaway and Conaway, 2013). It is important to mention that the Mediator complex is highly conserved between yeast and human (Conaway et al., 2005). The complex acts as an adapter protein between DNA binding TFs and the GTFs at the promoter which is important for transcription initiation as it functions as a bridge between the transcriptional regulators at the enhancers and RNA Pol II at the promoter (see Figure 10). Even though, it was always speculated that Mediator interacts directly with RNA Pol II, it was shown only recently that Mediator not only interact with the RNA Pol II subunit RPB3 but that this interaction is important for RNA Pol II recruitment *in vivo* (Soutourina et al., 2011). Besides RNA Pol II, it was also shown that Mediator can interact with several GTFs at the promoter including TFIID, TFIIA, TFIIB, TFIIE, TFIIF and TFIIH to recruit these factors to the PIC (Johnson et al., 2002; Johnson and Carey, 2003; Baek et al., 2006; Jishage et al., 2012; Eychenne et al., 2016). In fact, it is well established that Mediator is important for global RNA Pol II transcription by interacting with several proteins at the promoter using different interaction domains (Plaschka et al., 2015). However, recent studies also suggest that Mediator plays an even more important role than only in transcription activation. There is evidence that it also influences transcription elongation by helping to overcome the influence of elongation inhibiting factors (Malik et al., 2007; Jishage et al., 2012), by working as a platform to recruit positive elongation factors and pre-mRNA processing factors (Donner et al., 2010; Mukundan and Ansari, 2011) and by controlling the phosphorylation of the C-terminal domain (CTD) of RNA Pol II (Jiang et al., 1998; Boeing et al., 2010; Takahashi et al., 2011). Altogether, this shows how important the Mediator co-activator complex is for RNA Pol II transcription.

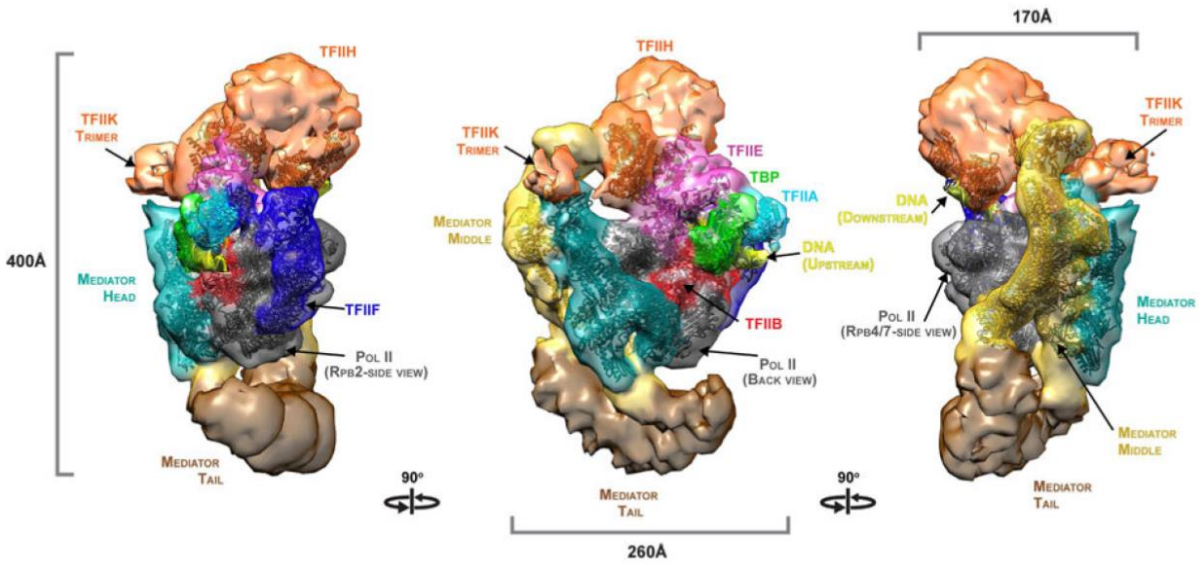


Figure 10: Structure of the Mediator-PIC complex. *Surface representations of the cryo-EM maps in which the interactions of the mediator with the PIC are shown. Three views are displayed with successive 90˚ rotation on the vertical axis. From (Robinson et al., 2016).*

1.3.2 Histone modifications and chromatin modifying enzymes

1.3.2.1 Transcription related modifications and enzymes

As already mentioned before, the chromatin landscape plays an important role in the accessibility of the DNA for TFs and other factors like chromatin remodelers. However, how is this accessibility achieved to ensure that the chromatin is “open” enough to enable the binding of proteins like TFs? One possibility to modulate chromatin accessibility is by covalent posttranslational modifications of histones. There are many different histone modifications with diverse effects on transcriptional activity like acetylation (ALLFREY et al., 1964), methylation (MURRAY, 1964), phosphorylation (Kleinsmith et al., 1966) and ubiquitination (Goldknopf and Busch, 1977). It is important to mention that these modification mainly happen at the N-terminal tail of the histones although modifications in their globular domains are also reported (Kouzarides, 2007). These modifications are performed by specific chromatin modifying enzymes which can either add (called “writers”) or erase (called “erasers”) histone marks. Therefore, proteins that can recognize these histone modifications and trigger a response are called “reader”.

One of the best characterized modification is histone acetylation which is generally associated with transcription activation. Acetylation of lysine residues in the histone H3 and H4 tails can boost transcriptional activation in two ways. First, it was proposed that lysine acetylation neutralizes the positive charge of the histone which diminishes histone-DNA interaction and therefore increases the accessibility of the DNA. Second, histone acetylation can serve as a platform to recruit “readers” to the modified chromatin which can further promote transcription (Lee and Workman, 2007). Acetylation of histones is performed by histone acetyltransferases (HATs) which, together with acetyl-coenzyme A (AcCoA) as a co-factor, attach an acetyl group to the ε-amino group of lysine residues. HATs can be divided into four major groups depending on the differences in their catalytic domains: Gcn5/PCAF, MYST, p300/CBP and Rtt109 HATs. The Gcn5/PCAF and MYST HATs are the most highly studied and have homologs from yeast to human, whereas the p300/CBP family of HATS is metazoan specific and the Rtt109 HAT is fungal specific. On the other side of HATs are the histone deacetylases (HDACs) which are needed to erase the acetylation modification. They can be divided into 5 groups (class I, IIa, IIb, III and IV) according to their phylogenetic distance (Gregoretti et al., 2004).

In contrast to acetylation which is involved in transcription activation, the role of histone methylation is not so clear. Both the lysine and arginine residues of histones are the target for methylation modification. However, it appears that lysine methylation is present in active as well as inactive chromatin. For instance, heterochromatic regions are enriched in H3K9 di- and trimethylation and silenced loci harboring developmental genes show H3K27 dimethylation. On the other hand, genes that are actively transcribed show lysine methylation at H3K4, H3K36 and H3K79 (Sims and Reinberg, 2006). Thus, H3K4me2/3 happens at the 5’ end of transcribed genes whereas H3K36me2/3 and H3K79me2/3 are concentrated in gene bodies (Petty and Pillus, 2013; Zentner and Henikoff, 2013). The enzymes responsible for lysine methylation are histone lysine methyltransferases (HKMTs). They consist of eight classes (HKMT1-8) for which every class of HKMTs has their specific lysine residue target on histone 3 (H3) and/or histone 4 (H4) (Allis et al., 2007). Furthermore, there are also four classes of histone arginine methyltransferases (HRMTs I-IV) which are also grouped by the modification they are able to perform. On the opposite side, there are also two classes of histone lysine demethylases that can remove lysine methylations: lysine

demethylase-like family and Jumonji C-terminal domain family. Until now, histone arginine demethylation remains only a speculation as no biochemical pathway could be found to perform arginine demethylation (Bannister et al., 2002).

1.3.2.2 Phosphorylation of histone variant H2AX

Besides histone modifications involved in transcription regulation, there are also other very important modification of histones which can be recognized by other processes in the nucleus. One of these modification, which is not directly concerning transcription, is the phosphorylation of the histone variant H2AX (γH2AX). This modification is generally known as a marker for DNA damage and in particular for DNA double-strand breaks (DSBs) and a highly important starting point for DNA damage repair. After DNA damage, is the phosphorylation of H2AX spreading over a region of up to 2 Mbp from the initial DSB which acts as an amplification step to sense the break. Thus, the phosphorylation of H2AX is one of the first steps to activate the signaling cascade of the DNA damage response (DDR) by attracting chromatin remodelers and other proteins involved in the DNA damage repair (Fernandez-Capetillo et al., 2002; Giglia-Mari et al., 2011). The levels of H2AX in the chromatin varies between 2-25% of the whole H2A pool depending on the cell type (Kinner et al., 2008). In contrast to most other histone modifications, occurs the phosphorylation near the C-terminal end of the protein at the SQ motif of the histone on serine 139 (S-139). The three kinases known to be responsible for the phosphorylation of H2AX are ATM, ATR and DNA-PKcs which all belong to the phosphatidylinositol-3-OH-kinase-like family of protein kinases (PIKKs). Thus, which kinase is carrying out the phosphorylation can depend on the type and the state of the DNA damage (Andegeko et al., 2001; Burma et al., 2001). Early phosphorylation directly at the DSB could be performed by DNA-PKcs as it interacts and gets activated by interacting with a factor which binds directly the broken DNA ends of the DSB (Stiff et al., 2004; Wang et al., 2005). However, this would result in a reduced phosphorylation range which is required for the spreading of the γH2AX signal. Thus, it was indicated that ATM would be the most suitable kinase to induce the spreading of the phosphorylation as it is activated and recruited through local chromatin modifications associated with DNA damage and therefore is able to modify several H2AX molecules within this chromatin domain (Bakkenist and Kastan, 2003).

On the other hand, if the DSB is not induced randomly but by UV damage or replicative stress, it is detected by the kinase ATR which therefore phosphorylates H2AX (Ward and Chen, 2001; Limoli et al., 2002; Hanasoge and Ljungman, 2007). However, the activity of ATR is dependent on single stranded DNA and therefore does not always represent a DSB. The presence of replicative stress in form of stalled replication forks can result in single stranded DNA which can give rise to the formation of γH2AX foci (Paulsen and Cimprich, 2007). Thus, the presence of γH2AX foci especially in S-phase is not always a sign for a DSB. Nevertheless, H2AX modification needs to be reverted after the repair of the DNA damage which could be managed in two different ways: Either by replacing γH2AX with new H2AX in the nucleosome or by de-phosphorylation of the present histone variant. Several phosphatases have been suggested to be involved in the process including the phosphatases 2A (PP2A) and 2Cγ (PP2Cγ) (Chowdhury et al., 2005; Kimura et al., 2006). However, for γH2AX generated during replication by ATR, it was shown that the de-phosphorylation is performed by PP4-phosphatase complex containing PP4C, PP4R2 and PP4R3b (Chowdhury et al., 2008). On the other hand, the exchange of H2AX with H2A was shown to be realized by the Facilitates chromatin transcription (FACT) complex (Heo et al., 2008).

1.3.3 The SAGA co-activator complex

As mentioned earlier, the Spt-Ada-Gcn5 acetyltransferase (SAGA) complex is a perfect example of a huge multisubunit co-activator complex containing 18-20 subunits with a size of 2 MDa, which harbors many different activities. SAGA was studied to a great extent in *Saccharomyces cerevisiae,* but it is important to mention that the complex is highly conserved from yeast to human concerning components, modules and function. The SAGA complex exhibits different modules with two of them having distinct enzymatic activities that can perform acetylation and deubiquitination of histones and other non-histone substrates (Baker and Grant, 2007; Nagy and Tora, 2007; Rodríguez-Navarro, 2009). However, as for many multisubunit complexes, not all functions of certain subunits are known until now. Additionally, it is still unclear how the different functions of the modules are coordinated and if they act in a more antagonistic or cooperative manner to regulate cellular processes.

The modular structure of human SAGA (hSAGA) includes an activator recruitment module called TRRAP, a TBP interaction unit composed of SUPT3H, the

deubiquitination module (Dub module) including USP22, ENY2, ATXN7L3 and ATXN7, an architecture unit with TADA1, SUPT7L, SUPT20H as well as several TBP-associated factors (TAF) and last but not least the acetylation module including GCN5 (general control nonderepressable 5), TADA2B and TADA3 (Koutelou et al., 2010) (see also Table 1). The coactivator subunit GCN5 has a high homology to the acetyltransferase family member PCAF (p300/CBP-associated factor) (Martinez et al., 2001). The first histone acetyltransferase that was discovered was the p55 protein from *Tetrahymena thermophile* which turned out to be an orthologue of the yeast coactivator protein Gcn5 (Brownell et al., 1996). The complex is recruited to the genetic loci by the interaction of specific TFs with TRRAP. The acetylation module can afterwards catalyzes the acetylation of histone H3 which will loosen the histone-DNA interaction to facilitate the binding of other TFs and the establishment of the PIC (Balasubramanian et al., 2002). The TBP interaction unit can also assist in TBP recruitment, PIC formation and transcriptional activation (Mohibullah and Hahn, 2008). Furthermore, it is possible that SAGA is not only promoting gene activation but is also involved in transcription elongation by accompanying RNA Pol II and acetylating as well as removing nucleosomes during gene expression (Govind et al., 2007). Additionally, it was suggested that the DUB module also enables elongation by deubiquitination of histone H2B which allows for the phosphorylation of the C-terminal domain of RNA Pol II by recruiting the Ctk1 kinase (Wyce et al., 2007).

***Table 1**: Composition of the SAGA complex*

Module	Subunits
HAT	GCN5/PCAF TADA2b TADA3 SGF29
DUB	USP22 ATXN7/-L1/-L2 ATXN7L3 ENY2
Link with activators	TRRAP
TBP regulation	SUPT3H
Structural core	TAF5L TAF6L TAF9/TAF9b TAF10 TAF12 TADA1 SUPT7L SUPT20H

Subunit composition of the SAGA complex in *H. sapiens*. The subunits are grouped according to the module/function in which they are involved.

Considering that SAGA would need specific TFs to be recruited to chromatin in the first step, it was often questioned if this co-activator would have an effect on global transcription or if it would act only at a few specific genes. Especially as overall results from yeast indicated that only 10% of all genes would be regulated by SAGA (Huisinga and Pugh, 2004). However, recent studies in yeast and human cells showed that the HAT activity of SAGA is capable of acetylate histone H3 genome wide on all actively transcribed genes. Additionally, also the DUB activity was shown to be active on the transcribed regions of expressed genes (Bonnet et al., 2014). Furthermore, analysis of the level of newly synthesized mRNA in yeast SAGA mutants showed that SAGA is indeed needed for global mRNA synthesis at all genes transcribed by RNA Pol II (Baptista et al., 2017). These results indicate that SAGA is an important co-activator

for global RNA Pol II transcription and that it functions at multiple steps during the transcription process.

1.3.4 Chromatin remodeler complexes

There are different mechanisms to promote DNA accessibility for TFs and other factors involved in transcription. Beside the already mentioned mechanism to posttranslational modify histones to modulate chromatin folding and the incorporation of non-allelic histone variants to alter nucleosome stability, there is also the possibility to re-position, evict or alter the composition of nucleosomes by using chromatin remodeler complexes (Swygert and Peterson, 2014). Chromatin remodeling factors are multisubunit complexes which use the energy of ATP hydrolysis to facilitate the removal or deposition of histones and therefore to create other TFs access to the DNA (Clapier and Cairns, 2009). The first identified chromatin remodeler complex was the budding yeast Swi/Snf complex (Peterson and Herskowitz, 1992). Since then, four families of remodeler complexes could be identified in eukaryotes: SWI/SNF, INO80, ISWI and the CHD family. Each family has their own characteristic ATPase subunit, all of them related to the DEAD/H superfamily of DNA helicases. Besides their ATPase subunit, they all have individual accessory subunits containing interaction domains to facilitate the binding to specific TFs and/or posttranslational modified histones (Hargreaves and Crabtree, 2011).

The SWI/SNF family of remodeling enzymes contain a binding motif called bromodomain to facilitate interactions with acetylated lysine residues on histones (Dürr and Hopfner, 2006). However, how these complexes catalyze the movement of the DNA around histones by ATP hydrolysis after binding to chromatin is still not completely clear. One model for SWI/SNF remodelers is the “loop recapture” model (Figure 11). This model states that the hydrolysis of ATP generates a loop of DNA which creates new histone contacts with neighboring linker DNA (Strohner et al., 2005). Interesting to mention is that the SWI/SNF complex is proposed to have a general impact on transcription by working together with specific transactivators and the histone acetyltransferase GCN5 (Biggar and Crabtree, 1999). The remodeler complexes from the CHD and ISWI families do not contain bromodomains but in contrast chromodomains or PHD fingers to be able to bind specifically to methylated histones (Marfella and Imbalzano, 2007). In contrast to SWI/SNF, the smaller ISWI

remodeler complexes need to bind as a dimeric motor to DNA to enable the bidirectional translocation of DNA over the nucleosome. This makes sense as the role of these remodeler complexes lies in nucleosome spacing (Gangaraju and Bartholomew, 2007; Blosser et al., 2009). There are also chromatin remodeler complexes that harbor chromatin modifying subunits to modify histones themselves like the INO80 familiy of remodelers. They comprise of a subunit to perform deubiquitination of histones H2A and H2B which could play a key role in transcriptional activation and DNA repair (Yao et al., 2008).

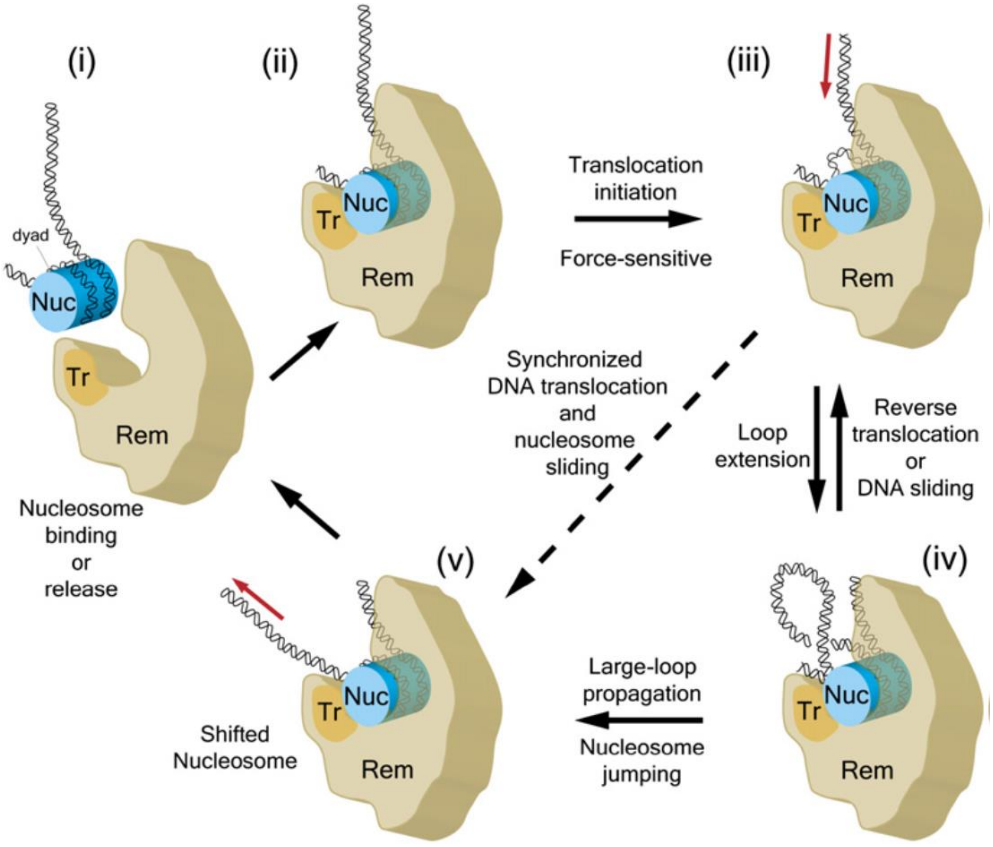


Figure 11: Model for nucleosome remodeling by generating gene loops. *(i) Unbound state of chromatin remodeler (Rem) and nucleosome (Nuc). (ii) Binding of the remodeler to the nucleosome trough a pocket. (iii) The ATPase/translocase subunit (Tr) engages the nucleosomal DNA and forms a small bulge near the dyad. (iv) Processive translocation generates intranucleosomal DNA loops which can either result in active reverse translocation, DNA sliding or nucleosome jumping. Translocation can also lead to immediate nucleosome sliding (dashed line). (v) Release of the nucleosome from the chromatin remodeler. From (Zhang et al., 2006).*

1.4 RNA Pol II transcription initiation

The chromatin landscape was modified and the promoter is accessible but how is transcription initiated with RNA Pol II? In metazoan cells none of the multisubunit RNA polymerases can initiate transcription at promoters on their own. Also RNA Pol II requires cis- as well as trans-regulatory elements to recruit it to the promoter. Besides the already mentioned elements at the enhancer including activators and co-activators, there exist also general transcription factors (GTFs) which help to load RNA Pol II to the promoter and to initiate transcription. These GTFs comprise of the factors TFIIA, B, D, E, F and H which all have their specific function in the initiation of transcription. In the following chapter, all factors involved in the assembly of the PIC will be discussed including RNA Pol II itself.

1.4.1 Assembly of the pre-initiation complex (PIC)

After the cooperative work of activators and co-activators at the promoter, RNA Pol II and the GTFs can assemble to form the PIC. Initially, the GTFs were characterized *in vitro* to be essential to initiate transcription with RNA Pol II. Since then, many studies investigated the role of every GTF and their interplay between each other to find out how the PIC is assembled. From all this work, two pathways were hypothesized: the sequential assembly pathway and the RNA Pol II holoenzyme pathway.

The most commonly known pathway is the sequential pathway in which the PIC assembles in a stepwise manner. In short, the formation of the PIC can be divided into the following steps (Shandilya and Roberts, 2012; Grünberg and Hahn, 2013; Sainsbury et al., 2015) (see Figure 12):

(i) Specific binding of TFIID including TBP to the TATA-box inducing DNA bending.
(ii) Binding of TFIIA and TFIIB to the TBP-DNA complex to stabilize the interaction.
(iii) Formation of the core PIC by recruiting the RNA Pol II-TFIIF complex to the existing upstream promoter complex.
(iv) Subsequent binding of TFIIE and TFIIH to complete the PIC (closed conformation).

(v) ATP-dependent melting of the DNA to form a “transcription bubble” (open conformation).

(vi) Initiation of RNA synthesis with a length of at least 6-10 nucleotides to avoid abortive transcription.

(vii) Further RNA synthesis to a length of about 25 nucleotides and afterwards dissociation of the initiation complex and formation of the RNA Pol II elongation complex.

On the other side, the RNA Pol II holoenzyme pathway was described after purification experiments revealed that a preassembled RNA Pol II containing holoenzyme can be purified together with several GTFs, chromatin remodelers and chromatin modifying enzymes (Ossipow et al., 1995). In more detail, these studies showed that RNA Pol II can be purified together with several GTFs (including TFIIB, TFIIF, TFIIE and TFIIH), chromatin remodelers (like the SWI/SNF complex) and Mediator subunits, even if the results of the purification vary from study to study. Interestingly, TFIIA and TFIID could not be found associated with the RNA Pol II holoenzyme. This is in good agreement with the fact that these GTFs would be needed to load and stabilize the holoenzyme on the DNA due to their specific promoter binding abilities (Koleske and Young, 1994).

It is important to mention that even if these pathways occur *in vitro*, there are no evidences for a preferred pathway *in vivo*. Both pathways are also not mutually exclusive and could both exist in living cells. Furthermore, this process is very dynamic and therefore there could exist some kind of “middle way” between these two pathways mentioned.

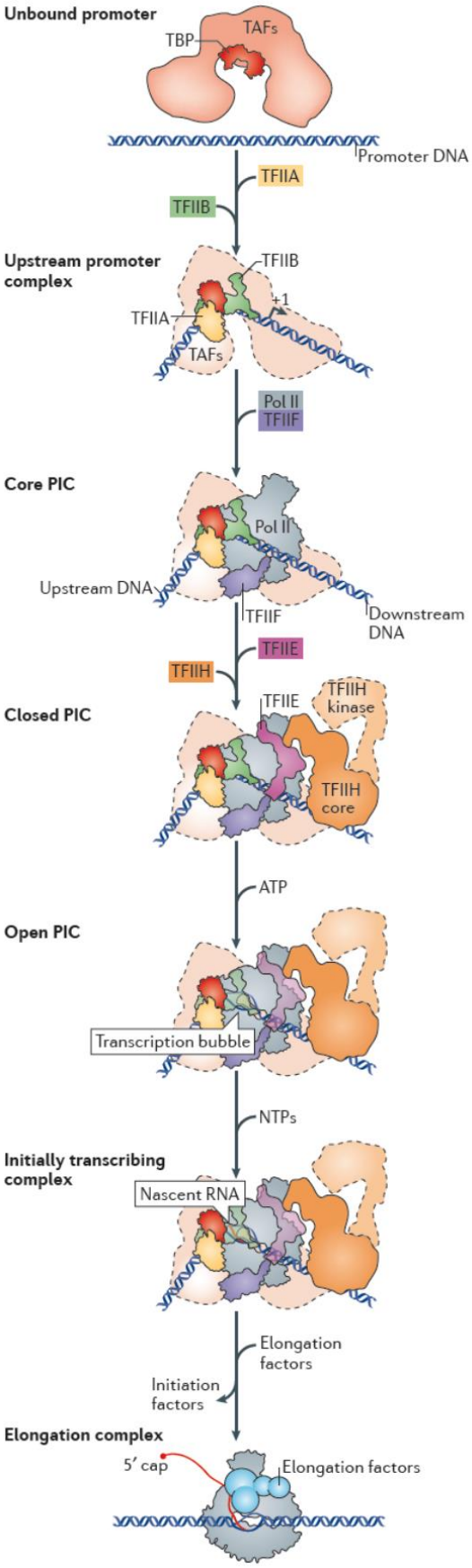


Figure 12: Legend on the next page →

Figure 12: Schematic representation of RNA Pol II transcription initiation following the model of sequential PIC assembly on promoter DNA from GTFs and RNA Pol II. *The first step includes binding of TFIID/TBP to the promoter resulting in bending of the DNA. The TBP-DNA complex is afterwards stabilized by TFIIA and TFIIB. This upstream promoter complex is bound by the RNA Pol II – TFIIF complex which results in the formation of the core PIC. Further binding of TFIIE and TFIIH finalizes the PIC formation. In the presence of ATP, the "transcription bubble" is formed by the opening of the DNA and RNA synthesis starts. Lastly, the dissociation of the initiation factors leads to the formation of the RNA Pol II elongation complex which is bound by several elongation factors. From (Sainsbury et al., 2015).*

1.4.2 Structure and function of RNA polymerase II (RNA Pol II)

RNA Pol II is one of three nuclear RNA polymerases in metazoan cells together with RNA Pol I and RNA Pol III. This multisubunit complex is completely responsible for the transcription of messenger RNA (mRNA) from most protein coding genes in eukaryotic cells. To finally produce mRNA, RNA Pol II interacts with a plethora of different factors like DNA and the general transcription factors through the different steps of transcription including initiation, elongation and termination. Structural studies on RNA Pol II have given new insights about how the complex is build up and also how the interactions with different factors is accomplished.

The RNA Pol II complex consists of 12 subunits (RPB1-12) with a size of > 0.5 MDa. The whole complex is highly conserved from yeast to human in sequence and structure and it is even possible to substitute certain subunits in the yeast complex with their mammalian counterparts. Certain subunits of RNA Pol II like RPB4, RPB7, RPB9 and the unstructured C-terminal domain (CTD) of RPB1 are exclusively present in this complex without any homologous subunits in the other RNA Pols. However, 5 of the subunits (RPB5, RPB6, RPB8, RPB10 and RPB12) are shared between RNA Pol II and the other polymerases RNA Pol I and RNA Pol III (Thomas and Chiang, 2006).

The crystal structure of the "core" RNA Pol II complex containing 10 subunits of the complex (missing RPB4 and RPB7) was resolved by (Cramer et al., 2001). The complex is composed of four different mobile modules named "core", "jaw-lobe", "shelf" and "clamp" all of which can be divided into several sub-modules (see Figure 13 and Table 2). The biggest module is the core module which includes the subunits RPB3, RPB10, RPB11 and RPB12 as well as RPB1 and RPB2 which together form the active

center of the polymerase. The jaw-lobe can be divided into the submodules “upper jaw” (containing RPB1 and RPB9) and the “lobe” which is a part of RPB2. The shelf module includes the “lower jaw” and the “assembly” domain which are build up by domains of RPB5 and RPB6 as well as domains of RPB1 called “foot” and “cleft”. The remaining clamp module is composed of domains of RPB1 and RPB2 and includes the “clamp core” and the “clamp head”. This shows that every subunit of RNA Pol II contains different important domains that play a role in a certain module which results in a highly mobile structure (see Figure 13A and Table 2). In the middle of these structure is RPB1 whose domains are present in all of the four modules. At the center of the complex is a huge cleft that is formed by all four mobile modules and was shown to have two different conformation: The open and the closed conformation. The open cleft is the conformation where the straight DNA strand can enter RNA Pol II from one side and gets in contact with the active center at the base of the cleft. The DNA strand is than passing through RNA Pol II and exits the complex through a gap composing of a region called the “wall” and an open clamp structure. It was also shown that the upper jaw domain and in particular RPB9 gets in contact with the DNA following this path and the subunit is potentially involved in the TSS selection as mutants of RPB9 have shown to be defective in TSS selection (see Figure 13B). It is also important to mention that the 10 subunit complex was shown to be elongation-competent but is not able to perform complete transcription initiation without RPB4 and RPB7 (Cramer et al., 2001).

***Table 2**: Composition of the RNA Pol II structural modules*

Module	Subunits
Core	RPB1 RPB2 RPB3 RPB10 RPB11 RPB12
Jaw-lobe	RPB1 RPB2 RPB9
Shelf	RPB1 RPB5 RPB6
Clamp	RPB1 RPB2
Heterodimer subunit	RPB4 RPB7

Structural module composition of the RNA Pol II complex. The subunits are grouped according to the module in which they are present.

A later study was successful to obtain the “complete” RNA Pol II structure which is initiation-competent and includes the two missing subunits (Bushnell and Kornberg, 2003). It was shown that a submodule of RPB4/RPB7 can associate reversibly with the core complex and binds to a pocket formed by RPB1, RPB2 and RPB6 at the base of the clamp module (see Figure 13C). This binding induces a conformational change which switches the clamp into the closed conformation. This indicates that single stranded DNA can enter RNA Pol II before the RPB4/RPB7 submodule is binding through the open clamp conformation and afterwards the clamp is closed to trap the single stranded DNA inside RNA Pol II. Furthermore, the block in the closed conformation suggests that double stranded DNA is never entering RNA Pol II. Additionally, it was shown that the RPB4/RPB7 sub-module can act as a binding platform for other factors as well as for RNA that is exiting the elongating polymerase. Further structural information about RNA Pol II and its interactions with several other

factors including GTFs, co-activators or the DNA that were gathered over the years are of high importance to understand the function of RNA Pol II, the formation of the PIC or the transition from initiation to elongation (He et al., 2013; Fishburn et al., 2015; Murakami et al., 2015; Louder et al., 2016; Plaschka et al., 2016; Hantsche and Cramer, 2017).

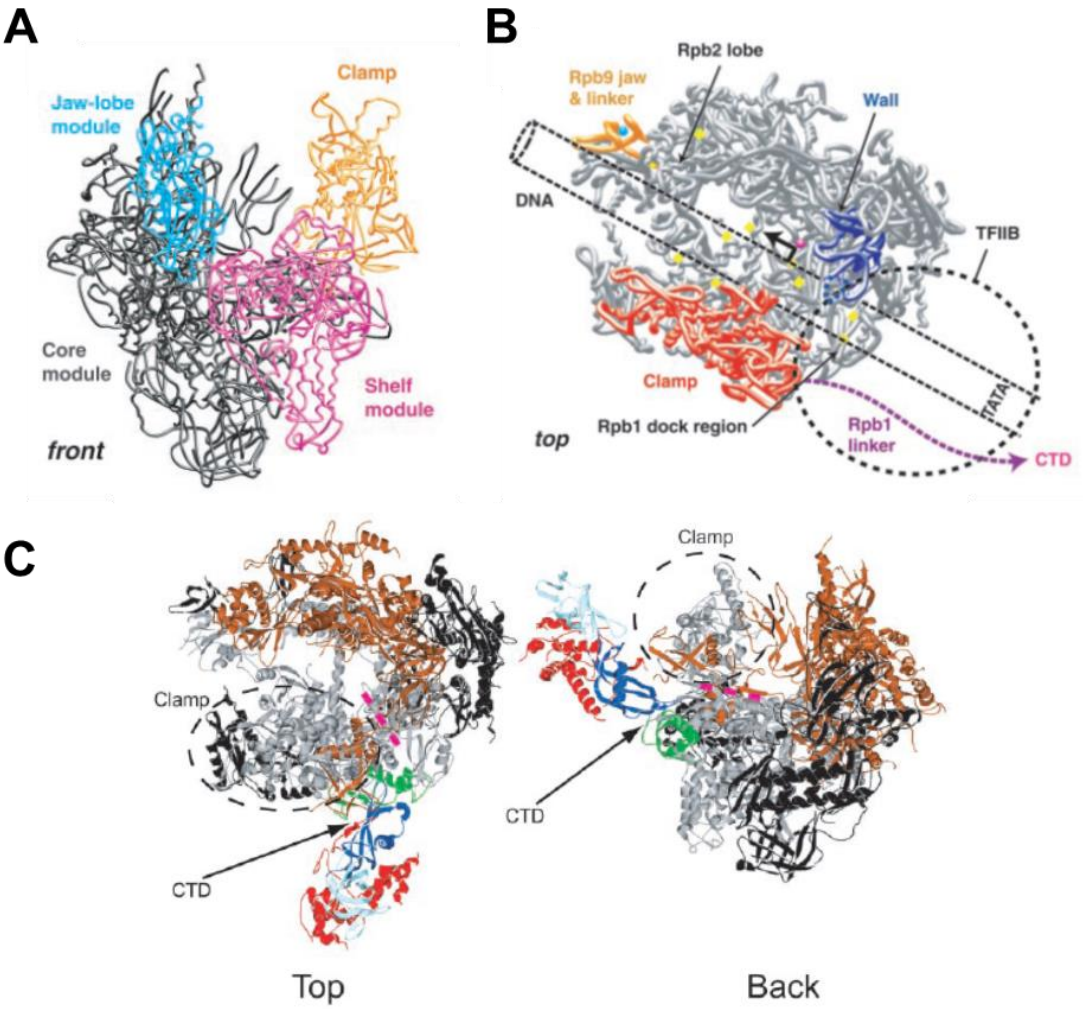


Figure 13: Legend on the next page →

Figure 13: RNA Pol II structure. A *Backbone traces of the RNA Pol II core, jaw-lobe, clamp and shelf modules of the 10 subunit core structure shown in grey, blue, yellow and pink, respectively.* ***B*** *Top view of the 10 subunit core complex with a DNA duplex indicated with a dashed cylinder. The regions of RPB9 involved in start site selection are shown in orange and the location of mutations that can affect start site selection are shown in yellow. The wall and clamp regions are shown in blue and red, respectively. The location of the GTF TFIIB is indicated by a dashed circle.* ***C*** *Backbone model of the complete 12 subunit RNA Pol II complex in top and back views. The subunits are color coded with RPB1 in gray, RPB2 in bronze, RPB4 in red, RPB6 in green, the N-terminal half of RPB7 in dark blue, the C-terminal half of RPB7 in light blue and the remaining subunits in black. The locations of the clamp, the RNA exit groove 1 (pink dashed line) and the CTD are indicated. Adapted from (Cramer et al., 2001; Bushnell and Kornberg, 2003).*

As mentioned before, in the middle of the complex is the largest subunit or RNA Pol II called RPB1 (Figure 14A). Besides its domains that are part of the different modules of the complex, it also harbors an unstructured C-terminal domain (CTD). The sequence of the CTD contains a tandem consensus of the seven amino acids Tyrosine-Serine-Proline-Threonine-Serine-Proline-Serine (Y^1 S^2 P^3 T^4 S^5 P^6 S^7) which is repeated 52 times in vertebrate RPB1. However only 21 out of the 52 repeats follow the consensus correctly whereas the rest show several amino acid substitutions mainly at the positions 2, 4, 5 and/or 7 (Corden et al., 1985). The CTD is mainly conserved at the N-terminal half of the sequence with the tyrosine at position one and the proline at position six being the most conserved residues. The length of the CTD is different between the species, as for example yeast CTD has only a length of 26 repeats. The CTD repeat is the target of a plethora of different posttranslational modifications including phosphorylation, glycosylation, ubiquitination and methylation (Kelly et al., 1993; Sims et al., 2011). In fact, taken into account the differences in the heptad sequence, there exists a huge number of combinations concerning the phosphorylation pattern with the possibility of different resulting conformations for the CTD. Furthermore, all five of the hydroxylated amino acids can be phosphorylated but the phosphorylation of the two serine residues serine 2 (Ser2) and serine 5 (Ser5) were shown to be the most abundant modifications (Schüller et al., 2016; Suh et al., 2016). RNA Pol II including an unphosphorylated CTD is designated as RNAP IIA whereas a RNA Pol II complex with phosphorylated CTD at Ser2 or Ser5 is known as RNAP IIO

(Cramer, 2004; Hirose and Ohkuma, 2007; Sikorski and Buratowski, 2009). Interestingly, the full CTD is not required to ensure cell viability, as studies have shown that around 50% of the natural number of heptads is enough (West and Corden, 1995). However, requirements for the CTD length can vary from species to species.

Several kinases have been identified to be able to phosphorylate the CTD most notably the kinases from the CDK family CDK7 and CDK9. Human CDK7 was initially discovered as TFIIH-associated kinase and is responsible for the phosphorylation of Ser5 during initiation (Feaver et al., 1991; Lu et al., 1992). In contrast, CDK9, or also called P-TEFb in mammalian cells, is known to phosphorylate Ser2 in the CTD of elongating RNA Pol II to overcome pausing of the polymerase near the promoter (Marshall and Price, 1995; Marshall et al., 1996). However, CDK7 and CDK9 (P-TEFb) are also able to phosphorylate Ser7 and Thr4 respectively (Figure 14B) (Akhtar et al., 2009; Hsin et al., 2011). Another kinase of the CDK family is CDK8, a subunit of the Mediator complex, which can phosphorylate both Ser2 and Ser5 and was identified as a negative regulator of TFIIH activity during transcription initiation (Liao et al., 1995; Sun et al., 1998).

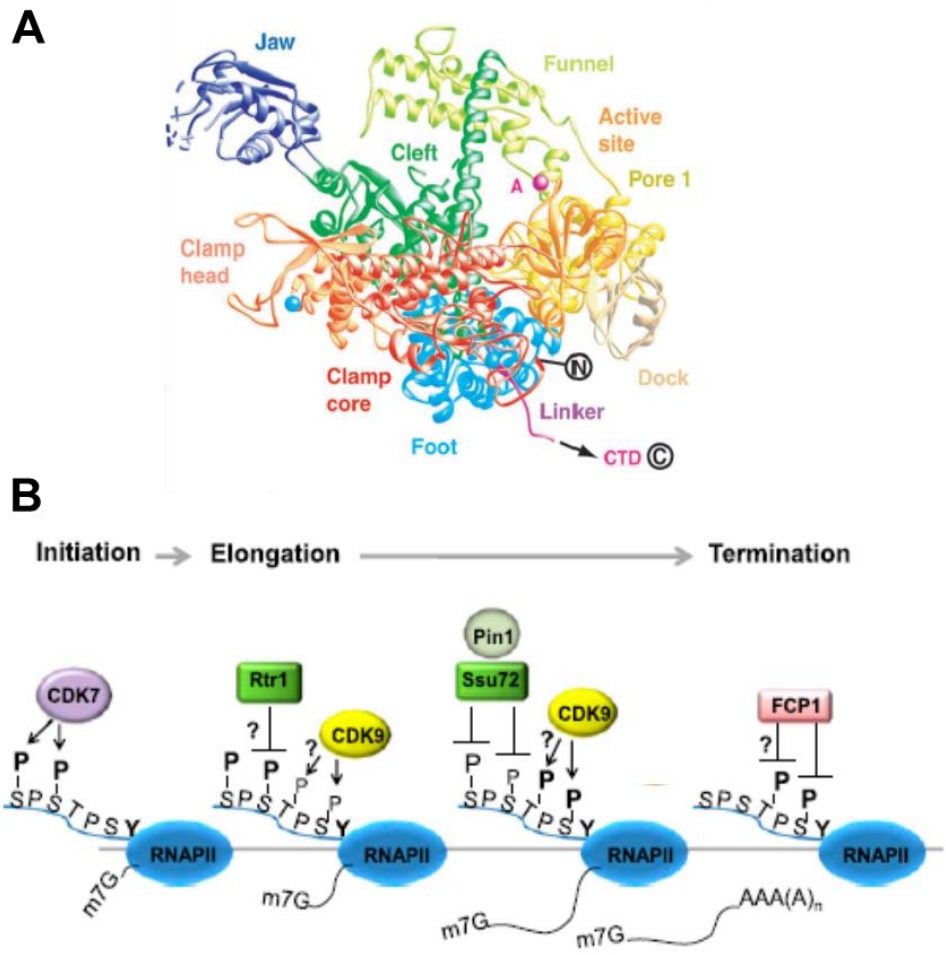


Figure 14: Structure of RPB1 and the dynamic modifications of the CTD during the transcription cycle. A *Ribbon diagram showing the structure of the subunit RPB1. The parts of RPB1 involved in the different RNA Pol II modules are indicated as well as the N-terminus and the CTD.* ***B*** *During transcription initiation, CDK7 is phosphorylating the residues Ser5 and Ser7 of the CTD. For the elongation process these marks are removed gradually by phosphatases (Rtr1 and Ssu72 with Pin1) and CDK9 is phosphorylating Ser2 and probably Thr4. During transcription termination are the marks at Ser2 and Thr4 removed by the phosphatase FCP1 to regenerate RNA Pol II for a new round of transcription. Modified from (Cramer et al., 2001; Hsin and Manley, 2012).*

On the other hand, the removal of a certain type of phosphorylation at the CTD at a given time is highly important to ensure the correct continuation of transcription. This requires the activity of dedicated phosphatases. Thus, two major phosphatases are known to dephosphorylate the CTD during transcription and are highly conserved from yeast to human: Fcp1 and Ssu72 (Figure 14B). The TFIIF-associating CTD

phosphatase 1 (Fcp1) was initially described in HeLa cells to be able to dephosphorylate both Ser2 and Ser5 even if it shows a preference for Ser2 (Chesnut et al., 1992; Cho et al., 2001; Ghosh et al., 2008). This indicates that Fcp1 is responsible for the turnover of RNAP IIO back to RNAP IIA, so that the polymerase can be recycled for the next transcription round. On the other side, Ssu72 was first identified as a suppressor of defective TFIIB but was later characterized as a component of the yeast cleavage and polyadenylation factor (CPF) (Sun et al., 1996; Dichtl et al., 2002; He et al., 2003). It assists in the 3' end formation of polyadenylated and non-polyadenylated RNA but is also responsible for the dephosphorylation of Ser5 and Ser7 as its activity peaks at the promoter and the 3' end of genes (Ganem et al., 2003; Krishnamurthy et al., 2004).

The phosphorylation pattern of the CTD is dynamically changing during the transcription process. At the step of transcription initiation, Ser5 and Ser7 are phosphorylated by the kinase subunit of TFIIH CDK7. After transcription initiation, the elongating RNA Pol II is paused downstream of the TSS by the binding of the negative elongation factors DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) (Yamaguchi et al., 1999; Wu et al., 2003). To overcome the pausing, the kinase P-TEFb is phosphorylating the two factors as well as the CTD at Ser2 and Thr4 which leads to the dissociation of the negative elongation factors and the continuation of elongation (Renner et al., 2001; Fujinaga et al., 2004; Cheng and Price, 2007). Additionally, the phosphorylation at Ser5 and Ser7 are removed gradually during elongation and the amount of phosphorylation of Ser2 increases with a saturation peak at around 600 nucleotides downstream of the TSS, regardless of the gene length. Moreover, Ser2 phosphorylation starts to decrease around 100 nucleotides downstream of the poly(A) addition site (Mayer et al., 2010) (Figure 14B). The phosphorylated CTD during elongation serves as an interaction platform for different proteins with RNA Pol II and the growing mRNA molecule. These proteins are involved in processes like histone modifications, 5' capping, mRNA splicing or termination of transcription (Hsin and Manley, 2012).

1.4.3 General transcription factors (GTFs)

1.4.3.1 TFIID

TFIID is a general transcription factor which can recognize and bind promoter sequences to initiate the assembly of the PIC. The complex comprises of TBP and 13 TAFs (TAF1-13) with a total size of around 1.2 MDa. The subunits of TFIID are conserved from yeast to human but metazoan cells also include cell type specific variants of TBP and several TAFs which lead to the formation of alternative TFIID complexes (Tora, 2002). These TFIID variant complexes can recognize a specific subset of promoters to activate their gene expression (Dikstein et al., 1996; Hansen et al., 1997; Goodrich and Tjian, 2010). Therefore, the composition of TFIID complexes in metazoan cells can vary depending on the developmental stage or the cell type (Müller et al., 2010; Maston et al., 2012).

The overall structure of the TFIID complex was determined using electron microscopy and displays an asymmetric tri-lobed structure (Andel III, 1999; Brand, 1999; Grob et al., 2006; Cler et al., 2009). Stoichiometry analysis using yeast TFIID revealed that six of the TAFs are present two times in the complex (TAF4, TAF5, TAF6, TAF9, TAF10 and TAF12) whereas the rest of the TAFs and TBP are only present once (Sanders et al., 2002). However, several studies have shown that a functional and symmetric core-TFIID complex exist *in vivo* which consists of pairs of TAF4, TAF5, TAF6, TAF9 and TAF12 (Wright et al., 2006; Bieniossek et al., 2013). This core structure can be achieved through the histone fold domains (HFD) present in TAF4, TAF6, TAF9 and TAF12 which are resulting in heterodimers of TAF6-9 and TAF4-12 and the addition of two WD40-repeat-containing TAF5 proteins. However, several TAFs which are not included into the core-TFIID structure are also containing HFDs (TAF3, TAF8, TAF10, TAF11 and TAF13). The symmetry of the core-TFIID complex is broken as soon as a sub-complex composing of TAF8-TAF10 is entering core-TFIID which results in the asymmetric 7TAF complex. This complex serves as a transition between core-TFIID and the complete holo-TFIID and as TAF10 is transported into the nucleus by TAF8, indicates that the formation of these complex highly depends on the synthesis rate of TAF8 and the time needed for the nuclear transport. The incorporation of the remaining TAFs and TBP will lead to the asymmetric holo-TFIID structure but further experiments need to be performed to test if other transition complexes or sub-

complexes exist and if they have specific functions (Figure 15) (Bieniossek et al., 2013).

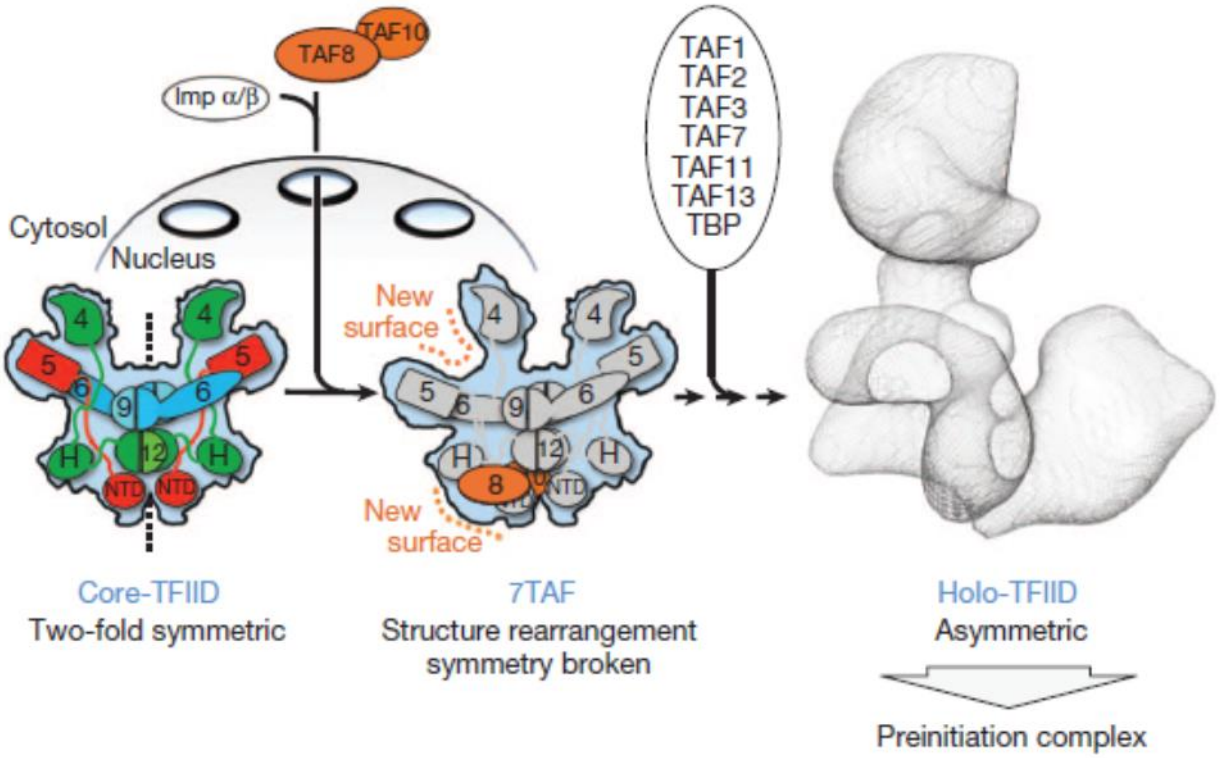


Figure 15: holo-TFIID assembly. *The core TFIID structure is symmetric including two copies of TAF4, TAF5, TAF6, TAF9 and TAF12. The import of a TAF8-TAF10 submodule by importins (imp) and its incorporation breaks the symmetry of the core-TFIID complex resulting in the 7TAF complex. This complex harbors two distinct halves and new binding surfaces for other subunits indicated by dashed lines. Incorporation of TBP and the remaining TAFs in single copies leads to the formation of the holo-TFIID structure (grey mesh) which is involved in PIC formation. From (Bieniossek et al., 2013).*

Concerning transcription initiation on a promoter containing a TATA-box, the promoter sequence is generally recognized by TBP. This is why TBP is also known as one of the key elements of promoter recognition. The structure of TBP is highly conserved and displays a bipartite saddle-like structure, where TATA-box binding is achieved by the concave part of TBP (see Figure 16). This binding will afterwards lead to a bending of the DNA of 90˚ which results in the asymmetric platform needed for

PIC assembly. The binding of TBP to DNA is highly regulated, especially by the binding of other proteins which inhibit either the recruitment of other GTFs to TBP or block the binding of TBP to DNA already from the start. Some of these negative factors are NC2, BTAF1 (Mot1 in yeast) and even the TAF1 N-terminal domains (TANDs) of TAF1. That shows that even TFIID can modulate TBP activity itself. The TAND domain of TAF1 can bind to different parts of TBP hindering its ability to detect DNA (Kotani et al., 1998; Bagby et al., 2000; Liu et al., 2008; Anandapadamanaban et al., 2013). However, binding of TFIIA can displace the inhibitory domains of TAF1 to stabilize the interaction of TBP with the promoter (Ozer et al., 1998). This shows how important GTFs like TFIIA and TFIIB are for the regulation of TBP and that they are serving as positive factors for initiation. Another mechanism for TBP binding inhibition is through the binding of BTAF1 to TBP (also known as B-TFIID) which was identified *in vitro* using the purified yeast counterpart of BTAF1 called Mot1 (Moyle-Heyrman et al., 2012). Briefly, after an initial formation of a TBP-DNA complex and bending of the DNA, the negative factor Mot1 (or BTAF1) can bind to the complex and after ATP-hydrolysis involving the ATPase domain of Mot1 (BTAF1), this ternary complex undergoes a conformational change which leads to the displacement of TBP from the DNA.

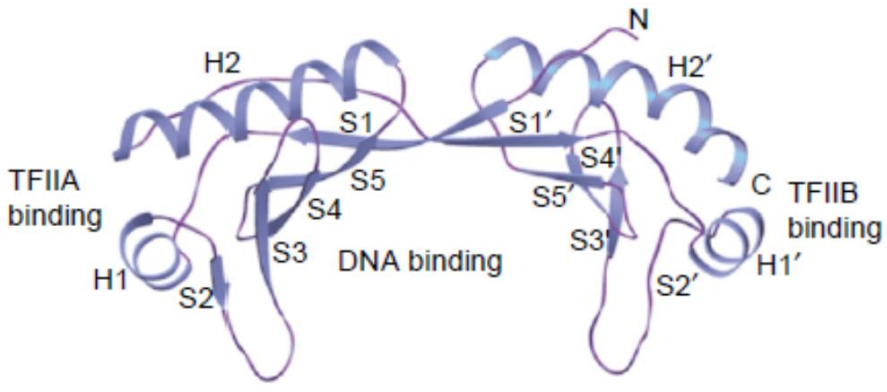


Figure 16: Ribbon diagram of the 3D structure of the TBP core domain. *The regions with which TFIIA and TFIIB are interacting with TBP are indicated. From (Davidson, 2003).*

However, it is known that a TATA-box is only found in around 10-20% of all yeast or human promoters (Yang et al., 2007). Thus, how is TFIID recruited to these TATA-less promoters? In fact, it was shown that several TAFs have the ability to recruit the TFIID complex to these promoter types. They can actually bind to other promoter motifs that were mentioned before in a previous section (section 1.1.2). It was shown that the INR motif can be recognized by TAF1 and TAF2 (Chalkley and Verrijzer, 1999) and the MTE and DPE sequences can be bound by TAF6-TAF9 (Burke and Kadonaga, 1997; Theisen et al., 2010).

Another possibility of TFIID recruitment to TATA-less promoter is its function as a co-activator and therefore the interaction with specific activators (Burley and Roeder, 1996). As mentioned before the TFIID composition can vary between different cell types and developmental stages which indicates that there are several variant TAFs included into specific TFIID complexes that can interact with specific activators to recruit TFIID and initiate transcription of specific genes. Recent studies examined the interaction of several human (p53, Sp1 and c-Jun) or yeast (Rap1) activators with TFIID. These activators are directly interacting with TFIID and after investigation it was found that in contrast with Mediator-activator interactions, TFIID shows no conformational changes after activator binding. Additionally, all the activators tested were binding on different locations on TFIID which indicates that different TAFs are required for the interaction with specific activators (Taatjes et al., 2002; Liu et al., 2009). One example for TFIID-activator interaction is the activator Rap1 which is required together with TFIID to express ribosomal protein genes in yeast. Rap1 binds through a network of interactions within TFIID including Taf4, Taf5 and Taf12. Even if TBP is not located near the binding site of Rap1, it was shown that an interaction of TFIID bound Rap1 with TFIIA serves as a bridge to contact TBP. This interaction results into a position change of TBP within the complex which could have different effects on TFIID which are not fully understood yet (Garbett et al., 2007; Papai et al., 2010):

(a) This position switch could stimulate an activator-depending binding of TFIID to the promoter.
(b) The TFIID-promoter interaction could be stabilized due to trapping of the DNA through the protein bridge.
(c) It could induce the recruitment of other PIC components like TFIIB or RNA Pol II to the promoter.

1.4.3.2 TFIIA

The auxiliary factor TFIIA is a heterodimer which harbors two domains including a 4-helix bundle and a 12-stranded β-barrel which enables the complex to bind the TATA-box as well as the underside of the TBP saddle. This structure explains the function of TFIIA during PIC formation as it is able to stabilize the binding of TBP to the promoter DNA (Imbalzano et al., 1994). The structure of yeast as well as human TFIIA-TBP-DNA is similar and resembles a “boot-shaped” TFIIA heterodimer which is not altering the initial TBP-DNA structure (Bleichenbacher et al., 2003). TFIIA is specific for RNA Pol II transcription and even if it is not essential for basal transcription, it can stimulate basal and activated transcription (Kang et al., 1995). As already mentioned above, besides it stabilizing function, TFIIA also competes with negative factors like NC2, BTAF1 (Mot1 in yeast) or the Taf1 TAND domain for TBP binding. Furthermore, the stability of the TFIIA-TBP-DNA complex depends on the DNA sequence that is bound (Stargell et al., 2001).

1.4.3.3 TFIIB

TFIIB is binding to the DNA immediately flanking the TATA-box which contain sequence elements already mentioned above called the BREs. As for TFIIA, it is involved in TBP-DNA binding stabilization but also serves as a platform to recruit RNA Pol II to the promoter. However, crystal structures of the TFIIB-RNA Pol II complex showed that besides the aforementioned functions of TFIIB, it also harbors some post-recruitment functions and is involved in TSS recognition and the initiation process.

The protein is the only GTF which harbors only one subunit. The structure contains an N-terminal loop, also called the “B-ribbon”, and a C-terminal “core” domain (Figure 17A) (Bushnell et al., 2004; Kostrewa et al., 2009; Liu et al., 2010). These two domains are responsible for the recruitment of RNA Pol II to the promoter by binding to the dock and wall domain of RNA Pol II to form the closed complex. After the establishment of the PIC complex, the conformation of the TFIIB-RNA Pol II complex is changing and the region between the B-ribbon and B-core domains will enter the RNA Pol II cleft and form two new elements, called the “B-reader” and “B-linker” domains (Kostrewa et al., 2009). The B-linker domain will help for the DNA melting 20 nucleotides downstream of the TATA-box which will lead to the sliding of the emerging

template strand into the RNA Pol II cleft. Afterwards, the B-reader will bring the template strand into position for initiation and contributes to find the INR sequence. Furthermore, TFIIB will stimulate the synthesis of a short RNA molecule of around 6 nucleotides and helps to stabilize this initiation complex by using the B-reader loop to block any further RNA synthesis. It is suggested that TFIIB also accompanies in DNA-RNA strand separation and guides the RNA to the RNA Pol II exit tunnel. Last, as soon as the newly synthesized RNA molecule reaches a length of around 12-13 nucleotides, it clashes with the B-ribbon domain and TFIIB is released (Figure 17B). This mechanism shows how important the interaction between TFIIB and nucleic acids is to accomplish a successful initiation-to-elongation transition (Pal et al., 2005; Kostrewa et al., 2009; Sainsbury et al., 2013).

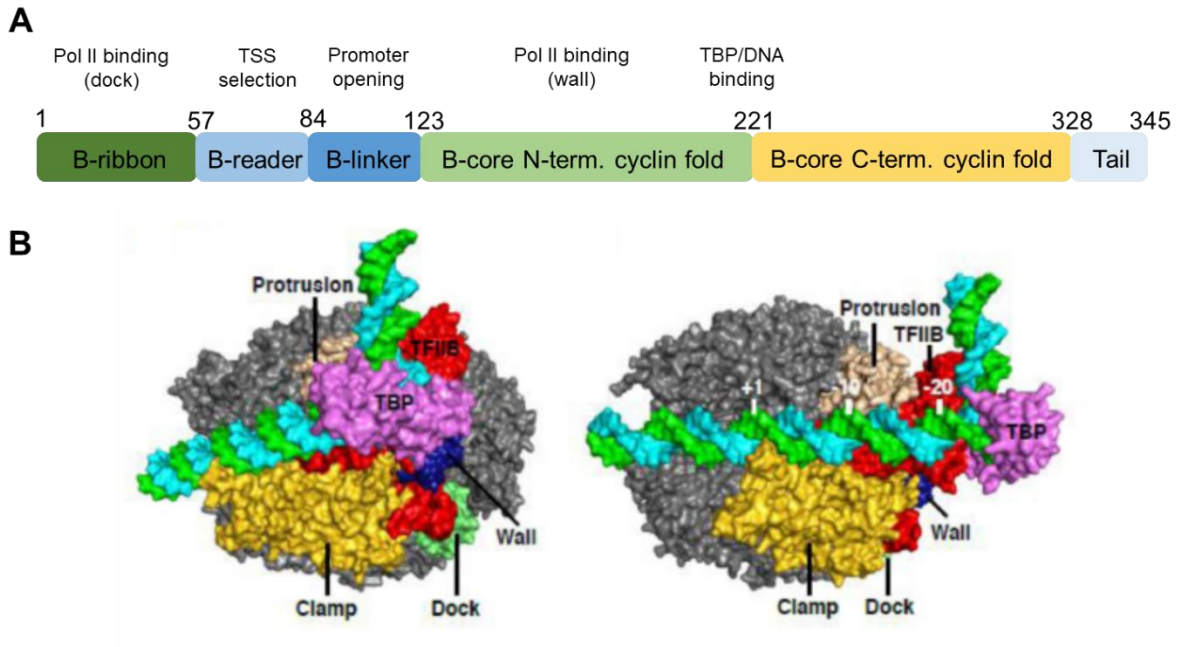


Figure 17: Structure and function of TFIIB. A *Organization of the TFIIB functional domains. Important domains for transcription initiation and interaction domains with other factors are indicated.* ***B*** *Structure of the closed promoter complex in top and side view, respectively. All factors are shown in surface representation with the RNA Pol II clamp in gold, the dock in lime, the wall in blue, protrusion in wheat and the rest of the complex in gray whereas TFIIB is shown in red and TBP in violet. The template and non-template strands of the DNA are displayed in cyan and green, respectively. Adapted from (Kostrewa et al., 2009; Liu et al., 2012).*

1.4.3.4 TFIIF

TFIIF is a heterodimer containing the subunits TFIIFα and TFIIFβ and was identified in mammalian cells due to its possibility to bind to RNA Pol II (Burton et al., 1986; Burton et al., 1988). It has several functions throughout the transcription initiation process. The complex assists in the recruitment of RNA Pol II to the promoter as well as inhibits non-specific binding of RNA Pol II to the DNA and also stabilizes TFIIB binding to the PIC (Čabart et al., 2011). After PIC assembly, TFIIF is known to support phosphodiester-bond formation and early RNA synthesis as well as for suppressing RNA Pol II pausing (Price et al., 1989; Yan et al., 1999). It is important to mention that transcription initiation can be accomplished *in vitro* without TFIIE and TFIIF but not if TFIIF is missing as well (Pan and Greenblatt, 1994).

1.4.3.5 TFIIE

As for TFIIF, the general transcription factor TFIIE is a heterodimer which consists of the two subunits TFIIEα and TFIIEβ (Ohkuma et al., 1990; Peterson et al., 1991). TFIIE is needed to recruit TFIIH to the PIC and acts therefore as a bridge between TFIIH and RNA Pol II (Maxon et al., 1994; Holstege et al., 1996). The main function of TFIIE (together with TFIIH) is to help to establish the open promoter conformation. Furthermore, it has been shown that TFIIE can bind to single stranded DNA to stabilize the open promoter complex (Kuldell and Buratowski, 1997; Yokomori et al., 1998).

1.4.3.6 TFIIH

TFIIH is a multisubunit complex with 10 subunits including the ATPase XPB, a core module consisting of six subunits including the ATPase XPD, a three-subunit kinase module of CDK7-cyclin H-MAT1 and the proteins p62, p52, p34, p8 and p44 (Gibbons et al., 2012; Murakami et al., 2012). After recruitment of the complex to the PIC by TFIIE, it is required for creating and stabilizing the open promoter conformation *in vitro* and *in vivo* but it is also involved in promoter escape. The complex harbors three subunits with catalytic activity: XPB and XPD are ATPases whereas CDK7 has a kinase activity. Promoter opening is dependent on the ATPase activity of the subunit

XPB but not on the activity of the subunit XPD (Conaway and Conaway, 1993; Schaeffer et al., 1993; Lin et al., 2005). TFIIH is also involved in the DNA nucleotide excision repair pathway where the helicase activity of XPD is required (Coin et al., 2007). XPB does not work as a classical helicase to unwind the DNA as it never binds directly to the DNA that needs to be unwound. Instead, it is the translocase activity of XPB that will help to load the DNA into the cleft of RNA Pol II (Grünberg et al., 2012). It was shown that XPB is scanning the DNA in a 5'-3' direction and translocate the DNA into the cleft of RNA Pol II which will lead to the unwinding of the DNA and the recognition of the TSS (Fishburn et al., 2015). Recently it was shown that inhibition of XPB ATPase activity has an impact on transcription whereas complete loss of XPB does not (Alekseev et al., 2017). As already mentioned in a previous section, the kinase module, especially the kinase CDK7, of TFIIH is capable to phosphorylate the CTD of RNA Pol II at serine 5 which is essential for the promoter escape of RNA Pol II (Serizawa et al., 1995). Interestingly, CDK7 alone is also catalytically active but phosphorylation of RNA Pol II CTD is only possible if the kinase is incorporated into TFIIH (Rossignol et al., 1997; Yankulov and Bentley, 1997).

1.5 Transcription elongation and termination

1.5.1 RNA Pol II pausing and elongation

What happens after transcription initiation is accomplished? As mentioned already in the previous section 1.4.2, RNA Pol II is known to be paused downstream of the TSS after initiation and this is regulated by binding of the negative elongation factors DSIF and NELF (Figure 18A). However, pause release can be achieved through the phosphorylation of the aforementioned factors by the P-TEFb complex which will result in the dissociation of the factors from RNA Pol II. Inhibition of P-TEFb by using the transcription elongation inhibitor flavopiridol results in a blockage of RNA Pol II entry into productive synthesis which shows how important the recruitment of P-TEFb is to release RNA Pol II from the pausing state (Ni et al., 2008; Rahl et al., 2010). The factor that is phosphorylated by P-TEFb in the DSIF complex is its largest subunit SPT5 (Lis et al., 2000; Yamada et al., 2006). Interestingly, SPT5 was also shown to act as a positive elongation factor by recruiting capping enzymes to the nascent transcripts (Wen and Shatkin, 1999). Recent studies also indicated that Ser7 phosphorylation of the RNA Pol II CTD by TFIIH is involved in maintaining the integrity of paused RNA

Pol II either at the 5' end of promoters or at the 3' termination site (Glover-Cutter et al., 2009). Several models have been proposed about the possible functions of paused RNA Pol II. One model states that paused RNA Pol II could be important to keep the promoter region nucleosome free and therefore the promoter active until GTFs or other regulatory factors can bind and RNA Pol II can be released. Another model suggests that pausing and the resulting nucleosome free promoter region could be bound quickly by activators or co-activators to ensure even more efficient activation of transcription. Furthermore, pausing could also serve as an additional regulatory step to have a combinatorial control through recruitment and release of RNA Pol II. Lastly, it could serve as another checkpoint to ensure that pre-mRNA processing factors, for example involved in 5' capping, have enough time to process the pre-mRNA before productive elongation can resume (Adelman and Lis, 2012).

After RNA Pol II is released from pausing the elongation of the transcript can start and different elongation factors are interacting with the polymerase to ensure productive elongation of the nascent transcript (Figure 18B). Some of these factors were already discussed before which either help to pause RNA Pol II or to release it from pausing. Other factors are either actively supporting elongation by stabilizing RNA Pol II on the DNA or helping to release arrested or stalled polymerase, or passively by modifying the CTD/histones or remodeling of the chromatin (Shandilya and Roberts, 2012). One of these elongation factors is TFIIS which can alleviate arrested RNA Pol II by stimulating the RNA Pol II mediated cleavage of the nascent transcript (Sims et al., 2004). Interestingly, other studies indicated that TFIIS could be also involved in transcription initiation (Kim et al., 2007). Important histone modifications for elongation are monoubiquitination of histone H2B and methylation of histone H4K36 performed by the elongation factors hPAF1 (including RNF20/40 and UbcH6) as well as Set2, respectively (Sims et al., 2004). Also the previously discussed phosphorylation of Ser2 on the CTD of RNA Pol II by P-TEFb as well as the removal of Ser5 by Ssu72 are facilitating productive elongation by providing the right platform for other proteins involved in pre-mRNA processing. Another class of elongation factors include histone chaperones like the FACT complex. It is involved in the re-deposition of evicted histones behind the transcription bubble to ensure that no initiation can occur within the open reading frame of the gene (Kaplan et al., 2003; Schwabish and Struhl, 2004).

Importantly, it was shown that the function of the FACT complex relies on the presence of monoubquitinated histone H2B (Pavri et al., 2006).

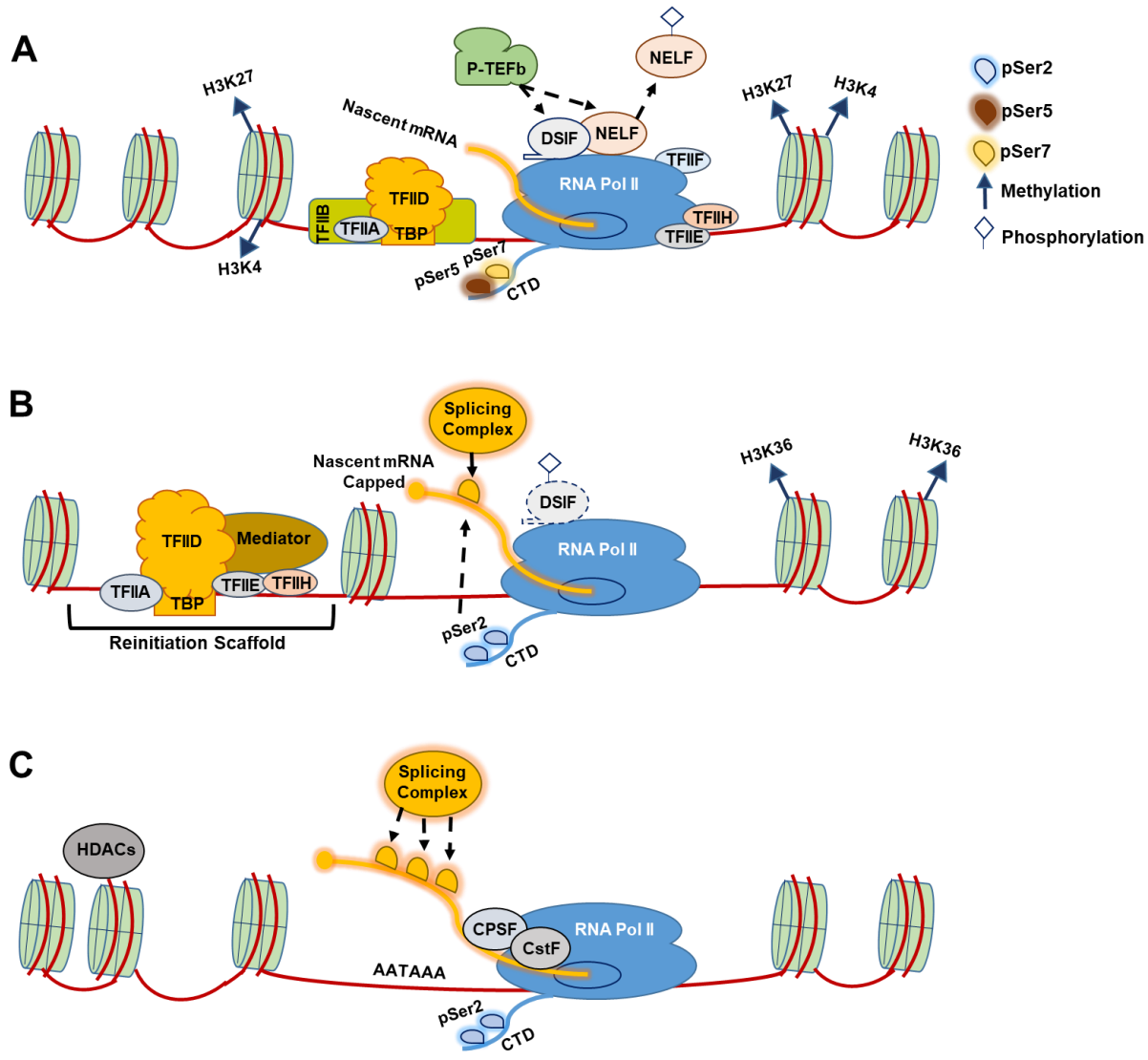


Figure 18: Legend on the next page →

Figure 18: Different steps of the transcription cycle after initiation. A RNA Pol II pausing. *The presence of the negative factors DSIF and NELF inhibit successful transcription initiation and induce RNA Pol II pausing. Nucleosomes flanking these paused promoters are enriched with specific histone modifications (H3K27 and H3K4 methylation). Paused RNA Pol II CTD is also enriched with phosphorylation on Ser7. The kinase subunit of P-TEFb called CDK9 can phosphorylate DSIF and NELF which leads to the dissociation of NELF from RNA Pol II and the start of transcription elongation whereas DSIF stays bound to the polymerase.* ***B Transcription elongation.*** *After promoter clearance the RNA Pol II complex is elongating the transcript while a part of the initiation complex stays bound at the promoter for a possible transcription reinitiation. The CTD of RNA Pol II is phosphorylated at Ser2 by CDK9 which serves as a platform to recruit capping enzymes and the splicing complex for co-transcriptional splicing of the transcript. Histone within the open reading frame (ORF) are enriched with H3K36 methylation.* ***C Transcription termination.*** *As soon as RNA Pol II reaches the poly A signal (AATAAA) at the gene terminal, 3' end processing and termination factors like CPSF and CstF are recruited. The DNA that was already transcribed by RNA Pol II is reassembled into chromatin by the action of histone chaperones and deacetylases (HDACs). Modified from (Shandilya and Roberts, 2012).*

1.5.2 Transcription termination and gene looping

The next question would be what mechanisms are involved at the end of the gene to terminate transcription and what happens with RNA Pol II afterwards? As soon as mRNA synthesis is completed, RNA Pol II dissociates from the DNA and this marks the end for transcription (Figure 18C). However, this event also serves as a new starting point for another round of transcription with the possibility to recycle RNA Pol II. There exists two well studied pathways for transcription termination which are the Nrd1-Nab3-Sen1-dependent pathway and the poly (A)-dependent pathway. The poly (A)-dependent pathway is used for the termination on protein coding genes where the mRNA precursor sequence ends on a poly (A) signal (5'-AAUAAA-3') followed by a G/U rich sequence at the 3' end. As soon as RNA Pol II transcribes the poly (A) signal, a reduction in its prosessivity can be detected which leads to the pausing of the complex further downstream. Again, the phosphorylated CTD of RPB1 on Ser2 serves as a platform to recruit several protein complexes involved in the termination process like the cleavage and polyadenylation specificity factor (CPSF) and the cleavage stimulatory factor (CstF) (Ahn et al., 2004; Kuehner et al., 2011). Both complexes are

recruited to the RNA Pol II CTD as well as to the transcribed poly (A) sequence on the pre-mRNA which induces first pausing and then eventually the release of RNA Pol II from the DNA as well as cleavage and polyadenylation of the transcript (Nag et al., 2007). However, RNA Pol II is also responsible for the transcription of long non-coding RNAs of various poorly understood functions. These Cryptic Unstable Transcripts (CUTs) show to have divers modes of 3' end processing but no poly (A) signal. The transcription processes involving the CUTs have shown to be terminated by the alternative Nrd1-Nab3-Sen1-dependent pathway and are getting rapidly degraded by the exosome (Vasiljeva and Buratowski, 2006). Importantly, most small nuclear RNAs (snRNAs) genes have a different mode of termination which includes the integrator complex. It was shown that for snRNA transcription termination, the phosphorylation of Ser7 on the CTD of RNA Pol II is required for the recruitment of the integrator complex and the following cleavage of the transcript (Egloff et al., 2007).

After transcription is terminated and RNA Pol II is released from the DNA template, the mRNA is further processed and afterwards exported to the cytoplasm for translation. RNA Pol II CTD modifications need to be cleared by phosphatases as already mentioned in a previous section to return to its unposphorylated form RNAP IIA. Afterwards, the resulting RNA Pol II complex is ready for the next transcription initiation round. It was shown that some of the GTFs can remain associated at the promoter as a partial PIC (Hahn, 2004; Sarge and Park-Sarge, 2005). This complex, stabilized by gene specific activators and co-activators, can serve as a platform to recruit RNA Pol II for several rounds of transcription reinitiation (Yudkovsky et al., 2000; El Kaderi et al., 2009). Several studies have also shown that promoter and terminal regions can interact with each other through several factors to achieve a phenomenon called gene looping. In human cells it was shown that TFIIB can interact with termination factors like CPSF and CstF which results in an interaction of promoter and terminator DNA and a possibly fast recycling of RNA Pol II as well as reinitiation of transcription (Calvo and Manley, 2003; Singh and Hampsey, 2007). An involvement of the GTF TFIIH in gene looping was also considered as it harbors the kinase for Ser7 phosphorylation of the CTD which was detected mainly at the termination site of snRNA genes but also to some degree in protein coding genes which could indicate a TFIIH driven recycling mechanism for RNA Pol II (Chapman et al., 2007; Glover-Cutter et al., 2009). It was also suggested that multiple genes could form some kind of

“transcription factories” with a high concentration of RNA Pol II molecules and transcription factors which transcribe RNA for several rounds at a specific area in the nucleus (Osborne et al., 2004).

1.6 Transcription-replication crosstalk

Transcription is not the only process that is taken place on chromatin, so what happens if the transcription machinery is encountering other mechanisms like replication? In some cases, the crosstalk between other processes and transcription can be very beneficial for the cell as in the case for transcription-coupled repair. However, often encounters between the transcription machinery and other mechanisms can have negative consequences. One of these processes is DNA replication in which the whole genome is replicated during the S-phase of the cell cycle before cell division to copy the genomic information for the daughter cell. Extensive research in the past has shown that conflicts between these two machineries can lead to genomic instability, replicative stress and DNA damage which are all hallmarks for cancer (Gaillard et al., 2013). Replication stress is defined by slowing down or stalling of the replication fork which hinders the progression of DNA synthesis and can induce DNA damage. However, as both processes, transcription and replication, are essential mechanisms for cell viability and proliferation, the cells have developed mechanisms for either preventing or resolving possible collisions and their consequences (Helmrich et al., 2013; García-Muse and Aguilera, 2016).

Transcription-replication-conflicts (TRCs) can take place in different ways depending on the functional states of the processes and their directionality on the chromatin. Therefore, key points of the severity and impact of the collisions seem to be depending mainly on the orientation of the machineries as well as on the type of the transcriptional block (Hamperl and Cimprich, 2016). Both machineries start replication or transcription from two distinct genomic locations which are origins and promoter, respectively. Additionally, they both are highly processive and comprise of a strict 5’-to-3’ polarity. Therefore, two different types of conflicts can occur either in co-directional orientation or in a head-on fashion (see Figure 19). It was shown that head-on collisions are much more severe to inhibit replication fork movement than co-directional conflicts. However, it is important to mention that direct contact of DNA and

RNA polymerase could not be detected until now. The first replisome factor that would encounter the stalled replication fork is the replicative helicase Mcm2-7 which would unwind the DNA strands ahead of the fork. However, it was shown that this mechanism is not functioning if a head-on TRC occurred (Mirkin and Mirkin, 2005; Prado and Aguilera, 2005; Srivatsan et al., 2010). Another explanation could be that positive DNA supercoils could accumulate between the two machineries and therefore induce replication fork stalling in the same way as it occurs after inhibition of Topoisomerase I and the resulting negative supercoiling (Tuduri et al., 2009; Bermejo et al., 2012). A co-directional collision cannot be influenced by one of the aforementioned challenges. However, also co-directional TRCs can be problematic if the transcription machinery is blocked during transcription. Several factors like RNA Pol II pausing, backtracking or transcription blockage due to a DNA lesion can lead to severe TRCs. Another transcriptional barrier is the formation of DNA-RNA hybrids, so called R-loops, which occur if the nascent RNA is hybridizing with the template strand (Gowrishankar et al., 2013). Although, it was shown that R-loops can have a physiological function, they can also block the continuation of the replication fork by inducing a co-directional TRC (Stirling et al., 2012).

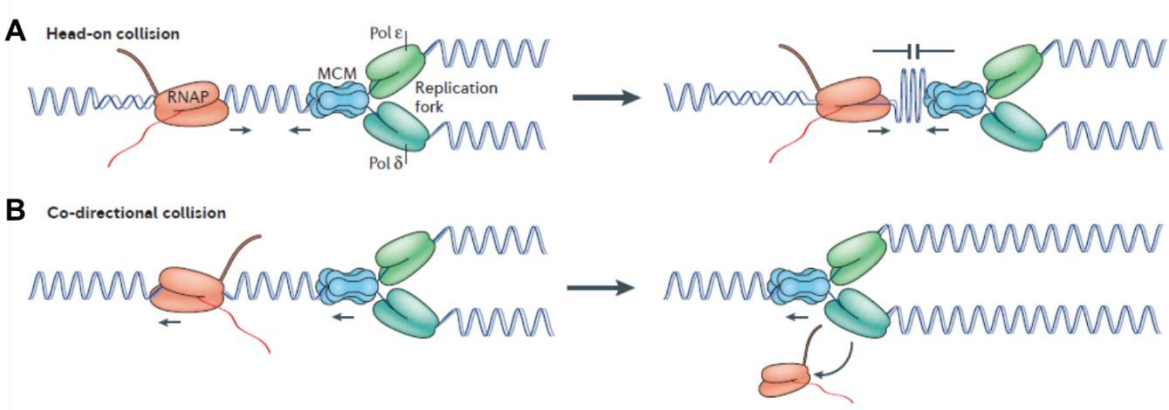


Figure 19: Types of transcription-replication collisions. A *The progression in opposite direction of a replication fork and an elongating RNA Pol II can lead to a head-on collision which induces blockage of the replication fork and can result in the formation of DNA breaks.* ***B*** *The progression of the replication fork and RNA Pol II elongation in the same direction can induce a co-directional collision if the replication fork moves faster than RNA Pol II or if transcription is paused or blocked. This blockage can be solved by the displacement of RNA Pol II from the DNA. Adapted from (García-Muse and Aguilera, 2016).*

However, how can these conflicts be avoided? One possibility is the redistribution of replication initiation factors by active transcription (Powell et al., 2015). Like this, the occurrence of replication initiation in regions of active transcription could be minimized. Additionally, during S-phase it is tried by the cell to compartmentalize replication and transcription into different regions of the genome and act within these territories for a distinct time. This enables the transcription of certain genes during early S-phase whereas other genes can be transcribed during late S-phase (Wei et al., 1998; Dimitrova, 2011).

There are also mechanisms to suppress conflicts which can be divided into two groups depending whether the replication or transcription machinery is taking action. Considering the transcription machinery, all mechanisms that can reactivate, destabilize or remove stalled RNA Pol II are helpful to avoid TRCs (Figure 20). Thus, the reactivation of backtracked RNA Pol II or the positive influence on elongation by the elongation factor TFIIS can help to counteract against collisions (Cheung and Cramer, 2011). Furthermore, it was shown that the human helicase RECQL5 can decrease the elongation rate and therefore reduce stalling or backtracking of RNA Pol II (Saponaro et al., 2014). DNA lesions induced through various DNA-damaging agents can also block RNA Pol II progression. In these cases the DNA damage response, or more specifically the transcription coupled nucleotide excision repair (TC-NER) pathway was shown to repair the lesion and also removes the blocked RNA Pol II from the DNA. Successful and rapid transcription termination can also decrease the chance of potential TRCs. Interestingly, different termination mutants showed an increase in R-loop levels at the termination site due to the unfinished transcription and the presence of the nascent RNA which can interact with template DNA (Mischo et al., 2011; Skourti-Stathaki and Proudfoot, 2014).

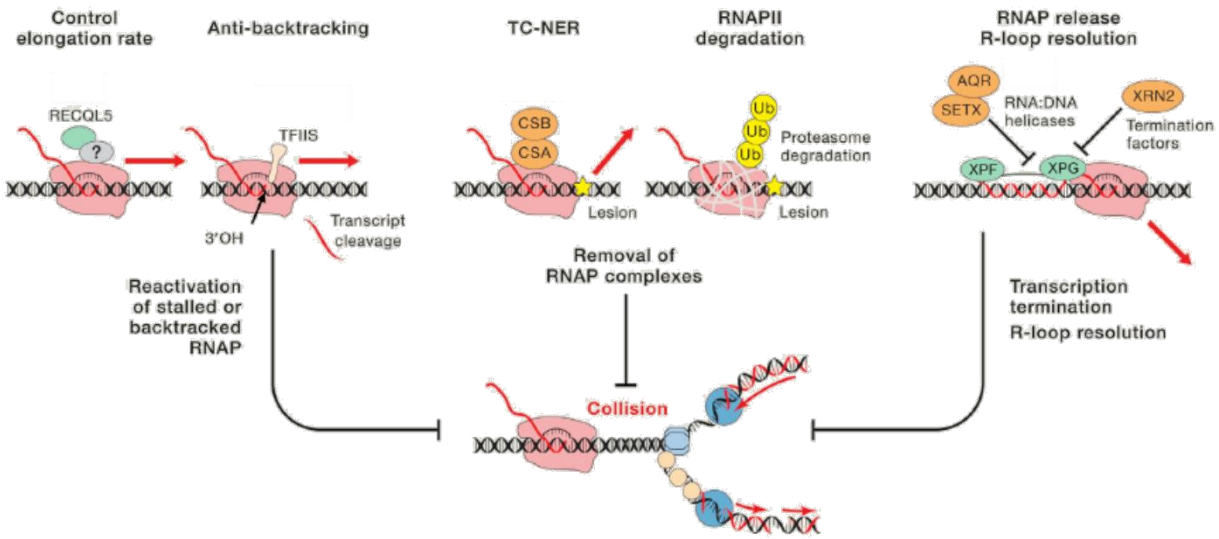


Figure 20: Pathways to resolve transcription roadblocks and avoid collisions. *The human RECQL5 helicase can reduce stalling or pausing by regulating the elongation rate of RNA Pol II. Other elongation factors like TFIIS can induce cleavage of backtracked transcripts to resume transcription (left). Blocked RNA Pol II due to DNA damage (yellow star) can be removed from the DNA by the TC-NER repair pathway or by proteasome-mediated degradation via poly-ubiquitiylation (middle). Successful transcription termination and resolution of R-loops is performed by Xrn2 exonucleases and RNA:DNA helicases like SETX or AQR. R-loops can be also recognized and resolved with the help of the TC-NER endonucleases XPF/XPG (right). From (Hamperl and Cimprich, 2016).*

On the other side, also factors involved in replication are helping to avoid TRCs. One option are auxiliary helicases which are traveling with the replication fork and can help to dislodge transcription complexes before a collision. Another level of regulation can occur through S-phase checkpoints. However, eukaryotic cells harbor hundreds to thousands of origins which requires multiple levels of regulation. Replication factors are licensing replication origins in the G1-phase and only these origins are replicated in S-phase. However, if replication is stalled, it is possible to restart replication at another origin which was not licensed before. Like this, the new origins can rescue the DNA synthesis. Another possibility is the ataxia telangiectasia-mutated and Rad3-related (ATR) dependent replication checkpoint. This kinase pathway can regulate origin firing, stabilizes replication forks and promotes fork repair and restart (Cimprich and Cortez, 2008; Ciccia and Elledge, 2010). However, there are evidences that TRCs are not always destructive. For example, it was shown that TRCs may have crucial

roles in the context of cellular morphogenesis and development (Blythe and Wieschaus, 2015).

2. Transcription visualization in vivo

The whole process of transcription is inheritably dynamic but most of what we know about the assembly, initiation and elongation of transcription and what was described in the sections before was often determined using biochemical techniques. While these findings have been heavily important to define key factors of transcription and their interactions with each other, they are not entirely suited to gain new insights about the kinetics of the transcription process (Levine et al., 2014). Moreover, the recent and rapid development of genome-wide high-throughput assays like chromatin immunoprecipitation (CHIP) or chromosome conformation capture experiments (Hi-C) also helped to provide new insights into TFs binding patterns across the genome as well as genome organization and chromatin architecture at the level of cell populations (Barski et al., 2007; Lieberman-Aiden et al., 2009). However, these techniques represent end point assays which only take into account a huge cell population and therefore cannot give any information about the 3D molecular structure and dynamics of the transcription process in individual living cells. Furthermore, the reliance of most of these techniques on purification approaches and *in vitro* reconstitutions raises the question how accurate the resulting data is to explain the processes happening in the complex environment of intact living cells. However, recent advances in the field of fluorescent labeling techniques in combination with cutting edge microscopy techniques can overcome some of the problems and can give new insights into the dynamical behavior of factors of the transcriptional machinery (Misteli, 2001; Mazza et al., 2012; Gebhardt et al., 2013). Thus, imaging techniques like fluorescence recovery after photobleaching (FRAP), fluorescence correlation spectroscopy (FCS), structured illumination microscopy (SIM) and single particle tracking (SPT) techniques can provide unique features to measure the assembly and dynamics of the transcription machinery in single living cells (Liu et al., 2015).

2.1 Transcription imaging and dynamics

Advances in biomolecular labeling techniques and microscopy modalities gave the opportunity to measure the dynamics of TFs in living cells. Early pioneering studies used on one hand fluorescent fusion proteins (FPs) (Tsien, 1998) and on the other side two highly important microscopy techniques to measure transcription dynamics: FRAP and FCS.

2.1.1 Florescence recovery after photobleaching (FRAP)

Briefly, in a typical FRAP experiment, a focused laser beam is bleaching the exogenously expressed FPs in a specific region in the nucleus. The fluorescent recovery in the bleached region is afterwards dependent on the diffusion and binding kinetics of the target proteins outside of the bleached area (K_{on} and K_{off}) as well as on the dissociation rates (K_{off}) and diffusion kinetics of the bleached FPs inside the target area (Figure 21A). FRAP has shown to be very effective to measure the residence times of TFs from several seconds to hours in living cells (Axelrod et al., 1976). In contrast, core histone subunits have shown to be highly stable with little exchange of molecules even after 1-2 hours. Interestingly, a fraction of the histone H2B exhibited a much faster exchange in the range of minutes which shows that the core histone complex is highly stable whereas H2B on the surface of active nucleosomes is exchanged more often (Kimura and Cook, 2001). FRAP experiments of RNA Pol II indicated an estimated elongation rate of the polymerase in a range from 0.4 kb/min to 4.3 kb/min. However, it is important to mention that FRAP cannot distinguish between initiating, elongating or pausing RNA Pol II which makes it difficult to suggest a definite elongation rate using this technique (Kimura et al., 2002; Yao et al., 2006; Boireau et al., 2007; Darzacq et al., 2007). Beside the difficulty to measure fast diffusion dynamics or residence times for subpopulations of bound proteins, FRAP has also the disadvantage that the results rely on protein overexpression and averaging over a large number of cells/measurements (Müller et al., 2010; Mazza et al., 2012).

2.1.2 Fluorescence correlation spectroscopy (FCS)

Another method for measuring diffusion rates of molecules is FCS which can be also used to obtain absolute molecular concentrations of molecules within a volume or to study molecular interaction dynamics in living cells (Magde et al., 1972). The technique is based on the observation of several single fluorescent molecules which are passing through a very small detection volume of a few femtoliter and the subsequent analysis of the diffusion using autocorrelation function calculations (Figure 21B). One advantage of this technique is that it is suitable for longer acquisitions as there is no selection of a specific molecule for observation but of several molecules that are passing through the detection volume. Additionally, the method is based on measuring fluorescence fluctuations instead of localization which enables for the resolving of fast diffusion dynamics. Moreover, by fitting the data to diffusion models, it is possible to distinguish between factors that are in chromatin-bound or -free states. However, as for FRAP also FCS is based on the averaging of measurements as well as on the models that are applied to analyze the data (Mazza et al., 2012). Additionally, to generate reliable fluctuations for detection, it is important that only a very low concentration of fluorescent target protein (<10 nM) is present in the nucleus even if this disadvantage can be partially overcome by combining FCS with photoactivatable proteins (White et al., 2016). In contrast to FRAP, is the detection of stably bound factors inside the detection volume problematic in FCS measurements as they generate no fluctuations and can get bleached quite quickly (Stasevich et al., 2010).

2.1.3 Single particle tracking (SPT)

However, both presented techniques are based on the averaging of several measurements from a cell population. Thus, how it is possible to study TFs dynamics in a single living cell on distinct fluorescently labeled molecules? Therefore, a more recent imaging approach namely SPT enables the observation of single molecules directly in motion within living cells (Figure 21C). Although, SPT is already an ancient technique that was originally used to study the movements of tiny objects like pollen under the microscope, the single molecule tracking in live cells was not possible until the development of fluorescence microscopy and protein labeling strategies. The first application of SPT was to study membrane proteins using fluorescently labeled antibodies (Ghosh and Webb, 1994). However, the analysis of single intracellular or even intranuclear proteins was not easy due to the lack of suitable labeling strategies to conquer the high packing density of the proteins inside the cell. Photoactivatable and photoswitchable FPs/dyes helped to overcome this problem due to the fact that their fluorescence could be either switched on/off or modified to another emission spectrum after excitation at a certain wavelength (usually 405 nm) to avoid fluorescent emission of all target proteins at the same time (Patterson and Lippincott-Schwartz, 2002; Lukyanov et al., 2005). Additionally, one of the most important breakthroughs was the development of different super-resolution microscopy techniques which enabled to achieve real single molecule resolution (see section 2.2) (Hell and Wichmann, 1994; Heintzmann and Cremer, 1999; Klar and Hell, 1999; Gustafsson, 2000; Betzig et al., 2006; Rust et al., 2006). Due to the possibility to resolve individual biomolecules, SPT has the ability to measure TF diffusion and binding kinetics, can be used to study subpopulation-associated structures and can investigate the different steps of multimolecular binding events in living cells (Liu and Tjian, 2018). However, also SPT has some drawbacks as, in contrast to FCS, the detection of fast moving molecules is difficult due to the induced motion blur effect in SPT. Additionally, longer acquisitions are problematic as the same molecule is imaged over time and therefore it will be bleached at a certain point during the acquisition (Chen et al., 2014).

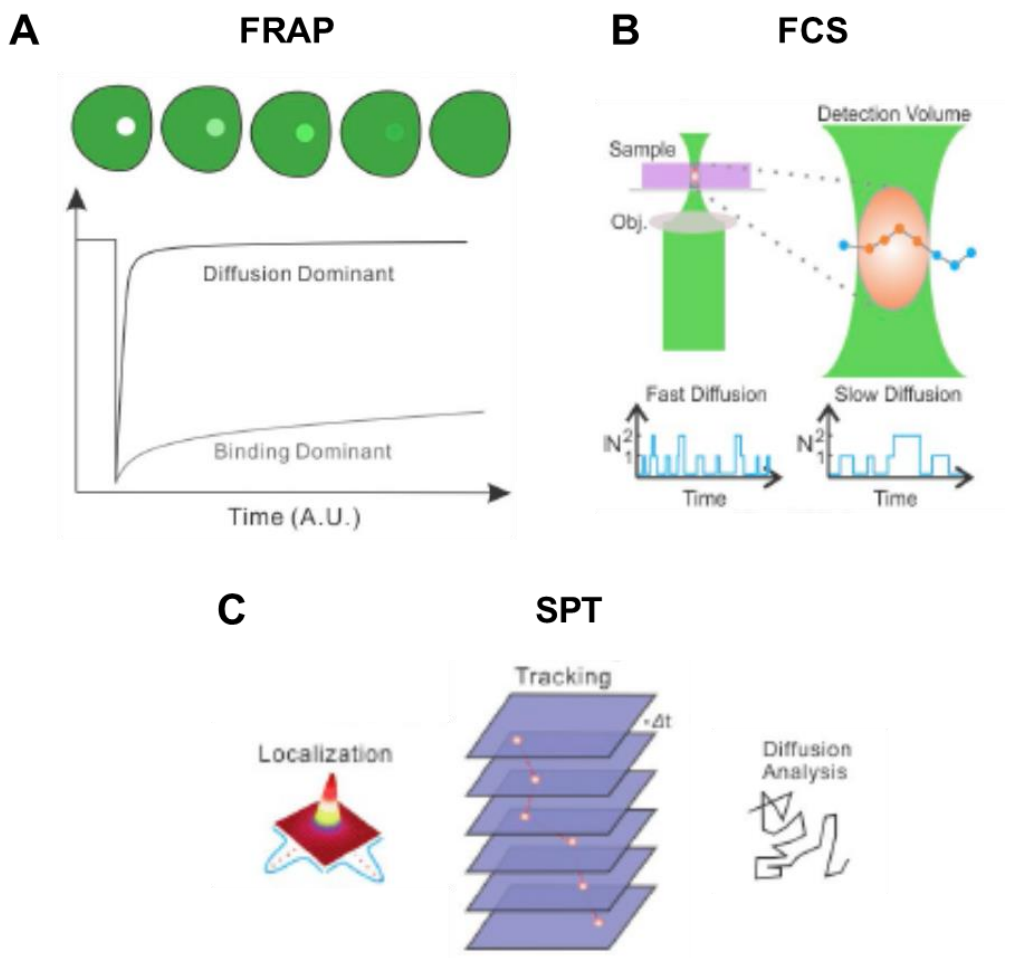


Figure 21: Different methods to measure and image TF dynamics in single cells. A *FRAP technique. A high intensity focused laser is bleaching an area inside of the cell. The fluorescence recovery can be measured within this area over time due to the dissociation of the bleached molecules and the diffusion of unbleached molecules into the region.* ***B*** *FCS method. The diffusion of single molecules is measured within a diffraction limited focal volume. Fast diffusion result in shorter temporal widths then slow diffusing molecules.* ***C*** *SPT technique. The position of a single fluorescent molecule is determined by localizing the centroid of the spot by Gaussian fitting. The position of the spot is afterwards followed across multiple frames to form single molecule trajectories which can be used to extract molecule dynamics. Adapted from (Liu and Tjian, 2018).*

2.1.4 Transcription factor dynamics

One of the most pressing unresolved question in understanding the dynamics of factors of the transcription machinery is how the molecules can navigate through the complex and packed environment of the nucleus to find their target site? It was shown in early FRAP studies that fluorescently tagged TFs have a highly rapid motion within the nucleus with diffusion coefficients ranging from 0.5 to 5 $\mu m^2 s^{-1}$ (Stenoien et al., 2001; Phair et al., 2004; Sprague et al., 2004). This diffusion behavior would make it possible for a TF to traverse and visit the full volume of a mammalian nucleus within a few minutes (Hager et al., 2009). Additionally, it was shown using artificial gene arrays that TFs have only short residence time at their genetic target site in the time scales of just a few seconds (McNally et al., 2000; Voss et al., 2011). Recent studies using SPT techniques confirmed these results by measuring the dynamics of the TFs Sox2 and Oct4 in the nucleus of living embryonic stem (ES) cells (Chen et al., 2014). It was shown that both TFs use a trial-and-error target search mechanism to find their binding site in the genome and that most of the molecules of Sox2 (97%) are in a stochastic motion in the nucleus, only colliding non-specifically with the DNA before they find their target site. Additionally, Sox2 also shows very short residence times on its target sites of around 12 seconds. All the results lead to a highly dynamic model for TF target site scanning and enhancer binding in which the TFs diffuse randomly through the nucleus where they bind occasionally the DNA at non-specific sites and shortly scan the DNA by local sliding, hopping or direct motor driven transport just to leave again until it finds its specific binding site (see Figure 22) (Phair et al., 2004; Hager et al., 2009). However, even there is the residence time only short which leads to a much more tunable on/off system for transcription, maybe to regulate the expression of the genes through the short interaction of TFs with the enhancer. In a recent study using a combination of FRAP and FCS, it was possible to show that also GTFs (TFIIB and TFIID) and co-activators (SAGA and ATAC) are highly dynamic with only transient associations with chromatin. Furthermore, it was shown that the ability of these complexes to interact with chromatin is dependent on the presence of histone H3K4 tri-methylation and therefore is regulated by active transcription (Vosnakis et al., 2017).

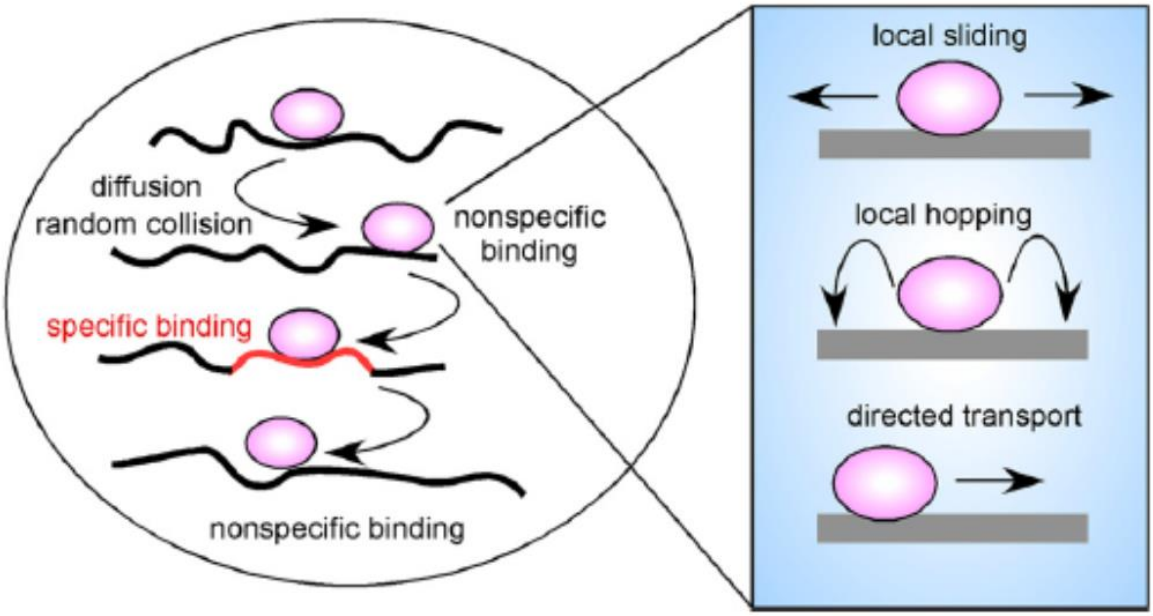


Figure 22: Types of TF movements in the nucleus of living cells. *A TF (purple) is diffusing randomly through the nucleus with transient interactions with the chromatin until a specific binding site is encountered in which prolonged binding occurs. The local motion at these binding sites can differ from sliding, local hopping or a directed movement of the TF on the chromatin. From (Hager et al., 2009).*

2.1.5 Transcription factor assembly *in vivo*

Another interesting question concerning *in vivo* imaging of transcription is where and how the transcription machinery is assembling within the nucleus or the cell? One subject within this field is the question how multiple TFs bind to an enhancer and if they do it in a random order or if a hierarchical order exists? Recent studies using single molecule analysis showed in ES cells that Sox2 is binding first to the target site before Oct4 is recruited (Chen et al., 2014). Furthermore, the deletion of a single Sox2/Oct4 composite site at a distal enhancer eliminates the chromatin accessibility for other TFs which indicates that a hierarchical mode of TF binding is present at least at this enhancer (Xie et al., 2017). Additionally, it was shown using lattice light sheet microscopy that the Sox2 stable binding sites form spatially restricted clusters in the nucleus which most likely correspond to enhancer clusters (Liu et al., 2014). The idea of clusters or so called “transcription factories” within the nucleus of fixed cells in which several enhancers, genes and proteins of the transcription machinery are concentrated was already postulated 25 years ago (Jackson et al., 1993). Later studies using early

fluorescence microscopy and electron microscopy described that genes are maybe transcribed in these transcription factories formed by clustered RNA Pol II molecules in the nucleus (Cook, 1999). However, live cell imaging studies using FRAP and super-resolution microscopy could not confirm that transcription is really happening within these factories as no stable chromatin bound RNA Pol II molecules could be detected within the clusters (Darzacq et al., 2007; Zhao et al., 2014). Another live-cell super-resolution microscopy study instead showed that RNA Pol II indeed forms subdiffraction-sized clusters within the nucleus but their lifetime is only a few seconds which is too fast for harboring transcribing RNA Pol II (Cisse et al., 2013). All these findings indicate that these RNA Pol II clusters exist in living cells but their function in transcription is still unknown. However, a recent study on the β-actin locus found that RNA Pol II cluster formation occurs right before mRNA production which suggests that these transcription factories are more likely involved in transcription initiation than in elongation (Cho et al., 2016). These studies indicate that the clusters in contrast of being RNA Pol II molecules on a gene transcribing it as “RNA Pol II trains” rather correspond to a high concentration of RNA Pol II molecules at the promoter to initiate transcription as long as the promoter is active. This would be consistent with a hypothesis in which TFs bind to already present enhancer clusters and due to their cooperative binding are inducing an increase of TF concentration at the enhancer clusters. This would favor the recruitment of chromatin remodelers, GTFs and RNA Pol II to the enhancer clusters which leads to the organization of RNA Pol II clusters or “transcription factories” for initiating transcription (Figure 23) (Liu and Tjian, 2018). However, even if co-localization between enhancer and RNA Pol II could be observed, it still needs to be proven if active transcription is really happening within these clusters (Liu et al., 2014).

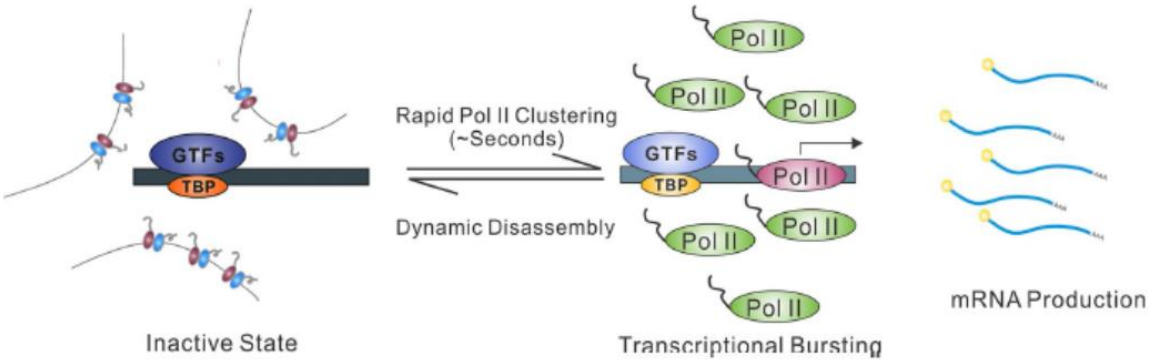


Figure 23: RNA Pol II clustering and transcription factories. *Live imaging data suggest that RNA Pol II is rapidly assembling and disassembling at the promoter after activation of transcription bursting. These resulting RNA Pol II clusters are mainly involved in transcription initiation as long as the promoter is still in the active state. From (Liu and Tjian, 2018).*

2.1.6 Liquid-liquid phase separation in transcription regulation

However, how are these clusters formed and why should they be advantageous for the cells to perform transcription? It was suggested that these clusters are formed by liquid-liquid phase separation in which proteins, DNA and other molecules self-organize in liquid-like droplets to form distinct compartments inside the nucleus. Several hypotheses are existing on how these liquid-like droplets are formed. A theoretical work suggested that the presence of TFs at the enhancer site could induce phase separation driven by the DNA itself like it was already described for nuclear bodies which are known to sequester target genes into specific microenvironments (Brown et al., 2008; Ching et al., 2013; Le Treut et al., 2016; Wang et al., 2016). However, clustering of cis-regulatory elements could not be confirmed until now at the single cell level. Another idea is that specific interactions of proteins with the DNA induce the generation of such phase separated compartments. It was shown that several sequence specific TFs contain simple repetitive and largely unstructured amino acid sequences (for example glycine- and proline-rich acidic repeats) that could serve as DNA-binding and activation domains to form highly dynamic phase separated compartments in the nuclei of living cells (Courey and Tjian, 1988; Patel et al., 2015; Shin and Brangwynne, 2017). These sequences are also known as low-complexity domains (LCDs) and were found to be present in many different TFs or proteins of the transcription machinery like TAF15, a member of the FET family of RNA binding

proteins (Bertolotti et al., 1996), or even the CTD of RNA Pol II (Kwon et al., 2013). Recently it was indicated that RNA Pol II CTD is able to perform phase separation and that these liquid-like droplets can incorporate further RNA Pol II molecules to form RNA Pol II clusters at active promoter. Furthermore, the ability of the CTD to form the droplets is dependent on its length. However, phosphorylation of the CTD by TFIIH and its kinase subunit CDK7 at Ser5 removes the RNA Pol II molecule from the droplet and the polymerase is able to engage into transcription (Boehning et al., 2018). This could indicate another transcription regulation step to be able to control the amount of RNA Pol II that can initiate at an active promoter but on the other side always having enough molecules present to induce transcription as often as needed. Another very recent study showed the implication of P-TEFb to generate phase separated compartments (Lu et al., 2018). As mentioned in a previous section, P-TEFb and more specifically its kinase subunit CDK9 is responsible for the phosphorylation of the RNA Pol II CTD at Ser2 to regulate successful elongation. However the P-TEFb complex consists of two subunits: CDK9 and CCNT1. In the study it was shown that deletion of the histidine-rich domain (HRD) of CCNT1 reduces the capacity of CDK9 to phosphorylate the CTD. Therefore, it was suggested that the HRD domain of P-TEFb is involved in liquid-like droplet phase separation which favors the interaction with RNA Pol II and this in turn is important to ensure phosphorylation of the CTD and successful transcription elongation. This could indicate another layer of transcriptional regulation which is dependent on successful phase separation. However, further studies need to be carried out to find out how big the impact of P-TEFb in the generation of phase separation really is. Nevertheless, these studies show how important phase separation could be to regulate transcription and maybe to achieve enhancer promoter contacts due to the high local concentration of TFs, GTFs and RNA Pol II within these liquid-like droplets. It could be possible that physical proximity is maybe more important than direct stable lock and key-like interactions.

2.2 Resolution revolution: Super-resolution microscopy techniques

Since the beginning of cell biological research has microscopy been an important tool to be able to understand different cellular functions like the distribution, interactions or modifications of various factors within the cells. Starting in the 17th century with Robert Hooke and Antoni van Leeuwenhoek and their studies of biological structures, microscopy developed over the centuries to improve imaging qualities and technical aspects. Among all these various improvements was the development of fluorescence microscopy. This technique had two important advantages as it enabled the specific labeling of the desired factor using fluorescent probes and it was possible to visualize these factors in living cells. However, fluorescence microscopy, in contrast to electron microscopy, was always restricted by limited optical resolution.

This restriction in resolution was due to the diffraction limit of light which was defined by Ernst Abbe and Lord Rayleigh in the 19th century (Abbe, 1873; Rayleigh, 1903). The resolution in an image is often defined by the distance that still allows to distinguish two objects without them overlapping. This can also be described as the point spread function (PSF) of the objects which defines how much a point-like object is spread out in the image and if the resolution is not high enough, than the PSFs of two objects will overlap. However, this explanation does not take into account that light acts as a diffracting wave. Therefore, the Abbe diffraction limit described that resolution can be defined by the wavelength of the light used to perform the imaging experiment as well as by the numerical aperture of the microscope objective. Thus, the highest resolution (or the best PSF) that can be achieved using the best setup with a large numerical aperture and perfect lenses is always diffraction limited and would be at around 200-300 nm in the lateral direction and 500-700 nm in the axial direction. Therefore, only cellular structures which are at least 200-300 nm apart from each other could be resolved with classical light microscopy. However, most of the processes and molecules inside the cells are much smaller in the range of 10-100 nm. Thankfully, over the last twenty years different super-resolution microscopy techniques were developed which are able to overcome the diffraction limit, to be able to resolve even single molecules inside living cells (Huang et al., 2008; Schermelleh et al., 2010). In the following section, three super-resolution microscopy techniques will be briefly described: single molecule localization microscopy (SMLM), stimulated emission

depletion (STED) microscopy and SIM. It is important to mention that the field of super-resolution microscopy is under constant development and more improved versions of the following techniques as well as combinations of different techniques are existing (Chang et al., 2016; Burri et al., 2017). However, this introduction will concentrate on the advantages and disadvantages of the three most commonly used and also commercially available techniques (see Figure 24).

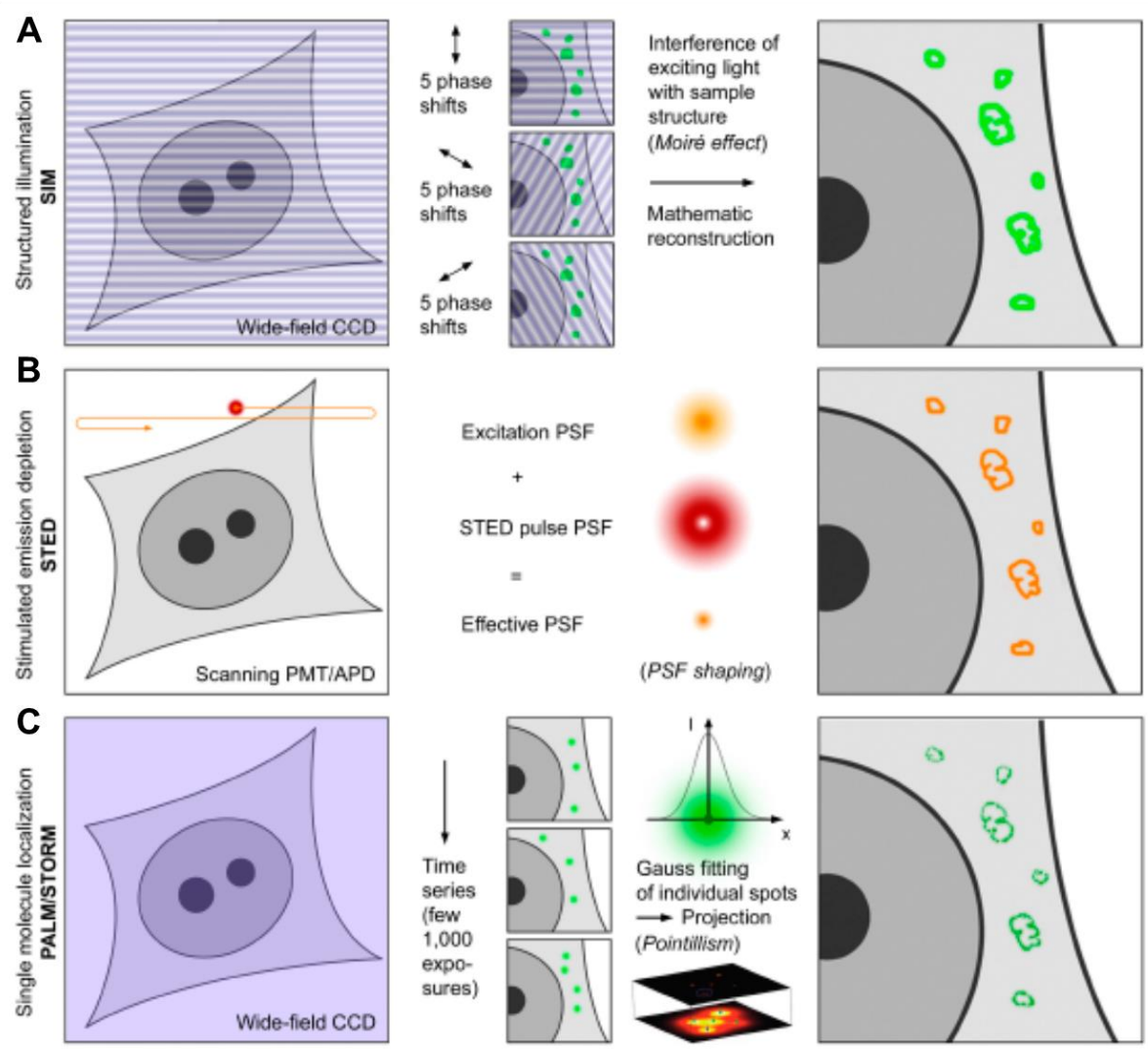


Figure 24: Legend on the next page →

Figure 24: Principles of the three main super-resolution imaging techniques. A *Structured Illumination microscopy (SIM). The sample is excited with a nonuniform wide-field illumination which is generated through laser light passing through an optical gating. This results in a stripe-shaped sinusoidal interference pattern which together with the sample information generates moiré fringes. Mathematical reconstruction of a total of 15 images per slice results in a high resolution image.* ***B*** *Stimulated emission depletion (STED) microscopy. In this technique, the sample is scanned with two different laser beams. The first laser beam is exciting the fluorophores whereas the second "donut shaped" laser beam is de-exciting a part of the signal back to the ground state. This modulation results in a PSF which is much smaller than the diffraction limit.* ***C*** *Single molecule localization microscopy (SMLM). This technique ensures that only a subset of fluorophores are excited at a given time point to be able to localize single fluorophores. This is achieved by using photactivatable proteins or photoswitchable dyes to activate the fluorescence at specific time points. Single molecule positions of thousands of single images are collected and reconstructed into one super-resolution image within the plane of focus. From (Schermelleh et al., 2010).*

2.2.1 Localization based super-resolution techniques

SMLM is general term for a group of super-resolution microscopy techniques which share the same principal in which the center of mass is located for every fluorescent molecule (PSF) in the image to be able to precisely fit their location in the picture. However, this was in the beginning only possible for samples with low amount of particles in close proximity (Bornfleth et al., 1998; Heilemann et al., 2002). This changed after the discovery of photoactivatable and photoswitchable proteins/dyes whose fluorescence can either be activated at a given time or switched to another wavelength by applying a fluorophore specific wavelength to the sample. Thus, by activating or switching these fluorophores stochastically over time made it possible to generate a localization map of single molecules inside the cells (Betzig et al., 2006; Hess et al., 2006; Rust et al., 2006). It is important to mention that only a few fluorophores are emitting at a specific time point, so to receive a full super-resolution image of all labeled molecules it is necessary to acquire many thousand images. These images are afterwards reconstructed to one super-resolution image. Depending on the labeling technique used, there are different SMLM techniques: photoactivatable proteins are used in photoactivated localization microscopy (PALM) and fluorescence photoactivation localization microscopy (FPALM) whereas photoswitchable proteins or

dyes are used in stochastic optical reconstruction microscopy (STORM). Many variations and improved versions of these techniques were developed over the years (Bornfleth et al., 1998; Egner et al., 2007; Fölling et al., 2008; Heilemann et al., 2008). The resolution of the resulting image is dependent on the amount of photons that were collected with more photons giving the better resolution. The resolution in the lateral direction is in the range of 10-20 nm.

The biggest advantage of the SMLM microscopy techniques is definitely the high resolution that can be achieved in contrast to all the other methods. Multicolor imaging using SMLM techniques has been reported but it highly depends on the distinct technique that is used as all SMLM techniques rely highly on the photophysical and photochemical properties of their fluorescent probes (Huang et al., 2008). STORM applications use photoswitchable probes where dye combinations with distinct excitation, emission and activation wavelengths are available (Bates et al., 2007). However, multicolor imaging using PALM and photoactivatable proteins is more challenging as the postactivation fluorescence of green emitting proteins is overlapping with the preactivation fluorescence emission of most red-emitting proteins. This can be overcome by using a reversible switching green fluorescent protein and a red fluorescent protein which can be activated only once, so that the green signal is acquired only after all red fluorescent proteins were imaged and bleached (Shroff et al., 2007; Shroff et al., 2008). 3D imaging is generally possible but as for multicolor imaging highly depends on the specific technique that is used as some of them can only achieve a few hundred nanometer z-stack due to technical restrictions. However, in combination with lattice light-sheet microscopy and new labeling approaches it was recently shown that multicolor imaging in a depth of 20 µm z-stack is feasible with the right setup (Legant et al., 2016). Another disadvantage of SMLM microscopy is the fact that the resulting raw images need to be reconstructed before a final super-resolution image is obtained which always induces the possibility to create reconstruction artifacts. As always, also the possibility for live imaging with SMLM is dependent on the distinct technique used. STORM and other stochastic methods using photoswitchable dyes need special imaging buffer containing oxidizing/reducing agents to induce the desired blinking of the dyes. However, these buffers are often toxic for the cells and therefore longer live acquisitions are not possible (Jones et al., 2011). New buffers are developed which try to use naturally present reducing agents

like thiol glutathione to overcome this problem (Klein et al., 2011; Benke et al., 2012). PALM application do not need these specific buffer to activate their proteins and is therefore better suitable for live tracking experiments (see also previous section). However, the long acquisition times of all SMLM techniques makes live imaging especially in 3D very challenging (see Table 3).

2.2.2 Stimulated emission depletion (STED) microscopy

STED microscopy is a super-resolution technique which modulates the PSF of the target molecules by performing controlled de-excitation of already excited fluorophores (Hell and Wichmann, 1994; Klar and Hell, 1999). In more detail, the fluorophores in the sample are first normally excited by the excitation laser. Afterwards, a second “donut” shaped laser with a stimulated emission depletion beam and zero intensity in the middle of the beam is scanning through the sample and is de-exciting the fluorophores which are slightly off the center of the PSF back to the ground state. This reduces the overall width of the PSF and therefore increases the resolution. The quality of the resulting image is dependent on the duration of the STED laser pulse, the correct timing of the pulse as well as on the quality of the zero intensity in the middle of the donut shaped laser (Klar et al., 2000; Dyba and Hell, 2002; Dyba et al., 2003). In contrast to the other super-resolution techniques is the resulting raw image directly a super-resolution image as no further reconstruction is needed. The resolution of the image is dependent on the size of the remaining fluorophores and can range from 50-70 nm in lateral direction.

Concerning the resolution power, STED can be seen as the intermediate technique between SMLM and 3D-SIM. Multicolor imaging, however, is quite challenging in STED microscopy as the fluorophore combination needs to be chosen carefully to ensure that the depletion laser is not exciting the other color (Donnert et al., 2007). 3D imaging is possible, but the labeling density in the sample needs therefore to be very high to counteract the bleaching that occurs during the long 3D acquisition of bigger z-stacks due to the interplay of the two lasers (Schmidt et al., 2008). One of the biggest advantages of STED microscopy is that the resulting raw image is directly the final super-resolution image and no further reconstruction of the data is needed. Live imaging is also difficult for longer acquisition periods as the two laser that are needed to perform STED are quite strong and therefore phototoxicity can affect cell viability.

However, live imaging approaches are possible as it was shown by a study in which the movement of individual synaptic vesicles could be observed in 2D in living neurons with a rate of 28 fps and a resolution of 60-80 nm (Westphal et al., 2008) (see Table 3).

2.2.3 3D Structured illumination (3D-SIM) microscopy

SIM is based on an adapted wide-field microscope which generates super-resolution images by illuminating the sample with a stripe-like or “structured” illumination pattern (Heintzmann and Cremer, 1999; Gustafsson, 2000). This pattern is generated through a movable optical gating and afterwards projected onto the sample via the objective. Several images of the sample are taken with different orientations of the illumination pattern and this combined with the fluorescent emission of the structures imaged in the sample are generating coarser interferences which are also known as moiré fringes. A mathematical reconstruction using computer algorithms and the 15 patterned raw images taken from every slice of the sample can generate a high resolution image of the fluorescently labeled structure in the cell. A 3D super-resolution image can be created by acquiring a z-stack using an extra excitation light modulation along the z-axis and a three-beam interference (Gustafsson et al., 2008; Schermelleh et al., 2008; Schermelleh et al., 2010). Thus, like this a twofold resolution gain beyond the classical diffraction limit can be achieved which corresponds to a resolution of around 110-130 nm in lateral direction and 250-300 nm in axial direction.

3D-SIM is achieving the lowest resolution of all three presented super-resolution techniques with only a twofold increase compared to the diffraction limit. However, the circumstance that the technique does not rely on specific photochemistry makes it one of the best techniques for multicolor imaging as all standard dyes and standard protocols can be used. Furthermore, 3D imaging is also easily possible with depth of up to several micrometer and due to the fact that 3D-SIM is a wide-field technique, the applied laser powers and consequently also the bleaching is much less compared to other techniques. It was possible, by using 3D-SIM, to visualize different structures and molecules of the nucleus including molecules from processes like transcription and replication in a multicolor three-dimensional image to study their overall nuclear organization (Markaki et al., 2010). Furthermore, it was also shown using 3D-SIM that the relative localization of polycomb repressive complex 2 (PRC2) and Xist RNA, two

molecules that were said to interact with each other to induce X chromosome silencing, are actually spatially separated from each other in the three-dimensional space (Cerase et al., 2014). Another recent study showed that 3D-SIM can be used to follow the fate of individual replication foci in 3D to the resolution of a single replicon (Chagin et al., 2016). However, one disadvantage of the technique is that the raw data has to be reconstructed before the final super-resolution image is obtained. Therefore, it is highly important that the labeling density is strong enough in the sample to avoid reconstruction artifacts in the final image. Software like SIMcheck were developed to test whether the resulting images contain artifacts and to give advice on how they can be avoided (Ball et al., 2015). 3D-SIM is also considered to be a powerful technique for live imaging due to the previously mentioned low laser power that is applied to the sample and therefore the low bleaching and phototoxicity that is occuring in the live sample. However, for now live imaging in 3D-SIM is still a challenging task as some technical improvements need to be made to ensure sample stability during acquisition and to accelerate image acquisition to be able to resolve faster processes (Schermelleh et al., 2010; Godin et al., 2014; Wegel et al., 2016) (see Table 3).

***Table 3**: Comparison of the three main super-resolution microscopy techniques.*

	SMLM	STED	SIM
Resolution	★	●	✖
Multicolor imaging	●	✖	★
3D imaging	✖	●	★
Image reconstruction	●	★	✖
Sample preparation	●	●	★
Cost and complexity	★	✖	●
Live imaging	●	●	★

Advantages and disadvantages of the three super-resolution techniques. Green star: good; yellow circle: medium; red cross: problematic. Adapted from (Wegel et al., 2016).

2.3 Fluorescent labeling strategies for imaging

2.3.1 Protein labeling strategies

Proteins inherit intrinsic fluorescence due to residues like tryptophan. However, this autofluorescence of proteins is very weak and not stable enough to perform real microscopic assays. Therefore, many techniques have been developed to label proteins with extrinsic labels like fluorescent proteins (FPs) or fluorescent dyes to be able to visualize them specifically under the microscope. However, every labeling strategy that is used nowadays has its advantages and disadvantages. Before a labeling strategy is selected, it is often important to think about what experiment needs to be performed to find out which properties the fluorescent labeling needs to have. In general, the ideal fluorescent label should be small, bright, stable and should not perturb the biological process that wants to be studied. Unfortunately, the perfect labeling technique does not exist, so compromises have to be made (Toseland, 2013).

2.3.1.1 Ectopic expression of fluorescent fusion proteins

The first fluorescent marker used was the green fluorescent protein (GFP) which was isolated from *Aequorea victoria* (SHIMOMURA et al., 1962; Chalfie et al., 1994). Since then, a whole collection of new fluorescent proteins have been developed with additional wavelengths, higher brightness, increased stability and specific imaging properties (Shu et al., 2006; Shaner et al., 2007; Lippincott-Schwartz and Patterson, 2008). However, certain FPs tend to oligomerize within the cells and it is generally said that FPs provide a lower quantum yield and photostability in comparison with fluorescent dyes even if new developed FPs try to overcome this obstacle. For imaging, these FPs were fused to the target protein either on the C- or N-terminus through recombinant cloning (Tsien, 1998). By introducing the plasmids coding for the fusion protein inside living cells, it was possible to express and study the target protein. One advantage was that the FPs were rather small (30 KDa) and therefore were said to show no effect on the behavior of the target protein. Additionally, due to the high amount of different FPs, multicolor imaging of several targets was possible. Like this, only the target protein was specifically labeled and the localization and dynamics of the protein could be measured. Another strategy is the fusion of the target protein to specific protein tags (like Halo- or SNAP-tag) instead of FPs (Los et al., 2008; Sun et

al., 2011). Afterwards, organic fluorophores can be added to the sample which covalently link to the protein tag. The organic fluorophores can enhance quantum yield and photostability, furthermore, a new fluorophore can be chosen for every acquisition. However, strong promoter were used to ensure a high expression of the fusion proteins which came with advantages but also with problems. The overexpression resulted into a high signal to noise ratio for imaging but the expression level is difficult to control and the high expression could also induce localization and functional defects of the target protein (Dean and Palmer, 2014). Furthermore, the most important disadvantage is that only an exogenously produced protein can be studied and not the endogenous counterpart. One strategy to overcome this problem was to specifically degrade the endogenous protein and replace it with the exogenously produced one (Darzacq et al., 2007).

2.3.1.2 Endogenous knock-in using the CRISPR/Cas9 technology

Another possibility to be able to label the endogenous protein is by inserting the fluorescent reporter sequence directly into the target protein gene by using genome editing. One such genome editing technology is the CRISPR/Cas9 system from archea and bacteria (Jinek et al., 2012). Briefly, by using target sequence specific guide RNAs (gRNA) and the Cas9 endonuclease it is possible to knock-in a specific sequence into the target locus (Mali et al., 2013; Qi et al., 2013). The result is a cell line which is producing the endogenous target protein with a FP or a protein tag. However, also this technique has its disadvantages. One would be that the technique is very time consuming and the efficiency for a homozygous knock-in is often very low. Furthermore, like for all fusion proteins is the overall quantum yield quite low in comparison with fluorescent dyes and it is very difficult to perform multicolor imaging of different targets as this would need to perform several knock-in with different tags. Nevertheless, it is a powerful technology which enables the imaging of endogenous proteins in living cells.

2.3.1.3 Antibody labeling strategies

Immunoglobulins or antibodies are of high importance in scientific research. They are used in many different techniques because of their ability to specifically bind a target protein. Therefore, it was logical that they will also be used for imaging. The most common use of antibodies for imaging is in immunofluorescence (IF). In this technique, cells are fixed using either organic solvents or aldehydes to preserve the cellular structures. Afterwards, the cells get permeabilized using nonionic detergents, so that the antibodies are able to get access to bind the target protein. After the application of a fluorescent secondary antibody which is binding to the target bound antibody, is the sample ready for imaging. Like this, it was possible to study the subcellular distribution of target molecules or co-localization of two different molecules. Thus, the advantages of the technique are that it easy to perform, the labeling is happening with fluorescent dyes and multicolor imaging is also easily possible. However, as all the labeling techniques, it also has its drawbacks. The biggest disadvantage is that it can be only applied on fixed cells. Furthermore, a validated and specific antibody for the target is needed. Additionally, it is known that the fixation and permeabilization procedure can induce artifacts (Schnell et al., 2012). It was shown that the protein localization, especially in the cytoplasm, can be altered by the fixation and permeabilization process. Therefore, live imaging controls should be always performed to ensure that the signal that is acquired is not just an artifact.

However, several studies tried to overcome these problems of fixation by introducing fluorescently labeled antibodies or antibody fragments (Fabs) directly into living cells. Like this, they could bind to their target protein inside the cell and could be tracked using different microscopy techniques. However, antibodies (150 KDa) are generally too big to cross the cell membrane, so different techniques were developed to transport them into the cells or by producing them directly inside the cells (Teng *et al.,* 2016). One strategy was the intracellular production of recombinant single-chain variable (scFv) fragment antibodies. Unfortunately, these “intrabodies” are often aggregating in the cells due to the reducing environment of the cytosol (Renaud et al., 2017). Previous studies developed methods like FabLEM in which Fabs are loaded into living cells using glass beads (Hayashi-Takanaka et al., 2009; Hayashi-Takanaka et al., 2011). Like this, it was possible to get new insights about the dynamics of specific histone modifications in living cells which is not possible with any genetically tagging

method but the technique suffered from low delivery efficiencies of the Fabs into the cells. Another recent study achieved the transport of fluorescent antibodies by creating pores in the cell membrane with a bacterial toxin called streptolysin O (Teng et al., 2016) but the cell viability decreased as the toxin and the additional membrane resealing step can be harmful for the cells. In addition, recently camelid derived nanobodies (VHHs) became quite popular because of their small size in contrast to conventional antibodies (15 kD). They can be cloned into plasmids and recombinant VHHs can be produced afterwards inside the cells (Rothbauer et al., 2006; Krah et al., 2016). The advantage is that due to their small size, they can diffuse freely in the whole cell to bind their target and cannot be affected by the reducing environment in the cytoplasm due to their missing disulfide bridges. However, as for the overexpression of FPs, it is difficult to control the expression of the VHHs in the cell which can lead to high background due to a large amount of unbound VHHs in the cell. Therefore, it can be difficult to distinguish between target-bound VHHs and free ones (Figure 25).

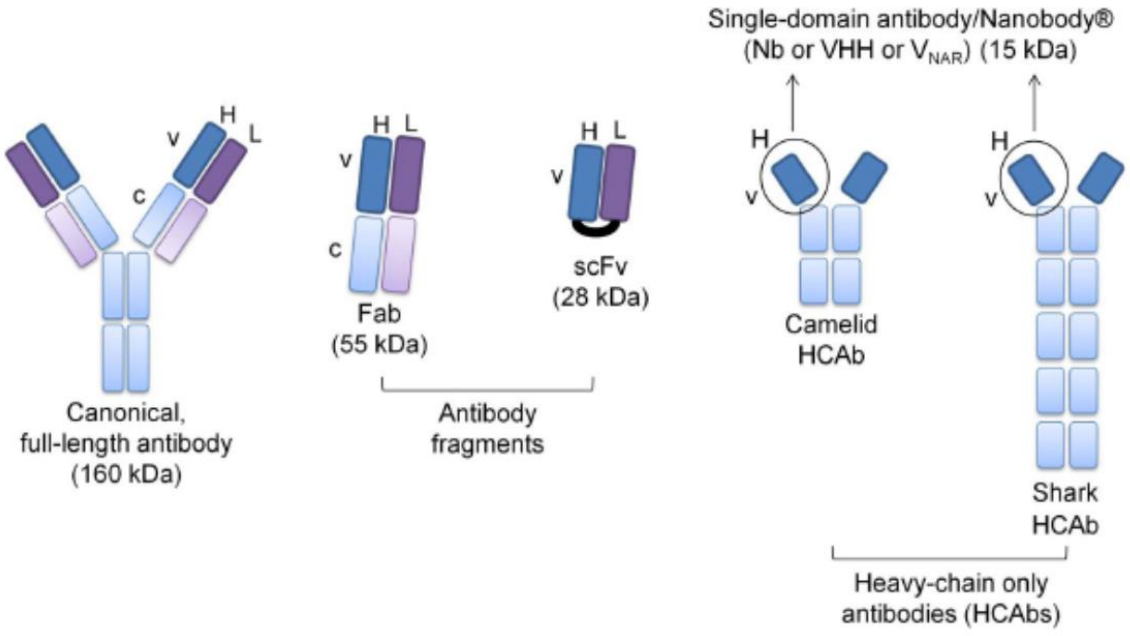


Figure 25: Legend on the next page →

Figure 25: Types of antibodies used for scientific research. *Traditional full-length antibodies have a size of around 150-160 kDa. In contrast antibody fragments like Fabs or scFv are already much smaller with sizes of around 55 kDa or 28 kDa respectively. The smallest antibodies are the single domain heavy chain only antibodies (VHHs; nanobodies) which are derived from camelids or sharks with only a size of around 15 kDa. From (Doshi et al., 2014).*

2.3.2 DNA and RNA labeling strategies

However, not only proteins are subject of fluorescent labeling. Also the labeling of chromatin and nascent RNA is of high importance to be able to localize specific genes in the nucleus or to validate the expression of genes. One of the most common way to label RNAs is by using single molecule mRNA fluorescence in situ hybridization (FISH) (Femino et al., 1998; Raj et al., 2008). In this technique complementary and fluorescently labeled probes are used to bind and label specific mRNA molecules in the cells which enables the quantitative analysis of the amount of certain mRNAs in a cell and therefore the expression of a specific gene. However this technique relies on the fixation of the sample to treat them with the fluorescent probes, so to be able to follow nascent transcription in living cells new labeling techniques were developed. One of these techniques is to use small single-stranded oligonucleotides called aptamers which work in a different way than the earlier mentioned FISH probes (Famulok et al., 2007). After binding to the target RNA, the aptamer undergoes a conformational change which allows for the binding of fluorophores to the aptamer. These fluorophores can be added to the medium and diffuse into the cells. Furthermore, they only emit light if they are binding to the RNA bound aptamer. Many different aptamers were developed over the past years with different specificities and imaging properties like “Spinach” or “Mango” (Paige et al., 2011; Autour et al., 2018). Another common method to follow nascent transcription in living cells is the labeling of RNAs with the help of reporter molecules. In this technique, a specific sequence is needed to be added to the studied gene which will give rise to a specific RNA sequence. This RNA will form a hairpin structure after expression which can be recognized by the fluorescently labeled reporter. The most known reporter is the MS2 coat protein (MCP), identified from bacteriophages and which can bind specifically to the MS2 hairpin structure of the mRNA (Bertrand et al., 1998). Like this the expression

of a specific gene can be followed in living cells by introducing the MS2 sequence into the gene and producing MCP-fluorescent fusion proteins inside the cell.

DNA labeling was generally performed after cell fixation by treating the cells with fluorescent dyes like DAPI or Hoechst (Latt and Wohlleb, 1975; Portugal and Waring, 1988). However, these molecules are labeling the whole chromatin and are therefore not suitable for the labeling of specific genes. Thus, as specific labeling strategies for living cells are not existing, the labeling of DNA is generally performed with the help of fluorescently labeled reporters. The most common used reporter system for gene labeling is the use of parts of the bacterial lactose operon. By adding a Lac operator before the gene of interest, the fluorescently labeled and exogenously produced LacI protein can bind to the Lac operator sequence in living cells to label the gene of interest. However, often several repeats of the operator need to be added before the gene, or a whole gene array needs to be established, to gain enough labeling density for successful imaging. More recently developed labeling techniques try to reduce the length of the target sequence for the reporter to go for single gene labeling. One of these studies uses a so called Anchor3 system to label specifically a single gene (Germier et al., 2017). After adding the Anch3 sequence directly before the gene of interest and the expression of the reporter protein OR3 (bacterial partition protein or ParB), the reporter is binding and accumulating/spreading at the Anch3 sequence which increases the labeling density to be able to visualize the single gene in living cells. Combination of the previously mentioned labeling techniques for a genetic locus and mRNA were also performed (Janicki et al., 2004; Rafalska-Metcalf et al., 2010). Like this it was possible to localize the gene array in the nucleus and during the same acquisition be able to measure the expression of the genes after activation due to the MS2 labeling system. Furthermore, the resulting cyan fluorescent protein could also be imaged in the cytoplasm (see Figure 26).

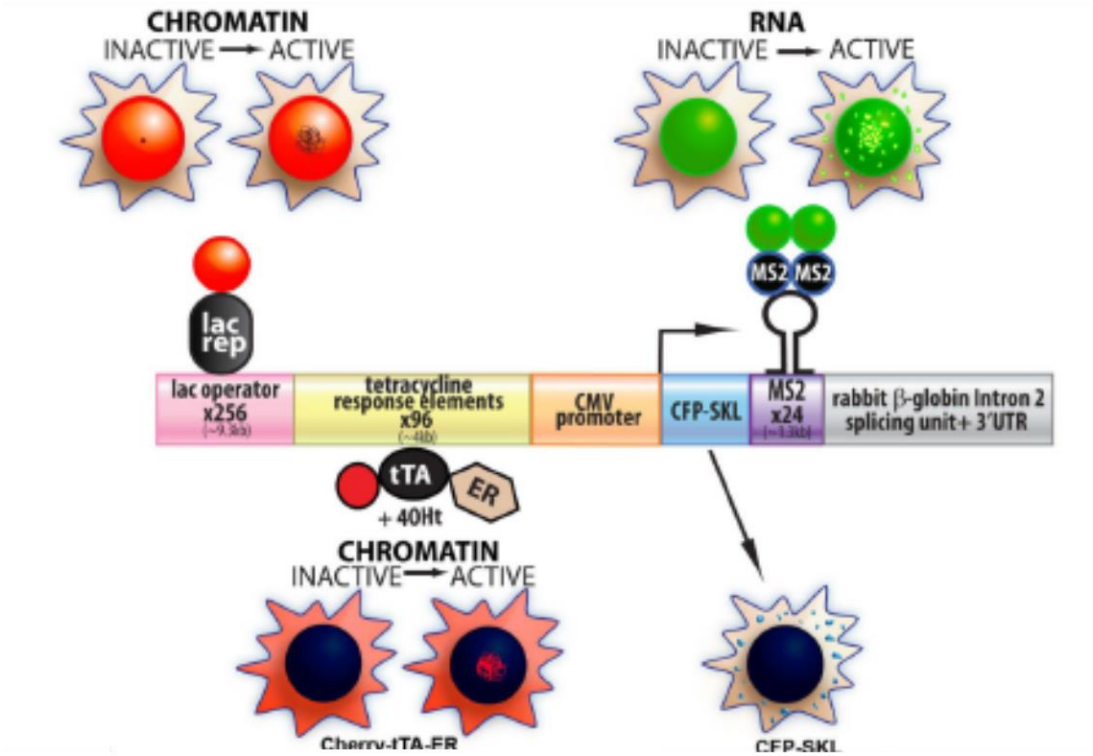


Figure 26: Single cell imaging system to follow transcription. *A 200 copy gene array was inserted into the genome of a cancer cell line (U2OS 2-6-3). This gene array includes different modules to image the transcription process in living cells. The gene array itself can be visualized by a fluorescently labeled lac repressor (lac rep) which binds to the lac operator. After transcription activation the gene array is de-condensing and the lac repressor spot is increasing. Transcription can be specifically activated by using the tetracycline transactivator (tTA) fused to a fluorophore (mCherry) and the estrogen receptor (ER) hormone binding domain. The Cherry-tTA-ER stays cytoplasmic until hydroxytamoxifen (4-OHT) is added to the cells which induces its entry into the nucleus and an accumulation at the gene array as well as activation of transcription. This leads to the transcription of the CFP-SKL gene and a MS2 repeat under the control of a CMV promoter. The resulting mRNA can be visualized by expressing the MS2-coating protein (MCP) in the cells which bind to the MS2 hairpins in the mRNA. The translated protein is a cyan fluorescent protein (CFP) fused to a peroxisome targeting signal (SKL) and can be visualized in the cytoplasm of the cells after transcription activation. Adapted from (Newhart and Janicki, 2014).*

AIMS

Aims of the work

Gene transcription in eukaryotic cells is controlled and regulated by a plethora of different proteins which preassemble in multiprotein complexes. In case of class II transcription, this process is controlled by several of these multisubunit factors and especially by RNA Pol II and different GTFs. However, it is important to note that most of the basic knowledge about transcription and transcription regulation derives from molecular biology, genetics and static binding experiments or by immunofluorescence (IF) experiments using fixed cells. Consequently, in living cells the knowledge about the dynamic movements, assembly or nuclear distribution of transcription factors involved in the subsequent steps of transcription is still limited. Thus, the possibility to visualize and track proteins in single living cells can give new insights about the natural behavior of that protein in the cell. The fluorescent labeling of the target proteins and complexes is in this case a crucial step to obtain reliable results. However, most of the proteins which are studied in fluorescence microscopy are overexpressed fluorescently tagged proteins (FTPs) which can behave differently to their endogenous counterparts. Furthermore, it is also well known that the function of transcription factors involved in chromatin-dependent processes is tightly linked to their interactions with diverse posttranslational modifications (PTMs) in the nuclear environment which cannot be visualized using FTPs. Thus, there is a demand for novel imaging and labeling approaches which enable the visualization of endogenously expressed proteins in single living cells.

Therefore the aims of my project were the following:

a) The development and implementation of a labeling technique to be able to visualize and track endogenous proteins and PTMs in living cells using intracellular antibodies and Fabs.

b) To study the assembly of multisubunit transcription factor complexes like TFIID *in vivo* to identify novel sub-modules.

c) To study the assembly and dynamics of endogenous RNA Pol II cluster in living cells to uncover which function they have for transcription.

d) The labeling of several endogenous transcription factors including RNA Pol II and different GTFs with intracellular antibodies to measure transcription dynamics in living cells.

RESULTS

Results

1. Cytoplasmic TAF2-TAF8-TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules (S. Trowitzsch, C. Viola, E. Scheer, S. Conic *et al.*; Nature Communications, 2015)

The GTF TFIID plays a central role in class II transcription by acting as a linker between cellular signaling events, regulatory DNA elements and the transcription machinery. Even if basal transcription at TATA-box containing promoters can be achieved without TFIID and only using TBP and several other GTFs *in vitro*, the full TFIID complex is important *in vivo* for the crosstalk with activators and the successful transcription from TATA-less promoters. Furthermore, several TFIID subunits are required to sense epigenetic modifications on nucleosomes and regulatory regions on gene promoter. However, while the general functions of the holo-TFIID complex as well as individual subunits of the complex are increasingly better understood, little is known about how the cell is assembling this essential multiprotein complex and if sub-complexes are existing with own functions in the nucleus or cytoplasm. The existence of a core TFIID complex in the nucleus of living cells provides evidence that holo-TFIID is assembled in a regulated fashion from stable sub-modules. Furthermore, several studies indicated the presence of a variety of different TFIID complexes with distinct subunit composition which are present in different cell types. Additionally, the dependence of certain TAFs to import other ones into the nucleus suggests that preassembled sub-modules already form in the cytoplasm of the cells. However, data confirming the existence of such cytoplasmic sub-complexes was lacking until now.

Therefore, in the following study we identified and characterized a novel endogenous TFIID sub-module in the cytoplasm of human cells consisting of TAF2, TAF8 and TAF10 by using different techniques like immunoprecipitation (IP), x-ray

crystallography or IF. Thus, the new TAF2-8-10 sub-complex was first characterized biochemically and structurally. It was shown that TAF8 plays a central role in the stabilization of the sub-module. Furthermore, it was demonstrated by x-ray crystallography that TAF8 and TAF10 are interacting with each other through a non-canonical histone fold domain. Additionally, a new interaction of TAF8 with TAF2 was reported mediated through several peptide motifs in the C-terminal domain of TAF8. Moreover, the formation of a putative nuclear import particle including the TAF2-8-10 complex and Importin α1 was described. Furthermore, we were able to show that the formation of this sub-complex is important for the incorporation of TAF2 into the core-TFIID complex in the nucleus.

Hence, to address aim **b)** of my project to study the assembly of transcription factors, I performed, together with Elisabeth Scheer, the *in vivo* characterization of the newly identified TAF2-8-10 sub-complex. The first question was if cytoplasmic TAF2, TAF8 and TAF10 are existing in the cytoplasm of the cells? Therefore, I performed IF experiments using polyclonal or monoclonal antibodies against TAF2, TAF8 or TAF10 and compared the cytoplasmic intensity of these IF signal with that of control experiments in which only secondary antibodies were used for the IF. The results showed an increased signal intensity in the cytoplasm of the TAF2, TAF8 and TAF10 IFs in comparison to the secondary antibody control. This suggested the presence of cytoplasmic TAF2, TAF8 and TAF10 in the cells. Another question was, if the localization of TAF2 is really dependent on TAF8 *in vivo*? Therefore, we knocked down endogenous TAF8 in HeLa cells using a RNA interference (RNAi) treatment of 48 hours and tested the change in the nuclear/cytoplasmic localization of TAF2 compared to control cells. I analyzed the cells using a specific IF staining of TAF2 and TAF8 and imaging by confocal microscopy. The short interfering RNA (siRNA) treatment of endogenous TAF8 led to the depletion of nuclear TAF2 and an enrichment of TAF2 in the cytoplasm. This indicated that the interaction of TAF8 with TAF2 is important for its import into the nucleus *in vivo*.

In conclusion, our results support the view of stable sub-complexes of TFIID in the nucleus and cytoplasm which can harbor potentially important functions on their own in the cell. It is also suggested that the formation of these sub-modules is important to regulate the import of transcription factor subunits into the nucleus and the stepwise assembly of the holo-TFIID complex. It is likely that such processes play important

roles in the regulation of assembly and activities of other multisubunit complexes involved in gene transcription.

These results were published on the 14th of January 2015 in *Nature Communications.*

Author's contributions:

Simon Trowitzsch – designed the study and interpreted the experiments. Produced and purified all proteins and characterized them together with Cristina Viola and Matthias Haffke. Determined the crystal structures. Wrote the manuscript.

Cristina Viola – designed the study and interpreted the experiments. Produced and purified all the proteins. Assisted in the writing of the manuscript.

Elisabeth Scheer – prepared the antibodies and performed the IPs. Helped in the characterization of the *in vivo* complex by performing knock down experiments.

Sascha Conic – Designed and interpreted the cell based experiments concerning the presence of cytoplasmic TAF2, TAF8 and TAF10 and the *in vivo* characterization of the TAF2-8-10 complex together with Laszlo Tora shown in the Figures 1d, 1e, 6c and Supplemental Figure 1b of the manuscript. Performed the IF staining and the confocal microscopy. Analyzed the microscopy data and performed the image processing. Assisted in the writing of the manuscript.

Virginie Chavant – performed MudPIT experiments. Assisted in the writing of the manuscript.

Marjorie Fournier – performed and analyzed MudPIT experiments. Assisted in the writing of the manuscript.

Ima-Obong Ebong and Carol V. Robinson – performed and analyzed the native mass spectrometry experiments. Assisted in the writing of the manuscript.

Gabor Papai, Christiane Schaffitzel and Patrick Schultz – prepared EM grids and collected and analyzed data and calculated reconstructions. Assisted in the writing of the manuscript.

Juan Zou and Juri Rappsilber – performed and analyzed CLMS experiments. Assisted in the writing of the manuscript.

Laszlo Tora – designed the study and interpreted the experiments. Wrote the manuscript.

Imre Berger - designed the study and interpreted the experiments. Wrote the manuscript.

nature COMMUNICATIONS

ARTICLE

Received 13 Aug 2014 | Accepted 2 Dec 2014 | Published 14 Jan 2015

DOI: 10.1038/ncomms7011 OPEN

Cytoplasmic TAF2–TAF8–TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules

Simon Trowitzsch[1,2], Cristina Viola[1,2], Elisabeth Scheer[3], Sascha Conic[3], Virginie Chavant[4], Marjorie Fournier[3], Gabor Papai[5], Ima-Obong Ebong[6], Christiane Schaffitzel[1,2], Juan Zou[7], Matthias Haffke[1,2], Juri Rappsilber[7,8], Carol V. Robinson[6], Patrick Schultz[5], Laszlo Tora[3] & Imre Berger[1,2,9]

General transcription factor TFIID is a cornerstone of RNA polymerase II transcription initiation in eukaryotic cells. How human TFIID—a megadalton-sized multiprotein complex composed of the TATA-binding protein (TBP) and 13 TBP-associated factors (TAFs)—assembles into a functional transcription factor is poorly understood. Here we describe a heterotrimeric TFIID subcomplex consisting of the TAF2, TAF8 and TAF10 proteins, which assembles in the cytoplasm. Using native mass spectrometry, we define the interactions between the TAFs and uncover a central role for TAF8 in nucleating the complex. X-ray crystallography reveals a non-canonical arrangement of the TAF8–TAF10 histone fold domains. TAF2 binds to multiple motifs within the TAF8 C-terminal region, and these interactions dictate TAF2 incorporation into a core-TFIID complex that exists in the nucleus. Our results provide evidence for a stepwise assembly pathway of nuclear holo-TFIID, regulated by nuclear import of preformed cytoplasmic submodules.

[1] European Molecular Biology Laboratory, Grenoble Outstation, 6 rue Jules Horowitz, 38042 Grenoble, France. [2] Unit for Virus Host-Cell Interactions, University Grenoble Alpes-EMBL-CNRS, 6 rue Jules Horowitz, 38042 Grenoble, France. [3] Cellular Signaling and Nuclear Dynamics Program, Institut de Génétique et de Biologie Moléculaire et Cellulaire, UMR 7104, INSERM U964, 1 rue Laurent Fries, 67404 Illkirch, France. [4] Proteomics Platform, Institut de Génétique et de Biologie Moléculaire et Cellulaire, UMR 7104, INSERM U964, 1 rue Laurent Fries, 67404 Illkirch, France. [5] Integrated Structural Biology Department, Institut de Génétique et de Biologie Moléculaire et Cellulaire, UMR 7104, INSERM U964, 1 rue Laurent Fries, 67404 Illkirch, France. [6] Chemistry Research Laboratory, University of Oxford, South Parks Road, Oxford OX1 3TA, UK. [7] Wellcome Trust Centre for Cell Biology, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JR, UK. [8] Institute of Bioanalytics, Department of Biotechnology, Technische Universität Berlin, 13353 Berlin, Germany. [9] School of Biochemistry, Bristol University, Bristol BS8 1TD, UK. Correspondence and requests for materials should be addressed to L.T. (email: laszlo.tora@igbmc.fr) or to I.B. (email: iberger@embl.fr).

Eukaryotic class II gene transcription is controlled by a plethora of proteins, which are preassembled in large multiprotein complexes, including RNA polymerase II, Mediator and the general transcription factors (GTFs)[1]. The sequential nucleation of GTFs and Mediator on core promoter DNA initiates regulated class II gene transcription[2]. The GTF TFIID plays a central role in this process by linking cellular signalling events with regulatory DNA elements and the transcription machinery[3]. Although a basal transcription system supporting initiation of transcription from TATA-box-containing promoters can be reconstituted with TATA-binding protein (TBP), TFIIA, TFIIB, TFIIE, TFIIF and TFIIH *in vitro*, TFIID is additionally required to respond to activators and for efficient transcription from TATA-less promoters[4,5]. In mammalian cells, most of the expressed protein-coding gene promoters are occupied by TFIID and loss of TFIID components leads to embryonic lethality[6–9]. TFIID subunits are implicated in crosstalk with epigenetic modifications on nucleosomes and regulatory DNA elements in promoter regions[10,11]. Structural analysis of TFIID by cryo-electron microscopy revealed the overall architecture of TFIID and provided important insights into subunit assembly and promoter recognition at low to medium resolution[12–16].

Canonical human TFIID consists of TBP and 13 TBP-associated factors (TAFs)[17]. Other non-canonical TFIID and TAF-containing complexes have been identified recently with key roles during spermatogenesis and stem cell development[18–20]. A central scaffold of canonical TFIID comprises two copies each of TAF4, 5, 6, 9 and 12, which were shown to form a symmetric core[12,21]. This core–TFIID complex was first identified in *Drosophila melanogaster* nuclei[21]. TAF3, 4, 6, 8, 9, 10, 11, 12 and 13 contain histone fold domains (HFDs), which stabilize discrete heterodimers (TAF3–10, TAF4–12, TAF6–9, TAF8–10 and TAF11–13) (refs 22–25). Among these HFD pairs, the TAF8–10 heterodimer plays a key role in the TFIID assembly pathway, is critical for the integrity of holo–TFIID and also fulfills essential functions in early embryonic development[6,8,26,27]. Binding of TAF8–10 to core–TFIID triggers a transition from symmetry to asymmetry, which was proposed to prime the recruitment of TAF1, 2, 3, 7, 11, 13 and TBP to complete holo–TFIID[12].

Evidence from genetic and biochemical studies showed that knockout of the TAF10 gene leads to impairment of mature TFIID assembly in F9 EC cells and to dissociation of TFIID in hepatocytes[6,26,27]. Biochemical data suggested that TAF8 and TAF10 interact strongly and specifically with each other via their HFDs[28]. Identification of human TAF8 uncovered high sequence similarities with the *Drosophila* protein PRODOS and the mouse TBN protein[8,28,29]. Mouse embryos carrying a mutation in TBN develop normally to the blastocyst stage but fail to develop further due to the lack of inner cell mass cells[8]. Interestingly, the same phenotype was also found in TAF10-knockout mice strongly suggesting that TAF8 and TAF10 are both involved in controlling embryonic development at similar stages[6]. The importance of this cooperative activity of TAF8 and TAF10 is supported by nuclear import assays, which showed that the transport of TAF10 from the cytoplasm to the nucleus depends on the nuclear localization signal (NLS) found at the carboxyl-terminal (C-terminal) end of TAF8 (ref. 30).

Human TAF2 (originally called either CIF150 or $TAF_{II}150$) has been previously described as an essential cofactor for TFIID-dependent transcription from promoters with initiator (Inr)-containing promoter elements[31–33]. Later it was suggested that a trimeric TBP–TAF1–TAF2 complex is minimally required for efficient utilization of the Inr and downstream promoter elements[11]. TFIID complexes containing or lacking TAF2 have been described[31,34] further suggesting that different types of TFIID complexes may exist in human cell nuclei. Recently, mutations in the TAF2-coding gene were shown to be associated with various neurological disorders[35,36]. Human TAF2 is predicted to adopt an aminopeptidase-like fold with an additional C-terminal unstructured region. Localization studies using immunopurified TFIID showed that TAF2 is an integral part of the central lobe of the holo-complex[13].

While general functions of individual TFIID subunits and the holo-complex are increasingly better understood, very little is known to date about how the cell assembles this essential multiprotein complex. The existence of physiological core–TFIID in the nucleus, containing a subset of TAFs, provides evidence that the holo-complex may be assembled in a regulated manner from stable, preformed partial TFIID subassemblies. The dependence of some of the TAFs on each other for nuclear import and the critical role of the TAF8–10 pair in functional remodelling of core–TFIID imply that discrete submodules preassemble also in the cytoplasm of cells. However, direct evidence for the presence of subassemblies in the cytoplasm is lacking to date.

By immunoprecipitating TAF-containing complexes from different cellular compartments, we identify a novel endogenous TFIID subcomplex formed by TAF2, 8 and 10 in the cytoplasm of human cells. We dissect cytoplasmic TAF2–8–10 biochemically and structurally. We elucidate the interactions that stabilize the complex and reveal a central role of TAF8 in its nucleation. By X-ray crystallography, we demonstrate a non-canonical histone-fold domain pair arrangement between TAF8 and TAF10. We report a novel interaction between TAF8 and TAF2, mediated by multiple peptide motifs in the TAF8 C-terminal region. Moreover, we describe the formation of a putative nuclear import particle comprising the TAF2–8–10 complex and Importin α1. Furthermore, we demonstrate that the TAF2–TAF8 interaction is not only crucial for formation of the cytoplasmic TAF2–8–10 complex, but also dictates incorporation of TAF2 into a physiological core–TFIID complex that exists in the nucleus.

Results

An endogenous cytoplasmic TAF2–8–10 complex. With the objective to better understand human TFIID assembly and in particular the incorporation of TAFs into holo–TFIID, we carried out immunoprecipitations from HeLa cell cytoplasmic and nuclear extracts. To test the role of TAF2 in the assembly process, we raised a polyclonal antibody using highly purified recombinant human TAF2 protein for the immunization procedure. We ascertained specificity of the purified antibody against recombinant TAF2 and endogenous TFIID by western blotting (Fig. 1a and Supplementary Fig. 1a). Using this antibody, we carried out co-immunoprecipitation experiments of endogenous TAF2 from the cytoplasm, where TAF2 is synthesized *de novo*, and from nuclear extracts, where TAF2 likely functions in the context of TFIID. To identify proteins that co-precipitated with TAF2 we subjected the immunoprecipitated samples to proteomics analysis by using the multidimensional protein identification technology (MudPIT). MudPIT analysis of proteins co-precipitated with TAF2 from the nuclear fraction revealed the full set of TFIID components (Fig. 1b and Supplementary Table 1). We observed differences in abundance of the individual TAFs, which may argue for the presence of distinct TAF2-associated TAF or TFIID-like complexes in the nucleus. Strikingly, MudPIT analysis of TAF2-associated proteins from the cytoplasmic fraction identified only TAF8 and TAF10, whereas none of the other TAFs could be detected (Fig. 1b,c). We confirmed the presence of TAF2, TAF8 and TAF10 in the cytoplasm of HeLa cells by immunofluorescence experiments (Fig. 1d,e and Supplementary Fig. 1b).

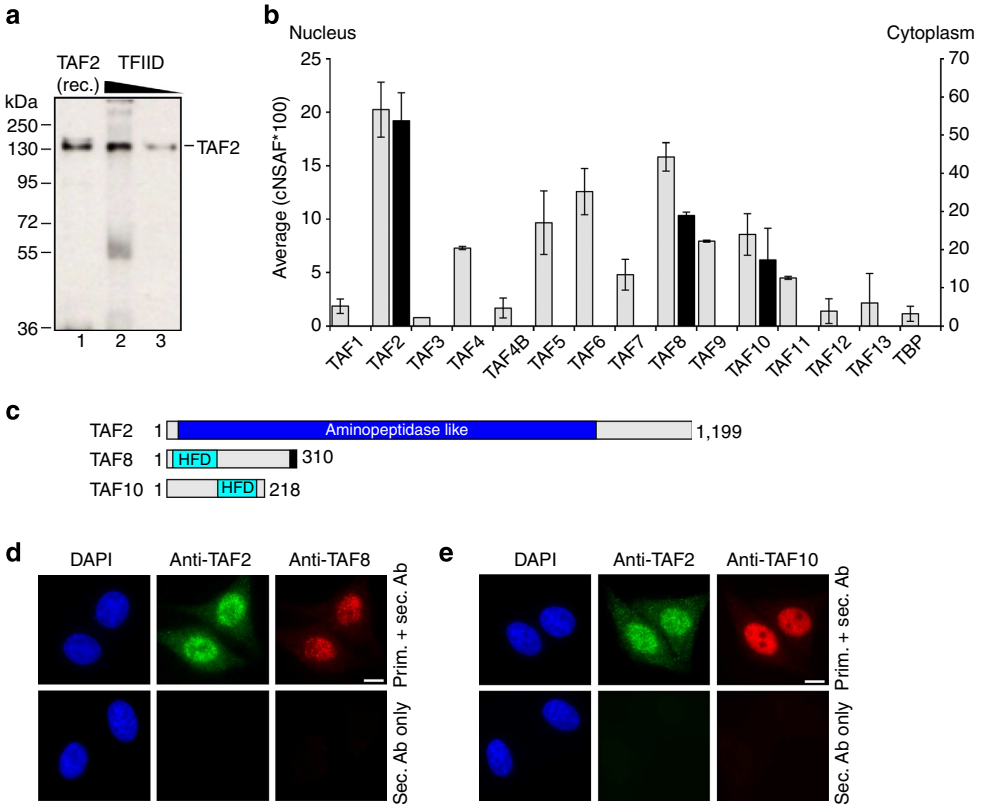


Figure 1 | A TAF2–8–10 complex exists in the cytoplasm. (**a**) Purified, polyclonal anti-TAF2 antibodies specifically recognize recombinant and endogenous TAF2. Recombinant (rec.) purified TAF2 (10 ng, lane 1) and immunopurified TFIID (300 and 150 ng; lanes 2, 3) were loaded on an 8% SDS-PAGE, blotted and analysed by western blot assay. Protein size markers are indicated. (**b**) Abundances of individual proteins co-immunoprecipitated from nuclear or cytoplasmic HeLa cell extracts (grey or black bars, respectively) using purified polyclonal anti-TAF2 antibodies were compared in units of normalized spectral abundance factors (NSAFs). Each column is the average of two independent experiments and error bars represent range of the data. (**c**) Domain organization of TAF2, TAF8 and TAF10 in a schematic view. Grey rectangles indicate predicted, unstructured regions. The NLS of TAF8 is shown as a black bar. Numbers indicate first and last amino acids in each protein. (**d**) Immunofluorescence microscopy of HeLa cells. Nuclei are visualized by 4′,6-diamidino-2-phenylindole (DAPI) staining (blue). TAF2 is displayed in green and TAF8 in red. The bottom panel shows images of control cells, which were treated with secondary antibodies only. Scale bar, 10 μm. (**e**) Immunofluorescence microscopy of HeLa cells as in **d**, but displaying TAF2 (green) and TAF10 (red).

These data suggest that a unique endogenous TAF2–TAF8–TAF10-containing TFIID building block exists in the cytoplasm.

TAF8 nucleates the TAF2–8–10 complex. To further analyse this TAF2–TAF8–TAF10 complex, we used highly purified recombinant human TAF2, TAF8 and TAF10 to reconstitute TAF2–8–10 *in vitro*. We produced recombinant TAF2 and the TAF8–10 pair separately in insect cells and tested complex formation by size-exclusion chromatography (SEC) experiments. SEC of a stoichiometric mixture of TAF2 and TAF8–10 showed a clear peak shift in retention volume towards earlier fractions as compared with the individual components (Fig. 2a). Analysis of the chromatographic fractions by SDS-polyacrylamide gel electrophoresis (SDS–PAGE) shows that all three polypeptides co-elute in the same fractions (Fig. 2a). We observed unusually high molecular weight estimates for the components TAF2 and TAF8–10, and also for the complete TAF2–8–10 complex, which exceed the calculated molecular weights of the proteins. These high estimates can be due to either oligomerization or elongated shapes of the specimens analysed. We therefore determined the oligomeric states of purified TAF2, TAF8–10 and the TAF2–8–10 complex by analytical ultracentrifugation sedimentation velocity and native mass spectrometry (MS) experiments. Sedimentation coefficients of 4.3 S, 2.3 S and 4.9 S were obtained for TAF2, TAF8–10 and TAF2–8–10, respectively (Fig. 2b). Continuous size-distribution analyses returned best-fit molecular weights of 140, 52 and 200 kDa. These values are in good agreement with monomeric TAF2, heterodimeric TAF8–10 and heterotrimeric TAF2–8–10 complexes, with subunit stoichiometries of 1:1 and 1:1:1 in case of the complexes.

Analysis of TAF2–8–10 by native MS revealed a predominant complex with an average molecular mass of 195,797 Da corresponding to a TAF2–8–10 heterotrimer containing one copy of each protein (Supplementary Fig. 2a and Supplementary Table 2). We subjected the TAF2–8–10 complex to collision-induced dissociation (CID) experiments in the mass spectrometer to probe for subunit interactions[37]. The resulting spectra reveal dissociation of the trimeric complex into TAF2–8 and TAF10 submodules (Fig. 2c and Supplementary Table 2). Notably, TAF2 and TAF10 do not interact under the conditions studied, since we did not observe a TAF2-10 species (Fig. 2c). We conclude from these data that TAF2, 8 and 10 assemble as a heterotrimeric complex with a 1:1:1 stoichiometry and that the complex is nucleated by TAF8 and stabilized by distinct TAF2–8 and TAF8–10 interactions.

TAF8 and TAF10 adopt a non-canonical histone fold dimer. We next dissected the interactions identified by CID. First, we determined the X-ray crystal structure of the TAF8–10 complex. Previous GST pull-down experiments suggested that the interaction between TAF8 and TAF10 is mediated by their HFDs, which are present in the amino-terminal (N-terminal) half of

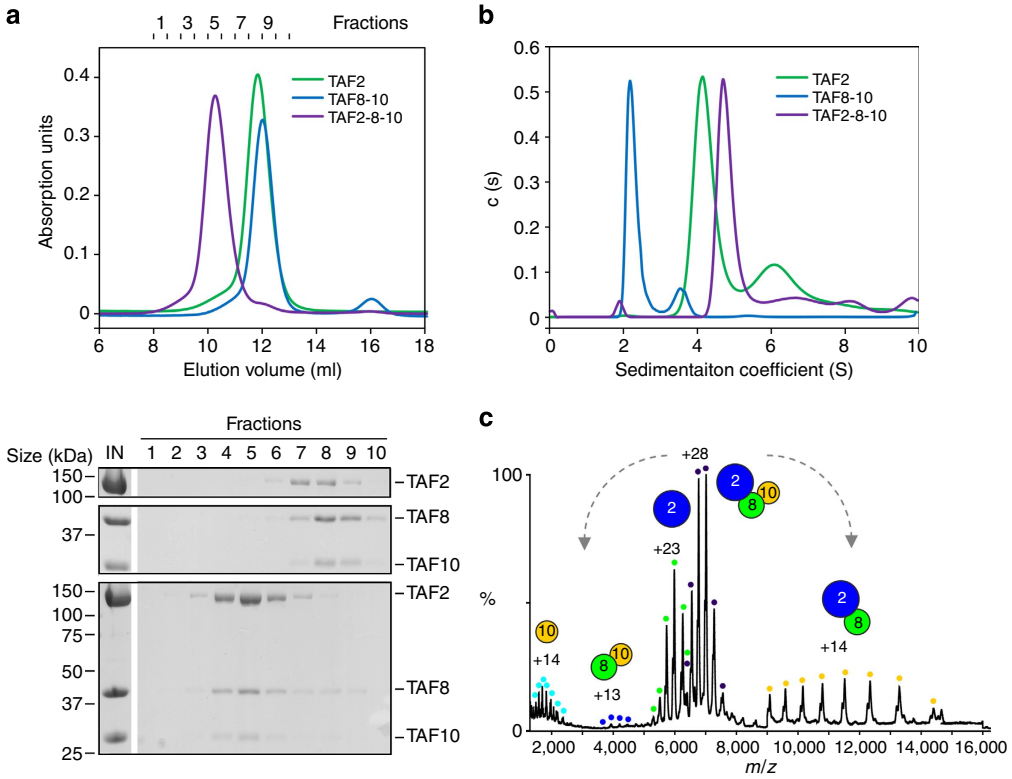


Figure 2 | Recombinant TAF2–8–10 complex. (**a**) TAF2, the TAF8–10 pair and a mixture of TAF2–8–10 were analysed by SEC. Elution profiles of TAF2 (green), TAF8–10 (blue) and TAF2–8–10 (purple) are plotted in relative absorption units at 280 nm versus elution volume (top). Fractions are numbered (top of graph). SDS-PAGE analyses of the eluted samples are shown (below). Molecular masses of protein standards are indicated on the left of gel sections. Protein denominations are shown on the right. First lane shows the SEC input (IN). (**b**) Absorbance c(s) profiles from sedimentation velocity analytical ultracentrifugation experiments are plotted for TAF2 (green), TAF8–10 (blue) and TAF2–8–10 (purple). (**c**) Mass spectrum of TAF2–8–10 complex electrosprayed from an aqueous ammonium acetate solution under high collision energy for subunit dissociation. The MS spectrum reveals peaks with corresponding masses for a TAF8–10 dimer (blue dots), TAF2 subunit (green dots) and a predominant TAF2–8–10 complex (purple dots) centred at 4,000, 6,000 and 7,500 *m/z*, respectively. At 12,000 *m/z* is a TAF2–8 dimer (yellow dots) resulting from the dissociation of the TAF10 subunit (light blue dots) from the intact TAF2–8–10 complex. Proteins and protein complexes are shown schematically as coloured circles.

TAF8 and the C-terminal half of TAF10 (refs 28,30). We co-expressed and purified full-length TAF8–10 complex in insect cells from a polyprotein construct[38], subjected the complex to limited proteolysis and defined the core complex to TAF8 residues 1–134 and TAF10 residues 98–218 (hereafter referred to as TAF8ΔC and TAF10ΔN, respectively; Supplementary Fig. 3a).

We prepared this TAF8ΔC–TAF10ΔN core complex, but only obtained crystals diffracting X-rays to 5–6 Å resolution. We therefore tested various N- and C-terminal deletion constructs of the two proteins in crystallization experiments. A complex of TAF8–10 comprising TAF8 residues 25–120 and TAF10 residues 112–212 yielded crystals, which diffracted incident X-rays to 1.9 Å resolution (Supplementary Fig. 3b). We determined the structure of this complex by the Sulfur-SAD method and refined the model to a crystallographic R value of 20.5% and a free R factor of 23.7% with excellent stereochemistry (Table 1). The final model includes TAF8 residues 28–120 and TAF10 residues 113–212 with the exception of a flexible loop in TAF10 comprising residues 178–191.

The crystal structure of the TAF8–10 complex reveals that the two proteins adopt atypical HFDs with three central α helices flanked by additional N- and C-terminal α helices (Fig. 3a). In our structure, TAF8 wraps entirely around the α2 helix of TAF10 markedly enveloping its interaction partner (Fig. 3a). Complex formation buries 2212.3 Å^2 with predominantly hydrophobic intermolecular contacts. As observed in other HFD interactions, the two opposing aromatic residues Y68 of TAF8 and F162 of TAF10 at the crossover of the α2 helices contact each other via hydrophobic stacking interactions and categorize the complex to the H3/H4 family of HFD-containing proteins (Fig. 3b and Supplementary Fig. 3c)[39]. The additional N-terminal α-helix of TAF10, αN and the C-terminal α-helix of TAF8, αC, contact each other on one side of the HFD in a head-to-tail fashion and significantly stabilize the complex by hydrophobic interactions centred on F119 of TAF10 (Fig. 3c).

Interestingly, the proteins TAF8 and TAF10 have similar L1 loop geometries, which are not found in other structures of related HFD-containing TAFs (Fig. 3d,e)[22,25]. In both proteins, a phenylalanine of loop L1 (F50 in TAF8 and F144 in TAF10) is embedded in a composite, hydrophobic cavity mainly formed by residues from helices α1/α2 of one protomer and helices α2/α3 of the other (Fig. 3d,e). The amino acids forming this hydrophobic cavity are remarkably similar in TAF8 and TAF10, suggesting an evolutionary interrelation between the two proteins (Supplementary Fig. 3d). To test the functionality of the TAF10 HFD and chimeric mutants thereof, we performed complementation assays in TAF10 null mouse F9 cells[40]. Interestingly, the human TAF10 HFD (residues 116–218) is fully functional in the complementation assay, whereas chimeric constructs, in which either the N-terminal region of TAF10 (residues 116–150) or the C-terminal region (residues 151–218) was substituted by sequences of the yeast TAF10 homologue, were not functional (Supplementary Fig. 3f,g).

Table 1 | X-ray data collection and refinement statistics.

	TAF8–10 Native	TAF8–10 S-SAD	Importin α1/TAF8-NLS Native
Data collection			
Space group	$P3_121$	$P3_121$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	51.32, 51.32, 144.40	51.30, 51.30, 144.70	54.27, 77.72, 128.57
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 90
Wavelength	0.98011	1.90745	0.93340
Resolution (Å)	44.44–1.91 (1.98–1.91)*	48.23–2.61	49.54–1.75 (1.81–1.75)
R_{merge}	2.92 (87.98)	2.00	5.00 (77.45)
$I/\sigma I$	23.4 (1.68)	48.43	19.18 (2.03)
Completeness (%)	93.91 (58.53)	100.00	99.60 (99.14)
Redundancy	4.80 (4.30)	20.91	4.60 (4.60)
Refinement			
Resolution (Å)	44.44–1.91		49.54–1.75
No. of reflections	16,793 (1,630)	12,943	55,423 (5,436)
R_{work}/R_{free}	20.5 (33.1)/23.7 (35.5)		15.3 (23.9)/18.0 (27.2)
No. of atoms	1,474		3,877
Protein	1,404		3,366
Ligand/ion	7		48
Water	63		463
B-factors			
Protein	58.2		30.9
Ligand/ion	59.9		59.6
Water	51.9		44.8
R.m.s. deviations			
Bond lengths (Å)	0.003		0.008
Bond angles (°)	0.681		1.158

R.m.s., root mean squared.
*Values in parentheses are for highest-resolution shell.

Primary sequence comparison with two other TAF10-interacting proteins, TAF3 and human SPT7L, shows that similar residues can be also found in their HFDs[30], arguing for a conserved binding mode of these proteins known to interact with TAF10 (Supplementary Fig. 3d). Our structure underscores that HFDs in TAFs can adopt a variety of conformations, which may differ significantly from the canonical histone pairs found in the nucleosome.

HFDs of TAF8 and TAF10 are dispensable for TAF2 binding. We next analysed the physical interactions between TAF2 and the TAF8–10 heterodimer. We first tested the effects of deleting the intrinsically unstructured regions of TAF8 and TAF10 on TAF2 binding in pull-down assays with purified proteins. As a control, full-length TAF8–10 was co-precipitated with TAF2 tagged with maltose-binding protein (MBP; Fig. 3f). Truncation of the N-terminal region of TAF10 (TAF8–TAF10ΔN, TAF10 residues 98–218) did not change the binding properties and still co-precipitated with MBP–TAF2. In contrast, a truncated complex of TAF8–10, in which the flexible C-terminal region of TAF8 was deleted (TAF8ΔC–TAF10, TAF8 residues 1–134), did not co-precipitate with MBP–TAF2 suggesting that the region that mediates binding to TAF2 resides in the C-terminal, low-complexity tail of TAF8 (Fig. 3f). We confirmed the interaction between TAF2 and the C-terminal tail of TAF8 by SEC. We utilized full-length TAF2 and a fusion protein of MBP with residues 105–310 of TAF8 and evidenced complex formation (Supplementary Fig. 4). These results are consistent with the CID data in native MS, which showed that only TAF8, and not TAF10, is directly interacting with TAF2.

TAF2 recognizes short motives in the TAF8 C-terminal region. We characterized the TAF2–TAF8 interaction further by means of a peptide array. We monitored the binding of His-tagged TAF2 to peptide arrays covering residues 105–310 of TAF8 (Fig. 4a). Densitometric analysis of the arrays indicated that TAF2-binding clusters around four distinct regions; a short N-terminal region I covering TAF8 residues 105–125, a less well-defined region II including residues 147–202 and regions III and IV spanning residues 207–238 and 282–310, respectively (Fig. 4a).

We next analysed the individual contributions of these four TAF8 regions to TAF2 binding by surface plasmon resonance (SPR) experiments. We generated N- and C-terminal deletion constructs of TAF8 and fused them to MBP (Fig. 4b). We monitored the association and dissociation phases of the MBP–TAF8 truncations on TAF2-charged sensor chips and compared binding kinetics at identical analyte concentrations. An MBP–TAF8-fusion construct spanning the entire C-terminal region (TAF8 residues 105–310) showed a maximal association level of about 85 response units (RU) with fast on and off rates (Fig. 4b). A shorter MBP-fusion protein lacking region I (TAF8 residues 141–310) showed similar kinetics but a reduced maximal association level of ~40 RU (Fig. 4b). MBP-fusion constructs with deleted regions I and II or IV (TAF8 residues 200–310 or 105–260, respectively) hardly interacted with immobilized TAF2 showing maximal association levels of less than 10 RUs (Fig. 4b). These data indicate that all four TAF2-interacting regions of TAF8 contribute cooperatively to the binding to TAF2.

On the basis of our peptide array and SPR results, we introduced TAF8 point mutants into the TAF8–10 polyprotein expression construct by substituting three triple amino-acid clusters spanning residues 185–187 (DVE), 222–224 (PYL) and 293–295 (PYL) with alanines. We produced and purified

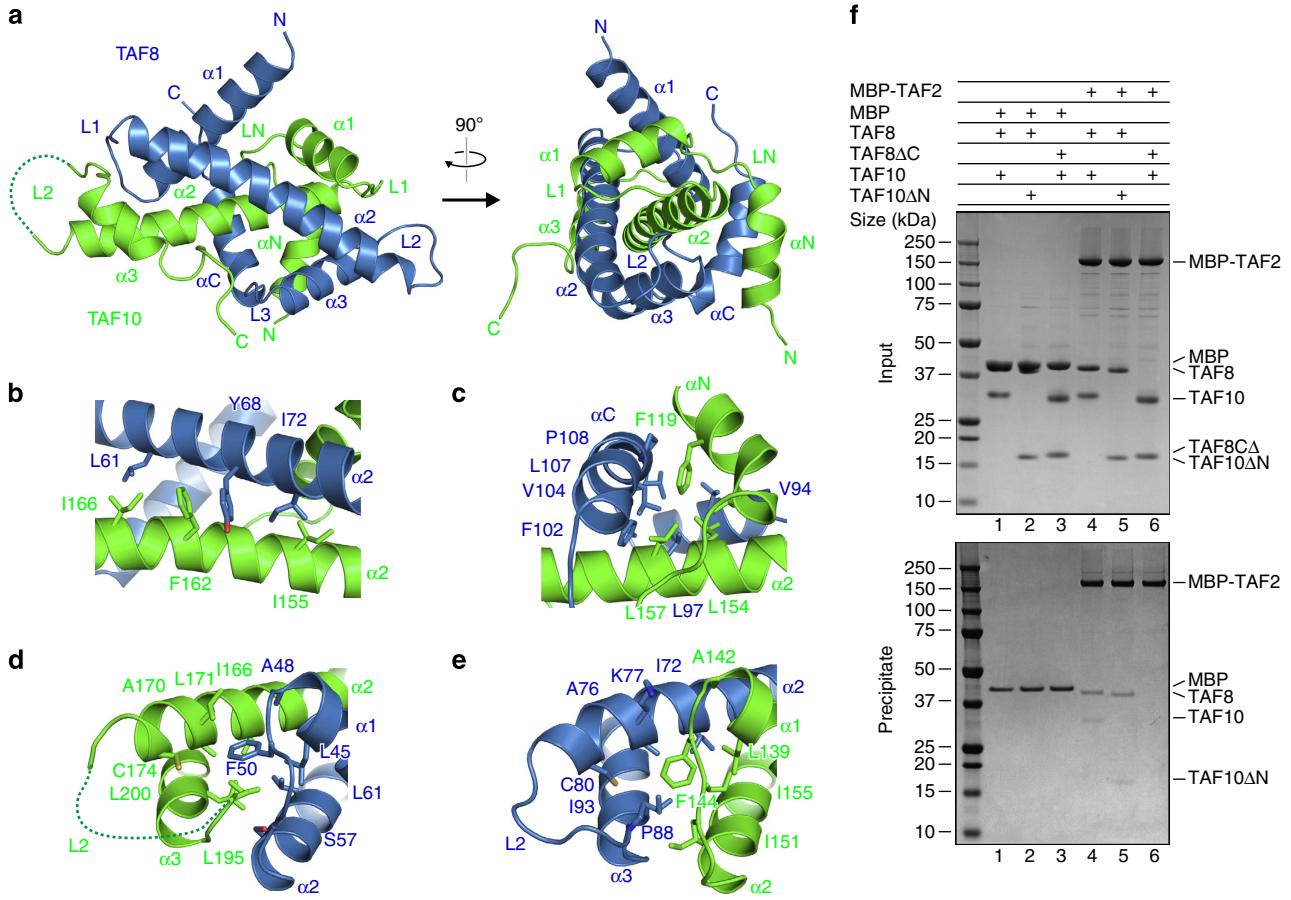


Figure 3 | TAF8–TAF10 interactions. (**a**) Crystal structure of human TAF8–10 complex is depicted in a cartoon representation. Two orientations related by a vertical rotation of 90° are shown. TAF8 is coloured in blue and TAF10 in green. The disordered L2 loop of TAF10 is represented by a dotted line. Secondary structure elements and loops are labelled. The TAF8–TAF10 complex adopts a non-canonical HFD pair. (**b**–**e**) Close-up views of the interactions between TAF8 and TAF10. Key interacting residues are highlighted. All structure drawings were generated with PyMOL (http://www.pymol.org/). (**f**) Pull-down experiments of TAF2 fused to MBP analysing the interactions with TAF8–10, TAF8–TAF10ΔN and TAF8ΔC–TAF10 HFD pairs (see main text for details). Unfused MBP is included as a control. Input samples (top) and samples precipitated on amylose resin (bottom) were resolved on 4–12% gradient gels. Protein identities are shown on the right.

wild-type TAF8–10 and the mutated TAF8–10 complex and analysed TAF2 binding via SEC (Fig. 4c). In contrast to wild-type TAF8–10, formation of a trimeric TAF2–8–10 complex was not observed with the three triple amino-acid cluster TAF8–10 mutant, corroborating the results that we obtained with our peptide array and SPR experiments (Fig. 4c).

TAF8 promotes TAF2 incorporation in TFIID. We showed recently that a TFIID subcomplex comprising TAF4, 5, 6, 8, 9, 10 and 12 (hereafter referred as 7TAF) can be formed *in vitro* by binding TAF8–10 to a physiological nuclear core–TFIID complex, which constitutes an important intermediate in holo–TFIID assembly[12,21]. We next asked whether the association of TAF2 to this 7TAF complex depends on the C-terminal region of TAF8, which we identified as responsible for TAF2 binding in the TAF2–8–10 complex. We produced and purified recombinant 7TAF and a 7TAFΔ complex, in which TAF8 is substituted by TAF8ΔC (Fig. 5a). We monitored binding of a mCherry-TAF2 fusion protein to 7TAF and 7TAFΔ complexes using SEC. We introduced the mCherry tag on TAF2 to unambiguously separate the protein from TAF4 in SDS–PAGE. In all SEC experiments, we used stoichiometric amounts of TAF2 in relation to TAF8–10 or the truncated TAF8ΔC–TAF10 complex. Interestingly, TAF2 could be fully incorporated into the 7TAF complex, whereas TAF2 did not interact noticeably with the 7TAFΔ complex, in which the C-terminal TAF2-interaction region of TAF8 had been deleted (Fig. 5b).

Next we mapped the position of TAF2 on 7TAF. To this end, we determined a three-dimensional model of negatively stained 7TAF complexes bound to TAF2 (hereafter referred as 8TAF complex) by single-particle electron microscopy and compared the resulting structure to the reconstruction of the 7TAF complex we had determined previously[12]. We observed major density differences clearly positioned on only one side of the particle, indicating TAF2 location (Fig. 5c and Supplementary Fig. 5a). Interestingly, our 8TAF complex reconstruction resembles a precursor to the characteristic clamp shape of holo–TFIID, in contrast to the less elongated shape of 7TAF and core–TFIID[10] (Supplementary Fig. 5b).

Next we sought to characterize possible alterations in the protein–protein interaction networks along the assembly pathway to holo–TFIID. In particular, we looked at the transition from 7TAF to 8TAF complexes on TAF2 binding by crosslinking and MS (CLMS) experiments. We crosslinked 7TAF and 8TAF complexes with the bifunctional reagent bis(sulfosuccinimidyl)-suberate, BS3 that targets mostly lysines[41] (Supplementary Fig. 6a). Crosslinked complexes were separated from non-crosslinked species by SDS–PAGE, in-gel digested and crosslinked peptides were assigned to ion masses observed by

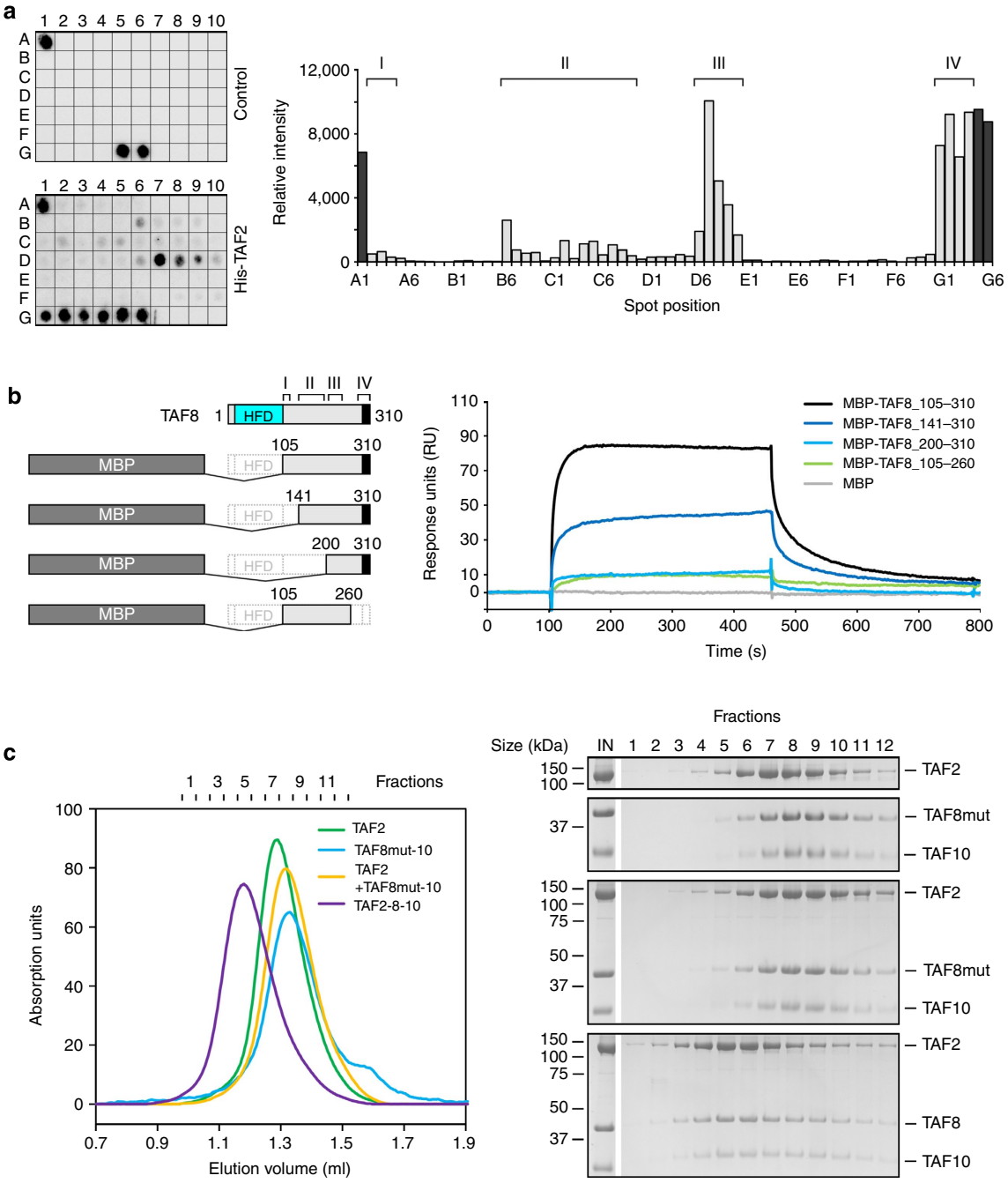


Figure 4 | TAF8–TAF2 interactions. (**a**) His-tagged TAF2 binding to overlapping peptides of the TAF8 C-terminal region (residues 105–310) spotted onto nitrocellulose membranes (spots A2-G4, left) was analysed by utilizing a peptide array. Bound TAF2 was visualized by luminol reaction and signal intensities were plotted for each spot after background subtraction (right). Spots A1, G5 and G6 served as positive controls. TAF2 protein was omitted for the control membrane. The four major binding regions (I–IV) are indicated above the histogram. (**b**) SPR experiments with immobilized full-length TAF2 as ligand and MBP (control) as well as MBP fusions of TAF8 fragments 105–310, 141–310, 200–310 and 105–260 as analytes. TAF8 deletion constructs are schematically shown as bar diagrams (left). TAF2-interacting regions on TAF8 as identified in **a** are highlighted. SPR sensorgrams at identical analyte concentrations of 500 nM are plotted as RU versus time (right). (**c**) SEC analyses assessing the influence of TAF8 point mutations on TAF2 binding. Elution profiles for the indicated proteins and protein complexes are plotted on the left and SDS-PAGE analyses of each run are shown on the right. Molecular masses of protein standards are denoted on the left of the gels and protein names on the right.

MS. We identified 37 protein–protein crosslinks for the 7TAF complex and 37 protein–protein crosslinks for the 8TAF complex with an overlap of 21 crosslinked peptides between the two complexes (Fig. 5d, Supplementary Fig. 6b–e and Supplementary Table 3). Our data suggest that TAF9 plays a central role in 7TAF complex architecture by interconnecting TAF4, 5, 6, 8 and 12 (Fig. 5d). In our CLMS data, prominent crosslinks between TAF8 and TAF10 were not present, consistent with the paucity of lysines within crosslinking distance, and the partly buried location of the TAF8–10 dimer within the 7TAF complex[12]. In the 8TAF complex, we observed crosslinks of the C-terminal region of TAF8 with residues on TAF2, which are predicted to map to the surface (Fig. 5d). In addition to its proximity to TAF8, TAF2 is also positioned closely to TAF5, 6 and 9

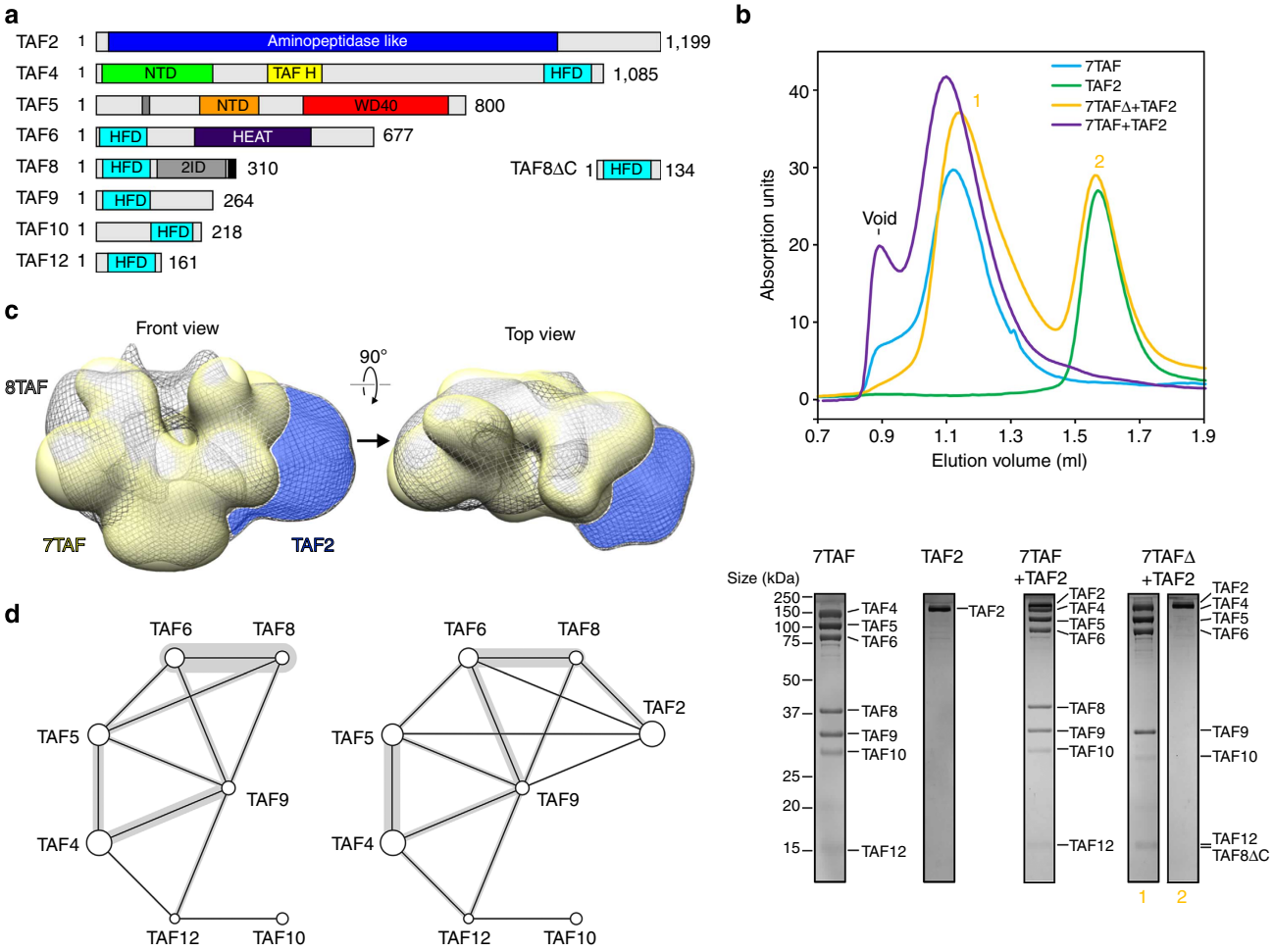


Figure 5 | TAF8 promotes TAF2 incorporation in TFIID. (**a**) TFIID components studied are shown as bar diagrams. Predicted low-complexity regions of the proteins are coloured in grey. A black bar denotes the TAF8 NLS. C-terminally truncated TAF8 (TAF8ΔC), which was used to reconstitute the 7TAFΔ complex, is depicted on the right. Numbers denote first and last amino acids for each protein. (**b**) Impact of the TAF8 truncation on TAF2 binding to 7TAF complexes. SEC elution profiles for indicated proteins and protein complexes are shown (top). Corresponding SDS-PAGE gel sections of peak fractions of each run are shown (bottom). Protein size markers are shown on the left; protein identities on the right. (**c**) Three-dimensional single-particle EM reconstruction of negatively stained 8TAF complex (grey mesh) superimposed on 7TAF complex (yellow, from ref. 12) is shown in two views related by a 90° rotation as indicated (arrow). Difference density attributed to bound TAF2 is highlighted in blue. (**e**) Protein-protein crosslink maps for the 7TAF complex (left) and the 8TAF complex (right) are shown. Circle sizes represent relative molecular weights of each protein. Black lines connect crosslinked proteins. Grey bars superimposed on black lines indicate crosslink frequencies (www.crosslinkviewer.org). Original images corresponding to the gel sections shown in Fig. 5b (bottom panel) are provided in Supplementary Figure 8.

and promotes crosslinking between TAF4 and 5 (Fig. 5d). Our data indicate that TAF2 is indeed anchored to the 7TAF complex via the TAF2-interacting region on TAF8 and that binding of TAF2 induces significant conformational changes that result in novel TAF–TAF interactions not present in the 7TAF complex.

TAF2–8–10 binds Importin α1 via the TAF8 NLS. Biochemical and cell biology experiments demonstrated that the C-terminal NLS within TAF8 is necessary for shuttling TAF8 and TAF10 from the cytoplasm to the nucleus in an Importin α/β-dependent fashion[30]. We asked whether the identified endogenous TAF2–8–10 complex would be capable of recruiting Importin α1 *in vitro* to form a nuclear import complex. To this end, we mixed highly purified TAF2–8–10 with a twofold molar excess of an Importin α1 variant lacking the Importin β-binding domain (Importin $\alpha1^{\Delta IBB}$). We observed efficient complex formation in SEC indicating that Importin $\alpha1^{\Delta IBB}$ was stoichiometrically incorporated into the TAF2–8–10 complex (Fig. 6a and Supplementary Fig. 7a). We also observed by SEC that TAF2 alone is not bound by Importin α1.

To define the binding region between Importin α1 to the TAF2–8–10 complex, we determined the X-ray crystal structure of the C-terminal NLS of TAF8 in complex with Importin $\alpha1^{\Delta IBB}$ at 1.75 Å resolution. The refined model has a crystallographic R value of 15.3% and a free R factor of 18.0% with good stereochemistry (Table 1). Importin α1 residues 72–497 and residues 297–305 of the TAF8 peptide could be unambiguously traced in the electron density map. The TAF8 peptide binds as a monopartite NLS via residues 297–302 (Fig. 6b). In previous Importin/NLS structures, asparagines N146, N188 and N235 of Importin α1 hydrogen bond to NLS main chain amide and carbonyl groups at positions P1, P3 and P5 (ref. 42) (Fig. 6b and Supplementary Fig. 7b). Importin α1 tryptophanes W142, W184 and W231 form apolar pockets, which accommodate the aliphatic moieties of lysine residues K300 and K302, and position the TAF8 NLS backbone via residues P297, K300 and K302 (Fig. 6b). The side chain of K299 of TAF8 is coordinated by the main chain

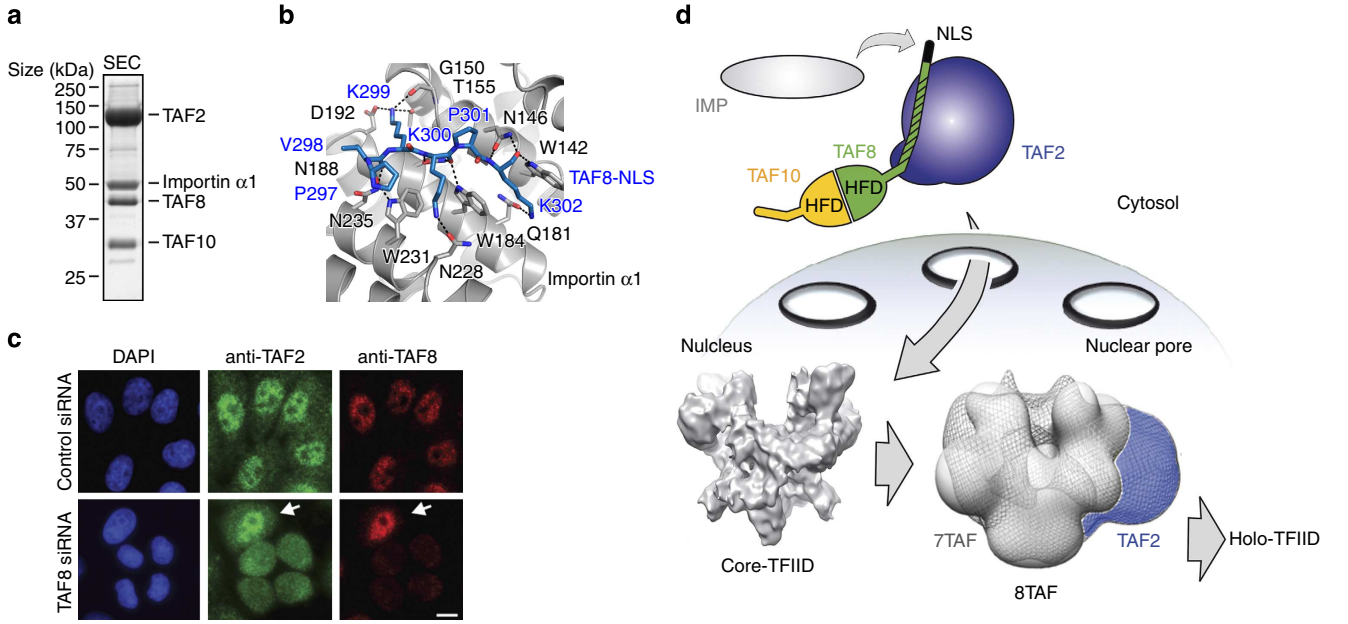


Figure 6 | Nuclear TFIID assembly from preformed submodules. (**a**) A complex consisting of TAF2, TAF8, TAF10 and Importin $\alpha1^{\Delta IBB}$ was formed from highly purified components. Importin $\alpha1^{\Delta IBB}$ was mixed in a twofold molar excess with purified TAF2–8–10 and the mixture purified by SEC. SDS–PAGE analysis of the peak fraction is shown. (**b**) Importin α1-TAF8 complex crystal structure. Magnified view of interacting residues of the major binding site of Importin α1 (grey) with residues of the NLS of TAF8 (blue). Importin α1 is shown in ribbon representation and the TAF8-NLS as a stick model. TAF8 residues R303 and R304 are not involved in contacting Importin α1 and are omitted for clarity. (**c**) Immunofluorescence microscopy of HeLa cells depleted of TAF8 (TAF8 siRNA) by RNAi or control cells (Control siRNA). Nuclei are visualized by 4′,6-diamidino-2-phenylindole (DAPI) staining (blue). TAF2 is displayed in green and TAF8 in red. Arrows point to a non-transfected cell. Scale bar, 10 μm. (**d**) Cartoon model of cytoplasmic TAF2–8–10 complex and nuclear holo–TFIID assembly. The NLS of TAF8 is filled in black. The TAF2-interaction domain within TAF8 is highlighted by shading. TAF2, blue; TAF8, green; TAF10, orange. The TAF2–8–10 complex resides in the cytoplasm, whereas the physiological symmetric core–TFIID complex is found in the nucleus[10,22]. The cryo-electron microscopy density envelope of core–TFIID complex is shown (adapted from ref. 12). On binding of Importin α1 (grey) to the TAF8 NLS, TAF2–8–10 translocates into the nucleus through a nuclear pore (arrow). In the nucleus, Importin α1 is released and TAF2, 8 and 10 associate with core–TFIID, to form intermediates including the asymmetric 8TAF complex along the holo–TFIID assembly pathway.

carbonyl group of G150, the hydroxyl group of T155 and the carboxylate of D192, whereas side chains of K300 and K302 of TAF8 are contacted by side chain carbonyl groups of N228 and Q181, respectively. We could also model a less well-defined short amino-acid segment at the minor binding site of Importin α1 (Supplementary Fig. 7c). To assess whether the minor binding site of Importin α1 plays a role in binding the NLS of TAF8, we determined the kinetic parameters for the Importin α1/TAF8–NLS complex formation by isothermal titration calorimetry. Using Importin $\alpha1^{\Delta IBB}$ as an analyte and an NLS peptide comprising TAF8 residues 288 to 310 as titrant, we obtained a 1:1 binding stoichiometry with a dissociation constant in the low micromolar range ($K_d = 10.4 +/- 0.8\,\mu M$; Supplementary Fig. 7d). In accordance with our crystal structure, the binding of the NLS of TAF8 to Importin α1 is driven by enthalpy involving mainly hydrogen bonds and van der Waals interactions (enthalpy change of $\Delta H = -18.5 +/- 1.3\,kcal\,mol^{-1}$ and entropy change of $-T\Delta S = 11.5\,kcal\,mol^{-1}$).

We next asked if the nuclear localization of TAF2 is dependent on the presence of TAF8 *in vivo*. Therefore, we knocked down endogenous TAF8 in HeLa cells by RNA interference (RNAi) treatment for 48 h and compared the nuclear/cytoplasmic distribution of TAF2 in TAF8 knockdown cells with control cells by immunofluorescence (Fig. 6c). Short interfering RNA (siRNA) treatment leads to depletion of TAF2 in the nucleus and to an enrichment of TAF2 in the cytoplasm, suggesting that the import of TAF2 is controlled by TAF8 (Fig. 6c).

Taken together, our data suggest the presence of a nuclear import particle in which the TAF2–8–10 complex is bound by the major binding site of Importin α1 via the NLS of TAF8, poised to shuttle into the nucleus (Fig. 6d). On release of Importin α1, the TAF2–8–10 module then combines with core–TFIID (Supplementary Fig. 2b,c). Association of these preformed TFIID submodules leads to conformational rearrangements in the resulting intermediate TAF complex which enables formation of the functional nuclear holo–TFIID (Fig. 6d).

Discussion

Elucidation of the structure and function of multiprotein complexes in gene regulation is an intense focus of current research efforts[43,44]. Whereas three-dimensional models of fully assembled multiprotein complexes derived from X-ray crystallography or single-particle cryo-electron microscopy provide a wealth of information on the architecture of such complexes, little is known about how the cell controls and regulates the ordered assembly of multiprotein gene regulatory complexes such as TFIID.

Several earlier studies described the existence of a variety of TFIID complexes with distinct subunit composition in different cell types[23,45–48] conveying a concept of modular TFIID assembly. To gain more insights into the regulated assembly of TFIID, and to try to understand how the regulated assembly of such complexes may contribute to gene regulation, we initiated a series of experiments to identify TFIID assembly intermediates in the cytoplasm, and different TFIID assemblies in the nuclei of human cells. In the framework of these experiments, we were also aiming to uncover the incorporation pathway of TAF2 in TFIID. We identified a novel TFIID building block comprising TAF2–TAF8–TAF10 in the cytoplasm of human cells. We also

characterized the interactions stabilizing this cytoplasmic complex in an integrated approach combining native MS, X-ray crystallography, SPR, peptide arrays and biochemical and biophysical methods. Our experiments indicate that TAF2 interacts with the C-terminal unstructured region of TAF8 *in vitro*, substantiating protein–protein interaction mapping experiments of *Saccharomyces cerevisiae* TFIID[49].

Previously, it was shown that TAF8 shuttles from the cytoplasm to the nucleus in an Importin α/β-dependent pathway and piggybacks TAF10 into the nucleus[30]. Owing to the lack of a NLS, TAF10 cannot translocate to the nucleus on its own and depends on the NLS of its interaction partner, TAF8 (ref. 30). Deletion mutants suggested that Importin binding resides in the extreme C-terminus of TAF8 (ref. 30). We show that a tetrameric complex consisting of Importin α1, TAF2, TAF8 and TAF10 can be assembled from purified components *in vitro* suggesting a co-import mechanism for the three proteins. Our TAF2 and TAF10 cellular localization experiments support this mechanism, indicating that knockdown of TAF8 by RNAi not only alters the cellular localization of TAF10, but likewise the localization of TAF2.

To define the precise modes of interaction between Importin α1, TAF8 and TAF10, we solved the X-ray crystal structures of Importin α1 bound to the NLS of TAF8 on one hand, and of the HFD pair formed by TAF8 and TAF10 on the other. The crystal structure of the TAF8–10 complex reveals atypical histone folds of the two proteins and shows a combination of symmetric and asymmetric structural elements. Both TAFs share characteristic conformations of their L1 loops, which give rise to pseudosymmetric structures at the extremities of their HFDs. Otherwise, the presence of additional αN and αC helices render the TAF8–10 complex asymmetric. The pseudosymmetric L1 loops are characteristic for the TAF8–10 complex, since similar arrangements cannot be found in the crystal structures of the *Drosophila* TAF6–TAF9 and the human TAF4–TAF12 complex[22,25] suggesting that the overall shape and precise geometry of the complex is important for integration into core–TFIID.

The crystal structure of Importin α1 with the NLS of TAF8 reveals that TAF8 has a classical short monopartite NLS, which is recognized by the major binding site of Importin α1. Interestingly, phosphorylation of conserved serine residues C-terminal to canonical NLSs of different nuclear proteins either enhance, or abolish, the binding affinity of different importins, thus regulating nuclear import[50–52]. Similarly, phosphorylation may also fine-tune nucleocytoplasmic shuttling of the TAF2–8–10 complex. Two serine residues predicted to be phosphorylated are located C-terminal to the NLS of TAF8. Therefore, the affinity of TAF8 to Importin α1 and consequently the nuclear import could be modified by phosphorylation. It will be interesting to see if such a regulatory mechanism by post-translational modification exists for the import of the TAF2–8–10 complex *in vivo*. Our crystal structure of Importin α1 bound to the TAF8 NLS further suggests that the cytosolic TAF2–8–10 complex together with Importin α1 constitutes an import particle responsible for delivering this building block into the nucleus. Likewise, our experiments indicate that this TAF2–8–10 building block is responsible for the incorporation of TAF2 in nuclear TFIID.

A stable TFIID core complex comprising two copies each of TAF4, 5, 6, 9 and 12 was identified in *Drosophila* and human cell nuclei[12,21]. Previously, we postulated that the binding of TAF8–10 causes a rearrangement of the symmetric TFIID core complex to an asymmetric particle, which is then capable of accommodating the remaining TAFs and TBP, each in single copy[12]. We propose that association of the TAF2–8–10 complex with the preassembled nuclear core–TFIID involves an intricate network of interactions between the TAF8 C-terminal tail and TAF2 on one hand, and the globular HFD pair of the TAF8–10 complex and core–TFIID on the other. Our current data suggest that TAF8–10 may function alike a chaperone to regulate nuclear import and integration of TAF2 into core–TFIID. Note, however, that in the cytoplasmic extracts, apart from TAF8 and TAF10, we did not detect any of the other TFIID components stably associated with endogenous TAF2. Therefore, we hypothesize that TAF1, 7, 11, 13 and TBP incorporate into the TFIID structure probably at a defined, later step, after TAF2–8–10 has been accreted. Moreover, notwithstanding the fact that we did not find either TAF1 or TBP associated with TAF2 in the cytoplasm in our co-IP coupled MS analyses, it still can be envisioned that the TAF2–8–10 complex is capable of nucleating the formation of the holo–TFIID complex, including the TAF1–TBP module, and thus promote transcription from Inr-containing core promoters.

Interestingly, TAF2-containing and TAF2-lacking, as well as TAF10-containing and TAF10-lacking, TFIID complexes have also been extracted from human cells[32,34,45]. Thus, in good agreement with the modular TFIID assembly concept, our observations suggest that the here characterized TAF2–8–10 building block would not always incorporate in all canonical TFIID complexes but, even in the nucleus, may exist as an independent regulatory entity. Future experiments will be required to elucidate the function(s) of holo–TFIID complexes versus complexes lacking TAF2–8–10. Along the same lines, it will also be interesting to test whether a TAF2–8–10 complex alone or in combination with core–TFIID can modulate transcription efficiency of Inr-dependent genes. Promoter architecture may at this junction control transcription regulation by gauging the assembly rate of holo–TFIID from building blocks[48]. From a pharmaceutical point of view, it is to date entirely unclear whether or not neurological disorders caused by mutations in the *TAF2* gene develop due to altered regulation of transcriptional activity or due to other currently unknown mechanisms. Future experiments will be required to elucidate if these TAF2 mutations may actually affect TAF8 binding and TFIID assembly.

Our results support the view that stable partial TFIID complexes—that potentially have important functions of their own—might exist in the cell. These complexes may represent functional cytoplasmic or nuclear modules, which assemble into holo–TFIID in a stepwise fashion. Also, our results point to an important role of cytoplasmic–nuclear transport in holo–TFIID formation. We anticipate that such processes will likewise play important roles in regulating the assembly and activities of many other multiprotein complexes that direct gene transcription.

Methods

DNA constructs. Cloning of TAF2, TAF8 and TAF10 expression constructs in MultiBac plasmids pPBac[38], pFL and pIDC[53] is detailed in the Supplementary Methods. Expression plasmids for subcomplexes TAF5–6–9 (pPBac-3TAF) and TAF4–12 (pDiFB-412) and Importin $\alpha1^{\Delta IBB}$ were described previously[12,54]. Truncated Importin α1 (residues 71–497) was generated by amplifying the coding region of Importin $\alpha1^{\Delta IBB}$ by PCR. All constructs were verified by DNA sequencing.

Sequence alignments. Alignments were generated using the ClustalW2 server[55] and plotted with ESPript (http://espript.ibcp.fr)[56]. Protein sequences for human TAF3 (UniProt accession number Q5VWG9) and human SPT7L (O94864) were retrieved from the UniProt server (www.uniprot.org).

Protein production and purification. MBP–TAF8-fusion proteins were produced in *E. coli* Rosetta (DE3) cells (Novagen) and purified by metal affinity chromatography using TALON resin (Clontech) followed by size-exclusion chromatography on a Superdex 200 16/60 column (GE Healthcare; detailed in Supplementary Methods). Importin α1 constructs (residues 60–529 or residues 71–497) were produced and purified as described[54], except that *E. coli* Rosetta (DE3) cells (Novagen) were used. Production and purification of core–TFIID and 7TAF complexes was performed as described[12].

Proteins TAF2, MBP–TAF2, mCherry-TAF2 and TAF8–10 complex were produced using the MultiBac system[53]. Expressed protein was captured via TALON resin (Clontech) from the cell lysate in batch. Proteins were further purified by ion exchange chromatography using a 5-ml SP-Sepharose HiTrap column (GE Healthcare) followed by gel filtration using Superdex200 10/300 or Superose6 10/300 columns (GE Healthcare; Supplementary Methods). Proteins were flash frozen in liquid nitrogen and stored at −80 °C in aliquots.

Binding experiments. SEC experiments were carried out with ÄKTA purifier or ÄKTA Micro systems (GE Healthcare) using Superdex200 10/300, Superdex200 PC3.2, Superose6 PC3.2 or Superose6 PC3.2 Increase columns. Binding experiments shown in Fig. 2a were performed in running buffer comprising 25 mM HEPES-NaOH, pH 7.5, 150 mM NaCl, 1 mM DTT (dithiothreitol). Runs in Figs 4c and 5b and Supplementary Fig. 2b,c were performed in buffer comprising 25 mM HEPES-NaOH, pH 7.5, 500 mM NaCl, 1 mM DTT.

Crystallization and structure determination. Screening for crystallization conditions was performed at the High Throughput Crystallization (HTX) laboratory (EMBL Grenoble, France; Supplementary Methods). Crystals of truncated TAF8–10 complex (TAF8 residues 25–120 and TAF10 residues 112–212) were refined manually by mixing equal volumes of protein solution containing 15–25 mg ml^{-1} TAF8–10 in 25 mM Tris-HCl, 150 mM NaCl at pH 7.5 and crystallization solution containing 1.4 M Na/K PO_4 at pH 7.6. Crystals grew in space group $P3_121$ with cell dimensions of $a=b=51.3$ Å and $c=144.8$ Å. Crystals were cryoprotected by adding 20% (v/v) glycerol and flash frozen in liquid nitrogen. Diffraction data were collected at 100 K on beamline PROXIMA 1 using a Pilatus 6 M detector (SOLEIL synchrotron, Gif-sur-Yvette, France) and were integrated and scaled using X-ray Detector Software (XDS)[57]. The structure of TAF8–10 was solved by the Sulfur-SAD method. A partial model could be built into the experimental electron density map by iterative rounds of density modification and automated structure building using programs Pirate and Buccaneer from the CCP4i suite[58]. The model was used to phase a high-resolution data set by molecular replacement using Phaser[58,59]. Diffraction data were corrected for anisotropy using the Diffraction Anisotropy Server (services.mbi.ucla.edu/anisoscale)[60] and an isotropic B of −11.99 Å^2. The TAF8–10 structure was built and refined using programs Coot[61] and Phenix[62], respectively, including TLS parameter and individual B-factor refinement.

Crystals of Importin α1 (residues 60–529) with a synthetic TAF8 NLS peptide (amino acids 297-PVKKPKIRRKKSLS-310 (Peptide Specialty Laboratory, Germany) were grown by mixing 2 µl of protein solution containing 8 mg ml^{-1} Importin α1/TAF8–NLS in 10 mM Tris-HCl, pH 8.0, 150 mM NaCl and 1 mM DTT with 1 µl reservoir solution containing 100 mM HEPES-NaOH, pH 7.1, 12% (w/v) polyethylene glycole 3,350 and 200 mM L-proline in sitting drop vapour diffusion plates. Crystals grew in space group $P2_12_12_1$ with cell dimensions of $a=54.3$ Å, $b=77.7$ Å and $c=128.6$ Å. Crystals were cryoprotected by the addition of 30% (v/v) ethylene glycole and flash frozen in liquid nitrogen. Data sets were collected at 100 K on beamline ID14-1 using an ADSC Quantum Q210 detector (European Synchrotron Radiation Facility ESRF, Grenoble, France). Diffraction data were integrated and scaled using XDS[57]. The structure of Importin α1/TAF8–NLS was solved by molecular replacement using Importin α1 (PDB ID 3RZ9) as a search model. The Importin α1/TAF8–NLS structure was built and refined with Coot[61] and Phenix[62], respectively, including TLS parameter, occupancy and individual B-factor refinements.

Surface plasmon resonance. Biosensor experiments were performed at 25 °C on a BIACORE 3000 (Biacore AB, Uppsala). TAF2 ligand was immobilized onto CM5 sensor chips (GE Healthcare) to a level of 2,500 RU using amine-coupling chemistry. Truncation mutants of TAF8 fused C-terminally to MBP were serially diluted into running buffer (25 mM HEPES-NaOH, pH 7.0, 300 mM NaCl, 0.01% (v/v) NP-40). For association phase, 150 µl of analyte at a concentration of 500 nM were injected at a flow rate of 25 µl min^{-1} and dissociation phases were monitored for 200 s by injecting running buffer only. Binding responses were recorded and responses from referencing sensorgrams were subtracted using BIAevaluation software (GE Healthcare). Data were globally analysed with the analysis software.

Isothermal titration calorimetry. Calorimetric experiments were conducted in duplicates with a MicroCal iTC$_{200}$ instrument (GE Healthcare) at 25 °C. Importin α1 (residues 71–479) and the TAF8 NLS peptide (residues 288-NPYLRPVKKP-KIRRKKSLS-310) were extensively dialysed against ITC buffer (25 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM β-mercaptoethanol) and used at concentrations of 39 µM and 1.5 mM, respectively. Protein concentrations were determined by absorbance spectroscopy at 280 nm with calculated extinction coefficients of 48,930 M^{-1} cm^{-1} for Importin α1 and 1,490 M^{-1} cm^{-1} for the peptide. TAF8 peptide (1.5 µl) was injected for 3 s with a spacing of 180 s between injections into 200 µl of Importin α1. Heat changes were recorded over 26 injections. Calorimetric titration data were integrated, corrected for heat of dilution of the TAF8 peptide alone and analysed using Origin software version 7.0 according to a one-site binding model. Binding stoichiometry (n), association constant (K_a), binding enthalpy (ΔH) and entropy change (ΔS) were deduced from fitted isotherms by nonlinear regression. Gibbs free energy difference was calculated using the equation $\Delta G=\Delta H-T\Delta S$.

Pull-down assays. MBP pull-down assays were performed by mixing 10 µg bait (MBP or MBP–TAF2) with 10 µg prey (TAF8–10, TAF8ΔC–TAF10, TAF8–TAF10ΔN) for 1 h at 4 °C in binding buffer (25 mM HEPES-NaOH, pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 2 mM β-mercaptoethanol). Protein mixtures were incubated with 20 µl Amylose resin (New England Biolabs) for 1.5 h at 4 °C. Resin was washed three times with binding buffer, once with washing buffer (25 mM HEPES-NaOH, pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 2 mM β-mercaptoethanol, 0.05% NP-40) and again three times with binding buffer. Proteins were eluted in 15 µl binding buffer containing 30 mM D-maltose and analysed by 4–12% Bis-Tris NuPAGE (Invitrogen).

Limited proteolysis experiments. The TAF8–10 complex (1 mg ml^{-1}) was treated with chymotrypsin at an enzyme-to-protein ratio of 1:10 (w/w). Samples were taken after 2, 5, 10, 20, 40 and 60 min and analysed by SDS–PAGE. To identify the TAF8–10 core complex, the proteolysed sample was loaded on a Superdex75 10/300 column (GE Healthcare) before N-terminal sequencing and MS analysis of comigrating polypeptides.

Peptide arrays. Pepscan libraries of the C-terminal region of TAF8 (residues 105–310) were immobilized on cellulose membranes via double β-alanine anchors and assembled using the SPOT technology (AG Molekulare Bibliotheken, Charité—Universitätsmedizin Berlin, Germany). Overlapping 20-mer peptides of TAF8 were synthesized by Fmoc (9-fluorenylmethoxycarbonyl) chemistry with an offset of three amino acids between neighbouring spots. Low-density hexa-Histidine peptides were used as controls. Pepscan membranes were blocked in blocking buffer (50 mM Tris-HCl, pH 7.6, 500 mM NaCl, 20% (w/v) sucrose, 3% (w/v) bovine serum albumin) for 1 h at 4 °C, washed with TBS (50 mM Tris-HCl, pH 7.6, 500 mM NaCl) and incubated for 1.5 h with His-tagged TAF2 (10 µg ml^{-1}) in blocking buffer or with blocking buffer alone. Membranes were incubated with mouse anti-His monoclonal primary antibody (Sigma-Aldrich, catalogue number H1029, dilution 1:3,000) and peroxidase-conjugated anti-mouse secondary antibody (Sigma-Aldrich, catalogue number A5906, dilution 1:10,000) in blocking buffer. Membranes were washed three times with TBS between each incubation step. Luminol solution (Pierce) was added and luminescence detected on a KODAK 4000MM photoimager. Images were analysed using the Dot Blot Analyzer tool in ImageJ.

Analytical ultracentrifugation. Sedimentation velocity experiments were performed in a Beckman XL-I analytical ultracentrifuge (Beckman Coulter). The purified proteins and protein complexes TAF2, TAF8–10 and TAF2–8–10 were loaded into sapphire-windowed cells with 12-mm optical path length and spun in an An-60Ti rotor (Beckman Coulter). Absorbance at 280 nm was measured for 16 h at 42,000 r.p.m. and 10 °C. The data were analysed in terms of continuous size-distribution (c(s)) with the Sedfit program[63], considering 200 particles with sedimentation coefficients, s, between 0.1 and 20 S. A partial specific volume of 0.73 and frictional ratios of 1.4 (TAF2, TAF2–8–10) and 1.6 (TAF8–10) were used. A regularization procedure with confidence level of 0.68 was applied. Sample densities and viscosities were determined with Sednterp[64] to 1.023 g ml^{-1} and 1.40 mPa.s (TAF2, TAF2–8–10) and 1.021 g ml^{-1} and 1.31 mPa s (TAF8–10).

Native MS. Purified TAF2–8–10 complex (30 µl) was buffer exchanged into 500 mM ammonium acetate buffer (pH 7.5) using Amicon spin concentrators (Millipore, 10 kDa MWCO). All MS experiments were performed on a Quadrupole Time-of-flight (Q-ToF) II mass spectrometer (Waters, Manchester, UK) in the positive ion mode[65]. For data acquisition, 2 µl of the sample was injected into the mass spectrometer with gold-coated capillary needles made in-house using a needle puller (Harvard Apparatus, Holliston, MA, USA). MS spectra were acquired using a capillary voltage of 1.7 kV and cone and collision voltages of 100 V. Time-of-flight and analyser pressures were at 5.6×10^{-6} and 4.2×10^{-4} mbar, respectively. Data sets were acquired and processed with MassLynx V4.1 software (Waters, UK) with minimal smoothing and no background subtraction. The recorded mass spectra were calibrated externally using 100 mg ml^{-1} caesium iodide in water.

CLMS analyses. 7TAF complexes were produced as described[12]. 8TAF complexes were reconstituted from purified 7TAF complexes by adding twofold molar excess of TAF2 and removal of unbound TAF2 by SEC. 7TAF (200 µg) and 8TAF complexes (200 µg) were crosslinked by BS3 (Bis-sulfosuccinimidyl suberate, Thermo Scientific) at complex/BS3 ratio of 1:5 (w/w) in crosslinking buffer (25 mM HEPES-NaOH, pH 7.6, 150 mM NaCl, 1 mM DTT) for 2 h on ice. The reaction was quenched by adding saturated ammonium bicarbonate solution followed by incubation on ice (45 min). Crosslinked samples were concentrated using spin concentrators (Millipore) and separated on NuPAGE 3–8% Tris-Acetate gels run in Tris-Acetate SDS running buffer (Invitrogen). Bands corresponding to crosslinked complexes were excised, crosslinked complex proteins reduced,

alkylated and trypsin digested following standard procedures. Crosslinked peptides were fractionated using SCX-StageTips following published protocols for linear peptides and desalted using C18 StageTips[41].

Mass spectrometry. Peptides were analysed on an LTQ Orbitrap Velos mass spectrometer coupled with an UltiMate 3000 Rapid Seperation LC system (Thermo Fisher Scientific). The column was packed into a spray emitter (75-µm inner diameter, 8-µm opening, 250-mm length; New Objectives) with C18 material (ReproSil-Pur C18-AQ 3 µm; Dr Maisch, Ammerbuch-Entringen, Germany) using an air pressure pump (Proxeon Biosystems). Mobile phase A consisted of water and 0.1% formic acid. Mobile phase B consisted of 80% acetonitrile and 0.1% formic acid. Peptides were loaded onto the column with 2% B at 500 nl min^{-1} flow rate and eluted at 300 nl min^{-1} flow rate in two steps: linear increase from 2% B to 40% B in 139 min; then increase from 40 to 95% B in 11 min. The eluted peptides were directly sprayed into the mass spectrometer. Peptides were analysed using a high/high strategy: both MS spectra and MS2 spectra were acquired in the Orbitrap. MS spectra were recorded at 100,000 resolution. The eight highest intensity peaks with a charge state of three or higher were selected in each cycle for ion trap fragmentation. Fragments were produced using CID with 35% normalized collision energy and detected by the Orbitrap at 7,500 resolution. Dynamic exclusion was set to 90s and repeat count was 1.

Data processing. The mass spectrometric raw files were processed into peak lists using MaxQuant (version 1.3.0.5) (ref. 41) at default parameters except for 'top MS/MS peaks per 100 Da' being set to 100. Search was conducted against TAF complex sequences using Xi software (version 1.3.355). Search parameters were MS accuracy, 6 p.p.m.; MS/MS accuracy, 20 p.p.m.; enzyme, trypsin; crosslinker, BS3 (including BS3 modification); max. missed cleavages, 4; fixed modification, carbamidomethylation on cysteine; variable modifications, oxidation on methionine; crosslinkable amino acids, N terminus, lysine, serine, tyrosine and threonine; fragments, b and y ions with loss of H_2O, NH_3 and CH_3SOH. The data have been validated by 5% FDR with manual validation. UniProt protein accession numbers of the protein sequences used to search the database were as follows: TAF2 (Q6P1X5-1; with a sequence variation $R^{785}G$; European Nucleotide Archive AAC68502.1), TAF4 (O00268-1), TAF5 (Q15542-1), TAF6 (P49848-1), TAF8 (Q7Z7C8-1), TAF9 (Q16594-1), TAF10 (Q12962-1), and TAF12 (Q16514-1). Crosslinks observed in artificially introduced sequences (for example, TEV cleavage sites in polyproteins or purification tags) were not included in the search. The MS data were deposited to the ProteomeXchange Consortium[66] via the PRIDE partner repository with the data set identifier PXD001454 (http://www.proteomexchange.org).

Antibody production and purification. TAF2 antibodies were generated by immunizing rabbits with purified TAF2. Antibody purification was done as described[67] with the following modifications: 2 mg of recombinant full-length human TAF2 were fixed on 400 µl Affi-Gel 10/15 beads (Bio-Rad) for 2 h at 4 °C with gentle agitation in PBS. Free active esters were blocked with 1 M ethanolamine HCl (pH 8) solution for 1 h at 4 °C under gentle agitation. The TAF2-bound gel was transferred to a column and washed four times with 10 volumes of PBS. Ten ml of rabbit polyclonal antibody sera raised against human TAF2 was applied twice and the column was washed with 10 ml of PBS before elution. Bound antibodies were eluted with 0.1 M glycine (pH 2.5) buffer. Fractions of purified antibody (500 µl) were collected and quickly neutralized by adding 50 µl 2 M Tris-HCl (pH 8.8) buffer.

Protein extract preparations and immunprecipitation and MudPIT analyses. HeLa cell nuclear extract (NE) preparations and immunoprecipitations were done as described[68] with minor modifications. Supernatant containing the cytoplasm was precipitated by adding stepwise 0.3 g ml^{-1} ammonium sulfate under agitation (4 °C, 30 min). Precipitated proteins were collected by centrifugation (30,000*g*, 4 °C, 20 min), resuspended and dialysed overnight.

Immunoprecipitated proteins were eluted from the protein G columns with 0.1 M glycine (pH 2.5) and quickly neutralized with 2 M Tris-HCl (pH 8.8). For MudPIT[69] analyses, protein mixtures were trichloroacetic acid precipitated, urea denaturated, reduced, alkylated and digested with endoproteinase Lys-C followed by modified trypsin digestion. Peptide mixtures were loaded onto a triphasic 100-mm diameter fused silica microcapillary column[70]. Loaded columns were placed in-line with a Quaternary Dionex Ultimate 3000 HPLC pump and a LTQ Velos linear ion trap mass spectrometer equipped with a nano-LC electrospray ionization source (Thermo Fischer Scientific). A fully automated 12-steps MudPIT run was performed during which each full MS scan (from 300 to 1,700 m/z range) was followed by 20 MS/MS events using data-dependent acquisition[69]. Proteins were identified by database searching using SEQUEST with ThermoProteome Discoverer 1.4 (Thermo Fischer Scientific)[71]. Tandem mass spectra were searched against a human protein sequence database (from the *Homo sapiens* 2013-04-03 Swissprot release). In all searches, cysteine residues were considered to be fully carboxyamidomethylated (+57 Da statically added) and methionine to be oxidized (+16 Da dynamically added). Relative protein abundance for each protein in a given sample was estimated by normalized spectral abundance factor[72]. Normalized spectral abundance factor values were calculated from the spectral counts of each identified protein. Larger proteins tend to contribute more peptide/spectra and, therefore, spectral counts were divided by protein length to provide a spectral abundance factor (SAF). SAF values were then normalized against the sum of all SAF values in the corresponding run allowing the comparison of protein levels across different runs. The MS proteomics data were deposited to the ProteomeXchange Consortium[66] via the PRIDE partner repository with the data set identifier PXD001427.

Immunofluorescence. Indirect immunofluorescence tests were performed as described[30] with the following modifications: cells were fixed with 4% paraformaldehyde for 15 min at room temperature (RT) and then permeabilized with 0.1% Triton-X100 for 20 min at RT, incubated for 1 h at RT with either an anti-TAF2 (rabbit polyclonal serum; 3038; described above; diluted 1:100) + anti-TAF8 (mouse monoclonal antibody (mAb) 1FR-1B6 (ref. 6); diluted 1:1,000) or anti-TAF2 + anti-TAF10 (mAb 6TA-2B11 (ref. 6); diluted 1:1,000) antibody mix followed by incubation (RT, 1 h) with secondary antibody mix including Alexa488-labelled goat anti-rabbit mAb (Life Technologies, catalogue number A-11034, diluted 1:3,000; detects anti-TAF2) and Alexa568-labelled goat anti-mouse mAb (Life Technologies, catalogue number A-11004, diluted 1:3,000; detects either anti-TAF8 or anti-TAF10). As negative control, cells were incubated with secondary antibodies only to quantify background signal. Cells were mounted using Vectashield mounting medium with DAPI (Vector laboratories Inc.). Images were analysed on a Leica widefield fluorescence microscope (DMRXA2) equipped with a CoolSnap HQ camera (× 63 or × 100 magnification). Fluorescence intensity measurements in the cell cytoplasm were performed using Fiji software; intensity values were normalized to background signals.

siRNA transfection. siRNAs targeting TAF8 (ON-TARGETplus SMARTpool siRNA J-015912-20, J015912-19, J015912-18, J015912-17; Dharmacon; Thermo-Sientific) and non-targeting control (D-001810-10-20) were transfected into HeLa cells using Lipofectamine 2000 transfection reagent (Invitrogen). Cells were fixed for immunofluorescence experiments 48 h after transfection.

Electron microscopy. *Specimen preparation*. 8TAF sample was stabilized by mild glutaraldehyde crosslinking (GraFix[73]). Two-hundred µl purified 8TAF complexes were loaded on a 4-ml centrifugation tube containing a 10 to 30% glycerol and a 0 to 0.15% glutaraldehyde gradient followed by centrifugation (34,000 r.p.m., 18 h, 4 °C) with a SW60 rotor (Beckman Coulter). Fractions containing stabilized sample were deposited onto a buffer exchange column (Zeba spin desalting columns, Pierce) to remove excess glycerol. Specimen was adsorbed onto a thin layer of carbon deposited on an electron microscopy grid and negatively stained for 45 s with 2% of uranyl acetate. Particles were imaged using a transmission electron microscope (Tecnai F20 G2, FEI) equipped with a field emission gun operating at 200 kV. Images were recorded under low-dose condition (total dose of 40–50 e $Å^{-2}$) on a 2,048 × 2,048 CCD camera (Ultrascan 1000, Gatan Inc., Pleasanton) at a magnification of 50,000 resulting in a pixel spacing on the specimen of 0.21 nm.

Random conical tilt reconstructions. The initial reference volumes were obtained by random conical tilt using XMIPP[74] and IMAGIC[75] software packages. Two consecutive images of the same area were taken at 45° and 0° tilt angles under low-dose conditions. A total of 1,546 tilt pairs were selected manually using XMIPP. Untilted images were aligned using iteratively refined two-dimensional class averages as references and multivariance statistical analysis and Hierarchical Ascendant Classification for clustering into 50 class averages with IMAGIC. Fifty volumes calculated from two-dimensional classes were aligned, clustered and averaged using XMIPP MLtomo to compensate for the missing cone, resulting in five random conical tilt (RCT) reconstructions.

Structure refinement. The best volume was used as reference for refinement cycles using a data set of 35,145 untilted molecular images windowed with the Boxer application of the EMAN2 software package[76] and coarsened by two resulting in a pixel spacing of 4.2 Å. Image sorting was found necessary to select the most homogeneous particles since part of the structure was flexible and prevented convergence. Sorting was performed by using first XMIPP then subsequently the RELION software package[77]. Final 3D reconstruction was performed in RELION with 2,361 sorted particles resutling in a structure with 37 Å resolution as estimated by the 0.5 Fourier Shell Correlation criteria. Images were prepared using Chimera software (http://www.cgl.ucsf.edu/chimera).

References

1. Thomas, M. C. & Chiang, C.-M. The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.* **41,** 105–178 (2006).
2. Rhee, H. S. & Pugh, B. F. Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* **483,** 295–301 (2012).
3. Albright, S. R. & Tjian, R. TAFs revisited: more data reveal new twists and confirm old ideas. *Gene* **242,** 1–13 (2000).
4. Hampsey, M. & Reinberg, D. RNA polymerase II as a control panel for multiple coactivator complexes. *Curr. Opin. Genet. Dev.* **9,** 132–139 (1999).
5. Juven-Gershon, T., Hsu, J. Y., Theisen, J. W. & Kadonaga, J. T. The RNA polymerase II core promoter—the gateway to transcription. *Curr. Opin. Cell. Biol.* **20,** 253–259 (2008).

6. Mohan, W. S., Scheer, E., Wendling, O., Metzger, D. & Tora, L. TAF10 (TAF(II)30) is necessary for TFIID stability and early embryogenesis in mice. *Mol. Cell. Biol.* **23,** 4307–4318 (2003).
7. Kim, T. H. *et al.* A high-resolution map of active promoters in the human genome. *Nature* **436,** 876–880 (2005).
8. Voss, A. K. *et al.* Taube nuss is a novel gene essential for the survival of pluripotent cells of early mouse embryos. *Development* **127,** 5449–5461 (2000).
9. Gegonne, A. *et al.* The general transcription factor TAF7 is essential for embryonic development but not essential for the survival or differentiation of mature T cells. *Mol. Cell. Biol.* **32,** 1984–1997 (2012).
10. Vermeulen, M. *et al.* Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131,** 58–69 (2007).
11. Verrijzer, C. P., Chen, J. L., Yokomori, K. & Tjian, R. Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. *Cell* **81,** 1115–1125 (1995).
12. Bieniossek, C. *et al.* The architecture of human general transcription factor TFIID core complex. *Nature* **493,** 699–702 (2013).
13. Papai, G. *et al.* Mapping the initiator binding Taf2 subunit in the structure of hydrated yeast TFIID. *Structure* **17,** 363–373 (2009).
14. Grob, P. *et al.* Cryo-electron microscopy studies of human TFIID: conformational breathing in the integration of gene regulatory cues. *Structure* **14,** 511–520 (2006).
15. Cianfrocco, M. A. *et al.* Human TFIID binds to core promoter DNA in a reorganized structural state. *Cell* **152,** 120–131 (2013).
16. Brand, M., Leurent, C., Mallouh, V., Tora, L. & Schultz, P. Three-dimensional structures of the TAFII-containing complexes TFIID and TFTC. *Science* **286,** 2151–2153 (1999).
17. Matangkasombut, O., Auty, R. & Buratowski, S. Structure and function of the TFIID complex. *Adv. Protein Chem.* **67,** 67–92 (2004).
18. Maston, G. A. *et al.* Non-canonical TAF complexes regulate active promoters in human embryonic stem cells. *Elife* **1,** e00068 (2012).
19. Goodrich, J. A. & Tjian, R. Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation. *Nat. Rev. Genet.* **11,** 549–558 (2010).
20. Müller, F., Zaucker, A. & Tora, L. Developmental regulation of transcription initiation: more than just changing the actors. *Curr. Opin. Genet. Dev.* **20,** 533–540 (2010).
21. Wright, K. J., Marr, M. T. & Tjian, R. TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. *Proc. Natl Acad. Sci. USA* **103,** 12347–12352 (2006).
22. Werten, S. *et al.* Crystal structure of a subcomplex of human transcription factor TFIID formed by TATA binding protein-associated factors hTAF4 (hTAF(II)135) and hTAF12 (hTAF(II)20. *J. Biol. Chem.* **277,** 45502–45509 (2002).
23. Xie, X. *et al.* Structural similarity between TAFs and the heterotetrameric core of the histone octamer. *Nature* **380,** 316–322 (1996).
24. Gangloff, Y. G. *et al.* Histone folds mediate selective heterodimerization of yeast TAF(II)25 with TFIID components yTAF(II)47 and yTAF(II)65 and with SAGA component ySPT7. *Mol. Cell. Biol.* **21,** 1841–1853 (2001).
25. Birck, C. *et al.* Human TAF(II)28 and TAF(II)18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family. *Cell* **94,** 239–249 (1998).
26. Metzger, D., Scheer, E., Soldatov, A. & Tora, L. Mammalian TAF(II)30 is required for cell cycle progression and specific cellular differentiation programmes. *EMBO J.* **18,** 4823–4834 (1999).
27. Tatarakis, A. *et al.* Dominant and redundant functions of TFIID involved in the regulation of hepatic genes. *Mol. Cell* **31,** 531–543 (2008).
28. Guermah, M., Ge, K., Chiang, C. M. & Roeder, R. G. The TBN protein, which is essential for early embryonic mouse development, is an inducible TAFII implicated in adipogenesis. *Mol. Cell* **12,** 991–1001 (2003).
29. Hernández-Hernández, A. & Ferrús, A. Prodos is a conserved transcriptional regulator that interacts with dTAF(II)16 in Drosophila melanogaster. *Mol. Cell. Biol.* **21,** 614–623 (2001).
30. Soutoglou, E. *et al.* The nuclear import of TAF10 is regulated by one of its three histone fold domain-containing interaction partners. *Mol. Cell. Biol.* **25,** 4092–4104 (2005).
31. Kaufmann, J., Ahrens, K., Koop, R., Smale, S. T. & Müller, R. CIF150, a human cofactor for transcription factor IID-dependent initiator function. *Mol. Cell. Biol.* **18,** 233–239 (1998).
32. Kaufmann, J., Verrijzer, C. P., Shao, J. & Smale, S. T. CIF, an essential cofactor for TFIID-dependent initiator function. *Genes Dev.* **10,** 873–886 (1996).
33. Hansen, S. K. & Tjian, R. TAFs and TFIIA mediate differential utilization of the tandem Adh promoters. *Cell* **82,** 565–575 (1995).
34. Martinez, E. *et al.* Novel cofactors and TFIIA mediate functional core promoter selectivity by the human TAFII150-containing TFIID complex. *Mol. Cell. Biol.* **18,** 6571–6583 (1998).
35. Najmabadi, H. *et al.* Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature* **478,** 57–63 (2011).
36. Halevy, A. *et al.* Microcephaly-thin corpus callosum syndrome maps to 8q23.2-q24.12. *Pediatr. Neurol.* **46,** 363–368 (2012).
37. Benesch, J. L. & Robinson, C. V. Mass spectrometry of macromolecular assemblies: preservation and dissociation. *Curr. Opin. Struct. Biol.* **16,** 245–251 (2006).
38. Nie, Y., Bellon-Echeverria, I., Trowitzsch, S., Bieniossek, C. & Berger, I. Multiprotein complex production in insect cells by using polyproteins. *Methods Mol. Biol.* **1091,** 131–141 (2014).
39. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **319,** 1097–1113 (2002).
40. Kouskouti, A., Scheer, E., Staub, A., Tora, L. & Talianidis, I. Gene-specific modulation of TAF10 function by SET9-mediated methylation. *Mol. Cell* **14,** 175–182 (2004).
41. Chen, Z. A. *et al.* Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* **29,** 717–726 (2010).
42. Conti, E., Uy, M., Leighton, L., Blobel, G. & Kuriyan, J. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell* **94,** 193–204 (1998).
43. Hahn, S. Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* **11,** 394–403 (2004).
44. Kochan, G. *et al.* Crystal structures of the endoplasmic reticulum aminopeptidase-1 (ERAP1) reveal the molecular basis for N-terminal peptide trimming. *Proc. Natl Acad. Sci. USA* **108,** 7745–7750 (2011).
45. Jacq, X. *et al.* Human TAFII30 is present in a distinct TFIID complex and is required for transcriptional activation by the estrogen receptor. *Cell* **79,** 107–117 (1994).
46. Müller, F. & Tora, L. The multicoloured world of promoter recognition complexes. *EMBO J.* **23,** 2–8 (2004).
47. Bell, B. & Tora, L. Regulation of gene expression by multiple forms of TFIID and other novel TAFII-containing complexes. *Exp. Cell Res.* **246,** 11–19 (1999).
48. Müller, F., Demeny, M. A. & Tora, L. New problems in RNA polymerase II transcription initiation: matching the diversity of core promoters with a variety of promoter recognition factors. *J. Biol. Chem.* **282,** 14685–14689 (2007).
49. Yatherajam, G., Zhang, L., Kraemer, S. M. & Stargell, L. A. Protein-protein interaction map for yeast TFIID. *Nucleic Acids Res.* **31,** 1252–1260 (2003).
50. Rona, G. *et al.* Phosphorylation adjacent to the nuclear localization signal of human dUTPase abolishes nuclear import: structural and mechanistic insights. *Acta Crystallogr. D Biol. Crystallogr.* **69,** 2495–2505 (2013).
51. Nardozzi, J. D., Lott, K. & Cingolani, G. Phosphorylation meets nuclear import: a review. *Cell Commun. Signal.* **8,** 32 (2010).
52. Kitamura, R. *et al.* Nuclear import of Epstein-Barr virus nuclear antigen 1 mediated by NPI-1 (Importin alpha5) is up- and down-regulated by phosphorylation of the nuclear localization signal for which Lys379 and Arg380 are essential. *J. Virol.* **80,** 1979–1991 (2006).
53. Fitzgerald, D. J. *et al.* Protein complex expression by using multigene baculoviral vectors. *Nat. Methods* **3,** 1021–1032 (2006).
54. Boivin, S. & Hart, D. J. Interaction of the influenza A virus polymerase PB2 C-terminal region with importin alpha isoforms provides insights into host adaptation and polymerase assembly. *J. Biol. Chem.* **286,** 10439–10448 (2011).
55. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23,** 2947–2948 (2007).
56. Gouet, P., Robert, X. & Courcelle, E. ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **31,** 3320–3323 (2003).
57. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132 (2010).
58. Winn, M. D. An overview of the CCP4 project in protein crystallography: an example of a collaborative project. *J. Synchrotron Radiat.* **10,** 23–25 (2003).
59. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D Biol. Crystallogr.* **63,** 32–41 (2007).
60. Strong, M. *et al.* Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **103,** 8060–8065 (2006).
61. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132 (2004).
62. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1948–1954 (2002).
63. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78,** 1606–1619 (2000).
64. Hayes, D., Laue, T. M. & Philo, J. *Program Sednterp: Sedimentation Interpretation Program* (Alliance Protein Laboratories, 1995).
65. Hernandez, H. & Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat. Protoc.* **2,** 715–726 (2007).
66. Vizcaino, J. A. *et al.* ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **32,** 223–226 (2014).

67. Muratoglu, S. *et al.* Two different Drosophila ADA2 homologues are present in distinct GCN5 histone acetyltransferase-containing complexes. *Mol. Cell. Biol.* **23,** 306–321 (2003).
68. Demény, M. A. *et al.* Identification of a small TAF complex and its role in the assembly of TAF-containing complexes. *PLoS ONE* **2,** e316 (2007).
69. Florens, L. & Washburn, M. P. Proteomic analysis by multidimensional protein identification technology. *Methods Mol. Biol.* **328,** 159–175 (2006).
70. McDonald, W. H., Ohi, R., Miyamoto, D. T., Mitchison, T. J. & Yates, J. R. Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: single-dimension LC-MS/MS, 2-phase MudPIT, and 3-phase MudPIT. *Intern. J. Mass Spectrom.* **219,** 245–251 (2002).
71. Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5,** 976–989 (1994).
72. Zybailov, B. *et al.* Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. *J. Proteome Res.* **5,** 2339–2347 (2006).
73. Stark, H. GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. *Methods Enzymol.* **481,** 109–126 (2010).
74. Sorzano, C. O. *et al.* XMIPP: a new generation of an open-source image processing package for electron microscopy. *J. Struct. Biol.* **148,** 194–204 (2004).
75. van Heel, M., Harauz, G., Orlova, E. V., Schmidt, R. & Schatz, M. A new generation of the IMAGIC image processing system. *J. Struct. Biol.* **116,** 17–24 (1996).
76. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157,** 38–46 (2007).
77. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180,** 519–530 (2012).

Acknowledgements

We thank the ESRF-EMBL Joint Structural Biology Group for access to ESRF beamlines and the EMBL-ESRF-ILL-INS Partnership for Structural Biology for access to structural biology instrumentation and protein analysis and characterization facilities. We are grateful to Andrew Thompson (SOLEIL) for support during data collection and structure solving. We also thank Aurelian Deniaud (EMBL) and Andrés Palencia (EMBL) for technical help with AUC, SPR and ITC experiments. S.T. was a European Commission (EC) Framework Program (FP) 7 Marie Curie post-doctoral fellow. I.B. is supported by the EC FP7 ComplexINC project (contract number 279039). C.S. is supported by a European Research Council (ERC) Starting Grant and LT by an ERC Advanced grant (BIRTOACTION, grant number 340551). I.B., L.T. and P.S. acknowledge support from the Agence National de la Recherche ANR Projet Blanc DiscoverIID and of the French Infrastructure for Integrated Structural Biology (FRISBI; ANR-10-INSB-05-01) and INSTRUCT as part of the European Strategy Forum on Research Infrastructures (ESFRI). J.R. is a Wellcome Trust Senior Research Fellow (grant number 084229). The Wellcome Trust Centre for Cell Biology is supported by core grant numbers 077707 and 092076, and the work was also supported by Wellcome Trust instrument grant number 091020.

Author contributions

S.T., C.V., L.T. and I.B. designed and interpreted the experiments. S.T., C.V. and M.H. produced and purified all proteins and performed characterizations. S.T. determined the crystal structures. E.S. prepared the antibodies and carried out all the anti-TAF2 IPs. V.C. and M.F. performed MudPIT experiments and M.F. and L.T. analysed them. I.-O.E. and C.V.R. performed and analysed native mass spectrometry experiments. S.C. and L.T. prepared and interpreted cell-based experiments. G.P., C.S. and P.S. prepared EM grids, collected and analysed data and calculated reconstructions. J.Z. and J.R. performed and analysed CLMS experiments. S.T., L.T. and I.B. wrote the manuscript with input from all authors.

Additional information

Accession codes: Atomic coordinates and structure factors are deposited in the Protein Data Bank (PDB 4WV4; 4WV6). Mass spectrometry data has been deposited in the ProteomeXchange database (PXD001454; PXD001427).

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Trowitzsch, S. *et al.* Cytoplasmic TAF2–TAF8–TAF10 complex provides evidence for nuclear holo–TFIID assembly from preformed submodules. *Nat. Commun.* 6:6011 doi: 10.1038/ncomms7011 (2015).

SUPPLEMENTARY INFORMATION

Cytoplasmic TAF2-TAF8-TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules.

Simon Trowitzsch[1,2], Cristina Viola[1,2], Elisabeth Scheer[3], Sascha Conic[3], Virginie Chavant[4], Marjorie Fournier[3], Gabor Papai[5], Ima-Obong Ebong[6], Christiane Schaffitzel[1,2], Juan Zou[7], Matthias Haffke[1,2], Juri Rappsilber[7,8], Carol V. Robinson[6], Patrick Schultz[5], Laszlo Tora[3,#] & Imre Berger[1,2,9,#]

[1] European Molecular Biology Laboratory, Grenoble Outstation, 6 rue Jules Horowitz, 38042 Grenoble, France

[2] Unit for Virus Host-Cell Interactions, Univ. Grenoble Alpes-EMBL-CNRS, 6 rue Jules Horowitz, 38042 Grenoble, France.

[3] Cellular Signaling and Nuclear Dynamics Program, Institut de Génétique et de Biologie Moléculaire et Cellulaire, UMR 7104, INSERM U964, 1 rue Laurent Fries, 67404 Illkirch, France.

[4] Proteomics Platform, Institut de Génétique et de Biologie Moléculaire et Cellulaire, UMR 7104, INSERM U964, 1 rue Laurent Fries, 67404 Illkirch, France.

[5] Integrated Structural Biology Department, Institut de Génétique et de Biologie Moléculaire et Cellulaire, UMR 7104, INSERM U964, 1 rue Laurent Fries, 67404 Illkirch, France.

[6] Chemistry Research Laboratory, University of Oxford, South Parks Road, Oxford, Oxfordshire, OX1 3TA, United Kingdom.

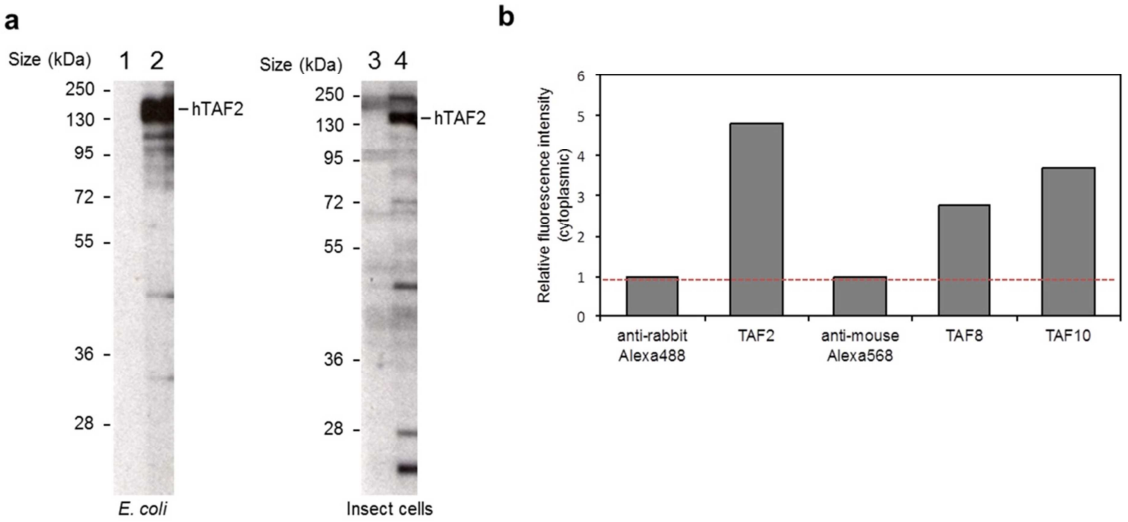
[7] Wellcome Trust Centre for Cell Biology, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JR, United Kingdom.

[8] Institute of Bioanalytics, Department of Biotechnology, Technische Universität Berlin, 13353 Berlin, Germany

[9] School of Biochemistry, Bristol University, Bristol BS8 1TD, United Kingdom

[#] Corresponding authors: Laszlo Tora (laszlo.tora@igbmc.fr) & Imre Berger (iberger@embl.fr)

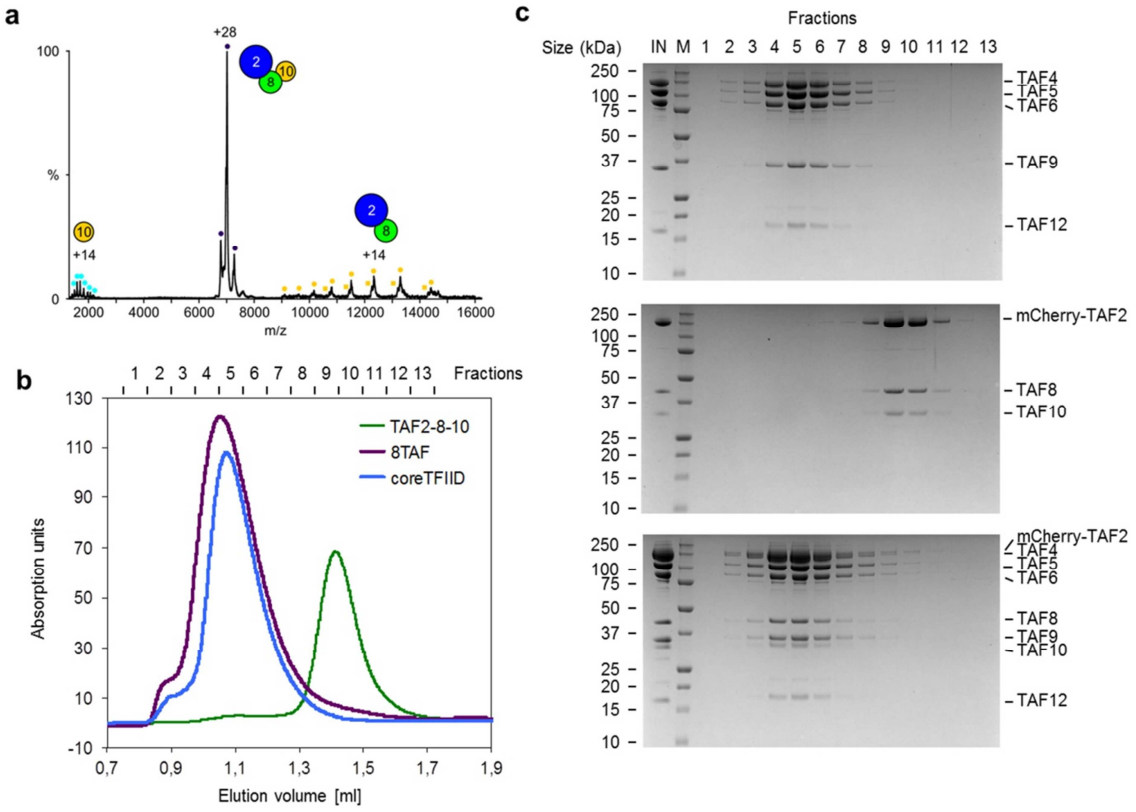
Supplementary Figure 1



Analysis of polyclonal anti-TAF2 antibodies and immunofluorescence (IF) measurements.

(**a**) Analysis of polyclonal anti-TAF2 antibodies. Crude extracts of *E. coli* (left) or baculovirus-infected insect cells (right) expressing 6His-tagged human TAF2 were resolved on SDS-PAGE. Proteins were blotted onto nitrocellulose membranes and incubated with pre-immune serum (lanes 1 and 3) or with the non-purified anti-TAF2 serum 3038 (lanes 2 and 4) taken from rabbits, which were immunized with recombinant human TAF2 protein. Protein size markers are indicated on the left of each blot. The polyclonal antibody recognizes recombinant TAF2 from *E. coli* and baculovirus-infected insect cells. (**b**) Quantification of fluorescence intensities in the cytoplasm of HeLa cells by IF. Cytoplasmic fluorescence intensities of control cells treated only with fluorescently labeled secondary antibodies (anti-rabbit Alexa488 and anti-mouse Alexa568) were compared to cytoplasmic fluorescence intensities of cells treated with anti-TAF2 + anti-TAF8 or anti-TAF2 + anti-TAF10 primary antibodies and the same set of secondary antibodies. Fluorescence intensities were normalized to the background controls.

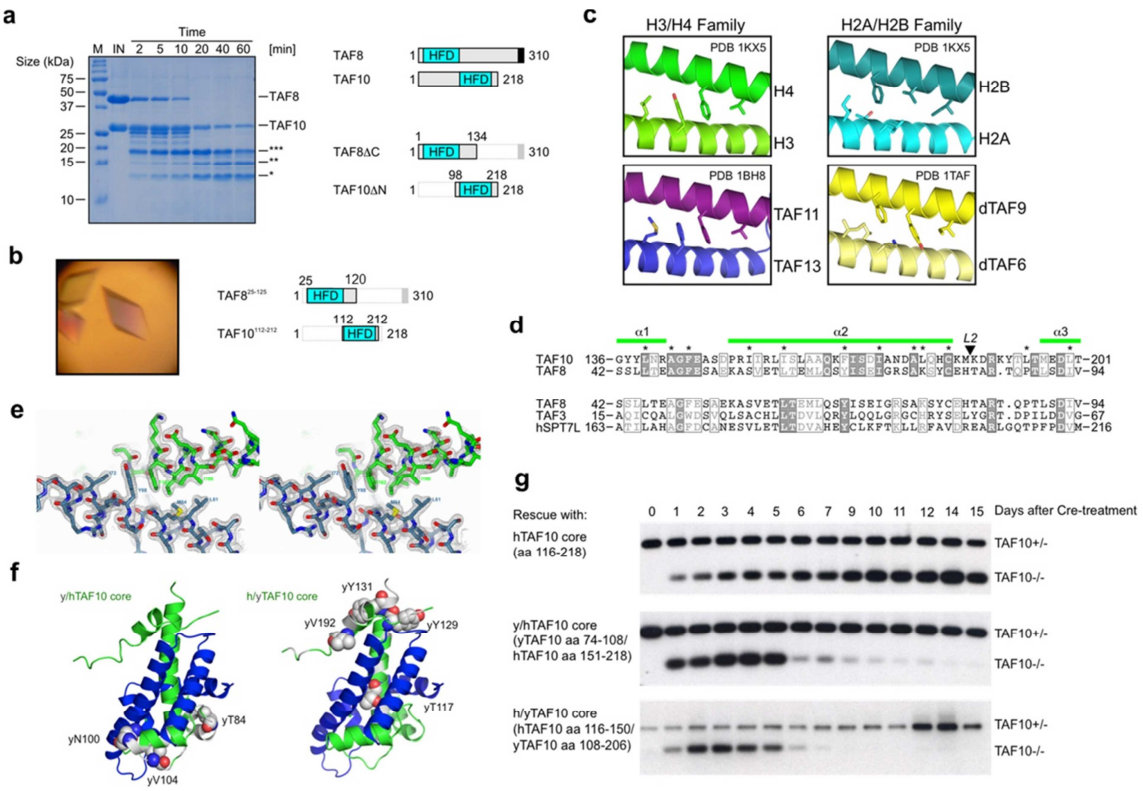
Supplementary Figure 2



Native mass-spectrometry of a recombinant TAF2-8-10 module and incorporation of the TAF2-8-10 module into core-TFIID.

(**a**) Recombinant TAF2-8-10 complexes were electrosprayed from an aqueous ammonium acetate solution. The TAF2-8-10 module (purple dots) centers on a charge state at 7000 m/z. Charge states at around 2000 m/z (light blue dots) and 12000 m/z (yellow dots) correspond to minor amounts of TAF10 and a TAF2-8 complex, respectively. Proteins and protein complexes are schematically shown as circles. (**b** and **c**) Binding analysis of the TAF2-8-10 module with core-TFIID using size exclusion chromatography (SEC). TAF2-8-10 module, core-TFIID (TAF4, 5, 6, 9, 12) and a mixture of TAF2-8-10 module and core-TFIID were analyzed. (**b**) Elution profiles of TAF2-8-10 module (green), core-TFIID (blue) and TAF2-8-10 mixed in stoichiometric molar ratio with core-TFIID (purple) are plotted in absorption units at 280 nm versus elution volume. Fractions are numbered (top of graph). (**c**) SDS-PAGE analyses of the eluted SEC fractions are shown. Molecular masses of protein standards are indicated on the left of gel sections. Protein denominations are shown on the right. IN, input sample.

Supplementary Figure 3

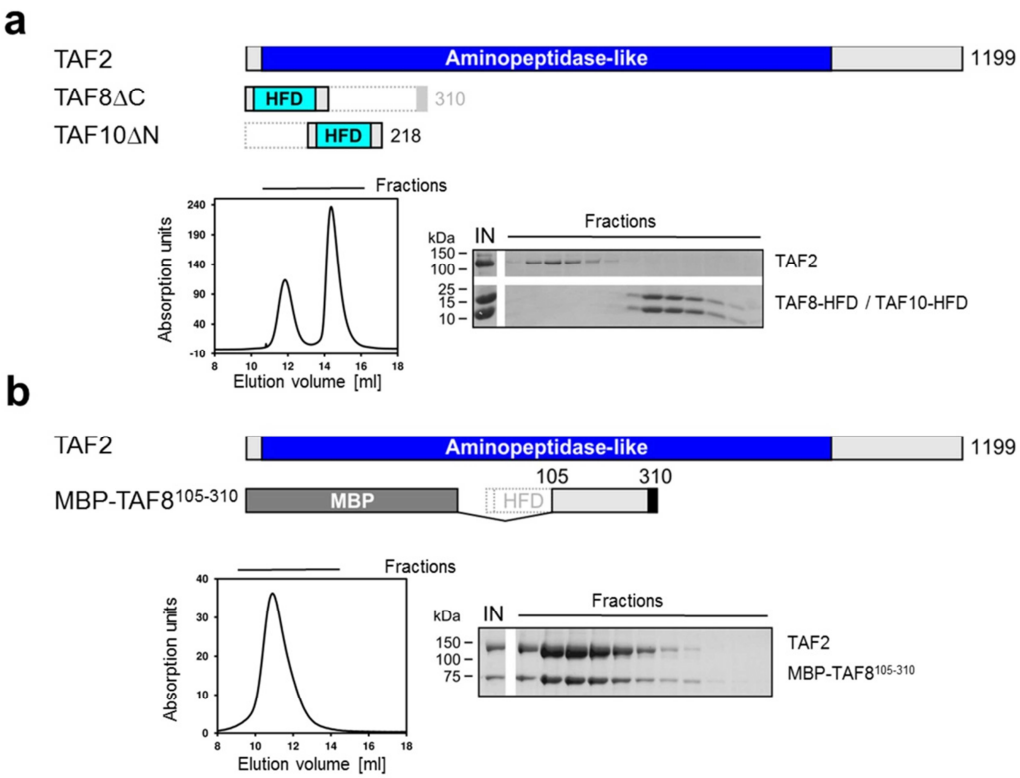


Structural analysis of a TAF8-10 complex.

(**a**) Time course of a limited proteolysis experiment with TAF8-10 using Chymotrypsin (left). Time points, protein size markers and protein identities are indicated. ***, TAF8 fragment spanning residues 1-159; **, TAF8 fragment spanning residues 1-134; *, TAF10 fragment spanning residues 98-218; IN, Input sample. HFD, histone fold domain. Bar diagrams of the proteins TAF8 and TAF10 are indicated as shown in Fig. 1c. Domain boundaries of the core TAF8-10 complex (TAF8ΔC and TAF10ΔN) are highlighted. (**b**) Image of crystals grown from a refined TAF8-10 construct (TAF8 residues 25-120 and TAF10 residues 112-212) with bar diagrams of the protein constructs. (**c**) Comparison of the central α helices of other histone fold-containing structures (PDB IDs 1KX5, 1BH8, 1TAF) showing an array of residues at the crossing of the helices. (**d**) Sequence alignment of the L1 loop regions of TAF8 and TAF10 (top). Putative L1 regions of TAF3 and human SPT7L are aligned to TAF8 (bottom). Start and end residues of the aligned sequences are indicated. Residues highlighted in Fig. 3d,e are marked by asterisks. Secondary structure elements are shown for TAF10 at the top of the alignment. Note that the L2 loop of TAF10 was removed for clarity (*L2* arrow). (**e**) Representative section of the $2F_o$-F_c electron density map (mesh) of the TAF8-10 crystal structure is shown in a stereo view, contoured at 1.5σ around the central helices of TAF8 (in blue) and TAF10 (in green). (**f**) Ribbon representations of models of the TAF8-10 complex with chimeric TAF10 molecules. The two chimeras comprise residues 74-108 of yeast and residues 151-218 of human TAF10 (left) or residues 116-150 of human and residues 108-206 of yeast TAF10 (right). Substituted yeast TAF10 residues are shown in space filling representation, colored in grey. Substituted yeast TAF10 residues which would give

rise to steric clashes, are highlighted. Color-coding is as in panel (e). (**g**) Conditional rescue experiments of TAF10$^{-/-}$ F9 embryonic carcinoma cells with TAF10 HFD and TAF10 human/yeast chimeric constructs spanning the TAF10 histone fold domain. Linearized plasmids encoding for human TAF10 (residues 116-218) and chimeric TAF10 as described in panel (f) were used to electroporate L^{-}/L2TAF10 F9 cells as described [1]. The excision of exon 2 is monitored by PCR analysis of the genomic DNA.

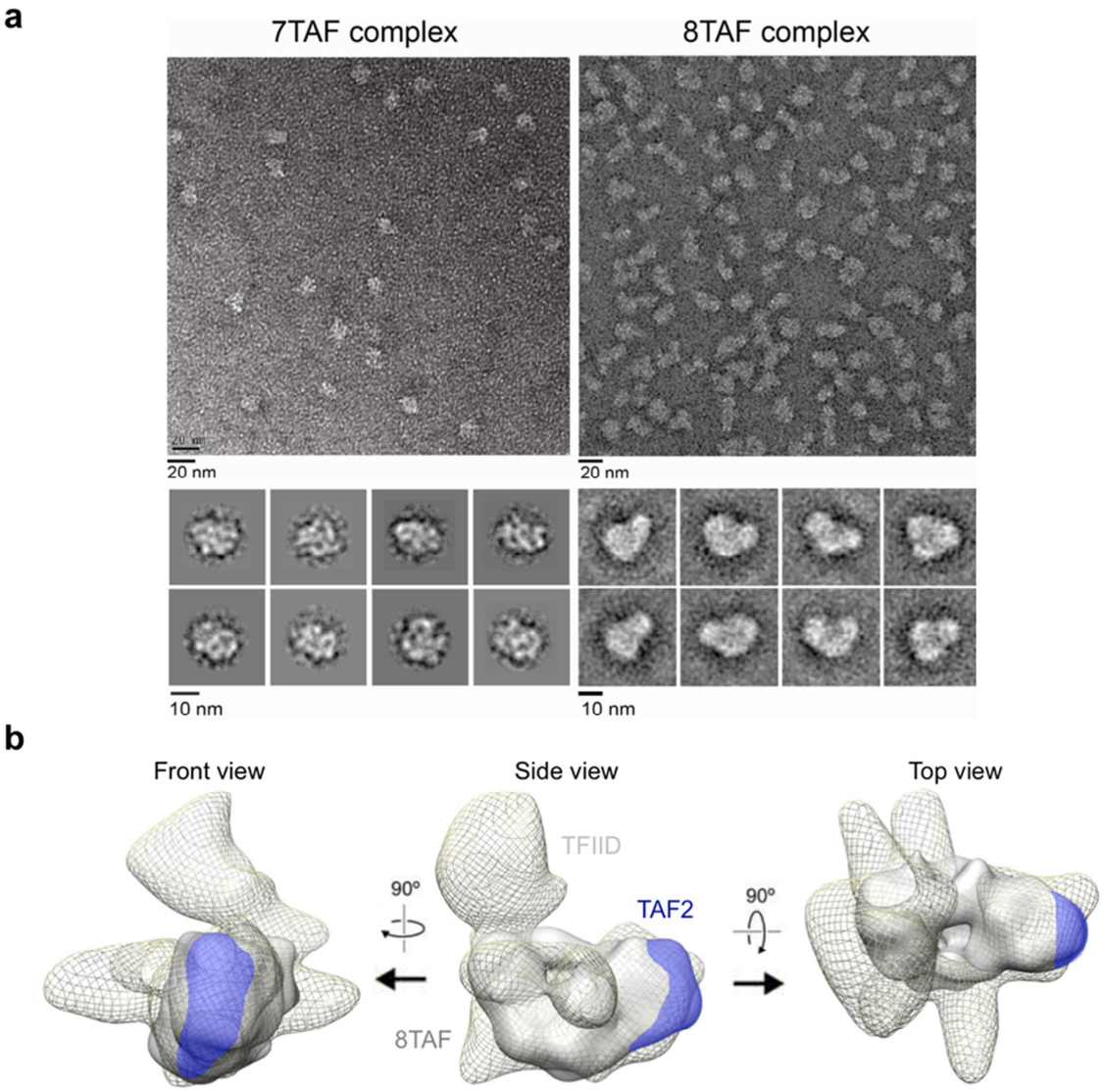
Supplementary Figure 4



TAF2 interacts with the C-terminal region of TAF8 but not with the core complex of TAF8-10.

(**a**) Binding analysis of TAF2 with the core construct TAF8ΔC-TAF10ΔN using gel filtration. The elution profile monitored at an absorption wavelength of 280 nm versus elution volume is shown on the left and the SDS-PAGE analysis of peak fractions is shown on the right. (**b**) Similar binding experiment as in (a) but with an MBP-fusion construct of the unstructured C-terminal region of TAF8 (TAF8 residues 105-310). Protein size markers and protein identities are indicated. IN, input sample.

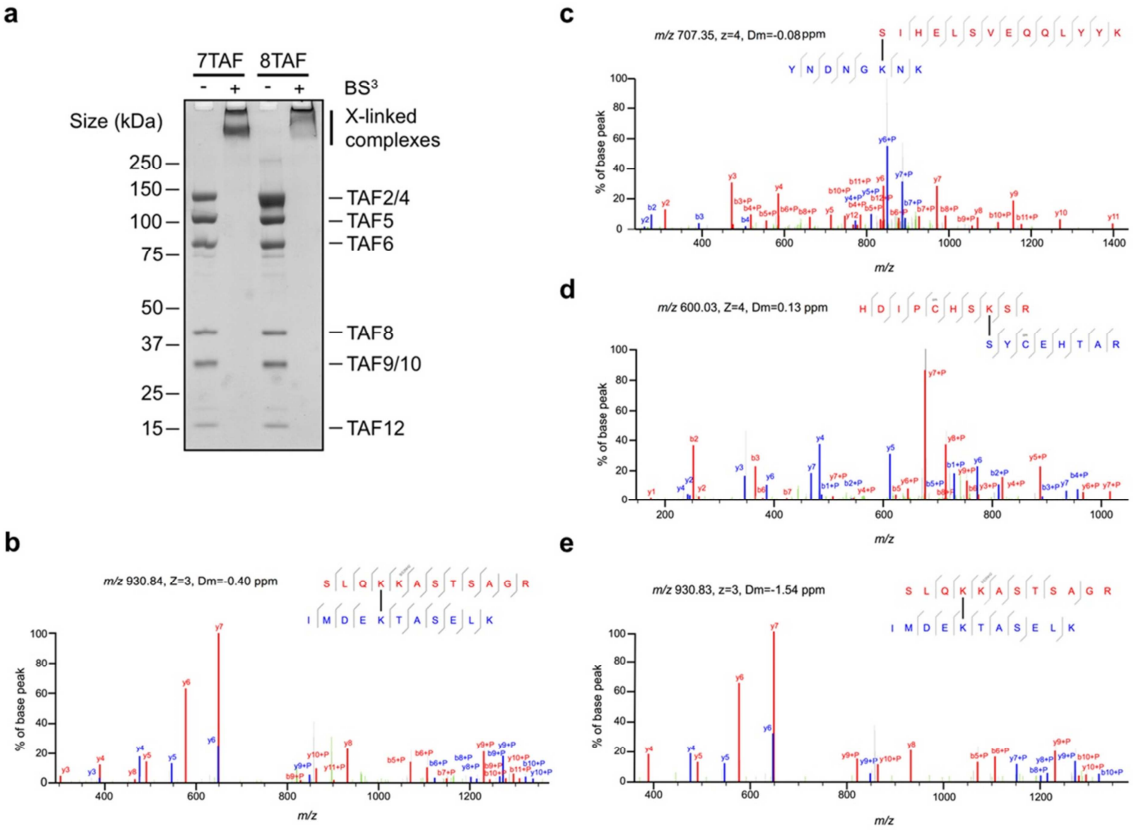
Supplementary Figure 5



Electron microscopy of 7TAF and 8TAF complexes

(**a**) Electron micrographs and 2D class averages of 7TAF and 8TAF complexes. A section of electron micrographs from 7TAF complex consisting of TAF4, 5, 6, 8, 9, 10 and 12 is shown on the left, with representative 2D class averages shown below. A similar section from 8TAF complex comprising TAF2, 4, 5, 6, 8, 9, 10 and 12 is shown on the right, with representative 2D class averages below. Scale bars are indicated. 8TAF complex has an elongated shape as compared to more compact 7TAF complex. Additional density corresponding to TAF2 is located at one side of the 8TAF complex, adopting flexible conformations. (**b**) 3D single particle EM reconstruction of negatively stained 8TAF complex (grey) superimposed on the EM density of the holo-TFIID complex (EMD-1195, grey mesh) is shown in three views, related by a 90° rotation as indicated. Density attributed to TAF2 in the 8TAF complex is highlighted in blue.

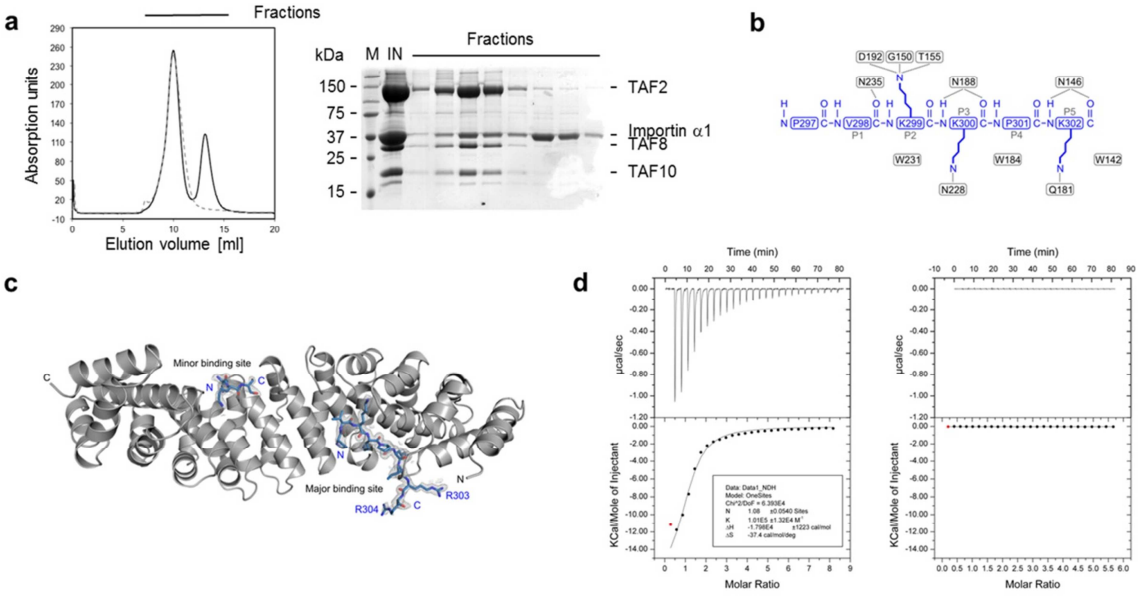
Supplementary Figure 6



Cross-linking of 7TAF and 8TAF complexes using bifunctional crosslinker BS^3 and analysis of cross-linked peptides by mass spectrometry.

(**a**) Cross-linking efficiency of 7TAF and 8TAF complexes was assessed on NuPAGE Novex 3-8 % Tris-Acetate gels (Invitrogen). Identical amounts of 7TAF and 8TAF samples before cross-linking (-) and after cross-linking (+) were loaded on each lane. Protein size markers are shown on the left and protein identities on the right. Cross-linked complexes are indicated. (**b-e**) Representative annotated high resolution spectra of cross-linked peptides derived from 7TAF or 8TAF complexes. (**b**) Linkage TAF9 (red) K134 – TAF5 (blue) K531 observed in SLQK(xl)KASTSAGR / IMDEK(xl)TASELK (*m/z* 930,84) from 7TAF. (**c**) Linkage TAF6 (red) S212 – TAF2 (blue) K786 observed in S(xl)IHELSVEQQLYYK / YNDNGK(xl)NK (*m/z* 707,35) from 8TAF. (**d**) Linkage TAF2 (red) K595 – TAF8 (blue) S78 observed in HDIPCHSK(xl)SR / S(xl)YCEHTAR (*m/z* 600,03) from 8TAF. (**e**) Linkage TAF9 (red) K134 – TAF5 (blue) K531 observed in SLQK(xl)KASTSAGR / IMDEK(xl)TASELK (*m/z* 930,83) from 8TAF.

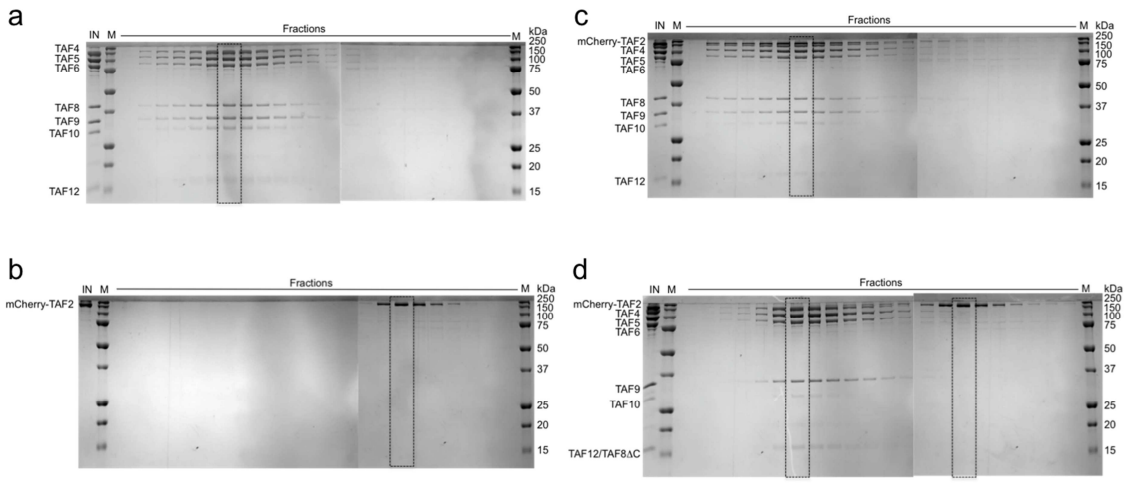
Supplementary Figure 7



Structural and biochemical characterization of the putative nuclear import particle comprising TAF2-8-10-Importin α1.

(**a**) Binding experiment as in Supplementary Fig. 4a, but with the TAF2-8-10 complex mixed with a two-fold molar excess of Importin a1$^{\Delta IBB}$. Elution profile of the mixture is shown as a black line. The dotted line shows the elution profile of the rechromatographed material pooled from the first peak (at around 10 ml). SDS-PAGE analysis of peak fractions is shown on the right. (**b**) Schematic representation of the interactions between Importin α1 and the NLS of TAF8. Residues engaged in salt bridges, van der Waals contacts or hydrogen bondings are indicated by dashed lines. Backbone amino and carbonyl groups of the NLS peptide are schematically drawn. Residue positions are indicated. (**c**) Structure of Importin α1 with an NLS peptide of TAF8. Importin α1 molecule (grey) is shown in cartoon representation and the TAF8 peptide as sticks in blue. The $2F_o$-F_c density map contoured at 1σ around the NLS peptide fragments is shown as a grey mesh. TAF8 residues R303 and R304, which are stabilized by crystal contacts but are not engaged in Importin α1 binding, are indicated. Major and minor NLS-binding sites on the Importin α1 molecule are denoted. (**d**) TAF8-NLS peptide binding to Importin α1 assayed by isothermal titration calorimetry. The upper panel shows the added heat to the cell over time with successive additions. The excess heat added per addition was integrated from the upper panel and plotted in the lower panel as a function of the ratio of the concentration of the NLS and Importin α1 in the cell. The right panel shows a control run without Importin α1 in the cell to assess heat of dilution of the peptide.

Supplementary Figure 8



TAF8 promotes TAF2 incorporation in TFIID. Original scans of Coomassie brilliant blue-stained polyacrylamide gels. Dashed boxes indicate sections of the gels that are shown in Fig. 5b. SDS-PAGE analyses of SEC runs using 7TAF (**a**), TAF2 (**b**), 8TAF (**c**), and 7TAFΔ + TAF2 (**d**) are shown.

Supplementary Table 1

MutPIT analysis of TAF2 co-immunoprecipitated proteins from nuclear and cytoplasmic HeLa cell extracts. TFIID subunits specifically enriched in TAF2 immunoprecipitations (IPs) as compared to control IP samples are indicated by protein sequence coverage (%), unique peptides and spectral counts.

	TAF2 IP (nuclear)			**TAF2 IP (cytoplasmic)**		
TFIID subunit	Sequence coverage (%)	Unique peptides	Spectral counts	Sequence coverage (%)	Unique peptides	Spectral counts
TAF1	13.19	14	56			
TAF2	33.19	32	573	30.94	25	424
TAF3	4.09	3	6			
TAF4	19.63	16	169			
TAF4B	9.98	6	19			
TAF5	38.63	21	131			
TAF6	36.78	20	207			
TAF7	26.65	7	44			
TAF8	48.71	10	100	36.77	6	55
TAF9	40.53	8	45			
TAF9B	23.90	5	15			
TAF10	27.06	3	34	20.64	2	15
TAF11	30.33	5	21			
TAF12	11.08	2	2			
TAF13	16.94	2	11			
TBP	7.37	2	12			

Supplementary Table 2

Native mass-spectrometry data of TAF2-8-10 complexes.

Protein / protein complex	Measured mass [Da]*	Calculated mass [Da]
TAF8	35028	34984
TAF10	23484 / 23751 (23618 ± 134)	23613
TAF8-10	57975 / 58976 (58475 ± 501)	58579
TAF2	136364 / 137526 (136945 ± 581)	137030
TAF2-8	171548 / 172754 (172151 ± 603)	172014
TAF2-8-10	195222 / 196372 (195797 ± 575)	195609

*Two series of peaks are observed in the spectra for TAF10 and TAF2 (and therefore also complexes containing these TAFs), likely due to post-translational modification. Mass averages are provided in brackets.

Supplementary Table 3

Intermolecular BS^3 protein-protein cross-links of 7TAF and 8TAF complexes.*

Cross-linked proteins (prot1-prot2)	7TAF						8TAF					
	prot1	aa	res	prot2	aa	res	prot1	aa	res	prot2	aa	res
TAF2-TAF5							TAF2	595	K	TAF5	531	K
TAF2-TAF6							TAF2	786	K	TAF6	212	S
TAF2-TAF8							TAF2	595	K	TAF8	78	S
							TAF2	786	K	TAF8	178	K
							TAF2	1110	K	TAF8	178	K
TAF2-TAF9							TAF2	595	K	TAF9	135	K
TAF4-TAF5							TAF4	887	K	TAF5	318	K
							TAF4	888	K	TAF5	292	K
	TAF4	***945***	***K***	***TAF5***	***518***	***K***	***TAF4***	***945***	***K***	***TAF5***	***518***	***K***
	TAF4	***955***	***K***	***TAF5***	***518***	***K***	***TAF4***	***955***	***K***	***TAF5***	***518***	***K***
	TAF4	958	K	TAF5	437	K						
	TAF4	***958***	***K***	***TAF5***	***440***	***K***	***TAF4***	***958***	***K***	***TAF5***	***440***	***K***
							TAF4	971	K	TAF5	407	K
TAF4-TAF9	***TAF4***	***929***	***K***	***TAF9***	***130***	***K***	***TAF4***	***929***	***K***	***TAF9***	***130***	***K***
	TAF4	***945***	***K***	***TAF9***	***108***	***K***	***TAF4***	***945***	***K***	***TAF9***	***108***	***K***
	TAF4	***955***	***K***	***TAF9***	***108***	***K***	***TAF4***	***955***	***K***	***TAF9***	***108***	***K***
	TAF4	958	K	TAF9	135	K						
	TAF4	989	K	TAF9	108	K						
TAF4-TAF12	***TAF4***	***868***	***K***	***TAF12***	***62***	***K***	***TAF4***	***868***	***K***	***TAF12***	***62***	***K***
							TAF4	868	K	TAF12	141	K
TAF5-TAF6	***TAF5***	***318***	***K***	***TAF6***	***8***	***K***	***TAF5***	***318***	***K***	***TAF6***	***8***	***K***
	TAF5	318	K	TAF6	48	K						
							TAF5	531	K	TAF6	389	K
TAF5-TAF8	TAF5	407	K	TAF8	178	K						
	TAF5	416	K	TAF8	178	K						
	TAF5	531	K	TAF8	178	K						
TAF5-TAF9	***TAF5***	***531***	***K***	***TAF9***	***134***	***K***	***TAF5***	***531***	***K***	***TAF9***	***134***	***K***
	TAF5	***531***	***K***	***TAF9***	***135***	***K***	***TAF5***	***531***	***K***	***TAF9***	***135***	***K***
TAF6-TAF8	***TAF6***	***65***	***K***	***TAF8***	***20***	***K***	***TAF6***	***65***	***K***	***TAF8***	***20***	***K***
	TAF6	***65***	***K***	***TAF8***	***78***	***S***	***TAF6***	***65***	***K***	***TAF8***	***78***	***S***
	TAF6	65	K	TAF8	178	K						
	TAF6	***110***	***K***	***TAF8***	***20***	***K***	***TAF6***	***110***	***K***	***TAF8***	***20***	***K***
	TAF6	166	K	TAF8	178	K						
	TAF6	169	K	TAF8	178	K						
							TAF6	179	K	TAF8	178	K
	TAF6	181	K	TAF8	178	K						
							TAF6	193	S	TAF8	178	K
							TAF6	195	K	TAF8	178	K
	TAF6	***212***	***S***	***TAF8***	***178***	***K***	***TAF6***	***212***	***S***	***TAF8***	***178***	***K***
	TAF6	238	K	TAF8	178	K						
	TAF6	342	K	TAF8	178	K						
	TAF6	361	K	TAF8	178	K						
	TAF6	367	K	TAF8	178	K						
TAF6-TAF9	***TAF6***	***65***	***K***	***TAF9***	***10***	***K***	***TAF6***	***65***	***K***	***TAF9***	***10***	***K***
	TAF6	***65***	***K***	***TAF9***	***24***	***K***	***TAF6***	***65***	***K***	***TAF9***	***24***	***K***
	TAF6	***65***	***K***	***TAF9***	***135***	***K***	***TAF6***	***65***	***K***	***TAF9***	***135***	***K***
							TAF6	110	K	TAF9	10	K
TAF8-TAF9	***TAF8***	***20***	***K***	***TAF9***	***10***	***K***	***TAF8***	***20***	***K***	***TAF9***	***10***	***K***
	TAF8	***178***	***K***	***TAF9***	***135***	***K***	***TAF8***	***178***	***K***	***TAF9***	***135***	***K***
TAF9-TAF12	***TAF9***	***62***	***K***	***TAF12***	***107***	***K***	***TAF9***	***62***	***K***	***TAF12***	***107***	***K***
	TAF9	***62***	***K***	***TAF12***	***114***	***K***	***TAF9***	***62***	***K***	***TAF12***	***114***	***K***
TAF10-TAF12	TAF10	177	K	TAF12	141	K						
							TAF10	189	K	TAF12	157	K

* Mass-spectrometry was used to identify thirty-seven unique cross-links in each sample. Cross-linked residues common to 7TAF and 8TAF are highlighted (bold and italic letters). prot stands for protein; aa for amino acid number; res for residue.

Supplementary Methods

DNA constructs. Coding sequences of full-length TAF8 (Uniprot accession number Q7Z7C8) and TAF10 (Uniprot accession code Q12962) were synthesized at GenScript (New Jersey, USA) as a polyprotein construct and cloned into pPBac vector from the MultiBac suite via restriction sites *Bst*EII and *Rsr*II. The triple alanine mutant of the TAF8-10 polyprotein construct was generated by substituting the *Bst*EII-*Apa*I fragment with a synthetic DNA fragment (GenScript) carrying mutated codons for TAF8 residues 185-187 (DVE to AAA), 222-224 (PYL to AAA) and 293-295 (PYL to AAA). ORFs coding for deletion constructs of TAF10 (with engineered N-terminal, Tobacco Etch virus (TEV)-cleavable deca-histidine tag) and TAF8 were subcloned into MultiBac transfer vectors pFL and pIDC.

The TAF2 coding sequence (UniProt accession number Q6P1X5 VAR_027855) was cloned into a modified pFL vector coding for an engineered N-terminal TEV-cleavable deca-His tag via restriction sites *Sal*I and *Hind*III. The mCherry-TAF2 construct was cloned by inserting the mCherry-coding sequence via the *Sal*I cleavage site into the pFL-HisTEVTAF2 vector. The MBP-TAF2 construct was generated in analogy to the mCherry-TAF2 construct. Transfer vectors were either first fused *in vitro* by Cre-LoxP recombination or directly integrated into the EmBacY baculovirus genome by *in vivo* Tn7 transposition using standard protocols.

For MBP-fusion constructs, coding sequences of truncation versions of the TAF8 protein were amplified via PCR from the synthetic polyprotein construct and cloned into pMAL-c vector (Novagen) with engineered C-terminal hexa-histidine tags via SLIC.

Protein Production and Purification. *E. coli* Rosetta^{TM}(DE3) cells (Novagen) were transformed with plasmids pMAL-c_TAF8_105-310, pMAL-c_TAF8_105-260, pMAL-c_TAF8_141-310, and pMAL-c_TAF8_200-310. Cells were grown in LB broth (Miller's) medium supplemented with 34 µg ml^{-1} Chloramphenicol and 100 µg ml^{-1} Ampicillin at 37 °C. Temperature was decreased to 20 °C at an optical density (OD_{600}) of 0.4. Protein production was induced at an OD_{600} of 0.8 by addition of

isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 µM. Cells were harvested 18 h post induction by centrifugation. Cell pellets were flash frozen in liquid nitrogen and resuspended in lysis buffer (50 mM HEPES-NaOH, pH 7.6, 300 mM NaCl, 5 mM Imidazole, 1 µM Leupeptin, 1 µM Pepstatin, 50 µg ml^{-1} lysozyme) supplemented with 1 tablet of cOmplete EDTA-free protease inhibitor cocktail (Roche). Crude extracts were prepared by sonication and cleared by centrifugation (Beckman JA-20 rotor, 45 min, 20000 rpm, 4 °C). All purification steps were performed at 4 °C on ÄKTA prime and purifier systems (GE Healthcare). Soluble extracts were passed over a 5 ml column of TALON® metal affinity resin (Clontech) equilibrated in lysis buffer. The resin was washed with 10 cv of washing buffer (50 mM HEPES-NaOH, pH 7.6, 300 mM NaCl, 5 mM Imidazole, 0.01% [v/v] NP-40, 1 µM Leupeptin, 1 µM Pepstatin) and 10 cv of washing buffer without NP-40. MBP-fusions were eluted by a linear imidazole gradient of 16 cv into elution buffer (50 mM HEPES-NaOH, pH7.6, 300 mM NaCl, 300 mM Imidazole, 1 µM Leupeptin, 1 µM Pepstatin). Peak fractions were pooled and concentrated to 1 ml via ultrafiltration in 15 ml, 10 MWCO spin concentrators (Millipore). Fusion proteins were further purified to homogeneity via size exclusion chromatography on a Superdex 200 16/60 column (GE Healthcare) in a buffer comprising 25 mM HEPES-NaOH, pH 7.6, 300 mM NaCl and supplemented with cOmplete EDTA-free protease inhibitor cocktail (Roche). Proteins were concentrated to ~ 10 mg ml^{-1}, flash-frozen in liquid nitrogen and stored at – 80 °C in aliquots.

Pellets of baculovirus-infected Sf21 cells expressing TAF2 constructs were resuspended in lysis buffer (50 mM HEPES-NaOH, pH 7.5, 500 mM NaCl, 10 mM Imidazole, 1 µM Leupeptin, 1 µM Pepstatin, supplemented with cOmplete EDTA-free protease inhibitor cocktail) and crude extracts were prepared by sonication and subsequently cleared by centrifugation (Beckman JA-25.50 rotor, 1 h, 25000 rpm, 4 °C). TAF2 constructs were captured from soluble extract via TALON® metal affinity resin (Clontech) in batch. The resin was extensively washed with lysis buffer and TAF2 constructs were eluted in lysis buffer supplemented with 250 mM Imidazole. Proteins TAF2 and mCherry-TAF2 were furthermore purified by ion exchange chromatography using a 5 ml SP-Sepharose HiTrap column (GE Healthcare) and eluted from the column by a linear salt gradient in a buffer comprising 25 mM HEPES-NaOH, pH 7.5, 150 to 1000 mM NaCl, 1 mM □-mercaptoethanol, 1 µM Leupeptin and 1 µM Pepstatin. Prior to ion exchange chromatography the His-tag

was removed by incubating the proteins with Tobacco Etch Virus (TEV) protease (produced in house). TAF2 constructs were finally polished by gel filtration using Superdex200 10/300 or Superose6 10/300 columns (GE Healthcare) in a buffer comprising 25 mM HEPES-NaOH, pH 7.5, 500 mM NaCl, 1 mM β-mercaptoethanol. Proteins were concentrated to ~ 5-20 mg ml^{-1}, flash-frozen in liquid nitrogen and stored at – 80 °C in aliquots.

TAF8-10 constructs were produced and purified essentially as described above for the TAF2 constructs with the exception that the lysis and SEC buffer contained 150 mM NaCl instead of 500 mM, Furthermore, protein complexes were subjected to SEC immediately after elution from the TALON resin. For crystallization purposes, the complexes were subjected to SEC in a buffer comprising 10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM dithiothreitol (DTT).

Supplementary Reference

[1] Kouskouti, A., Scheer, E., Staub, A., Tora, L. & Talianidis, I. Gene-specific modulation of TAF10 function by SET9-mediated methylation. *Mol. Cell* 14, 175-182 (2004).

2. Targeting the replisome with transduced monoclonal antibodies triggers lethal DNA replication stress in cancer cells (D. Desplancq, G. Freund, S. Conic *et al.*; Experimental Cell Research, 2016)

DNA replication is a process that is tightly linked with cell metabolism and division. Any failure of the DNA copying process during S-phase compromises genomic integrity and induces DNA breakage due to replication fork stalling. This circumstance is also used in cancer therapy by chemotherapeutic drugs to induce replication stress, DNA damage and finally apoptosis of the highly proliferating cancer cells. Two essential proteins at the replication fork are the DNA polymerase alpha which performs the DNA synthesis during replication and the DNA sliding clamp called proliferating cell nuclear antigen (PCNA) which acts in concert with the DNA polymerase as processivity factor and is generally known as the “maestro” at the replication fork. However, even if these two protein complexes can be considered as ideal targets for cancer therapy, they cannot be efficiently inhibited to an extent to promote DNA damage-induced cell death due to replication stress with the presently available molecules.

Therefore, in the following study and in accordance with aim **a)** of my project, we generated monoclonal antibodies against the trimeric form of PCNA and delivered them into living cancer cells to target endogenous PCNA. However, in contrast to using the intracellular antibodies to visualize and image endogenous PCNA in living cells, this study aimed for another possible application of intracellular antibodies by using specific inhibiting antibodies. Thus, these antibodies or Fabs can be used to induce replication stress in cancer cells which results in cell death. The inhibition of endogenous PCNA led to irreversible and extensive DNA damage in form of DNA double-strand breaks which induced cell death in a variety of different cancer cell types.

I generated two of the used inhibiting anti-PCNA antibodies called 4D6 and 2B6. Furthermore, I performed the confocal and 3D-SIM imaging of γH2AX which is a well described biomarker for double-strand breaks. IF experiments targeting γH2AX of cells

transduced with the anti-PCNA Fabs of 2H3 showed a highly speckled staining with many γH2AX foci and enlarged nuclei in contrast to control cells without antibody treatment. 3D-SIM imaging of these samples also showed that these foci correspond to clusters of around 10-15 individual spots. These results suggested that the antibody treatment induced a huge amount of DNA damage with thousands of γH2AX foci which is too much for the repair machinery to cope with and therefore this leads to cell death. Additionally, the 3D-SIM images revealed that the treated nuclei are not only increasing in size but also in volume.

In conclusion, this study showed that antibody-based intracellular targeting is a promising new approach to target key functions inside living cells. The antibodies and Fabs can be efficiently delivered into cancer cells by electroporation to target and inhibit the replisome. This application of intracellular antibody targeting by inducing replication stress could be used as a potential novel cancer treatment approach. Furthermore, the antibodies could also be used to identify and validate new functional accessible sites of intracellular targets for the development of new molecules for cancer therapy.

These results were published on the 9th of March 2016 in *Experimental Cell Research.*

Author's contributions

Dominique Desplancq – interpreted the experiments. Performed most of the cell based experiments. Wrote the manuscript.

Guillaume Freund – Performed experiments involving the comparison of chemical inhibition versus anti-PCNA antibody inhibition. Performed all experiments involving the MELC, HL60 and primary cell lines. Assisted in the writing of the manuscript.

Sascha Conic – generated the two anti-PCNA antibodies 4D6 and 2B6. Designed, together with Dominique Desplancq and Etienne Weiss, the experiment for antibody-mediated cell death induction shown in Figure 3A. Performed the confocal and 3D-SIM microscopy of the γH2AX IFs shown in Figure 4C and Supplemental Movie

S1. Analyzed the microscopy data and performed the image processing. Assisted in the writing of the manuscript.

Etienne Weiss – conceived the study and interpreted the experiments. Wrote the manuscript.

All the other authors contributed with expertise, reagents and assistance when necessary.

Experimental Cell Research 342 (2016) 145–158

Contents lists available at ScienceDirect

Experimental Cell Research

journal homepage: www.elsevier.com/locate/yexcr

Targeting the replisome with transduced monoclonal antibodies triggers lethal DNA replication stress in cancer cells

Dominique Desplancq [a,1], Guillaume Freund [a,1], Sascha Conic [b], Annie-Paule Sibler [a], Pascal Didier [c], Audrey Stoessel [a], Mustapha Oulad-Abdelghani [b], Marc Vigneron [a], Jérôme Wagner [a], Yves Mély [c], Bruno Chatton [a], Laszlo Tora [b], Etienne Weiss [a,*]

[a] Ecole Supérieure de Biotechnologie de Strasbourg, UMR 7242, CNRS/Université de Strasbourg, boulevard Sébastien Brant, 67412 Illkirch, France
[b] Institut de Génétique et de Biologie Moléculaire et Cellulaire, UMR 7104, CNRS/Université de Strasbourg, INSERM U964, rue Laurent Fries, 67404 Illkirch, France
[c] Faculté de Pharmacie, UMR 7213, CNRS/Université de Strasbourg, route du Rhin, 67401 Illkirch, France

ARTICLE INFO

Article history:
Received 6 January 2016
Received in revised form
29 February 2016
Accepted 6 March 2016
Available online 9 March 2016

Keywords:
Proliferating cell nuclear antigen
DNA polymerase alpha
DNA replication stress
Functional inhibition
Antibody transduction
Chemo-resistant cancer cells

ABSTRACT

Although chemical inhibition of the DNA damage response (DDR) in cancer cells triggers cell death, it is not clear if the fork blockade achieved with inhibitors that neutralise proteins of the replisome is sufficient on its own to overcome the DDR. Monoclonal antibodies to PCNA, which block the DNA elongation process *in vitro*, have been developed. When these antibodies were transduced into cancer cells, they are able to inhibit the incorporation of nucleoside analogues. When co-delivered with anti-PCNA siRNA, the cells were flattened and the size of their nuclei increased by up to 3-fold, prior to cell death. Analysis of these nuclei by super-resolution microscopy revealed the presence of large numbers of phosphorylated histone H2AX foci. A senescence-like phenotype of the transduced cells was also observed upon delivery of the corresponding Fab molecules or following PCNA gene disruption or when the Fab fragment of an antibody that neutralises DNA polymerase alpha was used. Primary melanoma cells and leukaemia cells that are resistant to chemical inhibitors were similarly affected by these antibody treatments. These results demonstrate that transduced antibodies can trigger a lethal DNA replication stress, which kills cancer cells by abolishing the biological activity of several constituents of the replisome.

1. Introduction

DNA replication is tightly coordinated with cell metabolism and division and any failure of its progression during S-phase generally compromises genomic integrity. The fundamental biological processes of DNA duplication occur at discrete nuclear foci that each harbour several synthetic units named replicons [30]. Super-resolution microscopy demonstrated that proliferating cells engage thousands of origins that fire together at a specific time during the S-phase [2,9]. Numerous specific proteins are dedicated to perform DNA synthesis at the replication forks and if their activity is hindered by DNA damage, this may lead to DNA replication stress and subsequently to chromosome instability, which may contribute *in fine* to severe pathologies, such as cancer [17,33,36].

Genome instability is a consequence of DNA lesions that can result from the action of a large panel of genotoxic compounds, of which ultraviolet and ionising radiations are the most-widely described agents. The toxic activity of carcinogens and chemotherapeutic drugs currently in clinical use is due to their ability to interfere with DNA integrity and synthesis [47]. The blockade of replication forks under these conditions, known as replication fork stalling, results in structural rearrangements at the fork (collapse) and DNA breakage [22,64]. Fork stalling is generally prevented by intricate mitotic and S-phase checkpoint pathways, that have evolved to respond to fork arrest and ensure replication completion [13,27]. It was recently shown that targeting this surveillance mechanism, which is stimulated upon formation of double-strand breaks (DSBs) following fork collapse, may represent a strategy of high therapeutic value for the treatment of cancer cells [10,24,55].

Abbreviations: DDR, DNA damage response; DSB, double-strand break; CRISPR, clustered regularly interspaced short palindromic repeats; CSK, cytoskeletal; DTT, dithiothreitol; EdU, 5-ethynyl-2′-deoxyuridine; EM-CCD, electron multiplying-charge coupled device; γ-H2AX, phosphorylated histone H2AX; HU, hydroxyurea; mAb, monoclonal antibody; OE-PCR, overlap extension-polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; PCNA, proliferating cell nuclear antigen; PIP, PCNA interacting protein; RPA, replication protein A; RSR, replication stress response; SIM, structured illumination microscopy; siRNA, silencing RNA; ssDNA, single-stranded DNA; TCEP, tris(2-carboxyethyl)phosphine; X-Gal, 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside

* Corresponding author.
E-mail address: eweiss@unistra.fr (E. Weiss).
[1] These authors contributed equally.

http://dx.doi.org/10.1016/j.yexcr.2016.03.003
0014-4827/

Indeed, when replication forks are halted, the replicative mini-chromosome maintenance helicase is thought to continue unwinding DNA downstream of the fork, thereby exposing single-stranded DNA (ssDNA). Deleterious endonuclease cleavage is prevented when replication protein A (RPA) coats the ssDNA leading to the activation of serine/threonine kinases, such as ATR and Chk1. Additional downstream actors are able to further delay progression of the cell cycle and promote DNA repair by phosphorylating the H2AX variant of histone H2A protein [29]. Thus, if the DNA damage response (DDR) is inhibited, DSBs accumulate and this leads to excessive genome instability and apoptosis. Another striking effect on viability is obtained when the cells are treated with substantial doses of genotoxic agents that react with DNA, such as alkylating and platinum components that block essential DNA metabolic functions at the fork level. It has been proposed that in this case apoptosis, that may or may not dependent on wild-type p53, is initiated when the DNA damage and/or DNA repair responses are overwhelmed [41,48]. Cell death can thus be provoked by targeting various proteins acting at the DNA replication-repair interface, and this seems to be an irreversible outcome for cells under intense DNA replication stress.

Proliferating cell nuclear antigen (PCNA) is the DNA sliding clamp that acts in concert with DNA polymerases to promote DNA synthesis in mammalian cells [34]. It consists of a homotrimer that encircles DNA and is an essential part of the replication machinery complex, also named replisome [63]. It was recently proposed that PCNA, which has been called the “maestro” at the replication fork [38], is not only an efficient DNA polymerase processivity factor, but may also serve as a marker of newly synthesised DNA that allows genome integrity checks and nucleosome assembly after synthesis [18]. Another component of the replisome that cannot be replaced at the fork is DNA polymerase alpha, the only replicative polymerase displaying DNA primase activity that is required for the initiation of the synthesis of the Okasaki fragments during lagging strand DNA replication. Because both PCNA and DNA polymerase alpha are essential for progression of the fork, they can be considered to be ideal targets for inducing acute replicative stress and consequent cell death in cancer cells. Although small molecules that bind to PCNA *in vitro* reduce the access of PCNA to chromatin in cells [12] and impair different mechanisms related to DNA lesion bypass repair [25], it is not clear whether such agents can trigger cytotoxic DNA replication stress efficiently on their own. Similarly, aphidicolin, a specific inhibitor of B-family DNA polymerases [4] that allows the synchronisation of cultured cells [11], apparently does not compromise the survival of cancer cells by extensive replicative stress. PCNA can also been targeted with peptides [42,59], but when this is the case, no significant replicative stress is achieved. This suggests that PCNA and DNA polymerase alpha, while being essential for S-phase, cannot be efficiently inhibited with the presently available molecules, at least at levels that promote DNA damage-induced cell death following fork stalling.

In an attempt to achieve specific inhibition of chromatin-bound PCNA leading to DNA replication stress, as generally observed with DNA-intercalating compounds such as cisplatin, we have generated mouse monoclonal antibodies (mAbs) directed against the trimeric form of PCNA and have delivered them into cultured cancer cells using the protocol developed by [16]. Here, we describe the biological effects of antibody fragments that neutralise PCNA or DNA polymerase alpha in cells. The inhibition of both targets irreversibly leads to extensive DNA double-strand breakage because the treated cells cannot withstand such a biological intervention. This effect was observed in a variety of cancer cells including the chemo-resistant HL60 cell line and forms the basis of a novel strategy for inducing a lethal replication stress without modifying the genome or introducing potential mutations that could allow the recovery and further cycling.

2. Materials and methods

2.1. Monoclonal antibodies

The anti-PCNA monoclonal antibodies were generated as described [16]. Briefly, BALB/c mice were immunized with bacterially-expressed human PCNA protein. The growing hybridomas were screened by ELISA and by cell staining using fixed HeLa cells. The hybridoma clone SJK 132-20 [52] was a generous gift of Dr. J. Gannon, LRI, South Mimms, UK. The characteristics of hybridomas 4E9 and 4C6 have been previously described [16]. All antibodies were purified from hybridoma supernatant on Protein G Sepharose (GE Healthcare, Vélizy-Villacoublay, France) and kept at 4 °C in PBS at a concentration above 5 mg/mL. The Fab fragments were obtained by digesting pure antibody samples with papain (Sigma-Aldrich, St-Quentin Fallavier, France) essentially as described [1], followed by size exclusion chromatography on Superdex S200 10/300 (GE Healthcare). They were subsequently concentrated with centrifugal filtration units (Merck Millipore, Molsheim, France) to obtain Fab samples of 5 mg/mL.

2.2. *In vitro DNA elongation assay*

DNA elongation was performed essentially as described [8]. In brief, the primer-template duplex (12.5 fmol) consisting of a ^{32}P-labelled 25-mer oligonucleotide annealed to a 226 nucleotide long single-stranded DNA was incubated with HeLa nuclear extract (20 µg) in 25 mM HEPES pH 7.5, 4 mM ATP, 4 mM dNTP, 2 mM DTT and 4 mM $MgCl_2$ for 10 min at 30 °C. The effect of the antibodies was tested by the addition of 1–4 µg of pure antibody to the extract before incubation. The elongated products were analysed by electrophoresis on a 10% polyacrylamide gel containing urea (final concentration: 7 M). The radioactive bands were visualized with a Typhoon FLA 9500 apparatus (GE Healthcare) and the amount of elongated products was calculated by analysing the recorded images with the ImageQuantTL software (GE Healthcare).

2.3. Production of PCNA mutants and gel filtration analysis

The plasmid pET-hisPCNA was generated by subcloning the human PCNA coding region into the vector pET15b. pET-hisPCNA (Y114A) is a derivative of pET-hisPCNA which has been modified by OE-PCR to exchange the Tyr residue at position 114 of the PCNA coding region with an Ala codon using oligonucleotides 5′ GAAAGTTTCAGACGCTGAAATGAAGTTG and 5′CAACTTCATTTCAGCGTCTG-GAAACTTTC. The 2-step amplified PCR product was cloned following restriction with NdeI and BamHI enzymes. The same strategy was used to generate pnCS-UbiPCNA(164-261). The ubiquitin coding region was amplified from the pET-Ubi vector using the T7 universal promoter primer and oligonucleotide 5′-TTTCACTCCGTCTTTGCCTCCACCACCACGTAGACGTAAGAC and the PCNA coding region between residues 164 and 261 was amplified from pET-hisPCNA using oligonucleotide 5′-CGTCTACGTGGTGGTGGAGGCAAAG ACGGAGTGA-AATTTTCTGC and the universal T7 terminator primer. The SOE-PCR product generated with these two fragments was inserted in the pnCS vector [45] using NdeI and BamHI restriction enzymes. The pETM41-PCNA (1-163) was obtained by subcloning the PCNA coding region encompassing residues 1–163 by PCR with oligonucleotides 5′-G CGCGAGGTCTCCCATGTTCGAG and 5′-GCGCGA-GCGGCC-GCCTATGCACA from pET-hisPCNA into the pET-M41 vector [51]. Expression of the different polypeptides was performed following transformation of the generated constructs into the *E. coli* BL21DE3 pLysS strain. The cells expressing MBP-PCNA(1-163) and

Ubi-PCNA(164-261) fusions were harvested after induction with IPTG during 4 h at 37 °C and kept at −20 °C until use. The his-PCNA and his-PCNA(Y114A) proteins were expressed by autoinduction during 24 h at 30 °C or by induction with IPTG during 16 h at 20 °C, respectively. Both polypeptides were recovered by lysing the pelleted bacteria in 20 mM HEPES buffer pH 7.2, 0.3 M NaCl, 0.5 mM TCEP, 5 μg/mL DNase I, 15 μg/mL RNase A, 0.2 mM PMSF, EDTA-free protease inhibitor cocktail (Complete™; Roche Life Science) and subsequent chromatography of the cleared lysate on a HiTrap nickel column (GE Healthcare), followed by a 16/600 Superdex column equilibrated in PBS containing 0.5 mM TCEP. Binding in solution of his-PCNA or his-PCNA(Y114A) to the anti-PCNA Fab molecules was performed by mixing the pure proteins (6 μM of monomer) with purified Fab samples in excess during 2 h at room temperature in PBS and subsequent loading of the mixture on a Superdex S200 10/300 column (GE Healthcare) equilibrated in PBS. For evaluating the simultaneous binding of different Fabs to his-PCNA, equimolar quantities of Fab were added sequentially to the PCNA preparation during 2 h before loading on the gel filtration column. The protein content of the collected fractions was analysed by SDS-PAGE and Coomassie blue staining. The peptide p21 (GRKRRQTSMTDFYHSKRRLIFSRYIRS; [14]) was purchased from Eurogentec (Angers, France).

2.4. Cell assays

The HeLa and U2-OS cells (laboratory stocks) were maintained in Dulbecco's modified Eagle's tissue culture medium (DMEM; Life Technologies) at 37 °C in a humidified 5% CO_2 atmosphere. The MelC [15] and HL60R [20] cells were grown in Roswell Park Memorial Institute medium (RPMI-1640; Life Technologies) under similar conditions. Both media were supplemented with L-glutamine (2 mM), gentamicin (50 μg/mL) and 10% heat inactivated fetal calf serum. Fresh cells were thawed from frozen stocks after 10 passages. Transduction experiments with purified antibodies by electroporation were performed as previously described [16]. Where indicated, the cells were treated with siRNA (Life Technologies) using the same electroporation protocol two days prior to antibody transduction. The number of cells remaining attached to the dish following treatment was determined by manual counting after dissociation with trypsin and staining with Trypan blue or by spectrometry with the PrestoBlue™ cell viability kit (Life Technologies) according to manufacturer's protocol. Alternatively, the harvested cells were fixed overnight with a cold ethanol solution after 2 washes with PBS containing 1% glucose and 1 mM EDTA. These cells were then stained with propidium iodide (Sigma-Aldrich; 50 μg/mL) mixed with RNaseA (10 μg/mL) and analysed with a FACSCalibur™ flow cytometer (BD Biosciences, Le Pont de Claix, France) using the CellQuest™ Pro software. For the EdU incorporation experiments, cells were grown on glass coverslips and EdU (final concentration: 0.1 μM) was added to the culture medium for 16 h. The incorporated EdU was revealed with the Click-iT EdU AlexaFluor 488 imaging kit (Life Technologies), according to manufacturer's protocol. Where indicated, the harvested cells were also subjected to genomic DNA extraction. Typically, 10^5 cells in 200 mM Tris–HCl pH 8, 5 mM EDTA, 10 mM NaCl, 1% SDS and proteinase K (200 μg/mL, Roche Life Science) were incubated for 5 h at 55 °C and, after addition of DNAse-free RNAseA (Sigma-Aldrich; 5 μg/mL), the samples were treated twice with phenol:chloroform:isoamyl alcohol (25:24:1). Extracted DNA was concentrated with ethanol and, after resuspension in TE buffer, analysed by agarose gel (0.8%) electrophoresis.

2.5. CRISPR/Cas9 assay

The pCas9_GFP plasmid (Addgene, Cambridge, USA) which harbours the CAS9-2A-eGFP coding region and a gRNA cloning cassette digested with the BbsI restriction enzyme (New England Biolabs, Ipswich, USA) and the following pair of synthetic oligonucleotides was ligated into the linearized vector: 5′-CACCGGGCCAGGTTGCGGTCGCAG-3′ and 5′-AAACCTGCGACCGCAACCTGGCCC-3′. The resulting vector, named pkG8, was sequenced on both strands and transfected in HeLa cells using jetPEI (Polyplus transfection, Illkirch, France) according to the manufacturer's instructions. 24 h post-transfection, the GFP-positive cells were isolated with a cell sorter (FACSCalibur™ ; BD Biosciences) and allowed to grow in complete DMEM medium for several days.

2.6. Western blot analysis

For the revelation of the PCNA polypeptides expressed in *E. coli*, induced bacteria were lysed in SDS gel-loading buffer and the soluble extracts were analysed by Western blotting. The PCNA fusions were detected by incubating the blot with 2B6 or PC10 antibodies (final concentration 0.1 μg/mL) and subsequently with sheep anti-mouse HRP conjugate (GE Healthcare). For the analysis of the HeLa proteins, soluble extracts (30 μg/lane) in RIPA buffer were used. γ-H2AX and actin were revealed with rabbit monoclonal antibody EP854(2)Y (Abcam, Paris, France) and rabbit polyclonal serum A206 (Sigma-Aldrich), respectively. The Chk1 protein was identified with mouse monoclonal antibody G-4 (Santa Cruz Biotechnology, Dallas, USA) and the phospho-Chk1 polypeptide was detected with rabbit monoclonal antibody 133D3 (Cell Signalling Technology, St-Quentin-en-Yvelines, France). Bound secondary antibodies (HRP-labelled or IR dye-labelled) were quantitated with ECL reagent (GE Healthcare) and analysis with the Image QuantLAS 4000 imager (GE Healthcare) or with the Odyssey infrared imager (Li-Cor, Lincoln, USA).

2.7. Microscopy

For the analysis by classical immunofluorescence microscopy, the cells were fixed with 4% paraformaldehyde for 20 min and, after permeabilisation with 0.2% Triton X100 for 5 min, incubated with the different antibodies diluted in PBS containing 10% fetal calf serum. The electroporated mAbs and Fab fragments were detected with AlexaFluor 488 or 568 labelled-anti-mouse immunoglobulins (Life Technologies). γ-H2AX and RPA32 were revealed with the rabbit monoclonal antibody EP854(2)Y (Abcam) and the rabbit polyclonal serum 7300-244A (Bethyl Laboratories), respectively. To detect PCNA in the electroporated or transfected cells, we used the rabbit polyclonal antibody Ab 15,497 (Abcam). After incubation, the coverslips were mounted with 4′,6′-diamino-2phenyl-indole (DAPI) Fluoromount-G (Southern Biotech, Birmingham, USA) and imaged with a Leica DM5500 microscope equipped with 63X and 100X objectives. The images were processed with ImageJ software. To analyse the cells in real time, we used a HeLa cell line that expresses constitutively H2B fused to GFP [28]. These cells were cultured in L-15 medium supplemented 2 mM L-Glutamine, 50 μg/mL gentamicin and 10% heat inactivated fetal calf serum. Highly inclined and laminated optical sheet (HILO) microscopy [54] was performed on an Olympus IX-71 inverted microscope. Samples were illuminated with a laser diode at 488 nm (Spectra Physics, Santa Clara, USA) and the signal was recorded on an EM-CCD camera (Hamamatsu, Massy, France). To minimize damage induced by the laser illumination, non-illuminated cells were imaged at each measurement. HILO images were processed with ImageJ software. Three-dimensional structured illumination microscopy (3D SIM) was performed on a DeltaVision OMX-Blaze V4 system (Applied Precision Imaging, Issaquah, USA) equipped with a Plan Apo N 60× oil immersion objective lens (Olympus), 4 liquid-cooled sCMOs cameras (pco Edge, Photometrics) and solid-state lasers. The 568 nm laser line was used during acquisition and the optical z-sections were separated by 0.125 μm with an exposure time of 35 ms. The raw images were processed and reconstructed using the DeltaVision OMX SoftWoRx software package

(v6.1.3, Applied Precision) as described [50]. The final image processing was performed using the Fiji/Image J software.

3. Results

3.1. Mabs that block DNA replication in vitro arrest DNA synthesis in cells

In a previous study, we analysed the behaviour of mAbs that were

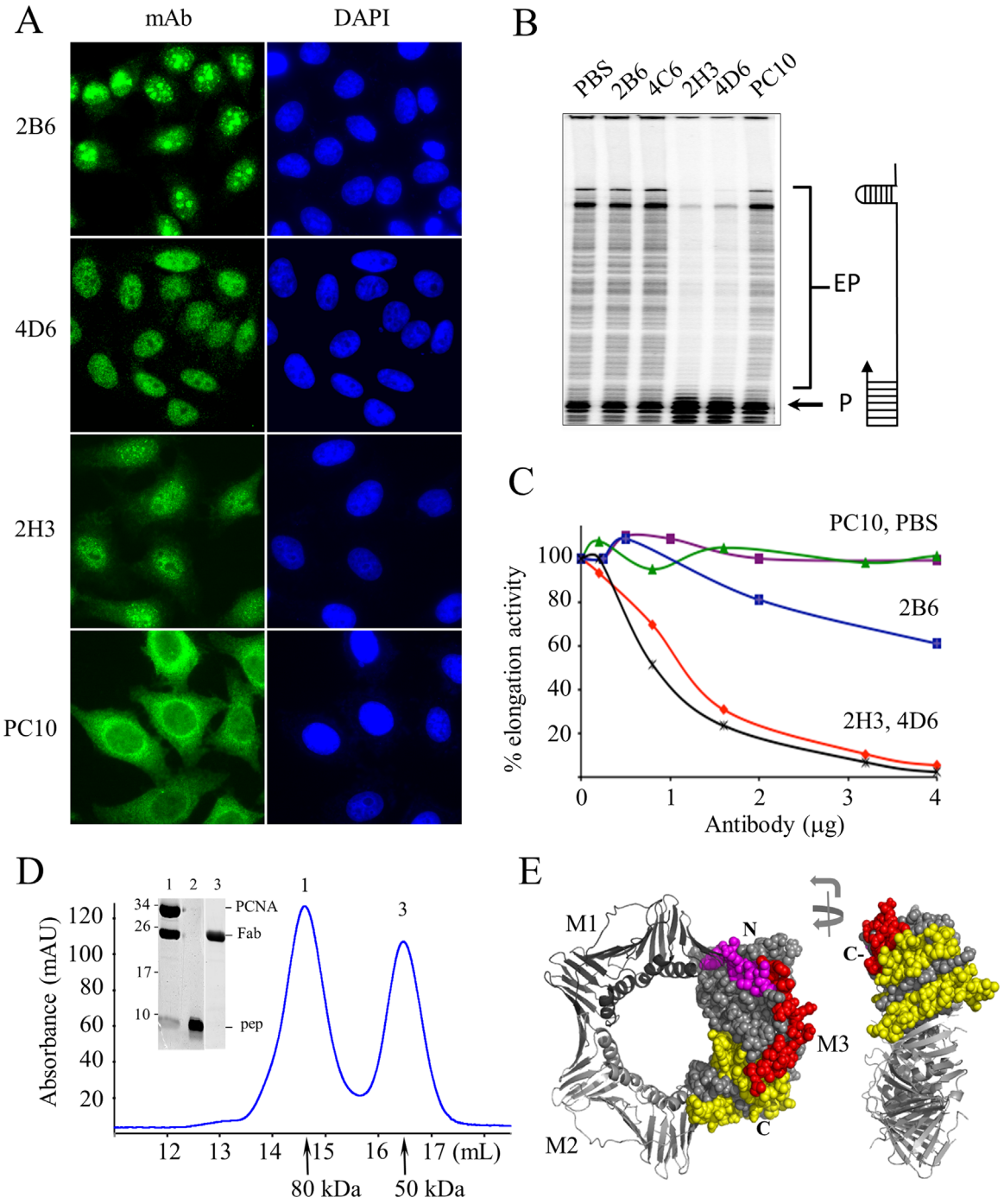


Fig. 1. Binding characteristics of the anti-PCNA antibodies. (A) Intracellular localization of delivered antibodies as probed by immunofluorescence microscopy. The pictures correspond to typical fields of HeLa cells transduced with the indicated antibodies (2 μg) at 72 h post-treatment. The delivered antibodies were revealed with Alexa Fluor 488 labelled goat anti-mouse immunoglobulins. Magnification: x 630. (B) *In vitro* neutralising activity of anti-PCNA antibodies as probed with DNA elongation assays. The indicated antibodies (4 μg) were mixed with HeLa nuclear extracts and incubated with ^{32}P-labelled primer/template hybrid (right). The elongated products (EP) were analysed on acrylamide gel and visualized as indicated in Section 2. PBS was used as control. (C) Inhibition of DNA synthesis with varying amounts of Fab. The analysis was done as in (B), except that the elongation products were quantitated. The curves summarize the data obtained in 2 independent experiments with the indicated antibodies. (D) Gel filtration analysis of PCNA/peptide/Fab complexes. PCNA was incubated with a 10-fold excess of p21 peptide and a two-fold excess of Fab 2H3 for 2 h before loading on the column. The elution profile is shown. Aliquots of peak fractions 1 and 3 were analysed by SDS-PAGE and Coomassie staining (inset). Pure p21 peptide (15 μg) was, in addition to calibration markers (in kDa, left) used as migration control (lane 2). (E) Schematic representation of the regions of PCNA (structure from PDB 1AXC) recognised by antibodies 2B6 and 2H3 (yellow), PC10 (magenta) and p21 peptide (red). The bound residues are highlighted in one of the 3 monomers (M). The N- and C-terminal ends of the monomer are indicated. Front and side views are shown on the left and right, respectively.

introduced into living cells by electroporation [16]. This method of delivery allows almost all cells to be transduced without any loss of viability. Here, we show that, among seven different antibodies that specifically react with the PCNA protein as probed by immunofluorescence microscopy, three (2B6, 2H3 and 4D6) bind to *de novo* synthesised PCNA in the cytoplasm of electroporated cells and are translocated into the nucleus in a piggyback fashion [16]; Fig. 1A). A typical staining of the replication foci was observed when the cells were washed with CSK buffer prior to fixation (not shown). This was not the case with the anti-PCNA reference mAb PC10 [60], suggesting that the three novel mAbs bind to regions of PCNA that are accessible in the cells. The epitopes of these new mAbs are likely away from residues 111–125, an immunodominant region of PCNA recognised by a number of mAbs such as PC10, which possess no inhibitory effect on DNA synthesis [46]. To test whether the piggybacked antibodies can interfere with DNA replication, we performed *in vitro* DNA elongation experiments using a classical DNA polymerization assay [3] and analysed their capacity to inhibit the functional association of DNA polymerase delta with PCNA. Almost no DNA elongation products were observed on gel following preincubation with 2H3 and 4D6 antibodies (Fig. 1B). Interestingly, 2B6 behaved almost as the irrelevant antibody 4C6 and the non-inhibitory antibody PC10, suggesting that it does not hinder the formation of active elongation complexes. We obtained similar results when the experiments were carried out with the corresponding Fab preparations (Fig. 1C; Supplementary Fig. 1), indicating that bivalent binding of either 2H3 or 4D6 antibodies to PCNA was not required for abrogation of DNA elongation *in vitro*.

To gain further insight into the region(s) of PCNA recognised by these two antibodies, we mixed equimolar amounts of Fab molecules with either wild-type PCNA (trimer; 90 kDa) or mutant PCNA ($PCNA^{Y114A}$), which accumulates essentially as a monomer upon overexpression in *Escherichia coli* [26], and assessed the amount of bound and free molecules by gel filtration (Supplementary Fig. 2). Both Fab 2H3 and 4D6 reacted with wild-type and mutant PCNA, whereas Fab 2B6 clearly bound only to the monomeric PCNA mutant and did not recognise the trimeric form of PCNA. Surprisingly, Fab 2H3 and 4D6 altered the trimeric structure of PCNA upon binding, because only 1:1 complexes of Fab and PCNA monomer of about 80 kDa were observed in the main peak. A similar observation was made when Fab PC10 was analysed, suggesting that binding of these antibodies destabilises recombinant trimeric PCNA. The presence of both antigen and Fab molecules in all peak fractions was verified by SDS-PAGE (not shown). Moreover, when either Fab 2H3 or Fab 4D6 were mixed with Fab PC10 and allowed to react with wild-type PCNA, complexes of about 130 kDa, which corresponds to the apparent molecular weight of two Fab molecules and one PCNA monomer, were observed (Supplementary Fig. 2B and C). Since this was not the case when Fab 2H3 was assayed together with Fab 4D6 or with Fab 2B6, it suggests that the three piggybacked anti-PCNA antibodies bind to a similar region that is distinct from that recognised by PC10. In addition, Fab 2H3 and 4D6 do not react with unfolded PCNA protein, in contrast to 2B6 and PC10. We roughly mapped the region recognised by 2B6 under these conditions using truncated PCNA fusions (Supplementary Fig. 2D) and found that it reacts with the C-terminal domain of PCNA by Western blotting.

To confirm that the piggybacked antibodies do not interact with the central part of PCNA, a cavity covered by the interdomain connecting loop recognised by PIP-box proteins, we performed gel filtration analysis of the complexes formed after mixing PCNA with Fab 2H3 and a peptide corresponding to the C-terminal end of p21 protein (PIP-box sequence, [14] both in excess. As shown in Fig. 1D, a major fraction (peak 1) corresponding to Fab 2H3 and p21 peptide interacting with the PCNA monomer, as evidenced by SDS-PAGE, was observed together with a shoulder (peak 3) corresponding to free Fab molecules. This was also the case when Fab 4D6 was used. Together, these results demonstrate that both 2H3 and 4D6 antibodies, which severely hinder the DNA elongation process, bind to a region located in the C-terminal domain of PCNA and distinct from that recognised by the majority of PCNA interacting components carrying a PIP box. It is likely that the region targeted by the piggybacked antibodies is located on the C-face also known as the front face (Fig. 1E) that binds DNA polymerases [18].

To examine if antibodies 2H3 and 4D6 could inhibit DNA elongation in transduced cells, we performed pulse-chase experiments with a modified thymidine analogue (EdU) that is efficiently incorporated into newly synthesised DNA and that can be labelled fluorescently in a highly specific manner after cell fixation. Bright staining of the nuclei was observed when the HeLa cells were pulsed during a period of 16 h after transduction with either phosphate-buffered saline or control antibody 4C6 (Fig. 2A). This was also observed after transduction with 2B6 antibody. However, almost no staining was obtained following transduction of HeLa cells with 2H3 and 4D6 antibodies (Fig. 2A), both of which were also strong inhibitors of DNA elongation *in vitro* (Fig. 1B). The same effect was observed in U2-OS cells (Supplementary Fig. 3). The results together indicate that both 2H3 and 4D6 mAbs can block DNA replication *in vitro* and in cells.

Next Fab molecules were prepared from the indicated mAbs and were tested in the above described cellular assay, in parallel with Fab molecules derived from antibody SJK 132, a previously described antibody that binds to DNA polymerase alpha and that inhibits DNA replication *in vitro* [52]. As for the complete antibodies, no effect was observed following transduction of Fab 2B6 or irrelevant Fab 4C6, whereas strong inhibition of DNA synthesis was observed with Fab 2H3 and Fab SJK 132 (Fig. 2A, right). Our results thus indicates that Fab molecules, originating from antibodies that severely impair DNA replication *in vitro* and that target essential proteins of the replisome, can also promote the arrest of replication forks in cells.

3.2. Inhibition of PCNA or DNA polymerase alpha by neutralising antibodies leads to cell death

When analysing the overall survival rate of HeLa cells treated with the anti-PCNA inhibitory mAbs or Fab molecules under the above conditions, we found that at 3 days post-transduction the cell number was reduced by approximately 50% when compared to that observed with the non-inhibitory molecules (Fig. 2B, red columns). We obtained similar results in U2-OS cells (Fig. 2B, blue columns), suggesting that the effects on proliferation of the delivered anti-PCNA inhibitory reagents are limited, likely due to the renewal and abundance of PCNA in transformed cells [40].

To investigate whether lowering the PCNA levels in cells transduced with anti-PCNA inhibitory antibodies affects the survival rate, HeLa or U2-OS cells were co-treated with both anti-PCNA siRNA and the antibody or Fab preparations. Preliminary experiments performed with a commercially available siRNA at a concentration of 20 nM and 0.5 nM in HeLa and in U2-OS cells, respectively, showed that 2 days of incubation was needed in both cases to observe a reduction of 60–70% in the PCNA level (Supplementary Fig. 4A). Because the cells did not loose viability under these conditions, they were subjected to transduction with either mAb or Fab 2B6, 4D6 and 2H3. Microscopy analysis after the third day of incubation following transduction and siRNA treatment showed that the anti-PCNA inhibitory antibodies or Fab molecules drastically affected the cell morphology and survival (Fig. 3A and Supplementary Fig. 4A). Indeed, nearly all cells were flattened before shrinking and dying, a phenomenon that was not observed following treatment with siRNA only or co-treatment with siRNA and the non-inhibitory antibody 2B6 or its Fab fragment. When

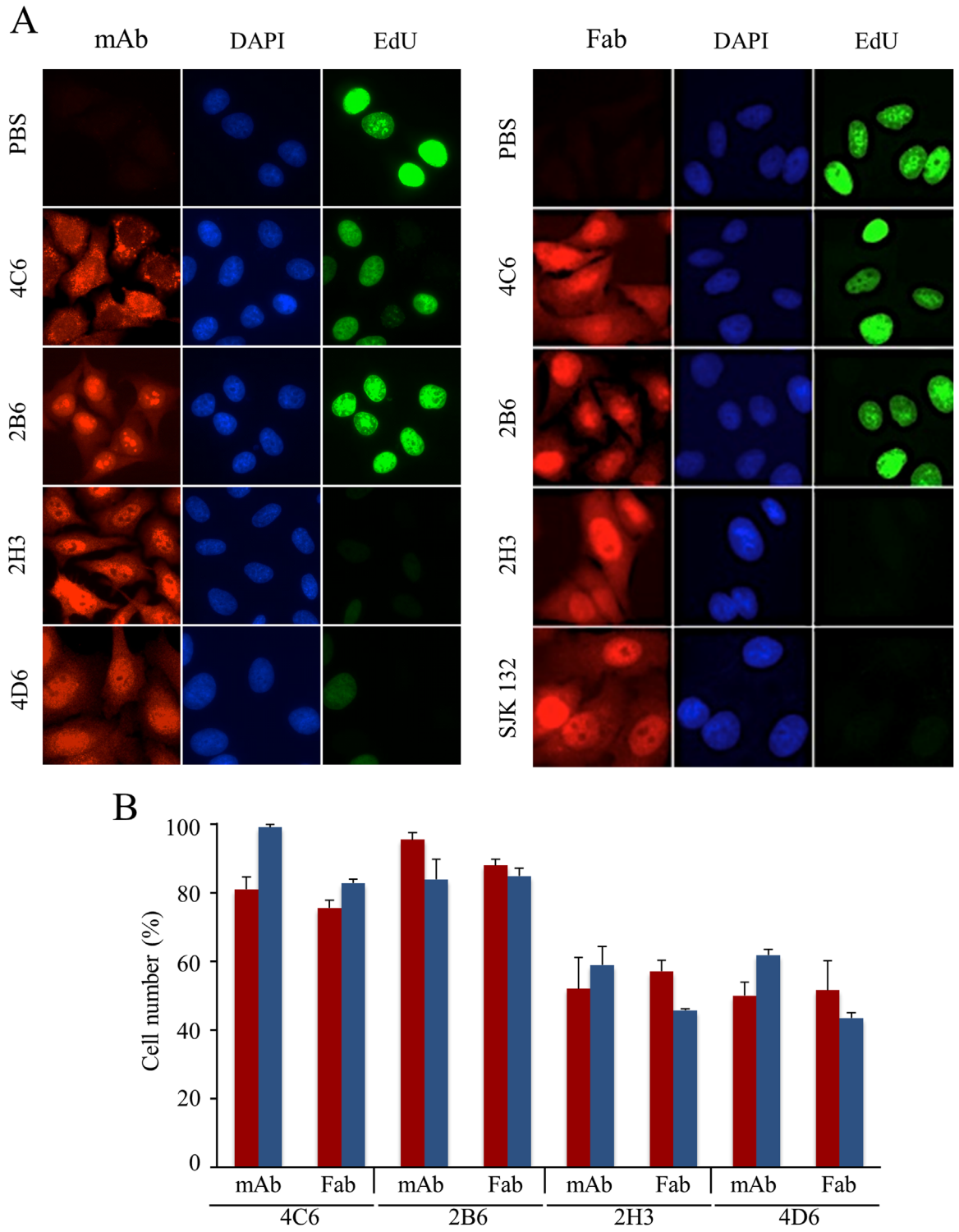


Fig. 2. Effect of the transduced antibodies or Fabs on DNA synthesis in cells and cell number. (A) HeLa cells were transduced with the indicated antibodies (20 μg; left) or Fab (5 μg; right) and 0.1 μM EdU was added to the culture medium 48 h or 24 h post-transduction, respectively. After incubation for 16 h, the cells were fixed and incorporated Edu was revealed with Alexa Fluor 488 azide (Section 2). The antibodies or Fab were detected in parallel by adding to the incubation mixture Alexa Fluor 568 labelled-anti-mouse goat immunoglobulins. The micrographs correspond to typical fields of cells analysed in 3 independent experiments. The anti-E6 4C6 antibody was used as a negative control. (B) Percentage of living cells after antibody or Fab treatment. HeLa (red) or U2-OS (blue) cells were transduced as indicated in A. At 72 h post-transduction, the cells remaining attached to the plastic were trypsinised and counted as indicated in Section 2. The percentage of cells was calculated by taking those treated with PBS in parallel as reference. The data presented are the means of three independent experiments and errors bars represent standard deviations.

cells were treated with siRNA and with either antibody 2H3 or 4D6 (or their corresponding Fabs), we observed strong cell enlargements, shape changes and abrogation of the cell cycle as evidenced by fluorescence-activated cell sorting (Fig. 3B and Supplementary Fig. 4C). The typical enlargement of both HeLa cell cytoplasm and nucleus is reminiscent of premature senescence [56] since blue-dyed precipitates were visible after fixation and incubation with the chromogenic substrate X-Gal (not shown). Interestingly, the increase of the HeLa cell size was also observed when they were treated with the anti-DNA polymerase alpha Fab SJK 132 or when

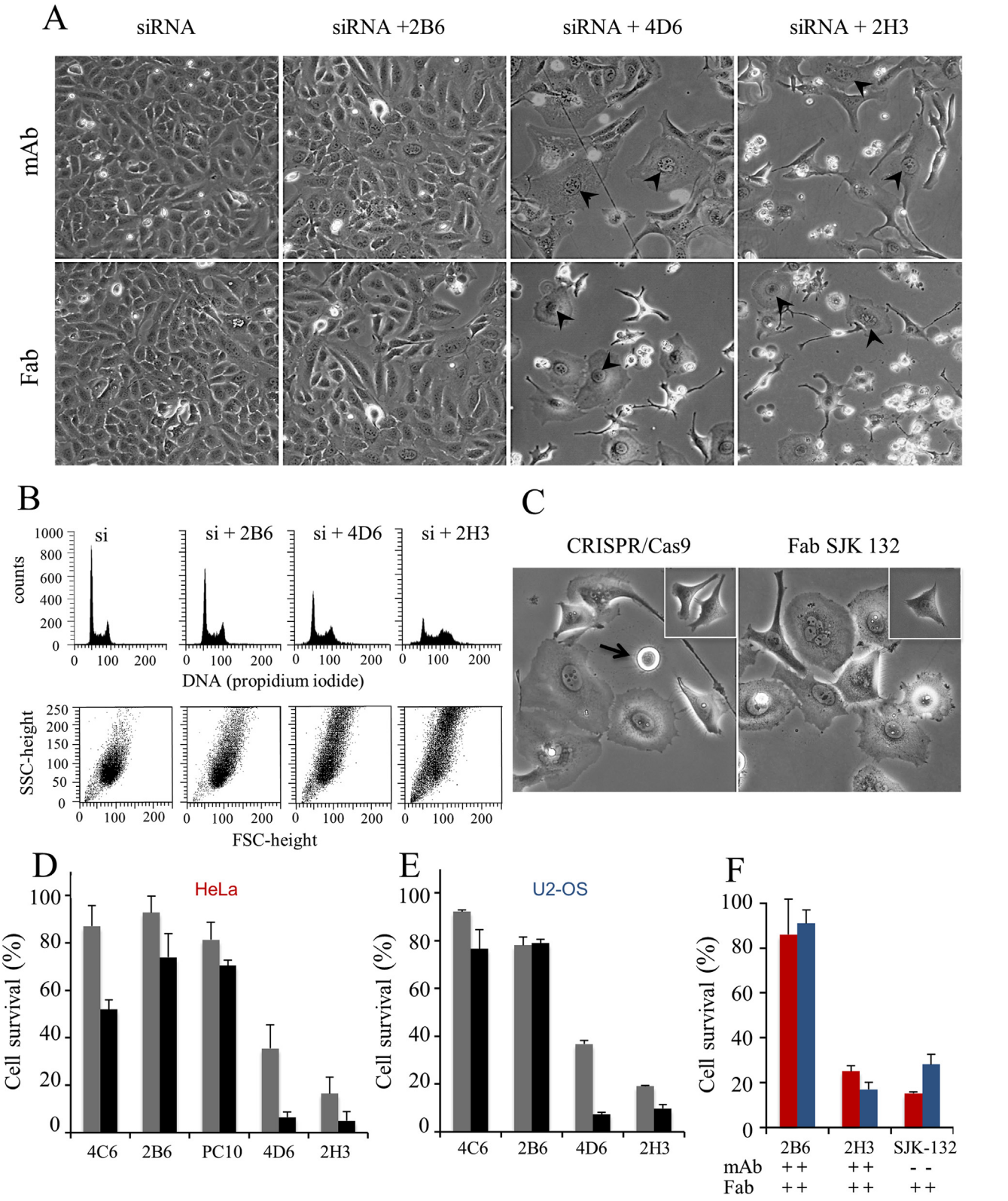


Fig. 3. Phenotypical cell changes after co-treatment with siRNA and antibodies. (A) Antibody-mediated cell death induction. HeLa cells were transfected with anti-PCNA siRNA for 48 h and subsequently transduced by electroporation with the indicated antibodies or Fab fragments. At 72 h post-transduction, the cells were observed with optical microscopy. The pictures show typical fields of the cells observed at the same magnification (x200). Flattened cells with an increased nuclear size following treatment with 4D6 and 2H3 reagents are indicated (arrows). PBS was used as antibody or Fab control. (B) FACS analysis of HeLa cells treated as in (A). After 120 h of incubation, the treated cells were trypsinised and subjected to FACS analysis after staining with propidium iodide. Approximately 2.5×10^4 cells were counted in each case. The upper panels correspond to the measurement of DNA content (cycling state) and the lower panels correspond to the measurements of forward- (FSC) and side- (SSC) scattered light, reflecting size and granularity of the cells, respectively. Si means siRNA (C) Cell morphology after PCNA gene disruption with the CRISPR/Cas9 technology or following transduction with Fab SJK 132. HeLa cells were transfected with plasmid pkG8 (Section 2) or transduced with Fab SJK 132 (5 μg). 24 h post-transfection, the cells expressing GFP were sorted by FACS and further incubated. The pictures show typical cells at 7 days post-transfection (left) or 3 days post-transduction (right). Cells treated with empty vector or PBS are also shown (insets). The flattened cells rounded up (arrow) before floating. Magnification: X 400. (D), (E) Cell survival rate after treatment with anti-PCNA siRNA and the indicated antibodies (gray) or Fab fragments (black). HeLa (D) or U2-OS (E) were treated as in (A) and the number of cells remaining bound to plastic after 3 days of incubation with the antibodies (gray bars) or the Fabs (black bars) was expressed as percentage of live cells, as compared to cells treated with PBS in parallel. (F) Cell survival rate after co-treatment with Fab and antibody. HeLa (red) and U2-OS (blue) were co-transduced with Fab (4 μg) and antibody (5 μg) as indicated. The number of cells remaining attached to the plastic at 3 days post-transduction was determined as in (D), (E). The data presented in each graph are the means of three independent experiments and error bars represent standard deviations.

the cells were incubated for 6–7 days following invalidation of the PCNA gene with the CRISPR/Cas9 method (Fig. 3C). In the latter case, expression of PCNA was impaired in cells attached to the plastic before they detached and floated (Supplementary Fig. 5). These results show that inhibition of PCNA or DNA polymerase alpha by means of neutralising antibodies or the corresponding

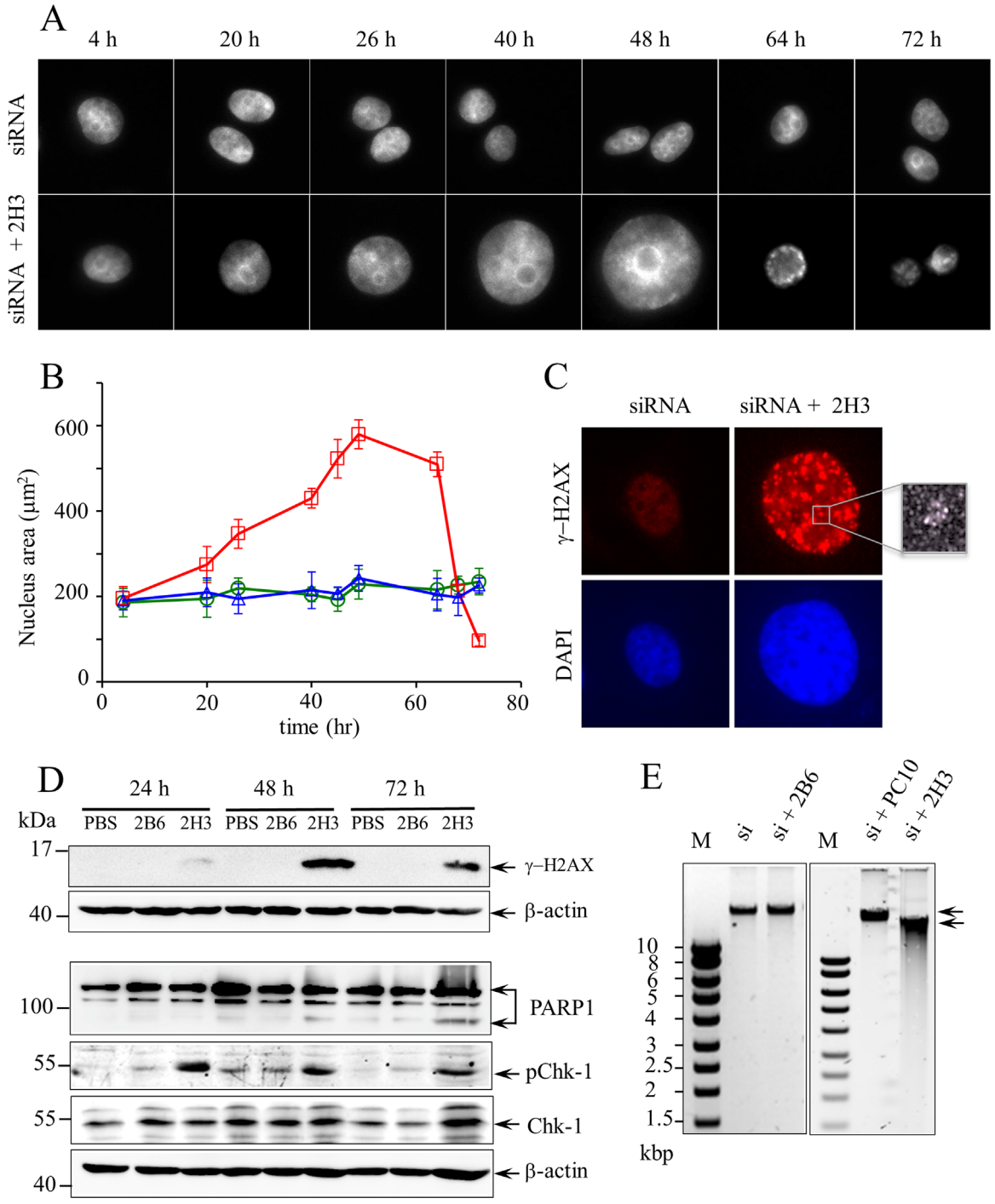


Fig. 4. Antibody-mediated DNA replication stress leads to cell death. (A) Analysis of the time-dependent variation of the nuclear size by fluorescence microscopy. HeLa cells expressing H2B-GFP were treated with siRNA + PBS (siRNA) or with siRNA and Fab 2H3 (siRNA + 2H3) as described in the legend of Fig. 3A. At the time indicated, typical cells were micrographed (Section 2). Magnification: x 400. (B) Statistical analysis of the nuclear area. HeLa cells were treated as in (A) and, at the indicated time, several dozens of cells were micrographed. The plot shows the variation of the area of the nuclei (mean values of 90 different cells per time point) of either siRNA- (blue) or siRNA + Fab 2B6- (green) or siRNA + Fab 2H3- (red) treated cells. (C) Analysis of the induction of γ-H2AX by immunofluorescence microscopy. HeLa cells treated as in A were fixed at 48 h post-transduction and incubated with anti-γ-H2AX rabbit monoclonal antibody and Alexa Fluor 568-labelled secondary immunoglobulins. The pictures show a typical cell in each case following counterstaining with DAPI. Magnification: x 1000. The square on the right corresponds to an enlarged image of a γ-H2AX focus as observed with 3D-SIM (see Movie S1). (D) Analysis of the DNA damage response by Western blotting. The cells co-treated with siRNA and Fab 2B6 (2B6) or Fab 2H3 (2H3) or no Fab (PBS), as in A, were harvested at 24, 48 and 72 h post-transduction. Crude extracts containing a similar amount of protein probed with relevant antibodies against the indicated polypeptides after electrophoresis on SDS gel and blotting. The migration of molecular weight markers is indicated. (E) Analysis of the integrity of genomic DNA by agarose gel electrophoresis. The genomic DNA of HeLa cells treated with siRNA (si) and the indicated Fab fragments was extracted 72 h post-transduction and samples containing a similar amount of DNA were loaded on agarose gel. After electrophoresis, the DNA bands were visualized by UV illumination following staining with ethidium bromide. DNA markers were migrated in parallel.

Fabs leads to a cell phenotype analogous to that observed for cells when PCNA expression is severely down-regulated.

Because the growth of the HeLa and U2-OS cells treated with the anti-PCNA antibodies or Fabs and siRNA was suppressed, we counted the cells that were still alive in the dishes at day 5. Compared to treatment with control antibodies or Fabs (PC10, 2B6 and 4C6), there was a dramatic reduction in the cell survival rate upon transduction with 2H3 and 4D6 (Fig. 3D and E). More than 90% of cells were floating in the culture medium in this case, indicating that blockade of PCNA with the transduced inhibitory antibodies is very effective when its expression level is significantly reduced. We obtained similar results when cells were transduced with Fab SJK 132 alone (Fig. 3F), confirming that DNA polymerase alpha is essential and that, likely due to its lower level than that of PCNA in cells, treatment with Fab alone was sufficient. Remarkably, treatment of either HeLa or U2-OS cells with both antibody 2H3 and Fab 2H3 in the absence of siRNA, which probably provides more anti-PCNA reagents that can bind to neosynthesised PCNA in the cytoplasm and to the nuclear fraction of PCNA, respectively, was comparable to treatment with Fab SJK 132 alone (Fig. 3F). Overall, these results show that targeting essential proteins acting at the replisome with antibody-based molecules that can block DNA synthesis arrests cell growth and subsequently leads to cell death.

3.3. *Anti-replisome antibody-induced cell death is a consequence of extensive DNA damage*

It is well established that inhibition of DNA replication can stress growing cells and that fork stalling can eventually lead to cytotoxic DSBs (see Introduction for references). To examine if this was true after treatment of cells with inhibitory Fabs that bind to either PCNA or DNA polymerase alpha, we first analysed, as a preliminary step, the fate of nuclei in HeLa cells constitutively expressing H2B-GFP following treatment with anti-PCNA siRNA and Fab 2H3 by fluorescence microscopy. Fig. 4A shows individual nuclei of typical cells observed during the 3 days of Fab treatment. Remarkably, their size sequentially increased until about 48 h post-treatment and then rapidly decreased to form condensed bodies by the end of day 3. We measured the area of these nuclei and found that the peak values were up to 600 μm^2 in 2H3-treated cells, whereas the areas in cells treated with either only siRNA or non-inhibitory Fab 2B6 together with siRNA remained relatively constant with an average value of 200 μm^2 (Fig. 4B). Because the fluorescent signal of H2B proteins present in the nuclei remained relatively uniform in each case until about 48 h post-treatment, we hypothesised that this enlargement of the chromatin staining may be due to randomly distributed DSBs that arise when replication forks are irreversibly stalled, as shown by the formation of RPA foci under similar conditions in HeLa cells (Supplementary Fig. 6B). The fact that chromatin was highly condensed in the cells at about 72 h post-treatment suggests that they were dying by apoptosis.

Phosphorylated H2AX (γ-H2AX) is a well-described biomarker of DSBs [29]. By monitoring its abundance in the enlarged nuclei with specific antibodies by immunofluorescence microscopy, we observed strong speckled staining in the entire nucleus upon 2H3 treatment. No such typical signal was visible in control HeLa or U2-OS cells (Fig. 4C and Supplementary Fig. 6C). The number of foci per nucleus varied, but most of the 2H3 mAb treated cells contained approximately 100–300 distinguishable spots, as recorded by image analysis (Section 2). Interestingly, a similar pattern of γ-H2AX was also obtained when HeLa cells were either transduced with Fab SJK 132 or when the analysis was performed following transfection with the PCNA-specific CRISPR/Cas9 plasmid (Supplementary Fig. 6A). We further examined the γ-H2AX signals in HeLa cells by super-resolution microscopy (3D-SIM) and found that the larger and intensively stained foci that were easily observed by conventional microscopy correspond to clusters containing an average of 10–15 individual spots (Fig. 4C, enlarged field on the right and Supplementary Movie 1). This analysis showed that the nuclei of either Fab 2H3- or SJK 132-treated cells each contained several thousands of γ-H2AX foci. In addition, it allowed us to determine that not only was the area of the nucleus increased, as evidenced by DAPI staining, but that the volume of the nucleus was enlarged by up to nearly 3.5-fold (Supplementary Movie 1). We confirmed the strong induction of γ-H2AX following transduction with Fab 2H3 by Western blotting (Fig. 4D, upper panel). Moreover, in these cells and almost not in cells treated with the non-inhibitory Fab 2B6, we detected phosphorylated Chk1, another relevant marker of DNA replication stress and ATR activation, at about 24 h post-transduction (Fig. 4D). By day 3, when cells began to die, cleaved PARP1 polypeptides were also detectable, suggesting that apoptosis was initiated in these cells. We also analysed the migration profile on gel of genomic DNA extracted from HeLa cells treated with siRNA and the control Fab 2B6, or siRNA+Fab PC10 or siRNA+Fab 2H3 (Fig. 4E). In contrast to that observed with non-inhibitory molecules, DNA did not migrate as a single broad band but as a smear following the treatment of cells with Fab 2H3, demonstrating that numerous cuts arise in these treated cells (Fig. 4E). The same result was obtained when Fab 4D6 was used instead of Fab 2H3 or when Fab SJK 132 was used (not shown). Together, these results suggest that blockade of the progression of a substantial number of replication forks, if not all, with anti-PCNA or anti-DNA polymerase alpha antibodies transduced into HeLa or U2-OS cells induces massive DNA breakage that cannot be repaired and thereby provokes lethal cytotoxic effects.

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.yexcr.2016.03.003.

3.4. *Targeting the replisome in primary and chemo-resistant cells*

Having shown that antibodies targeting the replisome promote cell death, we investigated whether this approach could be used to treat primary melanoma cells or, alternatively, cancer cells that are resistant to chemicals that promote DNA replication stress. We used the MelC cell line, a previously described cell line isolated from a metastatic lymph node of a stage III melanoma patient [15], and HL60R cells, a multi-drug-resistant form of human promyelocytic HL60 leukaemia cells [20]. MelC and HL60R cells are sensitive and resistant, respectively, to daunorubicin and etoposide, two well-characterised drugs that suppress topoisomerase II activity in mammalian cells (Fig. 5A). In preliminary experiments, we compared the sensitivity of these cells to the recently described T2AA small molecule, which binds to the region of PCNA that interacts with PIP box-containing proteins [44]. These cells were less susceptible to cell death than HeLa or U2-OS cells (Fig. 5B). Interestingly, in each case, even when high doses of the drug were used, we did not observe any enlargement of nuclei before cell shrinkage and subsequent detachment from the plastic. When MelC and HL60R cells were treated with siRNA+Fab 2H3 or siRNA+Fab 2B6, a clear effect was observed only with the inhibitory 2H3 molecule as almost all cells were dead at day 3 post-treatment (Fig. 5C). The nuclei of both cell lines were significantly enlarged at day 2 before shrinking at day 3–4, as observed with HeLa or U2-OS cells (see above). This was particularly visible under an optical microscope with the HL60R cell line. The cell diameter of 17+/− 2 μm in untreated cells was increased to 30+/− 1.5 μm following anti-PCNA treatment (Fig. 5D). In addition, under these conditions, there was a strong induction of γ-H2AX in both MelC and HL60R cells (Fig. 5C and D), further suggesting that the DNA

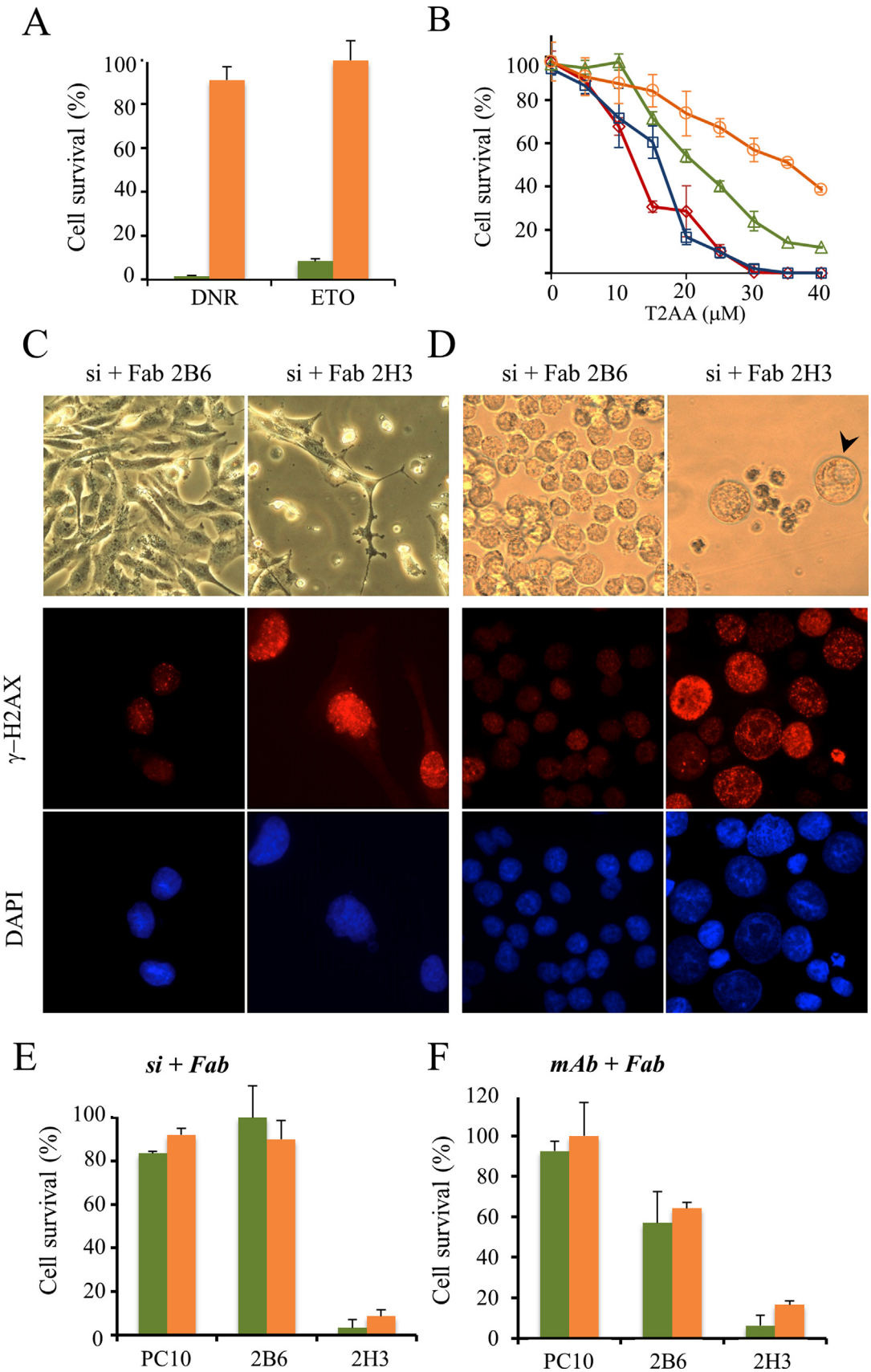


Fig. 5. Effect of the transduced anti-PCNA antibodies in the MelC and HL60R cell lines. (A) Assessment of chemo-sensitivity to daunorubicin (DNR) and etoposide (ETO). The drugs (1 μM) were added to sub-confluent cultures of MelC (green) and HL60R (orange) cells and the viable cells were counted 4 days post-treatment. Untreated cells were used as reference. (B) Assessment of chemo-sensitivity to T2AA. Varying concentrations of T2AA as indicated were added to sub-confluent cultures of HL60R (orange), MelC (green), HeLa (red) and U2-OS (blue) cells and the number of viable cells was determined as in A. (C) Effect of Fab 2B6 and Fab 2H3 on MelC cells. The cells were pre-treated with anti-PCNA siRNA for 2 days and subsequently transduced with the indicated Fab fragments. The upper panels show typical fields of the treated cells at 72 h post-transduction. The lower panels correspond to the detection of γ-H2AX and to the nuclear staining with DAPI of representative cells 48 h post-transduction. Magnification: x 400. (D) Effect of Fab 2B6 and Fab 2H3 on HL60R cells. The pictures correspond to cells treated and visualized as in C. (E), (F) Cell survival rate after co-treatment with siRNA and Fab (E) or antibody (mAb) and Fab (F). The amount of Fab and/or mAb used in these assays was identical to those indicated in the legend of Fig. 3. The percentage of viable cells at 3 days post-transduction was determined using cells treated in parallel with PBS as a reference. The green and orange columns correspond to MelC and HL60R cells, respectively. The data presented in each graph are the means of 3 independent experiments and the error bars represent standard deviations.

damage mechanism leading to cell death is also involved, as observed with HeLa or U2-OS cells (Fig. 4 and Supplementary Fig. 6). Treatment of both cell lines with preparations containing both mAb 2H3 and Fab 2H3, even in the absence of pre-treatment with siRNA, drastically affected their survival (survival rate of 5%; Fig. 5F). This indicates that the region of PCNA recognised by antibody 2H3 is also accessible in MelC and HL60R cells and that its occupancy with saturating amounts of inhibitory antibody molecules impairs DNA replication. Since HL60R cells contain substantial amounts of PCNA in the cytoplasm (V. Witko-Sarsat, personal communication), we verified the location of antibody 2H3 in these cells after fixation and found that it was exclusively localised in the nucleus at 2 days post-transduction (Supplementary Fig. 7A). Moreover, as observed in HeLa cells, treatment with antibody 2H3 led to severe genomic DNA fragmentation, which was comparable to that observed upon prolonged incubation of HL60R cells with 2 mM hydroxyurea (HU, Supplementary Fig. 7B), a potent DNA replication inhibitor that promotes DNA damage and DSB formation by depletion of the nucleotide pool [49].

We also investigated the performance of Fab SJK 132, which binds to DNA polymerase alpha in these cell lines, and compared the survival rate of cells transduced with this reagent with that observed following the addition of 1 μM etoposide to the cell culture medium. As expected, γ-H2AX was easily detectable after transduction with Fab SJK 132 and no signal was obtained after transduction with control Fab 2B6 or with Fab 4E9, which is derived from a non-inhibitory anti-DNA polymerase alpha antibody [16]. In particular, the intensity of γ-H2AX staining in MelC cells after 2 days of incubation was similar to that observed after the third day of treatment with either HU or etoposide alone (Fig. 6A). This was confirmed by Western blotting (Supplementary Fig. 8), suggesting that induction of γ-H2AX mediated by the Fab is faster than with chemicals that promote DNA stress. Notably, as observed with Fab 2H3, the size of HL60R cell nuclei was increased by about

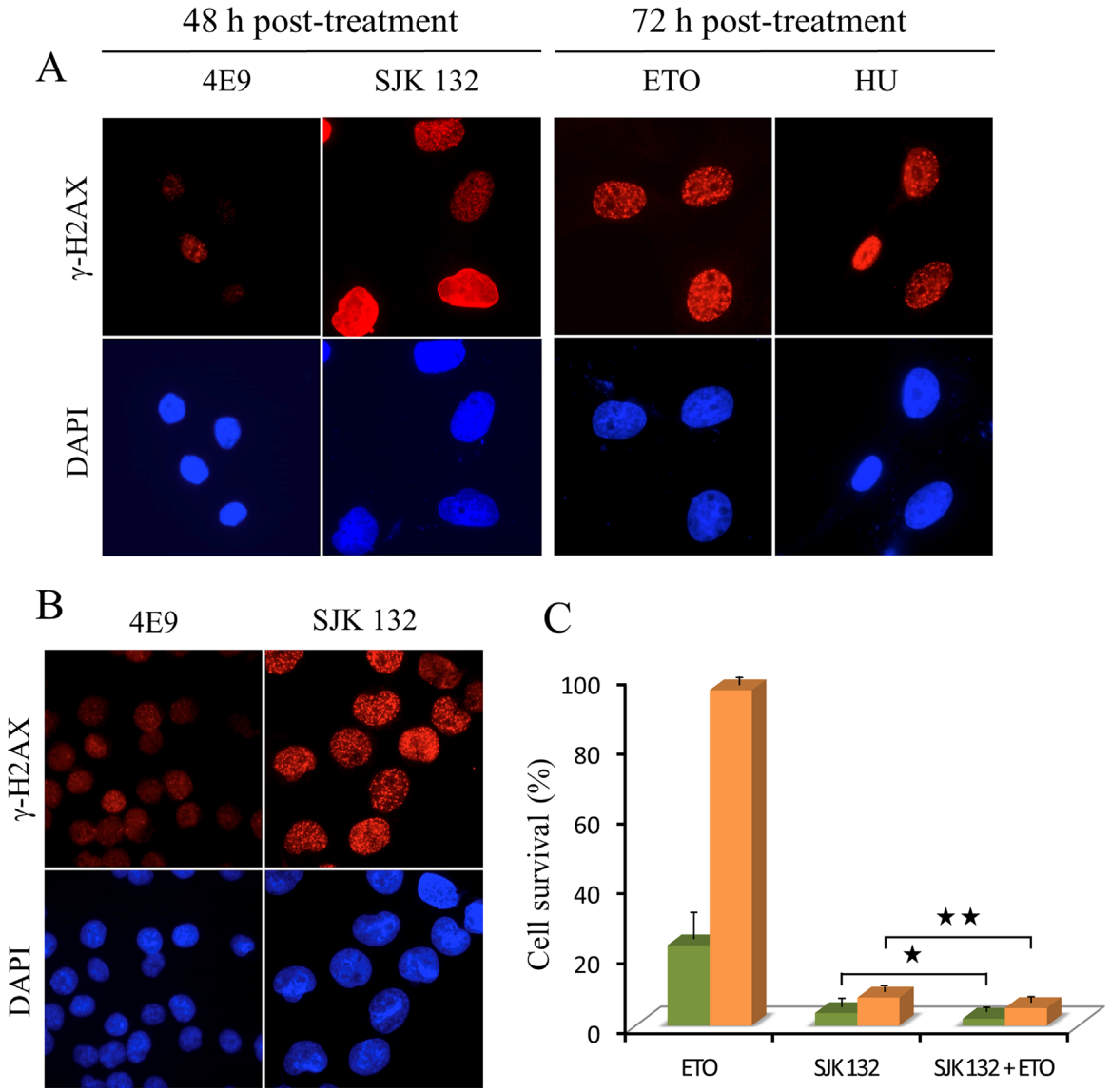


Fig. 6. Effect of Fab SJK 132 in the MelC and HL60R cell lines. (A) γ-H2AX induction following treatment of MelC cells with Fab 4E9, Fab SJK 132, etoposide (1 μM) and HU (2 mM). The cells were fixed at the time indicated and γ-H2AX staining was monitored by immunofluorescence microscopy with an exposure time of 500 ms. Typical cells after counterstaining with DAPI are show. Magnification: x 400. (B) γ-H2AX induction following treatment of HL60R cells with Fab 4E9 and Fab SJK 132 48 h post-transduction. The cells were analysed as in A. (C) Survival rate after co-treatment with Fab SJK 132 and 1 μM etoposide. MelC (green columns) or HL60R (orange columns) cells were either transduced with Fab SJK 132 or treated by addition of etoposide in the culture medium or transduced with Fab SJK 132 and treated with the drug. The number of cells remaining alive after 3 days of incubation was recorded and percentages were calculated using untreated cells as reference. The data presented are the mean values of 4 independent experiments. The *P* values were calculated with the Student's *t* test and are relative to the indicated measurements. ★, $P < 0.05$; ★★, $P < 0.01$.

2–3 fold following transduction with Fab SJK 132 (Fig. 6B), indicating that blocking either PCNA or DNA polymerase alpha in these chemo-resistant cells leads to a comparable cell phenotype.

To examine if the treatment with Fab SJK 132 is as efficient as treatment with etoposide in triggering cell death, we calculated the survival rate of MelC and HL60R cells after either incubation with etoposide alone or transduction. In parallel, we analysed whether co-treatment with Fab SJK 132 and etoposide is beneficial for killing the cells. In both cell lines, the percentage of cells remaining alive was lower after Fab transduction than upon etoposide treatment (Fig. 6C). Nearly all MelC cells were floating in the culture supernatant upon Fab SJK 132 transduction. Interestingly, induction of cell death with Fab SJK 132 in the chemo-resistant HL60R cell line was even more pronounced when etoposide was added after transduction (Fig. 6C). These results suggest that cytotoxic DNA replication stress induced by the anti-polymerase alpha Fab can be potentiated with a stress-promoting drug that targets another essential DNA replication actor (*i.e.*, topoisomerase II). It is worth mentioning that no significant induction of γ-H2AX or cytotoxicity was detectable after delivery of the non-inhibitory anti-DNA polymerase alpha Fab 4E9 (Fig. 6) and that the percentage of surviving MelC cells was comparable after either co-treatment with etoposide and Fab 4E9 or treatment with etoposide alone (not shown). Together, these results indicate that inhibition of replisome proteins by means of neutralising antibodies that act at the replication fork are efficient for triggering cell death not only in primary cells but also in cells with acquired chemo-resistance.

4. Discussion

Antibody-based intracellular targeting is a promising novel approach for targeting the action of key proteins inside cells [35]. However, it has not been possible until now to introduce immunoglobulins on their own in living cells, with the exception of a few polyreactive autoantibodies that bind preferentially to DNA [62]. Here, we show that mAbs and Fabs can be efficiently delivered to cancer cells by electroporation and that this robust system of transduction allows essential components of the replisome to be targeted in a very efficient manner. Among our newly isolated anti-PCNA mAbs, we identified two antibodies (also used as Fabs) which upon transduction, strongly inhibit DNA elongation *in vitro* and abolished DNA synthesis in different cancer cell lines. To our knowledge, this is the first report of the selection and characterisation of mAbs neutralising PCNA under physiological conditions. These antibodies likely recognise a conformational epitope located in the C-terminal region of PCNA, which is accessible in the *in vivo* context after PCNA neo-synthesis. The fact that after transduction, these newly developed anti-PCNA mAbs localised to the nucleus suggests that they remained bound to PCNA during its nuclear import and could inhibit the specific binding of PCNA to replicative DNA polymerases delta and epsilon that act at the fork. This is supported by the observation that DNA polymerase delta has been identified as the “acting polymerase” in the extracts used for the *in vitro* DNA elongation assay [3]. Another possibility is that loading of PCNA onto chromatin might be impaired by destabilisation of the trimeric active form of PCNA in the nucleus, since only complexes of Fab-PCNA monomers were observed by gel filtration. This would also impact the processivity of replicative DNA polymerases, because it is well established that the trimeric form of PCNA is essential for fork progression [26]. The fact that antibody 2B6 did not bind to the recombinant trimeric form, but was efficiently piggybacked into the nucleus, suggests that several forms of PCNA, not only the trimeric form, are imported into the nucleus. Moreover, to amplify the effects of the anti-PCNA antibodies, it was necessary to reduce the intracellular level of PCNA with siRNA, whereas the biological activity of DNA polymerase alpha was blocked when a similar amount of specific Fab was used alone. The abundance of DNA polymerase alpha molecules in mammalian cells is far lower than that of PCNA [58], which is in the range of 10^6 molecules per cell [40]. It seems that approximately 3×10^6 Fab molecules were delivered per cell when using 5 μg in the electroporation reaction mixture [16] and results not shown), suggesting that the delivered Fabs were present in the cell in excess only when targeting DNA polymerase alpha. The amount of antigen in the cell is thus a critical parameter for successful targeting with our protein delivery system. Furthermore, the Fab format of the antibody was most effective that because these molecules can diffuse rapidly into the nucleus to bind to the antigen. As both PCNA and DNA polymerase alpha antibodies were obtained from immunised mice, resulting in high affinity molecules [57], it was not surprising that monovalent binding was sufficient to neutralise fork progression and promote massive cell death in all tested cell lines.

Significant enlargement of both the cell cytoplasm and nucleus, that may correspond to senescence features [56], before ultimate death upon treatment with DNA-damaging compounds has not been extensively described before. It has been proposed that the induction of senescence and apoptosis could represent a barrier to tumorigenesis in injured pre-cancerous cells [5]. However, it is unclear whether this typical nuclear morphology is a direct consequence of extensive DNA double-strand breakage that leads to chromatin unfolding and expansion [31]. We systematically found that, soon after treatment with the inhibitory Fabs, the γ-H2AX levels increased considerably and the genomic DNA was significantly cleaved in cells having enlarged nuclei. This suggests that the binding of these antibodies promotes the collapse of the majority of forks in progression, a situation that may not be attainable when using clinically-relevant doses of genotoxic agents. Thus, enlargement of the nuclear volume may only occur above a certain threshold of DSB level. Interestingly, we could not observe this phenotype after addition of the inhibitor T2AA [44]. However, this phenomenon could easily be observed by microscopy after treatment with high doses of the DNA-intercalating drug cisplatin (not shown) or after prolonged incubation with HU or etoposide, two drugs that induce sustained replication stress [13]. Furthermore, by estimating the number of individual γ-H2AX foci after Fab treatment on super-resolution 3D-SIM micrographs, it appeared that this number is in the range of that reported for replication units in S-phase cells using similar microscopic approaches [7]. However, even if H2AX phosphorylation represents an amplified signal of the initial DSBs [29], it is possible that our antibody treatment halts the majority of on-going DNA replication processes in the cell (*i.e.*, almost all forks in progression would be blocked), a stress condition that cannot be rectified and that inevitably leads to apoptosis. This is in agreement with previous reports that cells containing large numbers of DSBs that overwhelm their capability for repair undergo death by activating one of the programmed death pathways [23,48]. In future studies, it might thus be interesting to determine, with regard to the target, the minimal dose of transduced Fabs that is required to reach this point of no return to cycling and to analyse how many unrepaired DSBs need to form to induce apoptosis signalling. This would be possible with our protein system of stress induction because irrelevant Fabs do not show intrinsic cytotoxicity and do not damage DNA chemically (see Figs. 5 and 6), presumably because they do not interfere wit DNA repair mechanisms. The observed systematic increase in nuclear size following the accumulation of DSBs under our Fab transduction conditions might therefore be the global result of numerous local chromatin expansions that occur in the vicinity of the DNA breaks that cannot be solved by the DSB repair

system [31,49]. Nevertheless, we also observed an increase in the area of the cytoplasm, in parallel with the abnormal nuclear shape (Fig. 3). It has been proposed that chromatin reorganisation in aging cells not only alters the expression of nuclear lamina proteins, but also correspond to the reorganisation of the endoplasmic reticulum and associated tubules, suggesting that cytoplasmic factors determine nuclear size [61]. It is thus possible that the changes observed in nuclear shape upon extensive DSB formation are not only a consequence of chromatin relaxation at DSB sites, but likely correspond to a global effect of the deregulation of gene expression upon chromatin fragmentation.

It was recently established that oncoproteins promote sustained DNA replication stress in cancer cells and that this stress, which results in genetic rearrangements, is essentially counteracted by the cellular replicative stress response (RSR) to maintain viability [17,21]. On the other hand, drugs that limit the activity of several essential actors of the RSR, such as ATR and Chk1 kinases, are very efficient for the selective killing of cancer cells [39,53,6], because they have almost no effect on normal cells that are, in principle, not under stress. Indeed, the RSR helps cancer cells to cope with replication stress, but unresolved DNA damage in this context remains nevertheless lethal, which suggests that the combined treatment of cancer cells with stress inducers and RSR inhibitors would be most beneficial for their elimination. We have not checked whether RSR inhibitors synergise with the Fabs described here, but it is possible that such dual treatment could lead to synthetic lethality at the cellular level, thereby necessitating a smaller amount of antibody to be delivered in order to observe cytotoxicity. If smaller amounts of Fabs are required, it might be possible to use recently developed protein delivery methods that are amenable to *in vivo* trials [43], instead of electroporation. Interestingly, a recent study that aimed to identify components of the cell that display synthetic lethal interactions with inhibition of ATR showed that DNA polymerase alpha is one of the best hits [37]. This confirms that this enzyme represents an excellent target for triggering lethal DNA replication stress. It is also worth mentioning that the killing effect of the antibodies could be demonstrated in both p53-positive (U2-OS) and p53-negative cells (HeLa, HL60R), thereby widening the applicability of the described approach to a large panel of cancer cells. Importantly we show also that cells resistant to chemicals that trigger replication stress are not resistant to antibodies that lead to the same type of stress, indicating that these cells are still susceptible to replicative stress. The antibody-based intervention may thus represent an ideal solution for treating cells with acquired chemo-resistance.

This study clearly shows that antibodies that neutralise essential replication factors *in vitro* are also inhibitory inside the nucleus. These antibodies are able to kill cancer cells *in vivo* by enhancing replicative stress and could be used as a potential novel cancer treatment approach by targeting components of the replisome complex. As suggested by [32], it may be possible to develop small molecules that target the epitopes recognised by such mAbs. Another possibility would be to use a multi-antibody delivery strategy for targeting the replisome with Fab fragments able to modulate the dynamics of RPA [19,56]. Such a multi-delivery strategy could possibly kill cancer cells by accelerating fork collapse and subsequent DNA breakage without using drugs for intervening in the replication stress response. Our results demonstrate that transduction with specific antibodies can inhibit the activity of essential replisome proteins in dividing cancer cells, suggesting that such an approach, which allows to discover and validate functional accessible sites of intracellular targets, may constitute a novel strategy for cancer therapy.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We thank Dr. V. Witko-Sarsat for stimulating discussions and for the gift the HL60R cell line, Dr. A. Coignard for the gift of the MelC cell line, Dr. G. Illuzzi and Dr. J.C. Amé for the gift of validated antibodies, Dr. C. Guilini for FACS assistance and Dr. A. Cordonnier for useful advice. We acknowledge the Imaging Core Facility (IMCF, University of Basel) and in particular A.I. Ferrand and K. Schleicher for the technical assistance provided on the OMX microscope. This work was supported by the Ligue contre le Cancer (CCIR-Grand Est), the University of Strasbourg (UdS), the Centre National de la Recherche Scientifique (CNRS) and the Agence Nationale de la Recherche (ANR-11-BSV5-010-02 Chromact; ANR-12-RPIB-0012-04 Oncovaccine, ANR-10-LABX-0030-INRT). LT is recipient of an European Research Council (ERC) advanced grant (ERC-2013-340551 Birtoaction).

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.yexcr.2016.03.003.

References

[1] Andrew S.M., Titus, J.A., 2003. Fragmentation of imunoglobulin G. Curr. Protoc. Cell Biol. 16, 16.4.
[2] D. Baddeley, V.O. Chagin, L. Schermelleh, S. Martin, A. Pombo, P.M. Carlton, A. Gahl, P. Domaing, U. Birk, H. Leonhardt, C. Cremer, M.C. Cardoso, Measurement of replication structures at the nanometer scale using super-resolution light microscopy, Nucl. Acids Res. 38 (2009) e8.
[3] N. Baldeck, J. Janel-Bintz, J. Wagner, A. Tissier, R.P. Fuchs, P. Burkovics, L. Haracska, E. Despras, M. Bichara, B. Chatton, A.M. Cordonnier, FF483-484 motif of DNA pol eta mediates its interaction with the POLD2 subunit of pol delta and contributes to damage tolerance, Nucl. Acids Res. 43 (2015) 2116–2125.
[4] A.G. Baranovskiy, N.D. Babayeva, Y. Suwa, J. Gu, Y.I. Pvlov, T.H. Tahirov, Structural basis for inhibition of DNA replication by Aphidicolin, Nucl. Acids Res. 42 (2014) 14013–14021.
[5] J. Bartkova, N. Rezaei, M. Liontos, P. Karakaidos, D. Kiestas, N. Issaeva, L. F. Vassiliou, E. Kolettas, K. Niforou, V.C. Zoumpourlis, M. Takaoka, H. Nakagawa, F. Tort, K. Fugger, F. Johansson, M. Sehested, C.L. Andersen, L. Dyrskjot, T. Orntoft, J. Lukas, C. Kittas, T. Helleday, T.D. Halazonetis, J. Bartek, Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints, Nature 444 (2006) 633–637.
[6] J. Benada, L. Macurek, Targeting the checkpoint to kill the cancer cells, Biomolecules 5 (2015) 1912–1937.
[7] M.C. Cardoso, K. Schneider, R.M. Martin, H. Leonhardt, Structure, function and dynamics of nuclear subcompartments, Curr. Opin. Cell Biol. 24 (2012) 79–85.
[8] A.M. Cordonnier, A.R. Lehmann, R.P. Fuchs, Impaired translesion synthesis in xeroderma pigmentosum variant extracts, Mol. Cell. Biol. 19 (1999) 2206–2211.
[9] Z. Cseresnyes, U. Schwartz, C.M. Green, Analysis of replication factories in human cells by super-resolution light microscopy, BMC Cell Biol. 10 (2009) 88.
[10] N.J. Curtin, Inhibiting the DNA damage response as a therapeutic manoeuvre in cancer, Br. J. Pharmacol. 169 (2013) 1745–1765.
[11] Z. Darzynkiewicz, H.D. Halicka, H. Zhao, M. Podhorecka, Cell synchronization by inhibitors of DNA replication induces replication stress and DNA damage response: analysis by flow cytometry, Methods Mol. Biol. 761 (2011) 85–96.
[12] K.L. Dillehay, W.L. Seibel, D. Zhao, S. Lu, Z. Dong, Target validation and structure activity analysis of a series of novel PCNA inhibitors, Pharmacol. Res. Perspect. 3 (2015) e00115.
[13] M. Dobbelstein, C.S. Sorensen, Exploiting replicative stress to treat cancer, Nat. Rev. Drug. Discov. 14 (2015) 405–423.
[14] C. Dong, Q. Li, S. Lyu, A.M. Krensky, C. Clayberger, A novel apoptosis pathway activated by the carboxyl terminus of p21, Blood 105 (2008) 1187–1194.
[15] G. Fregni, A. Perier, G. Pittari, S. Jacobelli, X. Sastre, N. Gervois, M. Allard, N. Bercovici, M.F. Avril, A. Caignard, Unique functional status of natural killer cells in metastatic stage IV melanoma patients and its modulation by chemotherapy, Clin. Cancer Res. 17 (2011) 2628–2637.

[16] G. Freund, A.P. Sibler, D. Desplancq, M. Oulad-Abdelghani, M. Vigneron, J. Gannon, M.H. Van Regenmortel, E. Weiss, Targeting endogenous nuclear antigens by electrotransfer of monoclonal antibodies in living cells, mAbs 5 (2013) 518–522.
[17] H. Gaillard, T. Garcia-Muse, A. Aguilera, Replication stress and cancer, Nat. Rev. Cancer 15 (2015) 276–289.
[18] R. Georgescu, L. Langston, M. O'Donnell, A proposal: evolution of PCNA's role as a marker of newly replicated DNA, DNA Repair (Amst) 29 (2015) 4–15.
[19] J.G. Glanzer, S. Liu, L. Wang, A. Mosel, A. Peng, G.G. Oakley, RPA inhibition increases replication stress and suppresses tumor growth, Cancer Res. 74 (2014) 5165–5172.
[20] S. Grimaudo, M. Tolomeo, A. Chimirri, M. Zappala, R.A. Gancitano, N. D'Alessandro, Selective induction of apoptosis in multidrug resistant HL60R cells by the thiazolobenzoimidazole derivative TBZ, Eur. J. Cancer 34 (1998) 1756–1763.
[21] T.D. Halazonetis, V. Gorgoulis, J. Bartek, An oncogene-induced DNA damage model for cancer development, Science 319 (2008) 1352–1355.
[22] A. Helmrich, M. Ballarino, E. Nudler, L. Tora, Transcription-replication encounters, consequences and genomic instability, Nat. Struct. Mol. Biol. 20 (2013) 412–418.
[23] S.A. Hills, J.F.X. Diffley, DNA replication and oncogene-induced replicative stress, Curr. Biol. 24 (2014) R435–R444.
[24] N. Hosoya, K. Miyagawa, Targeting DNA damage response in cancer therapy, Cancer Sci. 105 (2014) 370–388.
[25] A. Inoue, S. Kiuchi, A. Hishiki, Y. Shao, R. Heath, B.J. Evison, M. Actis, C. E. Camnan, H. Hashimoto, N. Fujii, A small molecule inhibitor of monoubiquitinated proliferating cell nuclear antigen inhibits repair of interstrand DNA cross-links, enhances DNA double-strand break and sensitizes cancer cells to cisplatin, J. Biol. Chem. 289 (2014) 7109–7120.
[26] Z.O. Jonsson, V.N. Podust, L.M. Podust, U. Hübscher, Tyrosine 114 is essential for the trimeric structure and the functional activities of human proliferating cell nuclear antigen, Embo J. 14 (1995) 5745–5751.
[27] R. Jossen, R. Bermejo, The DNA damage checkpoint response to replication stress: a game of forks, Front. Genet. 4 (2013) 26.
[28] T. Kanda, K.F. Sullivan, G.M. Wahl, Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells, Curr. Biol. 8 (1998) 377–385.
[29] A. Kinner, W. Wu, C. Staudt, G. Iliakis, H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin, Nucl. Acids Res. 36 (2008) 5678–5694.
[30] P. Kotsantis, R.M. Jones, M.R. Higgs, E. Petermann, Cancer therapy and replication stress: forks on the road to perdition, Adv. Clin. Chem. 69 (2015) 91–137.
[31] M.J. Kruhlak, A. Celeste, G. dellaire, O. Fernandez-Capetillo, W.G. Müller, J. G. McNally, D.P. Bazett-Jones, A. Nussenzweig, Changes in chromatin structure and mobility in living cells at sites of DNA double-strand breaks, J. Cell Biol. 172 (2006) 823–834.
[32] A.D. Lawson, Antibody-enabled small drug discovery, Nat. Rev. Drug. Discov. 11 (2012) 519–525.
[33] M. Macheret, T.D. Halazonetis, DNA replication stress as a hallmark of cancer, Ann. Rev. Pathol. 10 (2015) 425–448.
[34] N. Mailand, I. Gibbs-Seymour, S. Bekker-Jensen, Reulation of PCNA-protein interactions for genome stability, Nat. Rev. Mol. Cell Biol. 14 (2013) 269–282.
[35] A.L.J. Marschall, S. Dübel, T. Böldicke, Specific in vivo knockdown of protein function by intrabodies, mAbs 7 (2015) 1–26.
[36] A. Mazouzi, G. Velimezi, J.I. Loizou, DNA replication stress: causes, resolution and disease, Exp. Cell Res. 329 (2014) 85–93.
[37] D.L. Menezes, J. Holt, Y. Tang, J. Feng, P. Barsanti, Y. Pan, M. ghoddusi, W. Zhang, G. thomas, J. Holash, E. Lees, L. Taricani, A synthetic lethal screen reveals enhanced sensitivity to ATR inhibitor treatment in mantle cell lymphoma with ATM loss-of-function, Mol. Cancer Res. 13 (2014) 120–129.
[38] G.L. Moldevan, B. Pfander, S. Jentsch, PCNA, the maestro of the replication fork, Cell 129 (2007) 665–679.
[39] M. Murga, S. Campaner, A.J. Lopez-Contreras, L.I. Toledo, R. Soria, M. F. Montana, L. D'Artista, T. Schleker, C. Guerra, E. Garcia, M. Barbacid, M. Hidalgo, B. Amati, O. Fernandez-Capetillo, Exploiting oncogene-induced replicative stress for the selective killing of myc-driven tumors, Nat. Struct. Mol. Biol. 18 (2011) 1331–1335.
[40] S.N. Naryzhny, Proliferating cell nuclear antigen: a proteomic view, Cell. Mol. Life Sci. 65 (2008) 3789–3808.
[41] H. Offer, N. Erez, I. Zurer, X. tang, M. Milyavsky, N. Goldfinger, V. Rotter, The onset of p53-dependent DNA repair or apoptosis is determined by the level of accumulated damaged DNA, Carcinogenesis 23 (2002) 1025–1032.
[42] C. Olaisen, R. Müller, A. Nedal, M. Otterlei, PCNA-interacting peptides reduce Akt phosphorylation and TLR-mediated cytokine secretion suggesting a role of PCNA in cellular signaling, Cell Signal. 27 (2015) 1478–1487.
[43] V. Postupalenko, D. Desplancq, I. Orlov, Y. Arntz, D. Spehner, Y. Mely, B. P. Klaholtz, P. Schultz, E. Weiss, G. Zuber, Protein delivery system containing a nickel-immobilized polymer for multimerization of affinity-purified his-tagged proteins enhances cytosolic delivery, Angew. Chem. Int. Ed. Engl. 54 (2015) 10583–10586.
[44] C. Punchihewa, A. Inoue, A. Hishiki, Y. Fujikawa, M. Connelly, B. Evison, Y. Shao, R. Heath, I. Kuraoka, P. Rodrigues, H. Hashimoto, M. Kawanishi, M. Sato, T. Yagi, N. Fujii, Identification of small molecule PCNA inhibitor that disrupts interactions with PIP-box proteins and inhibits DNA replication, J. Biol. Chem. 287 (2012) 14289–14300.
[45] A.S. Rinaldi, G. Freund, D. Desplancq, A.P. Sibler, M. Baltzinger, N. Rochel, Y. Mély, P. Didier, E. Weiss, The use of fluorescent intrabodies to detect endogenous gankyrin in living cancer cells, Exp. Cell Res. 319 (2013) 838–849.
[46] G. Roos, Y. Jiang, G. Landberg, N.H. Nielsen, P. Zhang, M.Y. Lee, Determination of the epitope of an inhibitory antibody to proliferating cell nuclear antigen, Exp. Cell Res. 226 (1996) 208–213.
[47] W.P. Roos, B. Kaina, DNA-damage-induced cell death by apoptosis, Trends Mol. Med. 12 (2006) 440–450.
[48] W.P. Roos, B. Kaina, DNA damage-induced cell death: from specific DNA lesions to the DNA damage response and apoptosis, Cancer Lett. 337 (2013) 237–248.
[49] P. Rybak, A. Waligorska, L. Bujnowicz, A. Hoang, J.W. Dobrucki, Activation of new replication foci under conditions of replication stress, Cell Cycle 14 (2015) 2634–2647.
[50] L. Schermelleh, P.M. Carlton, S. Haase, L. Shao, L. Winoto, P. Kner, B. Burke, M. C. Cardoso, D.A. Agard, M.G.L. Gustafson, H. Leonhardt, J.W. Sedat, Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy, Science 320 (2008) 1332–1336.
[51] A.O. Sidi, K.O. Babah, N. Brimer, Y. Nominé, C. Romier, B. Kieffer, S.V. Pol, G. Travé, K. Zanier, Strategies for bacterial expression of protein-peptide complexes: application to solubilization of papillomavirus E6, Protein Exp. Purif. 80 (2011) 8–16.
[52] S. Tanaka, S.Z. Hu, T.S. Wang, D. Korn, Preparation and characterization of monoclonal antibodies against human DNA polymerase alpha, J. Biol. Chem. 257 (1982) 8386–8390.
[53] E.M. Taylor, H.D. Lindsay, DNA replication stress and cancer: cause or cure? Future Oncol. 12 (2016) 221–237.
[54] M. Tokunaga, N. Imamoto, K. Sakata-Sogawa, Highly inclined thin illumination enables clear single-molecule imaging in cells, Nat. Methods 5 (2008) 159–161.
[55] L.I. Toledo, M. Murga, O. Fernandez-Capetillo, Targeting ATR and Chk1 kinases for cancer treatment: a new model for new (and old) drugs, Mol. Oncol. 5 (2011) 368–373.
[56] L.I. Toledo, M. Altmeyer, M.B. Rask, C. Lukas, D.H. Larsen, L.K. Povlsen, S. Bekker-Jensen, N. Mailand, J. Bartek, J. Lukas, ATR prohibits replication catastrophe by preventing global exhaustion of RPA, Cell 155 (2013) 1088–1103.
[57] M.H.V. Van Regenmortel, Specificity, polyspecificity, and heterospecificity of antigen-antibody recognition, J. Mol. Recognit. 27 (2014) 627–639.
[58] M. Wang, M. Weiss, M. Simonovic, G. Haertinger, S.P. Schrimpf, M. O. Hengartner, C. von Mering, PaxDb, a database of protein abundance averages across all three domains of life, Mol. Cell Proteom. 11 (2012) 492–500.
[59] S.C. Wang, PCNA: a silent housekeeper or a potential therapeutic target? Trends Pharmacol. Sci. 35 (2014) 178–186.
[60] N.H. Waseem, D.P. Lane, Monoclonal antibody analysis of the proliferating cell nuclear antigen (PCNA). Structural conservation and the detection of a nucleolar form, J. Cell Sci. 96 (1990) 121–129.
[61] M. Webster, K.L. Witkin, O. Cohen-Fix, Sizing up the nucleus: nuclear shape, size and nuclear-envelope assembly, J. Cell Sci. 122 (2009) 1477–1486.
[62] R.H. Weisbart, G. Chan, G. Jordaan, P.W. Noble, Y. Liu, P.M. Glazer, R. N. Nishimura, J.E. Hansen, DNA-dependent targeting of cell nuclei by a lupus autoantibody, Sci. Rep. 5 (2015) 12022.
[63] J. Zech, J.Z. Dalgaard, Replisome components – post-translational modifications and their effects, Semin. Cell Dev. Biol. 30 (2014) 144–153.
[64] M.K. Zeman, K.A. Cimprich, Causes and consequences of replication stress, Nat. Cell Biol. 16 (2014) 2–9.

Supplemental Material Desplancq et al., 2016:

Supplementary material. Figure S1: Binding characteristics of the anti-PCNA Fab molecules. (A) Binding of Fab 2H3 and 2B6 to trimeric PCNA or monomeric PCNA^{Y114A}. Pure PCNA preparations were mixed with a two-fold excess of Fab molecules as indicated for 1 h at room temperature and the mixtures were subsequently deposited on a Superdex gel filtration column calibrated with standard proteins. Typical elution profiles when monomeric PCNA was mixed with 2H3 (red) or trimeric PCNA mixed with 2H3 (blue) or 2B6 (black) are shown. The peaks at about 90 kDa and 80 kDa correspond to trimeric PCNA and Fab/monomeric PCNA complexes, respectively. The peak at about 50 kDa corresponds to either monomeric PCNA or free Fab. (B) Epitope binning. PCNA was pre-mixed with either PC10 or 2H3 Fabs (in brackets) and then either Fab 2H3 or Fab 2B6 were added for 1 h at room temperature. The complexes were analysed by gel filtration as in A. No trimolecular complex with an apparent molecular weight of approximately 130 kDa was observed when Fab 2H3 and Fab 2B6 were used in this assay. The eluted complexes present in the peaks are schematically depicted. (C) Binding capacity of Fab 4D6. The indicated mixtures were performed and analysed as in B. (D) Expression of MBP-PCNA (1-163) and Ubi-PCNA (164-261) fusions in *E. Coli* and Western-blot analysis. Total extracts from induced cells overexpressing MBP-PCNA(1-163) and Ubi-PCNA(164-261) polypeptides (lanes 1 and 2, respectively) were analysed by SDS-PAGE and Coomassie blue staining. The migration levels of the overexpressed fusions are indicated (arrow). M, molecular weight markers. Both extracts were probed by Western blotting after transfer to nitrocellulose. The picture shows strips after incubation with PC10 or 2B6 antibodies and subsequent detection of the bound antibodies with HRP-labelled immunoglobulins.

Supplementary material. Figure S2: Analysis of the purified mAbs and Fab fragments by SDS-PAGE and Coomassie staining. 5 µg of the indicated molecules were deposited on gel. The migration profile of molecular weight markers is indicated on the left. H, heavy chain; Fd/L: Fd or light chain.

Supplementary material. Figure S3: Inhibition of DNA synthesis in U2-OS cells. U2-OS cells were transduced with the indicated antibodies (20 µg) as in legend of Figure 2. Inhibition of DNA synthesis and non-incorporation of EdU is shown by the absence of labelled nuclei (green).

Supplementary material. Figure S4: siRNA sensitivity and effect of the anti-PCNA antibodies in U2-OS cells. (A) Analysis of the PCNA levels after treatment with siRNA by Western blotting. HeLa or U2-OS cells were treated with increasing amounts of siRNA. The cells were harvested 72 h post-transfection and whole-cell extracts (30 µg) were analysed by SDS-PAGE and Western blotting. The presence of PCNA and actin on the blot were revealed with PC10, anti-actin polyclonal antibodies and IR dye-

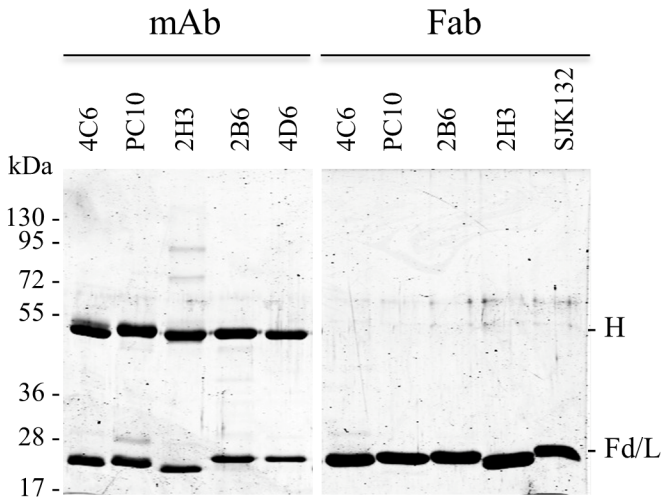
labelled anti-mouse or anti-rabbit immunoglobulins (Materials and Methods). (B) Antibody-mediated cell death induction. The cells were transfected with anti-PCNA siRNA for 48 h and subsequently transduced by electroporation with the indicated antibodies or Fab fragments. At 72 h post-transduction, the cells were observed by optical microscopy. The pictures show representative fields of the treated cells. Typical flattened cells with enlarged nuclei are indicated (arrows). PBS was used as antibody or Fab control. Magnification: x 200. (C) FACS analysis of U2-OS cells treated with either siRNA (si) or siRNA and 2B6 or 2H3 antibody. After 120 h of incubation, the treated cells were trypsinized and subjected to FACS analysis as indicated in the legend of Figure 3.

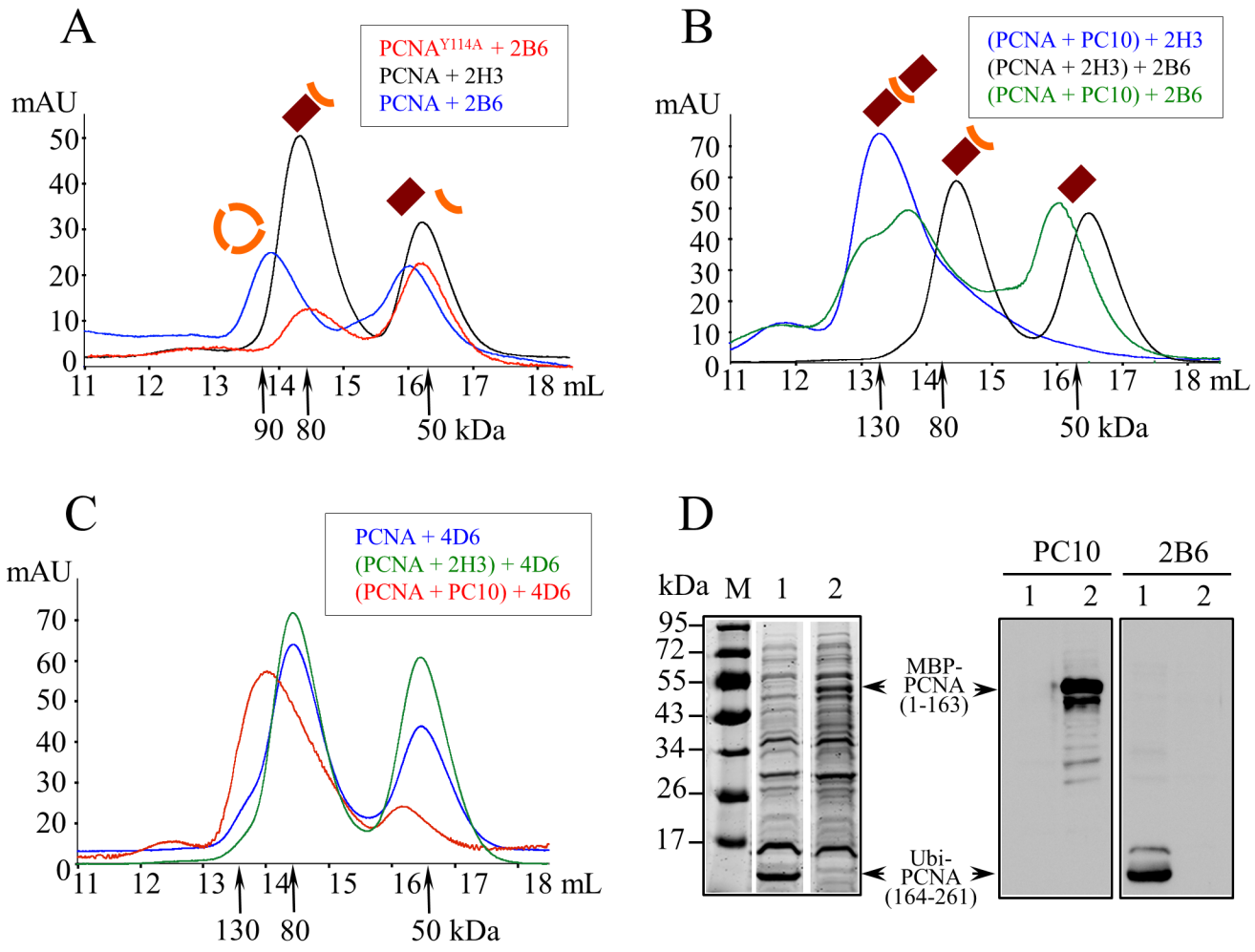
Supplementary material. Figure S5: Assessment of the down-regulation of PCNA. (A) HeLa cells were treated as indicated in the legend of Figure 3C. At the end of incubation, they were fixed and analysed by immunofluorescence staining after staining with anti-PCNA rabbit polyclonal antibodies. NT, not treated. Magnification: x 1000. (B) Western blot analysis of HeLa cells transfected with plasmid p□G8 (lane 2) or empty vector (lane 1). Crude extracts of approximately 5000 cells (6 µg) treated as in (A) were analysed by SDS-PAGE and blotting. PCNA and β-actin were revealed with PC10 and relevant anti-actin polyclonal antibodies, respectively.

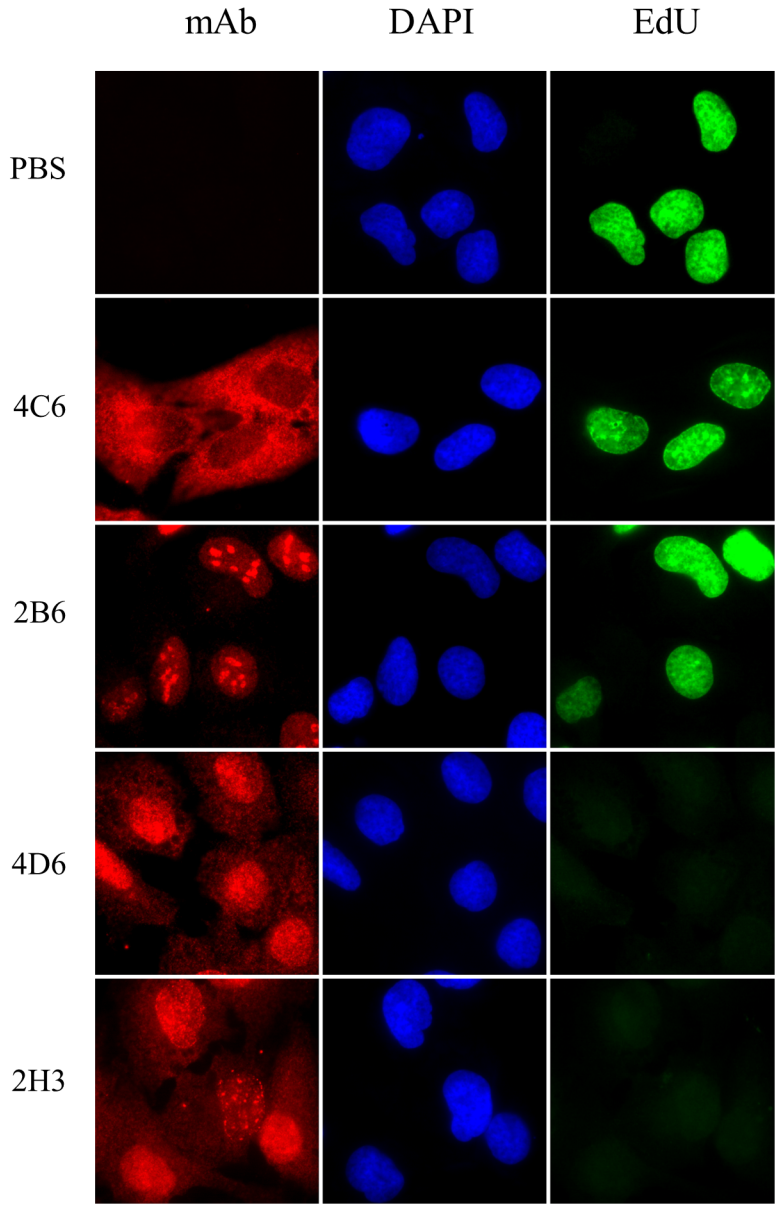
Supplementary material. Figure S6: Induction of γ-H2AX synthesis or RPA foci after PCNA gene disruption or transduction with either SJK 132 or 2H3 Fab fragments. (A) HeLa cells were treated as indicated in the legend of Figure 3C. After fixation, they were incubated with anti-γ-H2AX rabbit monoclonal antibody and Alexa Fluor 648-labelled secondary immunoglobulins. Before mounting for observation by immunofluorescence microscopy, DAPI was added. The pictures show representative fields of the observed cells. NT, not treated. Magnification: x 1000. (B) Relocalisation of RPA protein in HeLa cells after treatment with siRNA and Fab 2H3 at 72 h post-transduction. NT, not treated. (C) Detection of γ-H2AX in U2-OS as described in the legend of Figure 4C.

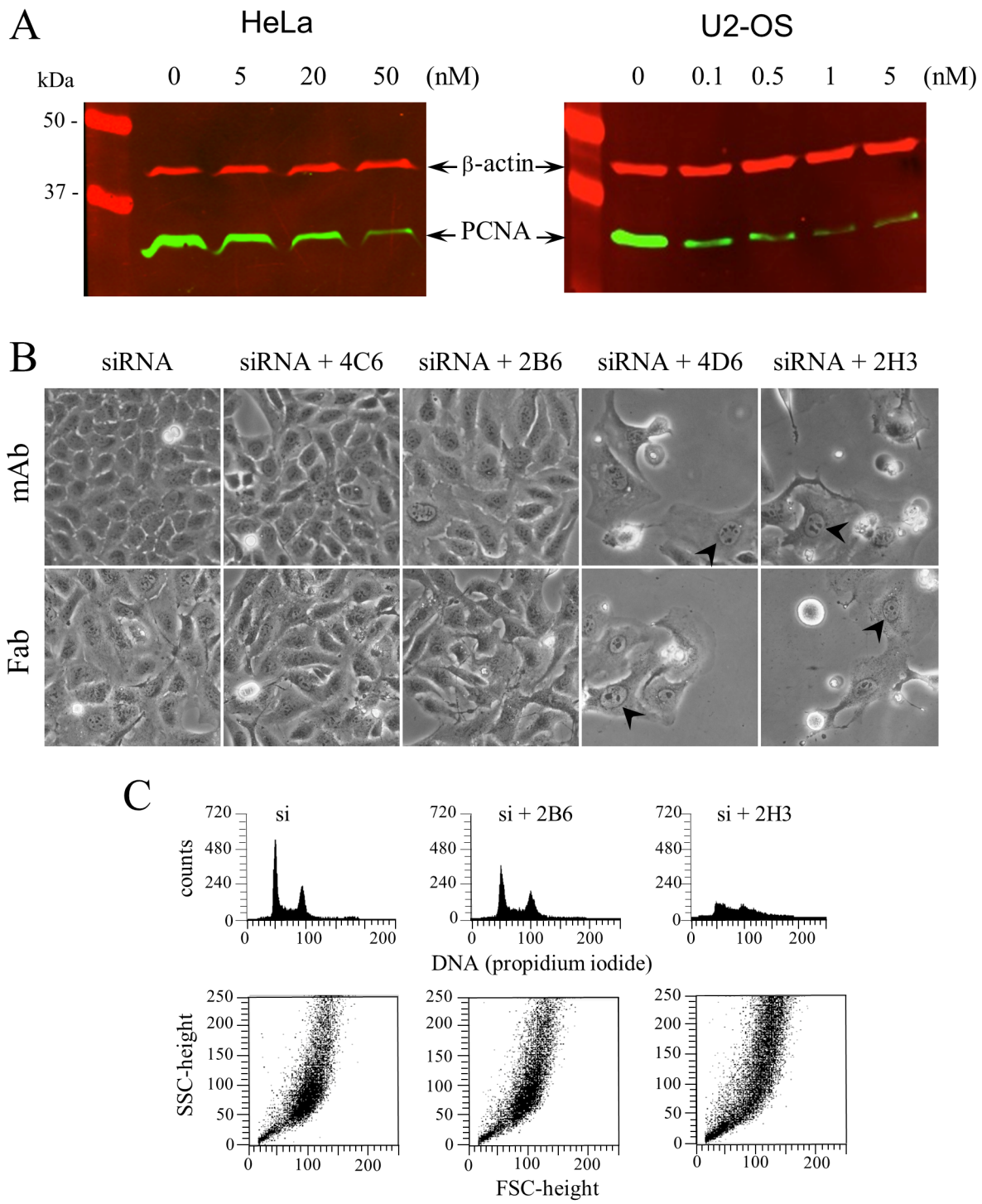
Supplementary material. Figure S7: Effect of 2H3 antibody in HL60R cells. (A) The cells were transduced with antibodies 2H3 and 4C6. 48 h post-transduction, the cells were fixed and the delivered antibodies were revealed with Alexa Fluor 488-labelled anti-mouse conjugate. The pictures show typical fields of cells observed by confocal microscopy. Magnification: x 630. (B) Analysis of the genomic DNA integrity. HL60R cells were treated with anti-PCNA siRNA and Fab 2B6 or Fab 2H3 and subjected to genomic DNA extraction at 72 h post-transduction (left). In parallel, HL60R were treated with 2 mM HU and genomic DNA extraction was performed at the indicated regular time points (right). The electrophoretic mobility of representative samples containing a similar amount of DNA was analysed on agarose gel. M, markers in kilobase pairs (kbp).

Supplementary material. Figure S8: Induction of γ-H2AX in MelC cells as probed by Western blotting. The cells were treated as indicated in the legend of Figure 6A and crude extracts containing a similar amount of protein (20 µg) were probed with relevant antibodies against the indicated polypeptides as indicated in the legend of Figure 4D

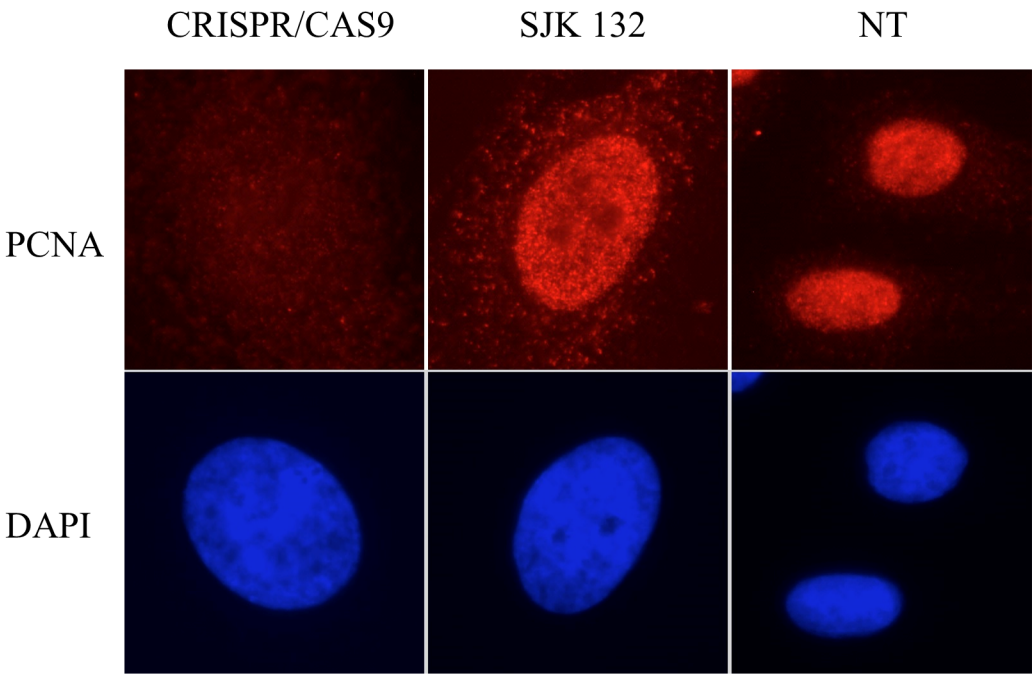

mAb
Fab
4C6
PC10
2H3
2B6
4D6
4C6
PC10
2B6
2H3
SJK132
kDa
130
95
72
55
36
28
17
H
Fd/L


A
PCNA^Y114A + 2B6
PCNA + 2H3
PCNA + 2B6
mAU
50
40
30
20
10
0
11 12 13 14 15 16 17 18 mL
90 80
50 kDa
B
(PCNA + PC10) + 2H3
(PCNA + 2H3) + 2B6
(PCNA + PC10) + 2B6
mAU
70
60
50
40
30
20
10
0
11 12 13 14 15 16 17 18 mL
130 80
50 kDa
C
PCNA + 4D6
(PCNA + 2H3) + 4D6
(PCNA + PC10) + 4D6
mAU
70
60
50
40
30
20
0
11 12 13 14 15 16 17 18 mL
130 80
50 kDa
D
kDa M 1 2
95
72
55
43
34
26
17
MBP-PCNA (1-163)
Ubi-PCNA (164-261)
PC10
2B6
1 2 1 2

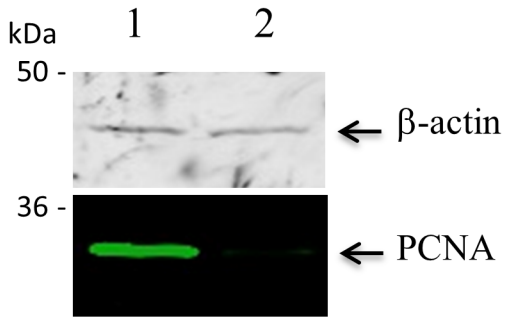

mAb
DAPI
EdU
PBS
4C6
2B6
4D6
2H3

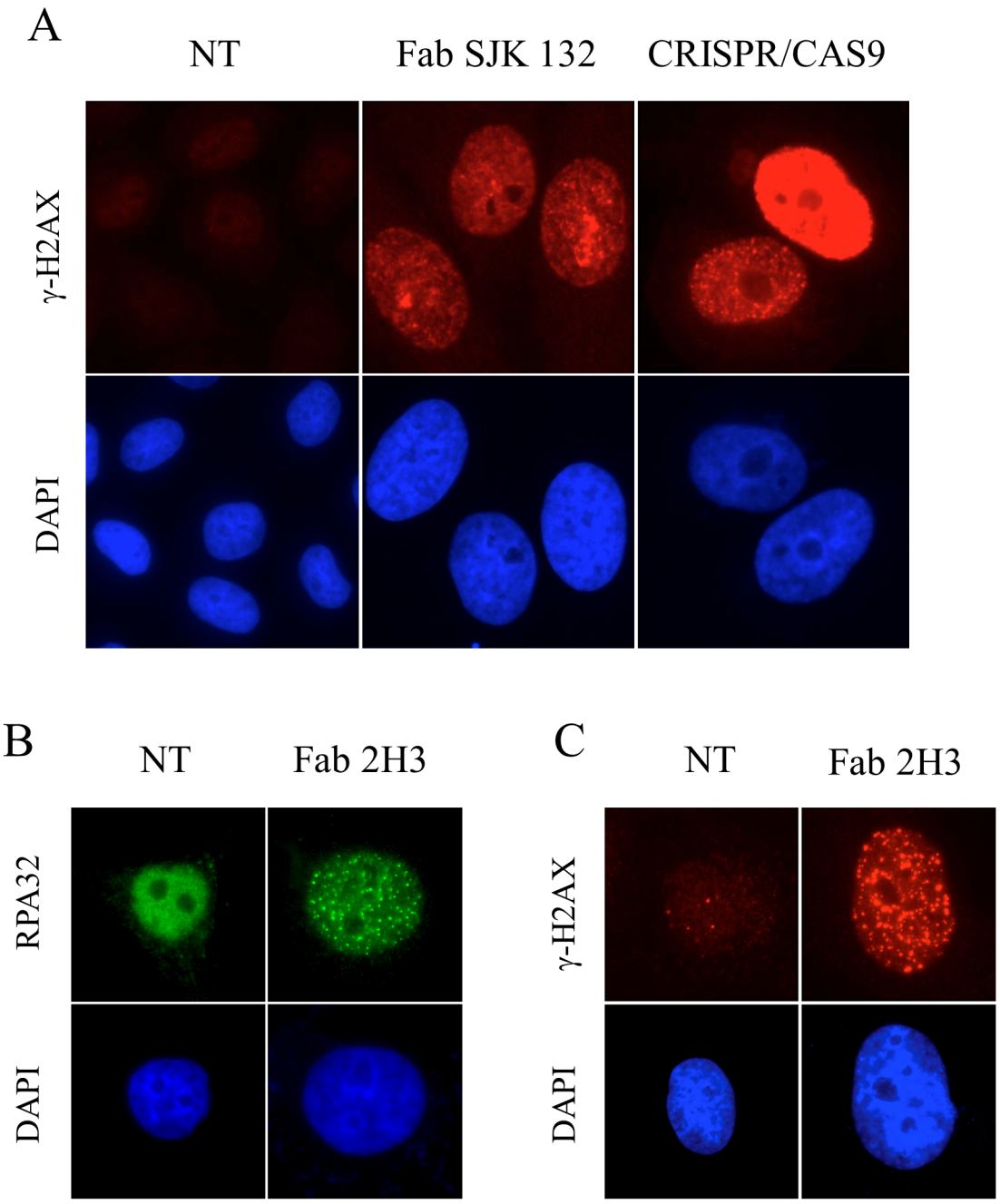

A
HeLa
U2-OS
kDa
0 5 20 50 (nM)
0 0.1 0.5 1 5 (nM)
50 -
37 -
β-actin
PCNA
B
siRNA
siRNA + 4C6
siRNA + 2B6
siRNA + 4D6
siRNA + 2H3
mAb
Fab
C
si
si + 2B6
si + 2H3
counts
720
480
240
0
100
200
DNA (propidium iodide)
SSC-height
250
200
150
100
50
FSC-height

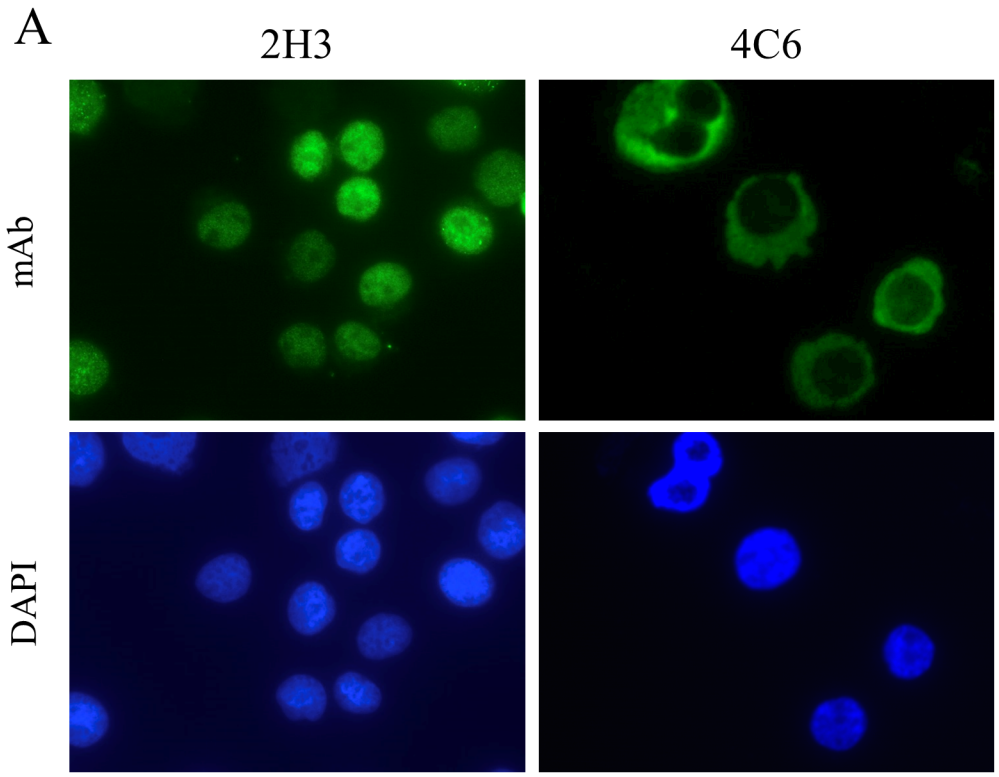
A

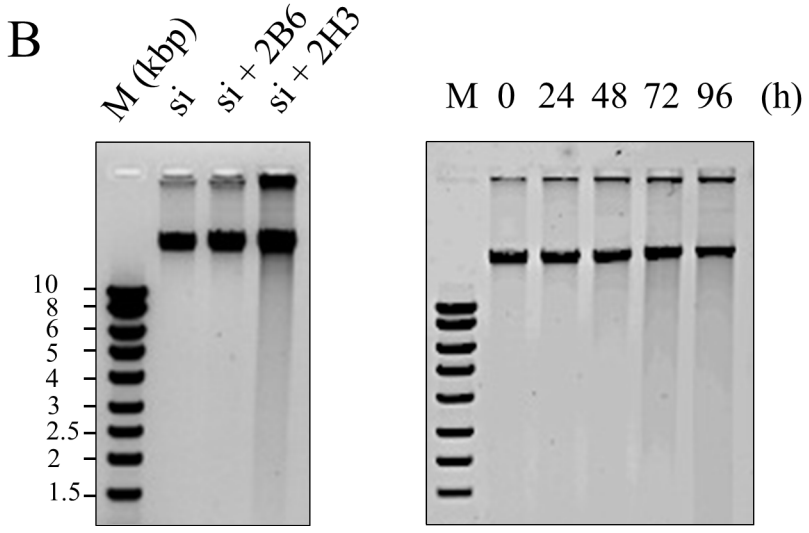


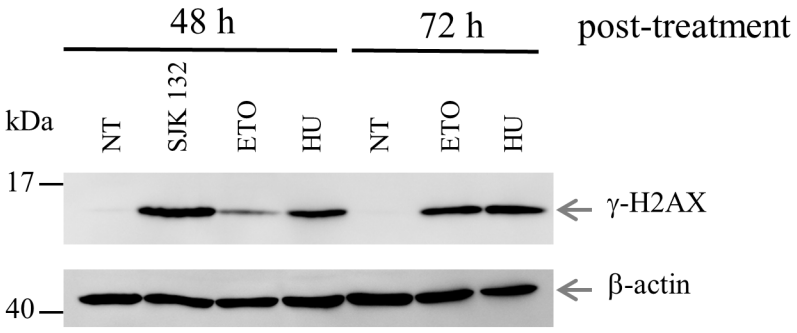
B




A
NT
Fab SJK 132
CRISPR/CAS9
γ-H2AX
DAPI
B
NT
Fab 2H3
RPA32
DAPI
C
NT
Fab 2H3
γ-H2AX
DAPI


A
2H3
4C6
mAb
DAPI


B
M (kbp)
si
si + 2B6
si + 2H3
10
8
6
5
4
3
2.5
2
1.5
M 0 24 48 72 96 (h)


48 h
72 h
post-treatment
kDa
NT
SJK 132
ETO
HU
NT
ETO
HU
17
γ-H2AX
40
β-actin

3. Imaging of native transcription factors and histone phosphorylation at high resolution in live cells. (S. Conic *et al.*; Journal of Cell Biology, 2018)

The fluorescent labeling of proteins to follow their spatiotemporal localization in real time was mainly achieved in the past by using transgenic or overexpression-based approaches. However, the specific labeling of endogenous factors or even PTMs in living cells is not yet routinely possible. The visualization of cellular structures and processes is typically performed either on fixed cells by using classical IF staining or in living cells by expressing exogenous fluorescent fusion proteins. Although these techniques showed to be very powerful to locate or track proteins inside the cells, they harbor some important drawbacks including fixation-related artifacts concerning IF staining or overexpression-related changes in the behavior of the fluorescent fusion proteins in contrast to their endogenous counterparts. Another possibility for live imaging of endogenous proteins is the knock-in of a fluorescent tag into the endogenous locus of the target protein using the CRISPR/Cas9 technology. However, these knock-in clones are often difficult to obtain as knock-in efficiencies are quite low and in addition this technique does not allow the specific labeling of PTMs. However, it was shown that the function of transcription factors and co-activator complexes which are involved in chromatin dependent processes are tightly linked to specific PTMs in the nuclear environment.

Consequently, there is a need for new imaging approaches to enable the specific labeling of endogenous target proteins and PTMs in living cells. Previous studies showed that intracellular targeting of proteins with antibodies or Fabs is possible (Hayashi-Takanaka *et al.*, 2011; Teng *et al*., 2016). However, these techniques suffered from low delivery efficiencies of the antibodies into the cell or low cell viability due to the harmful treatment to deliver the antibodies. In contrast, other studies including work that was presented already in this thesis showed that electroporation of antibodies results in high delivery efficiency and cell viability (Freund *et al*., 2013).

Having that, collaborators and I decided to develop a novel antibody-based imaging approach to label and track endogenous proteins and PTMs in living cells by using non-inhibiting antibodies and electroporation as delivery method.

- To address **aim a)**, I developed the novel versatile antibody-based imaging approach (VANIMA) which uses fluorescently labeled antibodies or Fabs. Several validation experiments were performed to ensure that the electroporated antibody is really binding to the target protein inside the living cells and that this binding does not inhibit the function of the target.
- To address **aim c)**, I studied the distribution of endogenous RNA Pol II and TAF10 using VANIMA and 3D-SIM microscopy. I also tested the changes of endogenous RNA Pol II clustering in the nucleus with or without transcription inhibition.
- To partially address **aim d)**, I performed live imaging tracking experiments of RNA Pol II and γH2AX using VANIMA and confocal as well as 3D-SIM microscopy.

Among the different obtained results, we were able to show that VANIMA can be used to label different transcription factors like RNA Pol II (through RPB1), TAF10 and TBP as well as a specific PTM in form of phosphorylated histone H2AX. Furthermore, we showed that the electroporated anti-RPB1 antibody is bound to the target inside the cell and that RPB1 can still incorporate into the RNA Pol II complex and bind to chromatin. Moreover, we tested that the transduced antibodies do not affect nascent transcription as well as cell proliferation and showed that they do not induce apoptosis. We performed 3D-SIM imaging of endogenous RNA Pol II and TAF10 and quantified the volume distribution of the RNA Pol II and TAF10 foci in the nucleus. Moreover, we analyzed the change in the foci volume of RNA Pol II and TAF10 foci with and without transcription elongation inhibition with flavopiridol. Interestingly, we observed that the number of larger RNA Pol II clusters was decreasing whereas the number of smaller foci was increasing after elongation inhibition. This suggests that these clusters correspond to transcription related RNA Pol II accumulations that dissociate after transcription is inhibited. Lastly, we tested if VANIMA can be used to track RNA Pol II clusters or γH2AX foci in living cells for hours (confocal microscopy) or for seconds (3D-SIM microscopy).

These results were published on the 12th of February 2018 in the *Journal of Cell Biology*.

Author's contributions

Sascha Conic – Designed the study. Performed all the experiments except the ones shown in Figure 3A, B and C as well as Supplemental Figure 1E. Analyzed and interpreted the data. Wrote the manuscript.

Dominique Desplancq – Characterized the anti-γH2AX antibody. Performed the experiment shown in Supplemental Figure 1E. Assisted in the writing of the manuscript.

Alexia Ferrand – Assisted in all 3D-SIM experiments presented in the study. Performed image processing of Figure 4, Figure 6A and Figure 7B and D. Assisted in the writing of the manuscript.

Veronique Fischer – Performed and analyzed the experiments shown in Figure 3A, B and C. Assisted in the writing of the manuscript.

Vincent Heyer – Designed and provided the vectors for the CRISPR/Cas9 knock-in experiments.

Bernardo Reina San Martin – Designed and provided the vectors for the CRISPR/Cas9 knock-in experiments.

Julien Pontabry – Programmed the plugin for Fiji that was used to analyze the data shown in Figure 5.

Mustapha Oulad-Abdelghani – Generated the anti-γH2AX and other antibodies used in the study.

Kishore Babu N. – Performed initial labeling and 3D-SIM experiments which started the project.

Graham D. Wright – Performed initial labeling and 3D-SIM experiments which started the project.

Nacho Molina – Analyzed and quantified the data shown in Figure 5. Wrote the manuscript.

Etienne Weiss – Conceived and designed the study. Wrote the manuscript.

László Tora – Conceived and designed the study. Wrote the manuscript.

JCB

Tools

Imaging of native transcription factors and histone phosphorylation at high resolution in live cells

Sascha Conic,[1,2,3,4] Dominique Desplancq,[5] Alexia Ferrand,[6] Veronique Fischer,[1,2,3,4] Vincent Heyer,[1,2,3,4] Bernardo Reina San Martin,[1,2,3,4] Julien Pontabry,[1,2,3,4,8] Mustapha Oulad-Abdelghani,[1,2,3,4] Kishore Babu N.,[9] Graham D. Wright,[7] Nacho Molina,[1,2,3,4] Etienne Weiss,[5] and László Tora[1,2,3,4,9]

[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire, [2]Centre National de la Recherche Scientifique, UMR7104, [3]Institut National de la Santé et de la Recherche Médicale, U964, [4]Université de Strasbourg, and [5]Institut de Recherche de l'ESBS, UMR 7242, Illkirch, France
[6]Imaging Core Facility, Biozentrum, University of Basel, Basel, Switzerland
[7]Institute of Medical Biology, A*STAR, Singapore, Singapore
[8]Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH), Institute of Epigenetics and Stem Cells, München, Germany
[9]School of Biological Sciences, Nanyang Technological University, Singapore

Fluorescent labeling of endogenous proteins for live-cell imaging without exogenous expression of tagged proteins or genetic manipulations has not been routinely possible. We describe a simple versatile antibody-based imaging approach (VANIMA) for the precise localization and tracking of endogenous nuclear factors. Our protocol can be implemented in every laboratory allowing the efficient and nonharmful delivery of organic dye-conjugated antibodies, or antibody fragments, into different metazoan cell types. Live-cell imaging permits following the labeled probes bound to their endogenous targets. By using conventional and super-resolution imaging we show dynamic changes in the distribution of several nuclear transcription factors (i.e., RNA polymerase II or TAF10), and specific phosphorylated histones (γH2AX), upon distinct biological stimuli at the nanometer scale. Hence, considering the large panel of available antibodies and the simplicity of their implementation, VANIMA can be used to uncover novel biological information based on the dynamic behavior of transcription factors or posttranslational modifications in the nucleus of single live cells.

Introduction

Although transgenic or overexpression-based approaches are well-established to follow the spatiotemporal localization (and in rare cases the activity) of different intracellular factors in real time, the detection of endogenous cellular factors in live cells is not yet routinely possible. Visualization of cellular structures and processes is typically performed by using immunofluorescence (IF) labeling of fixed cells or exogenous overexpression of fluorescently tagged proteins (FTPs) in live cells. In IF, specific labeling of proteins is typically achieved by incubating chemically fixed and permeabilized cells with primary antibodies followed by specific secondary antibodies conjugated to fluorophores. Despite many variables (e.g., permeabilization efficiency, protein denaturation, access to epitopes, and antibody quality), IF is routinely used for visualizing targeted, but immobile, proteins in fixed cells and tissues (Schnell et al., 2012; Teves et al., 2016). On the other hand, imaging of nuclear proteins in living cells is often achieved through exogenous expression of the protein of interest fused to a fluorescent protein tag (FP; Ellenberg et al., 1999; Betzig et al., 2006; Schneider and Hackenberger, 2017) or knock-in of an FP tag coding cDNA at the endogenous loci by the CRI SPR/Cas9 technology to create an endogenous FTP (Ratz et al., 2015). Although FTPs have proven to be very powerful, the continually developing FPs are suboptimal, when compared with dyes, because of the relatively limited quantum yield and low photostability. In addition, FTPs do not always behave as their endogenous counterparts (because of the FP tag) and/or their elevated levels when exogenously overexpressed (Burgess et al., 2012).

It has been well established that the function of transcription factors and coactivator complexes involved in chromatin-dependent processes are tightly linked to their mobility and interactions with diverse posttranslational modifications (PTMs) in the nuclear environment (Snapp et al., 2003; Kimura, 2005; Hager et al., 2009; Cisse et al., 2013; Vosnakis et al., 2017). Our current understanding of transcription regulation dynamics is often based on approaches, called fluorescence recovery after photobleaching and florescence loss in photobleaching, in which fluorescently tagged factors in the nucleus, or a whole cellular compartment, are bleached and the fluorescence redistribution is followed over time in live cells (Kimura et al., 1999, 2002; Dundr et al., 2002; Kimura, 2005;

Correspondence to Etienne Weiss: etienne.weiss@unistra.fr; László Tora laszlo@igbmc.fr

Supplemental material can be found at:
http://doi.org/10.1083/jcb.201709153

The Rockefeller University Press
J. Cell Biol.
https://doi.org/10.1083/jcb.201709153

Gorski et al., 2008; van Royen et al., 2011). Fluorescence correlation spectroscopy, is a microscopy technique where less than 200 molecules are measured, but also based on the detection and quantification of fluorescently tagged factors diffusing through a subfemtoliter observation volume (Macháň and Wohland, 2014). Moreover, single-particle tracking approaches combined with super resolution microscopy often rely also on protein tagging with FPs or photoactivable FPs (Beghin et al., 2017). Consequently, at present there is no simple approach to track accurately nontagged, native transcription factors or to detect the appearance and/or the disappearance of PTMs in the nuclear environment of living cells at high resolution. Thus, there is a demand for novel, powerful tools to gain insight in the dynamic behavior of endogenously expressed proteins in single live cells.

Fluorescently labeled antibodies poorly penetrate through the intact membranes of living cells, making it challenging to image intracellular endogenous proteins (Marschall et al., 2011). Methods have been described that attempted to overcome this through microinjection, osmotic lysis of pinocytic vesicles, loading with glass beads, or protein transfection by using various cationic lipids or polymers (Manders et al., 1999; Courtête et al., 2007; Röder et al., 2017). Recently, fluorescent labeling of proteins inside live mammalian cells has been achieved by using streptolysin O, a bacterial toxin, which creates pores in the cell membrane and allows the delivery of fluorescent probes (Teng et al., 2016). However, this method required additional steps to reseal the membrane pores. Many of these techniques require very specialized know-how and/or equipment, suffer from low efficiency, and/or are harmful for the cells. Significant effort has also been put into antibody engineering of single-chain variable (scFv) fragment antibodies, which can be expressed intracellularly as recombinant scFvs (intrabodies), but unfortunately many of these intrabodies have proven to be insoluble and aggregate in the reducing environment of the cytosol (Renaud et al., 2017). The delivery of nonlabeled mouse mAbs in human cells using electroporation and their subsequent detection in fixed cells has been described (Berglund and Starkey, 1989; Chakrabarti et al., 1989; Lukas et al., 1994; Freund et al., 2013; Marschall et al., 2014; Desplancq et al., 2016).

Because antibodies can be efficiently labeled with fluorophores by using conventional methods and reliably delivered into the cytoplasm by electroporation, we tested whether such probes, which do not need exogenous protein expression or genetic manipulations, can be used for the specific detection and tracking of endogenous nuclear factors in live cells. Here we describe a versatile antibody-based imaging approach (VANIMA) for conventional and super-resolution imaging and tracking of endogenous nuclear factors in live cells by means of fluorescently labeled antibodies or antibody fragments. Their intracytoplasmic delivery into cultured cells was achieved through a simple nontoxic and highly efficient electroporation step. By following the fate of these conventional and noninterfering probes in live cells, it was possible to uncover novel cell biological insights by tracking at nanometer scale native transcription factors (i.e., RNA polymerase II [Pol II], TATA binding protein [TBP], and TBP-associated factor 10 [TAF10]) and image the dynamics of phosphorylated histone H2AX.

Results

Proof of principle of VANIMA: Targeting RNA Pol II in single living cells

To visualize an endogenous nuclear target protein, we selected an mAb that was raised against the heptapeptide repeats present in the nonphosphorylated C-terminal domain (CTD) of the largest subunit (RPB1) of RNA Pol II, hereafter called anti-RPB1 mAb, which performed well in IF assays (Lebedeva et al., 2005). This mAb was first purified and randomly labeled with Alexa Fluor 488 fluorescent dye. Labeling efficiency calculations indicated that the anti-RPB1 mAb contained five to seven covalently linked dye molecules per mAb. To transduce the antibodies into cells, the cell membrane was shortly permeabilized by a brief electric shock with the use of a commercially available apparatus (see Materials and methods), enabling the antibodies to enter the cytoplasm. Once inside the living cells, the antibodies can be imaged by using various microscopy techniques. The labeled anti-RPB1 mAb was electroporated into a large variety of different mammalian or *Drosophila melanogaster* cell types with a delivery efficiency of ∼94–99% and a viability efficiency of 56–99% (Table S1). Approximately 6 h after electroporation, during which the cells attach to the culture dish, the labeled anti-RPB1 mAb was detected in the cytoplasm of human U2OS cells (Fig. 1 A and Video 1). Full-length mAbs are unable to enter the nucleus because of their large size (150 kD; Hayashi-Takanaka et al., 2011; Desplancq et al., 2016; Teng et al., 2016). However, interestingly, after ∼24 h the anti-RPB1 mAb-bound Alexa Fluor 488 signal was almost completely nuclear, indicating that the labeled anti-RPB1 mAb bound to newly synthesized target protein, RPB1, in the cytoplasm and was piggybacked into the nucleus (Fig. 1 A and Video 1). When we transduced 0.5, 1, 2, and 4 µg labeled anti-RPB1 mAb (corresponding to about between 5×10^4 and 4×10^5 antibody molecules per cell; Freund et al., 2013), we observed that with 4 µg electroporated anti-RPB1–Alexa Fluor 488 mAb the nuclear signal became saturated because at this concentration of mAb a cytoplasmic signal persisted 24 h after the transduction (Fig. 1 B). This indicated that with $\sim 4 \times 10^5$ molecules of antibodies per cell we have saturated all the available binding sites on the CTDs of RPB1 and that with between 2×10^5 and 4×10^5 molecules of antibodies per cell most of the endogenous Pol II molecules were labeled (Fig. 1 B). The fact that U2OS cells contain $\sim 9 \times 10^4$ molecules of Pol II (Zhao et al., 2014) further suggests that each RPB1 CTD may be bound by ∼2–4 molecules of anti-RPB1 mAb. Moreover, as each mAb is labeled with ∼5–7 molecules of dye, it means that each Pol II molecule can be visualized by 10–28 molecules of dye.

To test whether the electroporated anti-RPB1 mAb that was piggybacked to the nucleus by RPB1 (Fig. 1 A and Video 1) would stay bound to its target, we transduced U2OS cells with 0.5, 2, and 4 µg anti-RPB1 mAb. 24 h after transduction we lysed the cells, prepared whole-cell extracts, mixed the antibody-containing cell extracts with protein G Dynabeads, and tested whether the extracted anti-RPB1 mAb would still be bound to RPB1 (Fig. 1 C). Our experiment shows that the electroporated labeled anti-RPB1 mAb remains bound under these conditions and that all the cellular Pol II can be bound by the transduced labeled antibody.

As a large portion of Pol II is bound to the chromatin during transcription in the cells (Kimura et al., 1999), we tested whether the Alexa Fluor 488–labeled anti-RPB1-mAb would

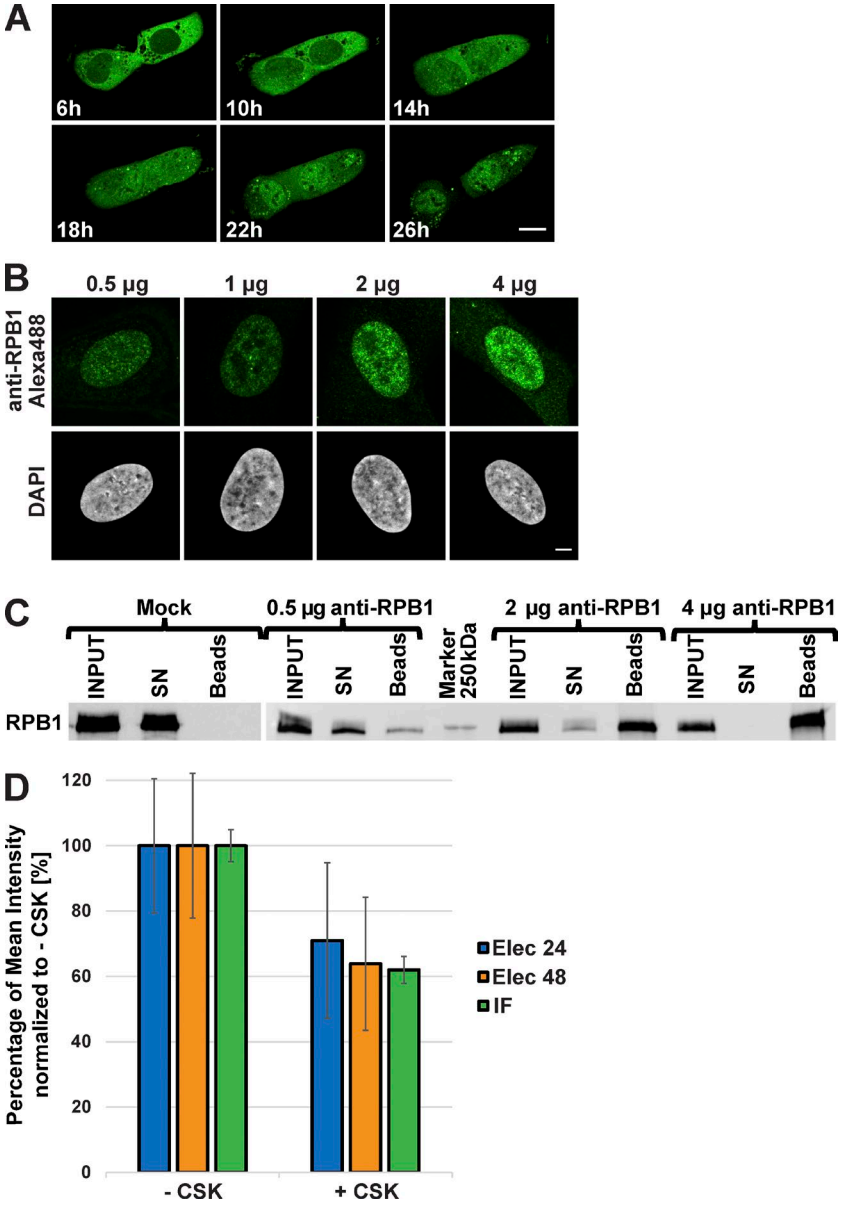


Figure 1. **Behavior of the anti-RPB1 mAb in U2OS cells. (A)** After transduction with Alexa Fluor 488–labeled anti-RPB1 antibodies, cells were imaged after 6 h of incubation and then every hour over a period of 20 h (see Video 1 for all time points). Bar, 15 µm. **(B)** Increasing amounts of Alexa Fluor 488–labeled anti-RPB1 mAb were transduced in U2OS cells and fixed 24 h after electroporation. A typical nucleus recorded in each case after counterstaining with DAPI is shown. Bar, 5 µm. **(C)** Binding capacity of anti-RPB1 mAb in U2OS cells. Cells were electroporated with 0 (mock), 0.5, 2, and 4 µg anti-RPB1 mAb and whole-cell extracts prepared 24 h after transduction (INP UT) were mixed with protein G beads. Bound and unbound material was analyzed by Western blotting. The blot shows the fraction of antibody-bound Pol II molecules adsorbed on the beads (beads) or left in the supernatant (SN), and detected with a secondary antibody. **(D)** After transduction with Alexa Fluor 488–labeled anti-RPB1 mAb (2 µg), cells were treated with or without CSK buffer. The histogram shows the mean fluorescence intensity of the nucleus of nontreated (–CSK) and CSK-treated (+CSK) cells 24 h (Elec 24h) or 48 h (Elec 48h) after electroporation. A classical anti-RPB1 mAb IF experiment was performed as additional control (IF). The +CSK signal is represented as the percentage of the mean intensity of the –CSK signal. Error bars represent the SD obtained with 10 recorded cells for each condition. All images were acquired by confocal microscopy on one single z plane.

also stay bound to the chromatin associated Pol II. To this end, 24 or 48 h after transduction anti-RPB1–Alexa Fluor 488 mAb–transduced cells were treated, or not, with a mixture of detergent and sucrose known as cytoskeleton (CSK) buffer, which is widely used to release soluble proteins from cells, including the nucleus (Cramer and Mitchison, 1995). Cells were then fixed, and the Alexa Fluor 488 signal was quantified from nontreated and CSK-treated cells 24 and 48 h after transduction. As a control, a classical anti-RPB1 mAb IF staining was performed. The quantification of IF detection of Pol II shows that in CSK-treated samples ∼60–70% of the total Pol II signal is bound to the chromatin. In agreement, the quantification of the electroporated anti-RPB1–Alexa Fluor 488 mAb signal indicated the presence of similar fraction of chromatin-bound endogenous Pol II (Fig. 1 D). These results further indicate that the transduced labeled anti-RPB1 mAb can bind to transcribing Pol II on the chromatin and that the electroporated mAb stays bound to its target during 48 h. These specific mAb-binding characteristics in cells suggest that VANIMA can be used for live-cell imaging experiments to characterize the behavior of transcription factors.

Imaging of several endogenous nuclear antigens with VANIMA

To further evidence the usefulness of the approach for imaging a range of nuclear factors, we have compared different transduced labeled mAbs (150 kD) with their corresponding Fab fragments (50 kD), because Fabs can freely enter the nuclei of cells (Hayashi-Takanaka et al., 2011). In these comparisons, different mAbs or Fabs were used, which were raised against different transcription factors (such as RBP1/Pol II, TBP, and TAF10). Our comparisons show that the labeled mAbs or their corresponding labeled Fab fragments perform similarly to label the endogenous transcription factors (Fig. 2, A and B; and Fig. S1 A). Importantly, labeled Fab fragments raised against nuclear proteins are reaching the nucleus 6 h after electroporation (Fig. S1 B), in contrast to mAbs that need ∼24–48 h to reach the nucleus by the piggybacking mechanism (Fig. 1 A).

Next, we verified whether the electroporated Alexa Fluor 488–labeled anti-TAF10 or anti-TBP mAbs would stay bound to their respective targets after electroporation and piggybacking in the nucleus. To this end cells were electroporated with in-

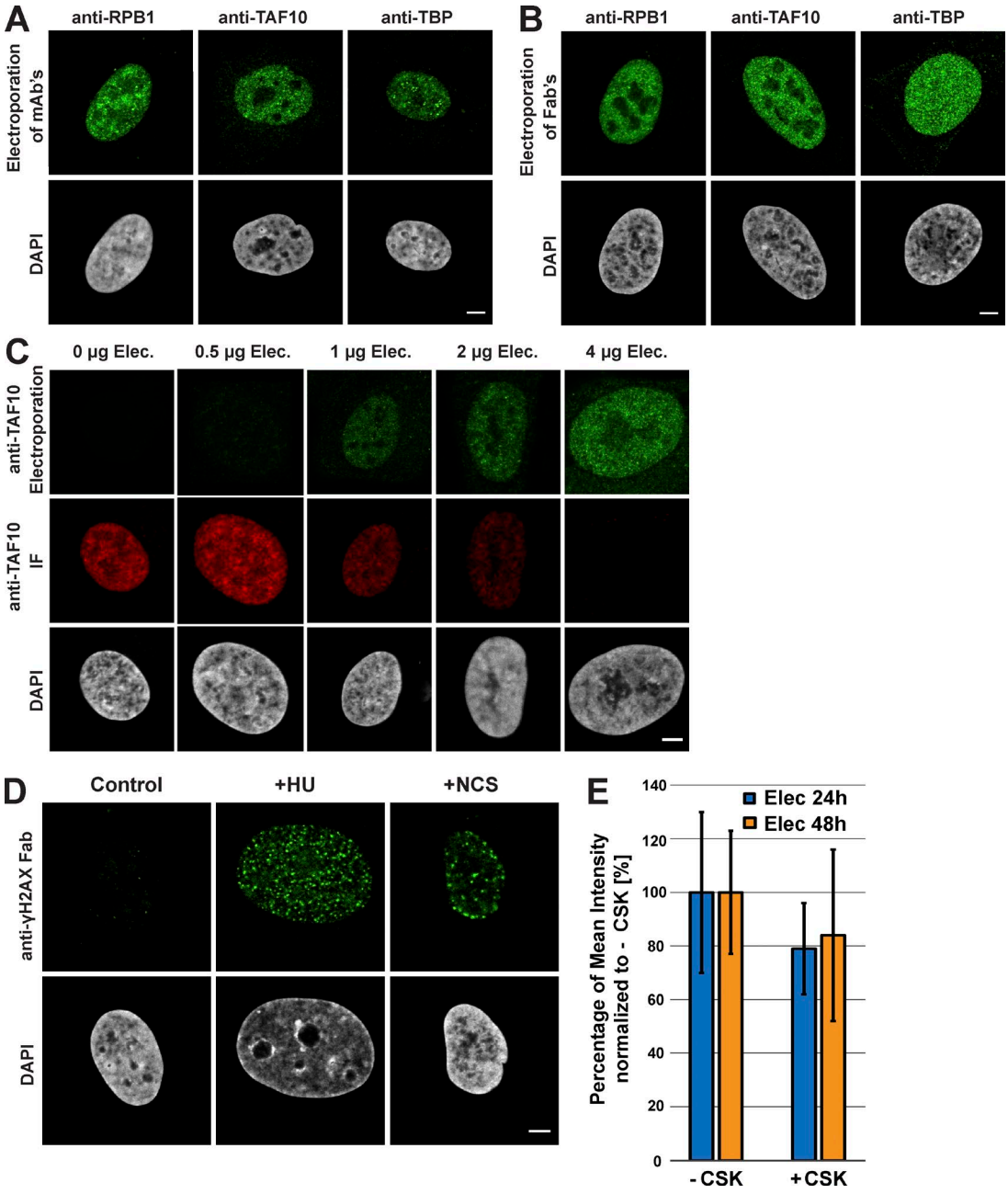


Figure 2. **Visualization of endogenous transcription factors and phosphorylated H2AX with VANIMA. (A)** The labeled mAbs binding specifically to the transcription factors RPB1, TAF10, and TBP were transduced in U2OS cells, and their localization in the cells was monitored by confocal microscopy 24 h after treatment. A single z plane is shown for each condition. The pictures represent a typical nucleus recorded in each case after fixation of the cells and subsequent counterstaining with DAPI. **(B)** Same as in A, except that the experiments were performed with the corresponding labeled Fab fragments. **(C)** Increasing amounts of Alexa Fluor 488–labeled anti-TAF10 mAb (green) were transduced in U2OS cells and fixed 24 h after electroporation (anti-TAF10 Electroporation). To verify binding of the antibody to TAF10, a competition assay was performed afterward by adding a constant amount (2 µg) of the same antibody but Alexa Fluor 568–labeled as IF antibody (red, anti-TAF10 IF; see also Fig. S1 C for quantification). DAPI staining is shown in gray. **(D)** The labeled Fab raised against γH2AX was transduced as in B, and its localization was recorded after treatment of the electroporated cells with either NCS (for 15 min) or HU (for 48 h). Control, nontreated cells. A typical nucleus is represented in each case. **(E)** After transduction with Alexa Fluor 488–labeled anti-γH2AX Fab (5 µg) and treatment with HU, cells were treated with or without CSK buffer before fixation. The histogram shows the mean fluorescence intensity of the nucleus of nontreated (–CSK) and CSK-treated (+CSK) cells 24 h (Elec 24h) or 48 h (Elec 48h) after electroporation. The +CSK signal is represented as the percentage of the mean intensity of the –CSK signal. Error bars represent the SD obtained with 10 recorded cells for each condition. Bars, 5 µm.

creasing amounts of Alexa Fluor 488–labeled antibodies, fixed 24 h after electroporation and subjected to IF with the same antibody but labeled with Alexa Fluor 568 dyes. These competition experiments and their quantifications show that when 10^6 cells were transduced with 4 µg antibodies, 24 h after electroporation the intracellular antibodies were still binding to all their target epitopes, as in these cells no significant IF signal could be detected (Fig. 2 C and Fig. S1 C).

It is noteworthy that electroporated mAbs raised against either a prokaryotic protein, and thus having no epitopes in the

human cell (such as the mAb against the maltose-binding protein [MBP]), or against a cytoplasmic target (such as the mAb against α-tubulin) do not enter in the nucleus (Fig. S1 D). All these results together suggest that both labeled mAbs and Fabs can be used for imaging nuclear antigens depending on the scientific question asked.

We also tested whether the transduced labeled antibodies would recognize chromatin-associated PTMs. To this end we used a Fab developed against γH2AX that is often considered a marker of DNA double-strand breaks (Siddiqui et al., 2015; Fig. S1 E). The histone variant H2AX, which can replace conventional histone H2A in nucleosomes, becomes phosphorylated on serine 139 (called γH2AX) upon DNA double-strand breaks. Note that when an epitope is generated only in the nucleus, such as histone PTMs, only labeled Fabs are adequate to detect these targets. Anti-γH2AX mAb was generated, and the corresponding labeled Fabs were transduced in control cells and in cells in which DNA damage was induced by hydroxyurea (HU) or neocarzinostatin (NCS) treatments (Fig. 2 D). As expected, Alexa Fluor 488–labeled Fab fragments could enter the nuclei of the cells and bind the serine 139 phosphorylated H2AX foci in the HU- or NCS-treated cell nuclei (Fig. 2 D), demonstrating that the transduced Fabs can bind to PTMs in the chromatin of live-cell nuclei. Next, we verified whether the electroporated Alexa Fluor 488–labeled anti-γH2AX Fab would stay bound to chromatin after electroporation and diffusion to the nucleus. To this end, cells were electroporated with anti-γH2AX Fab and treated with HU 6 h later, and soluble proteins were extracted with the CSK buffer 24 or 48 h after treatment. Cells were then fixed and the Alexa Fluor 488 signal (Fig. 2 E). These experiments further indicate that almost all the labeled anti-γH2AX Fab stays bound to chromatin and that at the indicated time points almost no unbound Fab could be detected.

To ascertain that our endogenous nuclear protein labeling approach with the use of the described antibodies would not interfere at a detectable level with the function of the target or cellular functions, we performed a series of tests 24 and 48 h after mAb electroporation. To verify whether the anti-RPB1, -TBP, or -TAF10 would inhibit transcription, RNA was isolated from electroporated cells and subjected to RT-qPCR analyses by using primers to amplify unspliced, and therefore newly synthesized, premRNA from Pol II target genes. The primers were designed to amplify sequences from introns to exons for several Pol II–transcribed genes (Table S2). As controls, cells were either transduced with an antibody targeting the bacterial MBP, which has no expected target in the human cells and therefore should not inhibit transcription. Cells were also treated with α-amanitin at a concentration that would inhibit Pol II transcription but not that of Pol I and Pol III. Our results show that the anti-MBP antibody and the other three mAbs tested did not significantly inhibit premRNA transcription of the tested Pol II genes, although α-amanitin almost completely abolished the transcription of the Pol II genes (Fig. 3, A and B). Next, we measured the cell cycle progression and the cell proliferation/replication capabilities of the antibody electroporated cells (Fig. 3, C and D). Both quantifications show that cell cycle progression and cell proliferation were not inhibited by the electroporation of the anti-RPB1, -TBP, -TAF10, or -MBP antibodies. Furthermore, apoptosis tests indicated that transduced antibodies did not induce significant cell death 24 h after their electroporation (Fig. 3 E). In conclusion, a noninterfering mAb recognizing a nuclear transcription factor should be suitable for VANIMA if after transduction it is piggybacked in the nucleus. Fabs can freely diffuse in the cell and only accumulate in the nucleus after transduction if bound to the nuclear target. In addition, both mAbs and Fabs should not inhibit significantly premRNA transcription, cell cycle progression, cell proliferation, or induce apoptosis.

Comparison to existing labeling techniques

We have also compared VANIMA to existing labeling techniques, such as IF, ectopic expression of GFP-fused transcription factors, or CRISPR/Cas9 knock-in technology. When using VANIMA and IF (Fig. S2 A) in parallel experiments, we obtained identical results on fixed cells, except that our approach does not necessitate a fixation step for the accurate detection of the targets (compare Fig. 2 A and Fig. S2 A). When comparing the labeling with transduced antibodies to the ectopic expression (overexpression) using GFP fusions of transcription factors, we observed as previously published that exogenously expressed GFP-RPB1 or CFP-TAF10 does not efficiently reach the nucleus or is excluded from the nucleus, respectively, in contrast to the endogenous counterparts (Soutoglou et al., 2005; Boulon et al., 2010; Wild and Cramer, 2012; Fig. S2, B and C). Moreover, ectopically expressed GFP-TBP was nuclear but excluded from the nucleoli of the cells (Fig. S2, B and C), suggesting that GFP-TBP does not enter the nucleoli despite TBP involvement in Pol I transcription (Hernandez, 1993). In contrast, the antibody-labeling method revealed the expected behavior of the endogenous nuclear transcription factors (compare Fig. 2 and Fig. S2, B and C). To be able to compare VANIMA to cells where a fluorescent tag has been expressed from the endogenous locus in fusion with a transcription factor, we knocked-in a Venus tag in frame at the 5′ end of the *TAF10* locus in U2OS cells using the CRISPR/Cas9 methodology. Stable Venus-TAF10 expressing heterozygous U2OS clones were generated, and the fluorescence obtained from these cells was compared with U2OS cells that were simply transduced for 24 h with an anti-TAF10 mAb labeled with Alexa Fluor 488. The comparison shows that electroporated cells give a signal largely overlapping with that obtained in Venus-TAF10 expressing cells but that the labeled mAb-bound TAF10 signal is brighter than Venus-TAF10 signal when using a confocal microscope (Fig. S2, D and E).

Analysis of Pol II, TAF10, and γH2AX distribution in subnuclear structures by super-resolution microscopy

To obtain high-resolution images of endogenous proteins and PTMs, we used super-resolution microscopy (Betzig et al., 2006). To be able to carry out multichannel detection and live-cell imaging the target-bound labeled mAbs and Fabs were visualized by 3D structural illumination (3D-SIM) super-resolution microscopy at ∼110 nm xy and ∼300 nm z resolution first in fixed cells (Schermelleh et al., 2008). By using 3D-SIM, the labeled mAbs and Fabs allowed the detection of well-defined individual spots of different sizes in the nuclei of U2OS cells (Fig. 4, A and B; and Videos 2–4). In agreement with previous studies (Markaki et al., 2010), the detection of Pol II, TAF10, and TBP by 3D-SIM seemed to be excluded from DAPI dense regions (Fig. 4, A and B).

We measured the nuclear distribution of Pol II and TAF10 molecules labeled with anti-RPB1 mAb-Alexa Fluor 488 and anti-TAF10 mAb-Alexa Fluor 488, respectively, using 3D-SIM. To quantify the number and sizes of the observed foci, we pro-

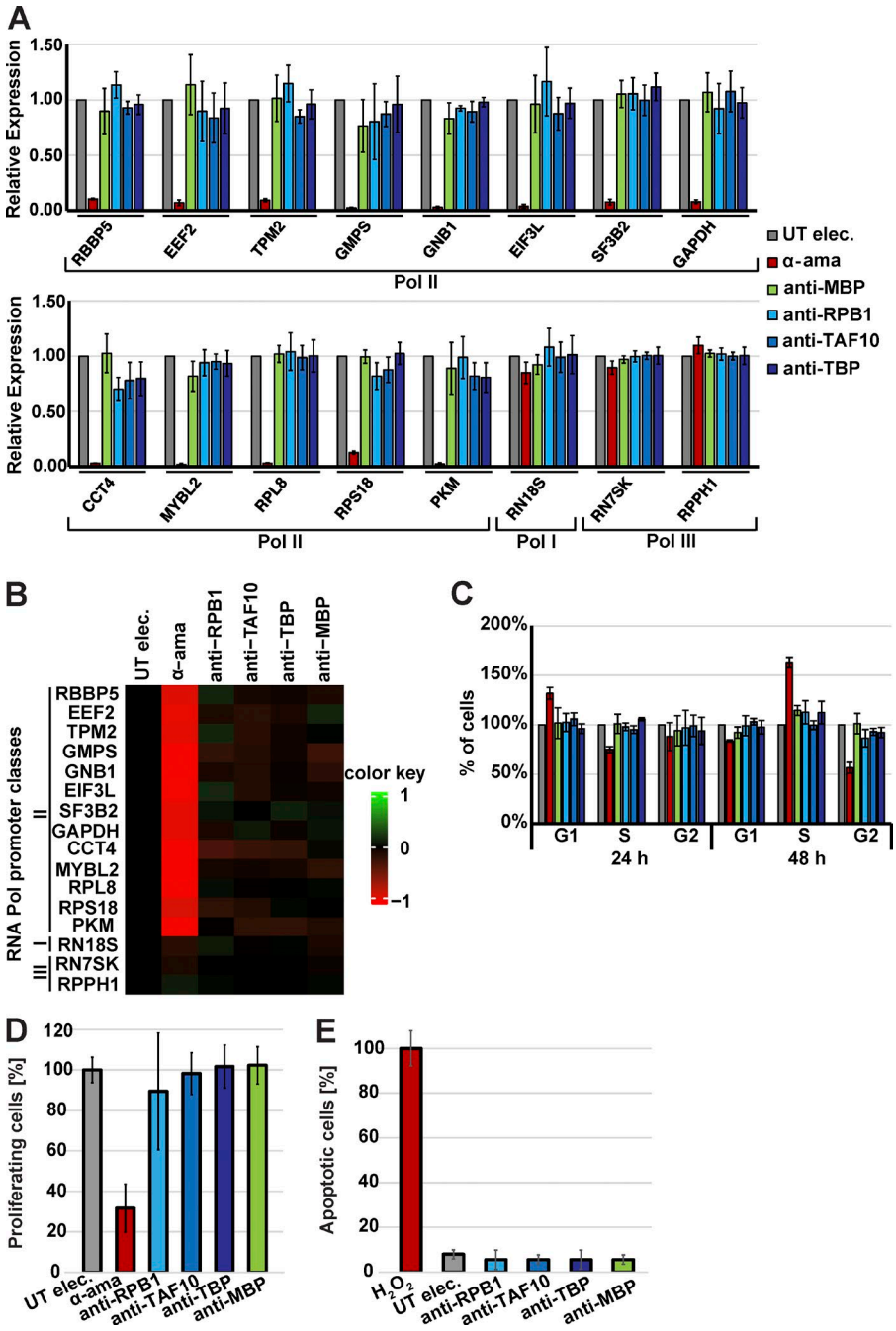


Figure 3. **The mAbs do not inhibit premRNA transcription, cell cycle progression, cell proliferation and do not induce apoptosis. (A)** U2OS cells electroporated but without antibodies (UT elec), electroporated and treated with α-amanitin (α-ama), electroporated with a control antibody binding to bacterial MBP (anti-MBP), or electroporated with the mAbs recognizing specifically RPB1, TAF10, or TBP (anti-RPB1, anti-TAF10, or anti-TBP). 24 h after electroporation, total RNA was isolated, and the expression of Pol I, Pol II, and Pol III genes was analyzed by RT-qPCR. Pol III transcripts were used for normalization. Newly synthesized RNA of the indicated genes was quantified with validated primer pairs (Table S2). The histograms correspond to the mean values obtained with three independent experiments. **(B)** The mean values of the three independent experiments shown in A are represented as a heatmap reflecting unchanged relative expression in black, up-regulation in green, and down-regulation in red. **(C)** U2OS cells were electroporated as in A, and cell cycle progression was monitored by propidium iodide staining and FACS analysis 24 or 48 h after electroporation. The cell cycle phases were normalized to cells electroporated without antibody. **(D)** U2OS cells were electroporated as in A, and their capacity of proliferation was monitored 24 h after transduction by EdU incorporation and FACS. The electroporated cells without the addition of antibody were used as control. The color code is as in A. **(E)** The cells were treated as in A, except an apoptosis test was performed 24 h after electroporation. Apoptosis induced by the addition of 10 µM H_2O_2 was taken as reference (100%). In each panel, the error bars represent the biological SD obtained from three independent replicates. UT, untreated cells.

cessed the images with Fiji/ImageJ and Matlab (see Materials and methods; Fig. 5, A–F). Our quantifications show that the size distribution of Pol II foci ranges from 10^{-3} µm³ to ∼1.6 × 10^{-2} µm³, with nearly 34% of the foci having the smallest volume (Fig. 5 A). TAF10 foci are in general smaller than those of Pol II, with 55% of the spots showing the smallest volume (Fig. 5 B). Interestingly, ∼3% of the Pol II foci are larger than 10^{-2} µm³, whereas only 0.4% of the TAF10 foci fall in this category (Fig. 5 C). To investigate the biological significance of the observed spot sizes, we have inhibited transcription with 2 µM flavopiridol (Flavo), a known inhibitor of Pol II transcription elongation (Chao et al., 2000). 1-h Flavo treatment significantly reduced the RPB1 CTD phosphorylation by pTEFb (Vosnakis et al., 2017). Interestingly, the Flavo treatment reduced the volume of bigger Pol II foci and consequently increased about twofold the percentage of smaller Pol II spots between 10^{-3} and 4 × 10^{-3} µm³ (Fig. 5 A). In addition, when the size distribution changes of the larger Pol II foci were considered (spots > 10^{-2} µm³) after Flavo treatment, the percentage of larger Pol II foci was decreased by a factor of 4 (Fig. 5 C). In contrast, the size distribution of the TAF10 foci was not affected by Flavo treatment (Fig. 5, B and C). Interestingly, the total number of Pol II foci increased after Flavo treatment and was followed by a parallel decrease in the mean cluster size of Pol II foci. In agreement with a scenario in which the large Pol II foci would dissociate in several smaller spots, the total volume of labeled spots did not change (Fig. 5, D–F). In contrast, transcription elongation inhibition did not influence the total number, mean cluster size, or total volume of TAF10 foci (Fig. 5, D–F), indicating that the observed Pol II cluster size shift reflected in vivo

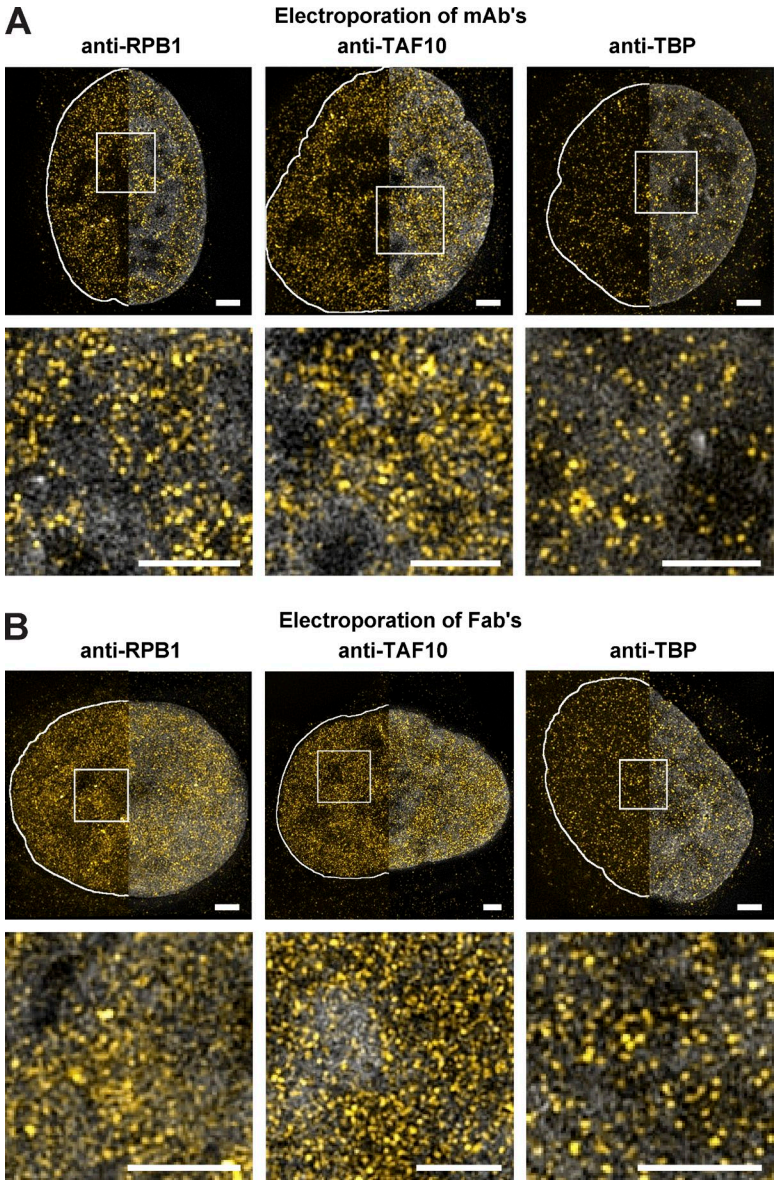


Figure 4. **Visualization of transcription factors with VANIMA by super-resolution microscopy. (A)** The labeled mAbs binding to the transcription factors RPB1, TAF10, and TBP (yellow) were transduced in U2OS cells, and their localization in the cells was monitored 24 h after transduction by 3D-SIM. The pictures show a typical nucleus recorded in each case after fixation and DAPI (gray) treatment (Videos 2–4). The Z maximum intensity projections of five slices show the labeled mAbs with (right half) or without (left half) DAPI counterstaining (gray). The solid white lines depict the nuclear contour. Bottom: Magnification of the white regions of interest, under the corresponding image. **(B)** The nuclei shown correspond to transduced U2OS cells as in A, except that transductions were performed with the corresponding labeled Fab fragments. Bars, 2 µm.

Pol II behavior changes after transcription inhibition. Using photobleaching techniques, it has been shown that, when transcription elongation is inhibited, total bound Pol II is released from the chromatin in general and becomes mobile (Kimura et al., 2002; Hieda et al., 2005; Vosnakis et al., 2017). Thus, our results show that when transcription elongation is inhibited by Flavo the larger Pol II foci dissociate, because Pol II molecules are released from these sites and become mobile.

To confirm the usefulness of delivered labeled antibodies in monitoring discrete nuclear structures labeled by various PTMs, we visualized and quantified the number of γH2AX-Fab–labeled foci before and after HU treatment using 3D-SIM (Fig. 6, A and B; and Videos 5 and 6). Our quantifications show that HU-induced DNA damage increased the number of γH2AX foci by ∼80-fold in treated cells (Fig. 6 B), suggesting that labeling with transduced Fab fragments allows precise analysis of chromatin modifications upon replication stress. The 3D-SIM experiments demonstrate that changes of individual nuclear structures, where transcription factors or specific PTMs are present or accumulate, can easily be revealed after different biological stimuli. Our approach can thus be used to uncover novel information concerning essential biological mechanisms.

Uncovering novel dynamic behaviors of transcription factors and PTM events by VANIMA by using high-resolution live-cell imaging

To test the adequacy of conventionally labeled antibodies for high-resolution live-cell imaging, we transduced anti-RPB1–Alexa Fluor 488 mAb into U2OS cells, and 24 h after transduction nuclei were imaged over a period of 2.5 h, taking images every 10 min by time-lapse confocal microscopy. These videos show that the larger Pol II spots/clusters, which can be easily detected at this resolution, are dynamically and constantly moving within the nucleus (Fig. 7 A and Video 7). To better visualize the shape and the movements of these larger Pol II clusters (ranging between 1 and 1.6×10^{-2} µm^3), they were imaged by using 3D-SIM over a short period. These live-cell measurements show that the larger Pol II-labeled foci are dynamic and are constantly associating and dissociating over time (Fig. 7 B and Video 8). In agreement with our nascent transcription ex-

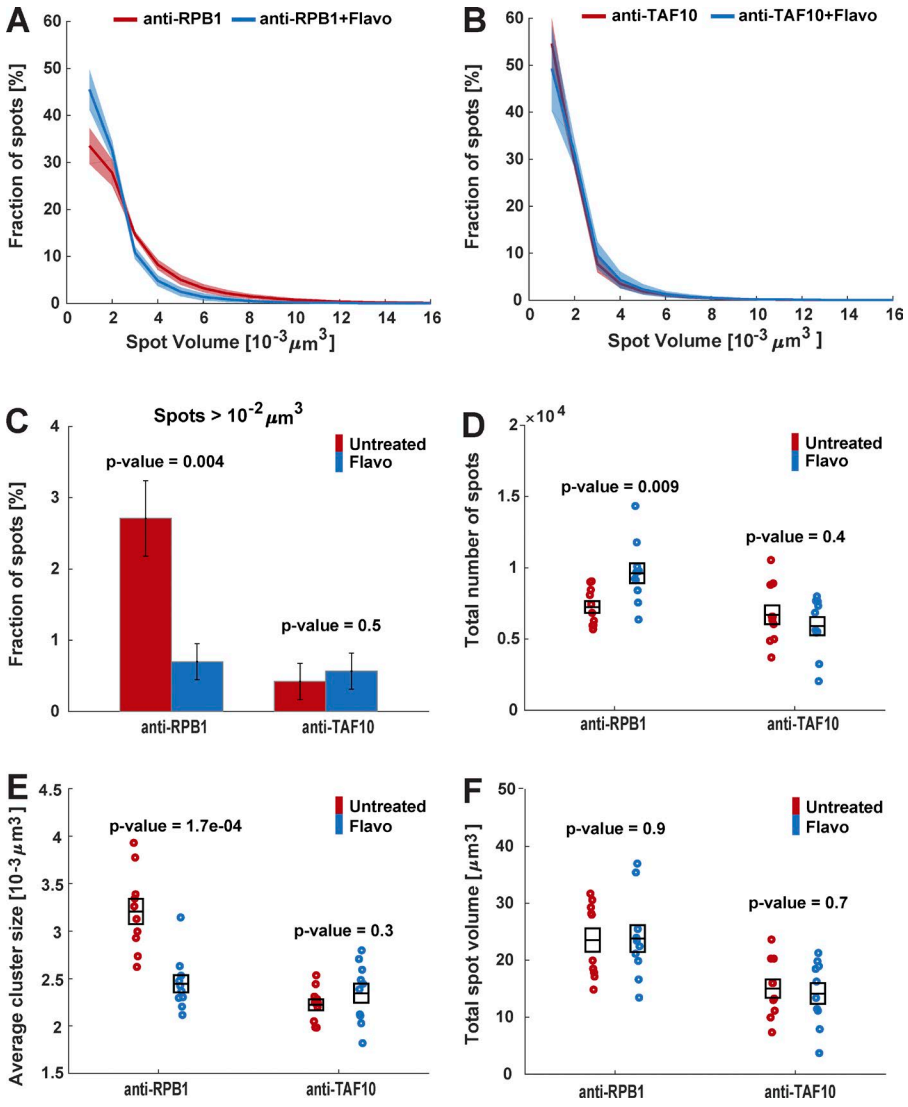


Figure 5. **Quantification of transcription factor distribution in single cells by using VANIMA and super-resolution microscopy. (A)** U2OS cells were transduced with Alexa Fluor 488–labeled anti-RPB1 mAb and then treated with Flavo (2 µM) for 1 h or not (Untreated). 24 h after treatment the cells were fixed and analyzed by 3D-SIM. The number of individual spots and their volume in individual nuclei were quantified by using Fiji/ImageJ and Matlab software. The graph shows the percentage of spots with a given volume in untreated (red) and treated cells with Flavo (blue) acquired from 10 individual cells for each condition. **(B)** Same treatment and analysis as in A, but an Alexa Fluor 488–labeled anti-TAF10 antibody was transduced. **(C)** Spot volumes were extracted from A and B, and the percentage of spots of RPB1 and TAF10 with a volume >10^{-2} µm^3 in the untreated (red) and Flavo (blue) treated cells is shown. The error bars represent the SE from 10 individual cells for each condition. **(D)** Total number of RPB1 and TAF10 spots in 10 individual nuclei for each condition are represented. **(E)** Mean cluster size of the RPB1 or TAF10 spots in 10 individual cells for each condition is shown. **(F)** Total spot volume of RPB1 and TAF10 in 10 individual nuclei for each condition is represented. All black boxes in D–F represent the means and their SEs for each sample. All p-values were calculated by using the two-sample *t* test.

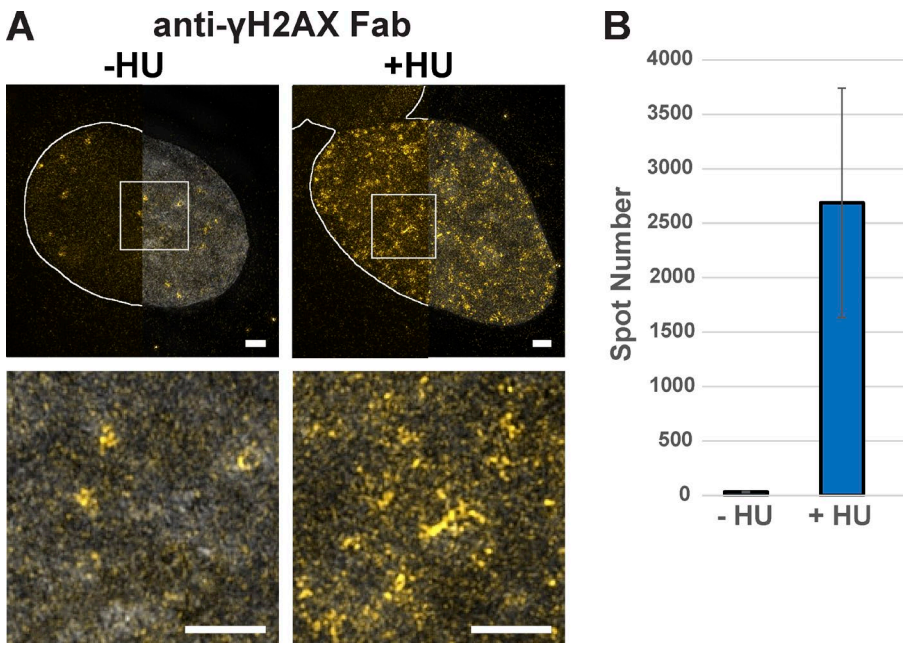


Figure 6. **Imaging of phosphorylated H2AX with VANIMA by super-resolution microscopy. (A)** The labeled anti-γH2AX Fab (yellow) was transduced in U2OS cells, and its localization in the nucleus was recorded by 3D-SIM after treatment with HU for 48 h (+HU) and staining with DAPI (gray). Untreated cells (–HU) were used as the control. The Z maximum intensity projections of 20 slices show the labeled anti-γH2AX Fab with (right half) or without (left half) DAPI counterstaining (gray). The solid white lines depict the nuclear contour. Bottom panels: magnification of the white regions of interest, under the corresponding image (Videos 5 and 6). Bars, 2 µm. **(B)** The number of spots presented in the nuclei as shown in A after quantification with Fiji/ImageJ software. Error bars represent the SD obtained with five recorded cells for each condition.

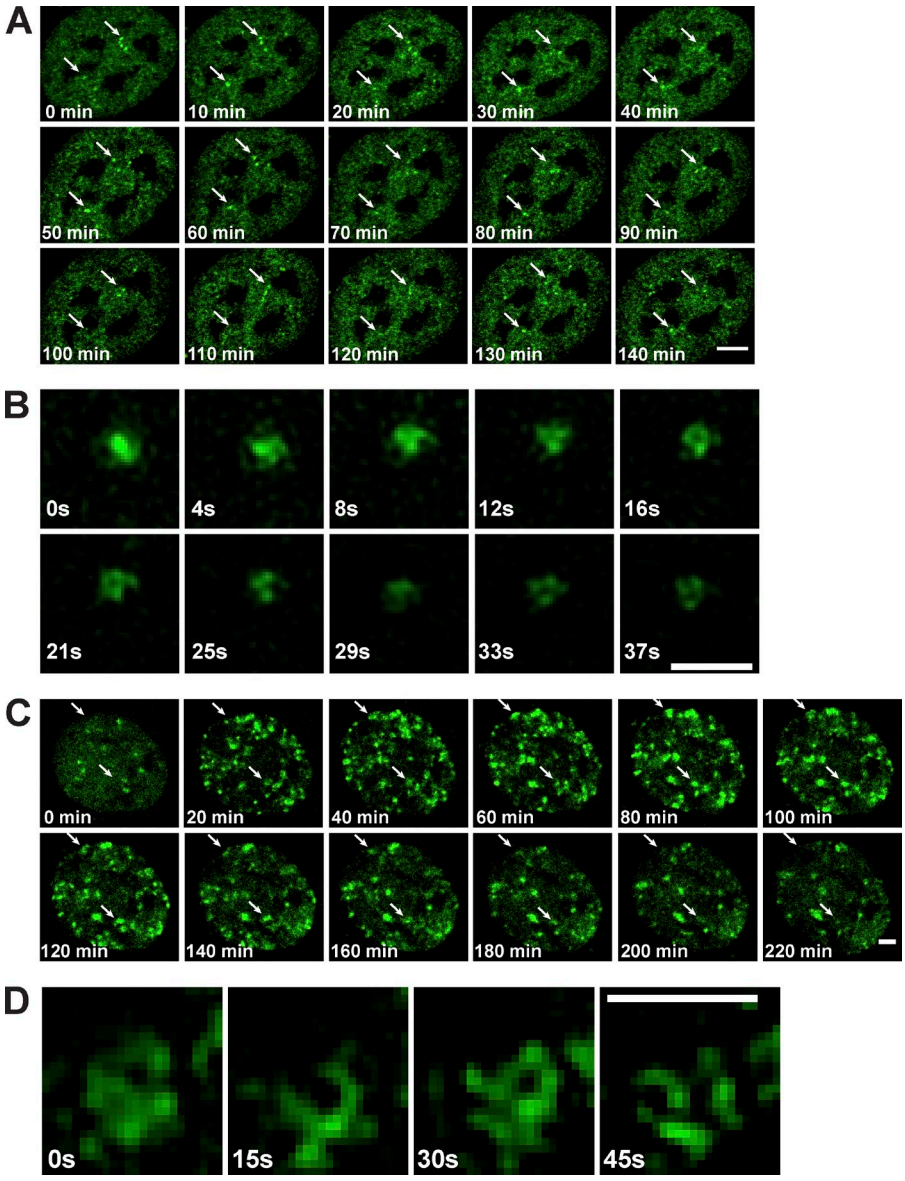


Figure 7. **Live imaging of transcription factors by using VANIMA. (A)** 24 h after electroporation, U2OS cells transduced with Alexa Fluor 488–labeled anti-RPB1 mAb were subjected to live-cell analysis by confocal microscopy focusing on one z section of individual nuclei. They were imaged over a period of 2.5 h and pictures taken every 10 min (Video 7). Arrows point to two larger Pol II cluster examples that move over time. Bar, 5 µm. **(B)** Imaging by 3D-SIM microscopy of an individual Pol II cluster observed in U2OS after transduction as in A. The images were taken over a period of 37 s every 4.1 s and show a maximum intensity projection of the 3D video (Video 8). Bar, 1 µm. **(C)** U2OS cells transduced as in A with the labeled anti-γH2AX Fab were subjected to live-cell analysis by spinning-disk confocal microscopy after the addition of NCS to the culture medium. Pictures were taken every 10 min over a period of 4 h (Video 9) and by focusing on a single z plane. The first time point (0 min) corresponds to the time of the drug addition. Arrows point to γH2AX clusters that appear and disappear over time. Bar, 5 µm. **(D)** Imaging of an individual γH2AX cluster by 3D-SIM microscopy observed in U2OS cells after transduction as in C. Images were recorded over a period of 45 s every 15 s (Video 10). The first time point (0 s) shown was taken 10 min after NCS treatment. Bar, 0.8 µm.

periments (Fig. 3 A), these observations suggest that the labeled mAb does not interfere with the transcription process.

Next, we visualized the induction of γH2AX-Fab labeled foci after NCS treatment by both confocal spinning disc microscopy (Fig. 7 C and Video 9) and 3D-SIM (Fig. 7 D and Video 10). These live-cell experiments demonstrate that the NCS-induced γH2AX foci form large clusters in a kinetic manner and that some of these clusters are stable in time, whereas others are increasing in size, suggesting that the Fab does not hinder the phosphorylation process. Thus, our antibody approach used for live imaging uncovered novel dynamic behaviors of transcription factors and PTM events of H2AX in real time.

Discussion

VANIMA is "right and fair"

Tens of thousands of full-length antibodies that specifically recognize targets with high affinity have been developed over the past decades and are available, mostly commercially. as research tools. Antibodies normally cannot cross intact cellular or subcellular membranes in living cells because of their large size and hydrophilicity (Marschall et al., 2011, 2014). Here we show that electroporation of labeled primary antibodies into live cells allows their efficient delivery into the cytoplasm of cells without significantly reducing their viability. Because full-length mAbs raised against nuclear proteins cannot enter the nucleus, the labeling observed in the nucleus over time can only be explained by the binding of the mAbs to their neosynthesized target and the subsequent import of the labeled mAb-antigen complex to the nucleus. Thus, VANIMA can be used for the characterization of cytoplasmic/nuclear turnover rates of newly synthetized nuclear proteins in live cells when using full-length mAbs. Moreover, the electroporation procedure allows the amount of delivered mAb or Fab to be tightly controlled for the specific and equimolar detection of target proteins (Fig. 1 B; Van Regenmortel, 2014) and hence can also be used for determining the abundance of the accessible antigens in the cell. It is important to note, however, that antibodies have to be charac-

terized for their noninterfering nature before they can be used for tracking native proteins or PTMs. It is likely that VANIMA can also be used with in vitro identified blocking antibodies to disrupt nuclear protein function in living cells.

The use of plasmid cDNA-based transfection assays to exogenously express FTPs is relatively rapid but suffers from the cell-to-cell variability and often protein overexpression (Fig. S2 B). This can be overcome by the generation of stable cell lines, expressing FP-tagged proteins to low levels, which could often take several months. To avoid exogenous protein expression, the genetic knock-in of FP tags into endogenous loci of cells with the use of the CRISPR/Cas9 technology can be used, but the characterization and genotyping of the knock-in could be labor intensive and time consuming because of relatively low efficiency. In addition, in the case of multicolor imaging, changing the colors of the knocked-in tags becomes again very time consuming, when compared with changing the dyes before conjugating them to the purified antibodies. In addition, nanobodies (VHH) derived from camelids, became popular recently for imaging because of their small size (15 kD). However, the generation of these recombinant cDNA expression tools, including their validation for imaging purposes, can be time consuming (Rothbauer et al., 2006; Rinaldi et al., 2013; Krah et al., 2016). Thus, our approach based on already available noninterfering antibodies is much faster and more reliable than any until now described antibody- or antibody fragment-delivery–based visualization method, while giving information on the behavior of endogenous targets.

VANIMA toward uncovering single-cell dynamic behaviors of transcription factors and PTM events in real time

The application of VANIMA to endogenous transcription factors and to a PTM of histone H2AX allowed the precise tracking of these targets in the 3D nucleus and in real-time. Thus, by using VANIMA, dynamic processes of fundamental biological mechanisms, also involving PTMs, can be visualized in nonfixed cells at high resolution. Our results suggest that the detected larger Pol II foci may contain several transcribing Pol II assemblies or Pol II "trains" (Tantale et al., 2016) possibly organized in topological associated domains and/or other control regions (Cisse et al., 2013; Zhao et al., 2014; Cho et al., 2016; Hnisz et al., 2017). The fact that the VANIMA-detected native Pol II foci became smaller when inhibiting transcription with a drug that inhibits transcription elongation is in agreement with previous studies that demonstrated by photobleaching techniques in the whole nuclear compartment that Pol II leaves the chromatin and becomes more mobile (Kimura et al., 2002; Hieda et al., 2005; Vosnakis et al., 2017). It is thus conceivable that the smaller spot size that we observed after Flavo treatment corresponds to "free" Pol II molecules. Note that previous studies visualizing exogenously expressed tagged RPB1 (α-amanitin resistant or not) after shorter Flavo treatment with different super-resolution techniques did not observe significant changes in Pol II spot size (Cisse et al., 2013; Zhao et al., 2014; Cho et al., 2016). Thus, it seems that VANIMA, through detecting endogenous factors, has an improved sensitivity when compared with previously reported RPB1-tagging–based imaging methods. Nevertheless, we also show that large Pol II foci are constantly forming, dynamically associating, and dissociating. By using VANIMA coupled to live 3D-SIM and/or other genome-labeling technologies, it will become possible to investigate, characterize, and dissect the function of the detected endogenous Pol II foci.

In addition, we have been able to monitor with high resolution an essential signal of nuclear DNA damage after insults with genotoxic drugs. In agreement with a recent study, we found that the phospho-H2AX foci correspond to clustered structures (Natale et al., 2017). Moreover, we show here for the first time that these clusters are spatially reorganized with time, likely because of the remodeling of the chromatin, which is necessary for the access of DNA repair proteins. The fact that some clusters come out of focus with time during the analysis is proof of the dynamic aspect of this histone modification. Because analyses with VANIMA are not restricted to endpoint experiments, it might be possible now to further highlight the precise cross talk between transcription and DNA repair. This will likely allow researchers to dissect how an injured cell manages the balance between death and survival.

Moreover, VANIMA coupled with 3D-SIM is suitable for high-resolution colocalization analyses by using up to four different colors. It may allow the in vivo colocalization of several factors within transcription complexes (such as Pol II and TBP in preinitiation complexes) and/or the colocalization of a defined transcription factor with visualizable genomic loci in live cells. These live colocalization studies would help elucidate dynamic nuclear processes based on the association and dissociation of regulatory factors with distinct labeled genomic locations or topological associated domains.

In conclusion, we have developed a strategy that is simple to implement for visualizing target antigens in their native form without fixation that can affect cell integrity (Schnell et al., 2012) and without causing any toxicity in the treated cells. Labeling of endogenous nuclear proteins with VANIMA strictly corresponds to the true antibody–antigen complexes that are taking place in the cell after antibody delivery. We believe that this approach can be used for live- and single-cell super-resolution detection of a large variety of factors and PTMs. Moreover, our method showing that labeled antibodies can be easily and efficiently delivered to cells, overcomes the previously frustrating antibody-delivery limitation issues in biomedicine. Thus, the cellular delivery of antibodies described in our study may also provide extremely useful tools against the fight of a variety of diseases.

Materials and methods

Cell culture

The human U2OS osteosarcoma cells (HTB-96; American Type Culture Collection [ATCC]) were maintained in DMEM supplemented with 10% FCS and 40 µg/ml gentamicin. Human foreskin fibroblast cells (SCRC-1041; ATCC) were cultivated in DMEM/F12 with GlutaMAX-I supplemented with 10% FCS, 15 mM Hepes, 100 UI/ml penicillin, and 100 µg/ml streptomycin. Mouse embryonic stem cells were maintained in DMEM supplemented with 15% FBS (Millipore), 100 UI/ml penicillin, 100 µg/ml streptomycin, 2 mM L-glutamine, 0.1 mM nonessential amino acids, 0.1% β-mercaptoethanol, 1,500 U/ml leukemia inhibitory factor and 2i inhibitors (Ying et al., 2008), 3 µM CHIR99021, and 1 µM PD0325901 (Axon Medchem) on plates coated with 0.1% gelatin solution in 1× PBS (PAN BIO TECH). All these cell lines were maintained in a 5% CO_2 atmosphere at 37°C. Schneider S2 cells (CRL-1963; ATCC) were cultivated by using SCHNEIDER medium containing 10% FCS (heat inactivated)

and 0.5% penicillin/streptomycin and were grown at 27°C. After electroporation the cells were cultivated for 24 h in their corresponding medium without any antibiotics.

Plasmids and transfection procedure

Four mammalian constructs were used for ectopic expression of fluorescent fusion proteins. The expression vectors for HA-GFP, GFP-hRPB1, CFP-hTAF10, and GFP-hTBP were described previously (Soutoglou et al., 2005; de Graaf et al., 2010; Vosnakis et al., 2017). The Flag-Venus microhomology-mediated end-joining (MMEJ) template (hTAF10-MMEJ) and the plasmid expressing three guide RNAs (one targeting the exon 1 of hTAF10 and two targeting the MMEJ template) and coexpressing Cas9-mCherry (hTAF10-Cas9) were assembled by Megawhop (Miyazaki, 2011) and golden gate cloning (Engler et al., 2009), respectively. For transfection, cells were plated into 12-well plates containing 18-mm-high precision cover glasses (Marienfeld) 1 d before transfection to achieve a confluency of ∼70–80%. They were transfected with 100 ng of the corresponding plasmid (GFP-hRPB1, CFP-hTAF10, GFP-hTBP, or HA-GFP) by using Lipofectamine 2000 reagent (Thermo Fisher) following the manufacturer's instructions. The cells were fixed 48 h after transfection for confocal imaging by using the protocol described in the section Sample preparation for imaging.

Antibodies and Fab fragments

The mouse mAbs against RPB1 (1PB-7G5 mAb), TAF10 (6TA-2B11 mAb), TBP (3TF1-3G3 mAb), and bacterial MBP (17TF2-1H4 mAb) were described previously (Lescure et al., 1994; Bertolotti et al., 1996; Zeder-Lutz et al., 1999; Lebedeva et al., 2005; Helmlinger et al., 2006). The anti–α-tubulin antibody was purchased from Sigma-Aldrich (clone DM1A). The anti-γH2AX antibody (14HH2-1H2 mAb) was generated by immunizing mice with the phosphorylated peptide (KATQA[phosphoS]QEY) as described previously (Muratoglu et al., 2003). Specificity of the new antibody was tested by ELISA (Fig. S1 E). Antibodies were purified by using preequilibrated Protein G Sepharose Fast Flow (GE Healthcare) in a batch purification for 2 h at 4°C. Afterward the Sepharose beads were transferred to a Poly-Prep Chromatography column (Bio-Rad) and washed for 20 column volumes with 1× PBS to remove any unspecific bound proteins. The antibodies were eluted in 1-ml fractions by using 0.1 M glycine, pH 2.7, and were directly neutralized with 70 µl of 1 M Tris-HCl, pH 8.2. The fractions containing most of the antibodies were pooled and dialyzed against 1× PBS before 10% glycerol was added to store the aliquoted antibodies at −80°C. Fab fragments of our mAbs were prepared by using the Pierce Mouse IgG1 Fab and $F(ab')_2$ Preparation kit (Thermo Fisher). Preparation was performed as written in the manufacturer's protocol by using a total amount of 1 mg mAbs and digesting them with ficin for 5 h at 37°C. Alternatively, the Fab fragments were prepared by digestion with papain (Sigma-Aldrich). The antibodies were cleaved into Fab fragments by addition of 400 ng papain per milligram of antibody. After incubation for 3 h at 37°C, the Fab fragments were separated from the Fc fragments and undigested antibody molecules by protein A Sepharose chromatography. Unbound Fab fragments were subsequently purified by size exclusion chromatography on a Superdex 75 10/300 (GE Healthcare) equilibrated in PBS. The recovered Fab were stored at 4°C at a concentration of 5 mg/ml.

Antibody labeling

All mAbs and Fab fragments were fluorescently labeled by using the same protocol. A solution containing 100 µg of antibodies or Fab fragments was dialyzed against 0.1 M sodium bicarbonate (Sigma-Aldrich) for 4 h at 4°C using DiaEasy dialyzing tubes (BioVision) to increase labeling efficiency by raising the pH of the antibody solution over a pH of 8. The labeling reaction was performed following the protocol of the Alexa Fluor Monoclonal Antibody Labeling kit (Thermo Fisher) to label 100 µg antibody or Fab fragment randomly with for example Alexa Fluor 488 dyes (A20181). Labeling efficiency was calculated by using the formula given in the manufacturer's protocol. The Alexa Fluor 488 dyes have a tetra-fluoro-phenyl ester moiety, which reacts with primary amines of proteins to form a covalent dye–protein conjugate. This labeling strategy results in a high labeling density with up to five to seven dyes per molecule of antibody.

Note of caution: To label antibodies or Fabs, we have used *N*-hydroxysuccinimide ester fluorophores that react with the amine group at the tip of the side chain of lysines. This is a conventional method of chemical labeling of proteins, which works fine with antibodies that do not harbor lysine residues in their binding site (paratope). If the quality of binding of the labeled Fab (that can be easily tested by IF) is affected by this technique and when the antibody is precious, we propose to set up a site-directed labeling, which consists in the preparation of $(Fab')_2$ fragments, which can be specifically labeled at the typical cysteine residues in the C-terminal of the Fab' (hinge region) with maleimide-activated fluorophores upon mild reduction. The scaffold cysteines present in the different IgG fold-domains of the Fab' are not accessible under these conditions. This method allows the addition of a maximum of two to three fluorophore molecules per Fab and preserves the antibody-binding site from any deleterious chemical alteration.

Electroporation procedure

Transductions were performed by using the Neon Transfection system (MPK5000; Thermo Fisher) and the corresponding Neon kits (MPK1096 or MPK10096; Thermo Fisher). To transduce 10^5 cells, the 10-µl Neon tips were used with 0.5–4 µg antibodies or Fab fragments; however, to transduce 1.2 x10^6 cells with 6–48 µg antibodies, the 100-µl Neon tips were used. The desired number of cells (depending on the number of transductions performed) were trypsinized and washed once with 4 ml 1× PBS before the pellet was resuspended in the supplied resuspension buffer. The volume corresponding to 1×10^5 or 1.2×10^6 cells was mixed with the labeled antibody or Fab solution and immediately transduced by using the following parameters: 1550 V, 3 pulses, and 10 ms per pulse. After transduction, the cells were transferred directly into 12-well plates (Corning) containing prewarmed medium without antibiotics. The medium was changed to antibiotic containing medium 24 h after transduction if the cells were used for live imaging; otherwise, they were fixed directly for fixed-cell imaging. Transduction efficiency was tested 24 h after electroporation of anti-RPB1 (1PB-7G5) mAb by counting 100 cells by using a confocal microscope to determine the percentage of cells showing a fluorescent signal in the nucleus. Cell viability after electric shock was determined by measuring the percentage of living cells before and after transduction by using a Countess II Cell Counter (Thermo Fisher) and Trypan blue staining of dead cells and normalization to the cell viability before electroporation.

Note of caution: The cells should not stay >20 min in the resuspension buffer, because the cell viability will decrease drastically. If many transductions need be performed, it can be advantageous to prepare several cell pellets and resuspend them one by one.

In the past, we tried classical electroporation with cuvettes to deliver antibodies inside cells, but this approach was not so successful, because the majority of the treated cells were dying after the electric shock (one pulse). The Neon apparatus used in this study corresponds to a novel electroporation device with a capillarity electrode. The design of the electrode in pipette (and not in cuvette) has been shown to produce a more uniform electric field within a small volume, which results in less toxicity to the cells without loss of transfection efficiency.

This apparatus is commercially available for DNA or siRNA transfection. However, we adapted the setting of several parameters (voltage, number of pulses, and efficiency of internalization) for optimal protein delivery. To our knowledge, this achievement allows nearly all treated cells to be transduced without loss of viability. Importantly, the same Neon electroporation apparatus has also been used successfully to deliver proteins in cells (Clift et al., 2017).

Sample preparation for imaging

For fixed sample preparation, the transduced cells were transferred to 12-well plates containing 18-mm-high precision cover glasses (Marienfeld). They were fixed 24 h after electroporation by using 4% PFA (Electron Microscopy Sciences) in 1× PBS prewarmed to 37°C for 5 min. Afterward the cells were washed twice for 5 min at RT with 1× PBS plus 0.02% Triton X-100, once with 1× PBS, once with 1× PBS plus 0.1% Triton X-100 for 20 min at RT, and then again twice for 5 min at RT with 1× PBS plus 0.02% Triton X-100 and once with 1× PBS. Next, the cells were incubated with a DAPI solution in dH_2O (1/2,500 dilution from 1 mg/ml stock solution; Sigma-Aldrich) for 30 s and afterward mounted with Vectashield (H1000, not containing DAPI; Vector Laboratories) if the samples were used for 3D-SIM microscopy. When samples were prepared for visualization with the use of confocal microscopy, they were directly mounted with Vectashield mounting medium containing DAPI (H1200; Vector Laboratories).

Because the target is already labeled with the transduced antibody, most of the washing steps mentioned in the section above are optional and are needed only if the signal-to-noise ratio during imaging is too low because of nontransduced antibodies, which can stick on the coverslip surface.

For classical IF, the cells were seeded as described before, but the day before the experiment to achieve a confluency of ∼70–80%. The fixation protocol was the same as for the transduced samples except that all wash steps are mandatory and there are additional incubation steps with the primary and secondary antibodies. After fixation as described above, the cells were permeabilized by using 1× PBS plus 0.1% Triton X-100 for 20 min at RT and then incubated with 2 µg primary antibody (anti-RPB1, anti-TAF10, or anti-TBP) diluted in 1× PBS plus 10% FCS for 1 h at RT. The negative control was incubated only with buffer missing any primary antibody. The cells were washed three times for 5 min at RT twice with 1× PBS plus 0.02% Triton X-100 and once with 1× PBS followed by an incubation with the Alexa Fluor 488–conjugated anti-mouse secondary antibody (Thermo Fisher) diluted 1/3,000 in 1× PBS plus 10% FCS for 1 h at RT. After three more washings for 5 min at RT, samples were mounted using Vectashield containing DAPI for confocal imaging. To eliminate all soluble proteins before fixation and to visualize chromatin-bound RPB1, the cells were treated with CSK buffer (100 mM NaCl, 3 mM $MgCl_2$, 10 mM Hepes, 300 mM sucrose, 0.3% Triton X-100, and protease inhibitor cocktail) before fixing with PFA for 10 min at RT.

For live-imaging the cells were transferred to µ-dishes (35-mm-diameter, high, glass bottom; ibidi) for confocal imaging or to µ-slides (8-well, glass bottom; ibidi) for 3D-SIM imaging after transduction containing prewarmed medium and incubated at 5% CO_2 and 37°C until imaging started. Before imaging the medium was changed to the described growth medium without phenol red for confocal imaging or Leibovitz's L-15 medium (Thermo Fisher) for 3D-SIM microscopy.

Transcription inhibition

Inhibition of transcription was achieved by treating U2OS cells either with α-amanitin (Molekula) or Flavo (Flavo hydrochloride hydrate; Sigma-Aldrich). Electroporated cells were incubated 6 h after transduction with 4 µg/ml α-amanitin overnight. Flavo treatment was performed 24 h after transduction by incubating the cells with 2 µM Flavo for 1 h.

DNA damage induction

For γH2AX imaging, DNA damage in the form of double-strand breaks was induced by using either HU (Sigma) or NCS (Sigma). For HU treatment, the cells were transduced with 2 µg anti-γH2AX Fab antibody and 12 h later treated with 2 mM HU for 48 h before the cells were fixed. To induce DNA damage with NCS, the cells were transduced as described before and 22 h later incubated with 100 ng/ml NCS for 15 min. Afterward the medium was changed to classical growth medium, and the cells were incubated for 2 h more before fixation. For γH2AX live imaging, the same protocol was followed except that 50 ng/ml (confocal microscopy) or 200 ng/ml (3D-SIM microscopy) NCS was added immediately before image acquisition.

Confocal microscopy

Confocal imaging of fixed samples was performed on an SP8UV microscope (Leica) equipped with a 561-nm DPSS laser, a 633-nm HeNe laser, a 405-nm laser diode, and a 488-nm argon laser. A 63× oil immersion objective (NA 1.4) was used, and images were taken by using the hybrid detector photon-counting mode. Confocal live imaging was performed on either an SP8X microscope (Leica) equipped with a white light laser (Leica) by using the 488-nm laser line or a Ti microscope (Nikon) equipped with a CSU-X1 confocal scanner (Yokogawa) and an Evolve back-illuminated EMCCD camera (Photometrics). 2D videos from the SP8X microscope were taken using a 63× oil immersion objective (NA 1.4) on photomultiplier tube detection mode and time intervals of 10 min. The Ti microscope 2D videos were taken using a 60× oil immersion objective (NA 1.4), an exposure time of 800 ms, and time intervals of 10 min. All images and videos were subsequently analyzed and processed by using Fiji/ImageJ software.

3D-SIM super-resolution microscopy and image analysis

3D-SIM was performed on a DeltaVision OMX-Blaze V4 system (GE Healthcare) equipped with a Plan Apo N 60× (1.42 NA) oil immersion objective lens (Olympus), four liquid-cooled sCMOS cameras (pco. edge 5.5, full frame 2,560 × 2,160; PCO) and 405-, 445-, 488-, 514-, 568-, and 642-nm solid-state lasers. The 405-, 488-, and 568-nm laser lines were used during acquisition, and the optical z sections were separated by 0.125 µm. For fixed cells, laser power was attenuated to 10 or 31.3%, and exposure times were typically between 75 and 400 ms. Live imaging of RPB1 or γH2AX was performed by using a laser power attenuated to 10 or 31.3% and an exposure time of 10–25 ms with time intervals of either 4.1 or 15 s and a total acquisition time of 45 s. The raw images were processed and reconstructed by using the DeltaVision OMX SoftWoRx software package (v6.1.3; Applied Precision).

For the 3D-SIM images in Fig. 4 and Fig. 6 A, the outline for the nucleus (DAPI channel) was defined after applying in Fiji a Gaussian blur (σradius 4), applying a threshold to match the nucleus outline ("mean algorithm"), and the outline was detected by using the "Analyze Particle" (with option "Include Holes"). The resulting outline was shown on the channel of interest, and the look-up table "Yellow Hot" has been applied to the image for a better visualization. The SIMcheck Fiji/ImageJ plugin (Ball et al., 2015) was used to check raw and reconstructed image quality. Channel intensity profiles, Fourier plots, motion and illumination variation, as well as modulation contrast to noise maps have been tested for all 3D-SIM images and are in general above the required thresholds.

The image processing and quantification was performed by using the Imaris software (Bitplane) for preparing 3D videos or Fiji/ImageJ software and in particular the 3D spot segmentation (Ollion

et al., 2013) as well as the 3D object counter (Bolte and Cordelières, 2006) for the quantification of the 3D images. In brief, the spots were segmented by finding local maxima in the image and afterward fitting a Gaussian distribution locally. As soon as the mask of each spot was available, factors such as spot number or volume could be computed. Finally, analysis of the spot data was performed by using Matlab (MathWorks). Distributions of spot volumes with the use of antibodies against RPB1 or TAF10 were computed by averaging the histograms of measured spot volumes >10 cells for each condition (Flavo-treated vs. untreated). In addition, the mean fraction of spots bigger than 10^{-2} μm^3 in each condition was reported. P-values were calculated by using the two-sample *t* test that allows to determine whether two population means are significantly different.

Flag-Venus hTAF10 knock-in

The knock-in of the Flag-Venus coding sequence at exon 1 of the hTAF10 gene was performed by using CRISPR/Cas9 and MMEJ (Nakade et al., 2014). In brief, U2OS cells were cotransfected with the hTAF10-Cas9 and hTAF10-MMEJ plasmids at a ratio of 2:1 by using FuGENE HD (Promega). After 48 h, cells that had taken up the Cas9 plasmid (mCherry positive) were sorted by flow cytometry (FACS ARIA; BD Biosciences) and cultured under limiting dilution conditions. Colonies were expanded and genotyped by PCR and tested for Flag-Venus tag insertion by IF. Sequencing of the PCR products confirmed the in-frame insertion of the Flag-Venus sequence. Note that all the three knock-in clones obtained were heterozygous.

Immunoprecipitation

For electroporation-immunoprecipitation (Elec-IP), 1.2×10^6 cells were transduced with 6–48 µg anti-RPB1 7G5 (corresponding to 0.5–4 µg antibody in 1×10^5 cells) 24 h before protein extraction. Cells treated with the same electric shock, but without any antibody, were used as a mock control. The cells were trypsinized and whole-cell protein extracts were produced by solubilizing the cell pellets in 40 µl RIPA buffer (50 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, and 0.1% SDS) and incubating them for 5 min on ice. The concentration of the extracts was determined by using a standard Bradford assay, and 30 µg extract was mixed with 100 µl of equilibrated protein G–coupled magnetic Dynabeads (Thermo Fisher) for an immunoprecipitation overnight at 4°C. Next, the Dynabeads were separated from the supernatant containing nonbound proteins and were washed three times with IP500 buffer (25 mM Tris-HCl, pH 7.9, 0.1% NP-40, 5 mM $MgCl_2$, 10% glycerol, 500 mM KCl, 2 mM DTT, and protease inhibitor cocktail) and two times with IP100 buffer (25 mM Tris-HCl, pH 7.9, 0.1% NP-40, 5 mM $MgCl_2$, 10% glycerol, 100 mM KCl, 2 mM DTT, and protease inhibitor cocktail) to remove any unspecific bound proteins. The beads with the bound antibody-protein complexes were stored in IP100 buffer. The input protein extracts, the supernatant of the Elec-IP, as well as the beads were analyzed afterward by Western blot.

Western blot analysis

Whole-cell protein extracts were prepared from cells washed twice with 1× PBS by using RIPA buffer (see the previous section). Elec-IP fractions were loaded on 4–15% precast SDS-PAGE gels (Bio-Rad) with Laemmli buffer. Protein transfer on nitrocellulose membranes was performed by using Mini Protean II tanks (Bio-Rad). Western blots were blocked by using 3–5% milk for at least 30 min before overnight incubation with the primary antibody against RPB1 (1PB-7G5 mAb, 1:1,000). Signal was detected by incubating for 1 h with HRP-conjugated secondary antibodies (1:10,000; Jackson ImmunoResearch) and revealed by using ECL (Thermo Fisher) and ChemiDoc Touch Imaging System (Bio-Rad).

Pre-mRNA transcription analysis

24 h before total RNA extraction, 1.2×10^6 U2OS cells were transduced with 24 µg anti-RPB1, anti-TBP, or anti-TAF10 antibodies. U2OS cells electroporated but without transduction of antibody were used as controls. Additionally, electroporated U2OS cells without transduction of antibody were treated with 4 µg/ml α-amanitin overnight as a positive control for transcriptional inhibition. As negative control, 24 µg an antibody targeting the bacterial MBP was transduced into U2OS cells. Total RNA was extracted by using Tri Reagent (Molecular Research Center, Inc.) and following manufacturer's instructions. Removal of genomic DNA contamination was achieved by using the TURBO DNA-*free* kit (Thermo Fisher). For reverse transcription, 3.2 µg of random hexamer primers (Thermo Fisher), dNTP Mix (Thermo Fisher), and Transcriptor Reverse transcription (Roche) were used following manufacturer's instruction. For qPCR, the cDNA samples were diluted and amplified by using SYBR Green 2× PCR Master Mix I (Roche) and a LightCycler 480 Instrument II (Roche) with the following program: one cycle of 5 min at 95°C for predenaturation, 45 amplification cycles with 10 s at 95°C for denaturation, 20 s at 65°C for primer annealing, and 20 s at 72°C for extension. Melting curves were determined between 65°C and 97°C followed by one cycle of cooling for 30 s at 40°C. Primer pairs used for qPCR are listed in Table S2. To quantify newly synthesized RNA Pol II transcripts, primer pairs amplifying from an intron to an exon were designed, therefore reflecting unspliced transcripts. The genes analyzed were selected randomly and represent genes of different chromosomes. However, because unspliced transcripts are a minority in total RNA extracts, the genes selected are mostly highly expressed genes. The obtained threshold-values were used to calculate the relative fold change by using the $\Delta\Delta C_T$ method by normalization to RNA Pol III transcripts (*RPPH1* and *RN7SK*) and taking into account primer efficiencies. The heatmap was based on the mean fold change, with the U2OS cells electroporated but without transduction sample set to zero change in expression and was generated by using R 3.4.3 and RStudio 1.1.383 and the ComplexHeatmap (Bioconductor) package.

Cell cycle analysis

For cell cycle analysis, 1.2×10^6 U2OS cells were electroporated with 24 µg anti-RPB1, anti-TAF10, or anti-TBP antibody. As controls, electroporated cells without any antibody were used. As positive control for transcriptional inhibition, electroporated cells without any antibody were treated with 4 µg/ml α-amanitin overnight. As negative control, 24 µg anti-MBP was transduced into U2OS cells. The cells were harvested 24 or 48 h after electroporation, washed with 1× PBS, and fixed in 70% ethanol. Fixed cells were stained with 15 µg/ml propidium iodide (Sigma-Aldrich) and treated with 75 µg/ml RNase A (Thermo Fisher) for 1 h before the FACS analysis. FACS analysis was conducted on a FACS Celesta (BD Biosciences) counting 10,000 cells per sample, and data analysis was performed by using FlowJo 10.2. The cell cycle phases were assigned manually.

Proliferation assay

Proliferation of U2OS cells after antibody transduction was tested by using the Click-it Plus EdU Flow Cytometry Assay kit (C10632; Thermo Fisher). A total amount of 1.2×10^6 cells was transduced with 24 µg anti-RPB1, anti-TAF10, anti-TBP, or anti-MBP antibody and incubated for 24 h at 5% CO_2 and 37°C. As controls transduced cells without any antibody were added either as positive control for normal proliferation or as negative control by adding 4 µg/ml α-amanitin (Molekula) overnight 6 h after transduction to see how proliferation was affected if transcription was inhibited. The cells were treated 24 h after transduction with 10 µM EdU for 1 h to test the proliferation capacity of the cells. Non-EdU treated cells for

every transduction were added as controls. The Click-it reaction with Alexa Fluor 488 was performed as described in the manufacturer's protocol. FACS analysis was performed on a FACS Celesta (BD Biosciences) counting 30,000 cells per sample. The positive control was used for normalization.

Apoptosis assay

To test if the cells would undergo apoptosis after transduction of antibodies, an APOPercentage apoptosis assay (Biocolor) was performed. U2OS cells (1×10^5) were transduced with 2 µg anti-RPB1, anti-TAF10, anti-TBP, or anti-MBP antibody and incubated for 24 h at 5% CO_2 and 37°C. As negative (0% apoptosis) control, electroporated cells without antibodies were used. As positive (100% apoptosis) control, cells were treated 20 h after transduction, without antibodies, with 10 mM H_2O_2, for 4 h to induce apoptosis. The apoptosis assay was performed as described in the manufacturer's protocol for the colorimetric assay. The results were normalized to the positive control.

Suitability of new antibodies for VANIMA

According to our experience, antibodies that recognize their epitopes in the intracellular context are the ones that have a good chance of working in VANIMA. We have observed that those antibodies that work fine when tested by IF also work in our live-cell imaging assays. This shows that the accessibility of the epitope in the intracellular context is the limiting factor and that likely all antibodies that are used for imaging in fixed cells will be excellent candidates for the VANIMA application. Within a set of 25 different antibodies that were all working in IF, only one was not adequate for VANIMA. In this case, we found that the epitope was hidden after neosynthesis in the cytoplasm and it became accessible only when the antigen was imported in the nucleus (Freund et al., 2013). After an assessment of the quality of the antibody in IF, it should be purified and labeled with fluorescent dyes as described in the Antibody labeling section. Depending on the localization of the target protein (nucleus or cytoplasm), a digestion of the antibody to Fab fragments could be considered. To identify the amount of antibody or Fab that needs to be electroporated to bind a suitable amount of target protein, a titration electroporation similar to the one shown in Fig. 1 B should be performed. It is important to note that amounts >10 µg antibodies or Fabs should be avoided because at this point the amount of protein electroporated starts to get toxic for the cells. Afterward, the binding of the antibody to the intracellular target should be verified by immunoprecipitation after electroporation as shown in Fig. 1 C or by performing an IF-electroporation comparison as shown in Fig. 2 C (and Fig. S1 C) depending on if the desired antibody has several or only one epitope on the target protein. The last step would be to verify if the antibody is blocking functions of the target protein or affecting the survival of the cells. A first indication is the viability of the cells after electroporation which should be, depending on the cell line used, >60–90% (see also Table S1). Other validation experiments would be to test the proliferation of the cells and the cell cycle progression or if apoptosis occurs (Fig. 3, B–E). Depending on the target protein, also more specific validation experiments should be considered as the premRNA transcription analysis for transcription factors (Fig. 3 A). After these validation tests, the antibody or Fab can be used for fixed- or live-cell imaging of endogenous proteins.

Online supplemental material

Fig. S1 shows different experiments to verify the efficiency (A), localization (B and D), target binding (E), and affinity (C) of different antibodies using VANIMA. Fig. S2 shows the imaging of transcription factors with classical labeling methods such as IF (A) or the genetic tagging with fluorescent tags (B–E). Table S1 shows electroporation of antibodies is highly efficient, keeping a high viability of the cells, and can be used in many different cell lines. Table S2 shows primers used to quantify RNA Pol II premRNA as well as RNA Pol I and Pol III transcripts. Video 1 shows the transport of labeled anti-RPB1 mAb from the cytoplasm to the nucleus of living cells. Videos 2–4 show nuclei of U2OS cells transduced with either labeled anti-RPB1, anti-TAF10, or anti-TBP mAbs analyzed by 3D-SIM microscopy. Videos 5 and 6 show nuclei transduced with anti-γH2AX Fab and treatment with or without HU analyzed by 3D-SIM microscopy. Video 7 shows confocal live-cell imaging of RNA Pol II using VANIMA. Video 8 shows 3D-SIM live-cell imaging of RNA Pol II clusters. Video 9 shows confocal live-cell imaging of γH2AX foci. Video 10 shows 3D-SIM live-cell imaging of γH2AX foci. Higher-resolution videos of the 3D-SIM videos can be obtained directly from the corresponding authors.

Acknowledgments

We thank T. Sexton, S.D. Vincent, and F. Müller for carefully reading the manuscript and helpful comments; L. Schermelleh for constant support; the Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC) cell culture service for help with the cells; and Y. Lutz, M. Koch, and E. Guiot from the IGBMC imaging platform. L. Tora thanks M. Featherstone for making the starting of this project possible at the School of Biological Sciences, Nanyang Technological University (Singapore).

This work was supported by funds from Conseil National de la Recherche Scientifique, Institut national de la santé et de la rechrche médicale, Université de Strasbourg, Ligue Contre le Cancer (Comité CCIRGE-BFC to E. Weiss), by the European Research Council Advanced Grant (ERC-2013-340551, Birtoaction to L. Tora), and a grant from a French State fund managed by the Agence Nationale de la Recherche (ANR-10-LABX-0030-INRT) under the frame program Investissements d'Avenir (ANR-10-IDEX-0002-02), and core founding from the Agency for Science, Technology, and Research for G.D. Wright.

The authors declare no competing financial interests.

Author contributions: S. Conic and V. Fischer performed experiments. S. Conic, E. Weiss, and L. Tora designed the experiments to develop the protocol. M. Oulad-Abdelghani generated the anti-γH2AX antibody, and D. Desplancq characterized it. K. Babu N. and G.D. Wright performed initial labeling and 3D-SIM experiments. S. Conic and A. Ferrand performed all the 3D-SIM experiments presented in the study. B.R. San Martin and V. Heyer designed and provided vectors for CRISPR/Cas9 knock-in experiments. N. Molina and J. Pontabry carried out image analyses. S. Conic, N. Molina, E. Weiss, and L. Tora wrote the manuscript.

Submitted: 28 September 2017
Revised: 15 December 2017
Accepted: 18 January 2018

References

Ball, G., J. Demmerle, R. Kaufmann, I. Davis, I.M. Dobbie, and L. Schermelleh. 2015. SIMcheck: A toolbox for successful super-resolution structured illumination microscopy. *Sci. Rep.* 5:15915. https://doi.org/10.1038/srep15915

Beghin, A., A. Kechkar, C. Butler, F. Levet, M. Cabillic, O. Rossier, G. Giannone, R. Galland, D. Choquet, and J.B. Sibarita. 2017. Localization-based super-resolution imaging meets high-content screening. *Nat. Methods.* 14:1184–1190. https://doi.org/10.1038/nmeth.4486

Berglund, D.L., and J.R. Starkey. 1989. Isolation of viable tumor cells following introduction of labelled antibody to an intracellular oncogene product

using electroporation. *J. Immunol. Methods.* 125:79–87. https://doi.org/10.1016/0022-1759(89)90080-X

Bertolotti, A., Y. Lutz, D.J. Heard, P. Chambon, and L. Tora. 1996. hTAF(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. *EMBO J.* 15:5022–5031.

Betzig, E., G.H. Patterson, R. Sougrat, O.W. Lindwasser, S. Olenych, J.S. Bonifacino, M.W. Davidson, J. Lippincott-Schwartz, and H.F. Hess. 2006. Imaging intracellular fluorescent proteins at nanometer resolution. *Science.* 313:1642–1645. https://doi.org/10.1126/science.1127344

Bolte, S., and F.P. Cordelières. 2006. A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* 224:213–232. https://doi.org/10.1111/j.1365-2818.2006.01706.x

Boulon, S., B. Pradet-Balade, C. Verheggen, D. Molle, S. Boireau, M. Georgieva, K. Azzag, M.C. Robert, Y. Ahmad, H. Neel, et al. 2010. HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Mol. Cell.* 39:912–924. https://doi.org/10.1016/j.molcel.2010.08.023

Burgess, A., T. Lorca, and A. Castro. 2012. Quantitative live imaging of endogenous DNA replication in mammalian cells. *PLoS One.* 7:e45726. https://doi.org/10.1371/journal.pone.0045726

Chakrabarti, R., D.E. Wylie, and S.M. Schuster. 1989. Transfer of monoclonal antibodies into mammalian cells by electroporation. *J. Biol. Chem.* 264:15494–15500.

Chao, S.H., K. Fujinaga, J.E. Marion, R. Taube, E.A. Sausville, A.M. Senderowicz, B.M. Peterlin, and D.H. Price. 2000. Flavopiridol inhibits P-TEFb and blocks HIV-1 replication. *J. Biol. Chem.* 275:28345–28348. https://doi.org/10.1074/jbc.C000446200

Cho, W.K., N. Jayanth, S. Mullen, T.H. Tan, Y.J. Jung, and I.I. Cissé. 2016. Super-resolution imaging of fluorescently labeled, endogenous RNA polymerase II in living cells with CRISPR/Cas9-mediated gene editing. *Sci. Rep.* 6:35949. https://doi.org/10.1038/srep35949

Cisse, I.I., I. Izeddin, S.Z. Causse, L. Boudarene, A. Senecal, L. Muresan, C. Dugast-Darzacq, B. Hajj, M. Dahan, and X. Darzacq. 2013. Real-time dynamics of RNA polymerase II clustering in live human cells. *Science.* 341:664–667. https://doi.org/10.1126/science.1239053

Clift, D., W.A. McEwan, L.I. Labzin, V. Konieczny, B. Mogessie, L.C. James, and M. Schuh. 2017. A method for the acute and rapid degradation of endogenous proteins. *Cell.* 171:1692–1706.e18. https://doi.org/10.1016/j.cell.2017.10.033

Courtête, J., A.P. Sibler, G. Zeder-Lutz, D. Dalkara, M. Oulad-Abdelghani, G. Zuber, and E. Weiss. 2007. Suppression of cervical carcinoma cell growth by intracytoplasmic codelivery of anti-oncoprotein E6 antibody and small interfering RNA. *Mol. Cancer Ther.* 6:1728–1735. https://doi.org/10.1158/1535-7163.MCT-06-0808

Cramer, L.P., and T.J. Mitchison. 1995. Myosin is involved in postmitotic cell spreading. *J. Cell Biol.* 131:179–189. https://doi.org/10.1083/jcb.131.1.179

de Graaf, P., F. Mousson, B. Geverts, E. Scheer, L. Tora, A.B. Houtsmuller, and H.T. Timmers. 2010. Chromatin interaction of TATA-binding protein is dynamically regulated in human cells. *J. Cell Sci.* 123:2663–2671. https://doi.org/10.1242/jcs.064097

Desplancq, D., G. Freund, S. Conic, A.P. Sibler, P. Didier, A. Stoessel, M. Oulad-Abdelghani, M. Vigneron, J. Wagner, Y. Mély, et al. 2016. Targeting the replisome with transduced monoclonal antibodies triggers lethal DNA replication stress in cancer cells. *Exp. Cell Res.* 342:145–158. https://doi.org/10.1016/j.yexcr.2016.03.003

Dundr, M., U. Hoffmann-Rohrer, Q. Hu, I. Grummt, L.I. Rothblum, R.D. Phair, and T. Misteli. 2002. A kinetic framework for a mammalian RNA polymerase in vivo. *Science.* 298:1623–1626. https://doi.org/10.1126/science.1076164

Ellenberg, J., J. Lippincott-Schwartz, and J.F. Presley. 1999. Dual-colour imaging with GFP variants. *Trends Cell Biol.* 9:52–56. https://doi.org/10.1016/S0962-8924(98)01420-2

Engler, C., R. Gruetzner, R. Kandzia, and S. Marillonnet. 2009. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One.* 4:e5553. https://doi.org/10.1371/journal.pone.0005553

Freund, G., A.P. Sibler, D. Desplancq, M. Oulad-Abdelghani, M. Vigneron, J. Gannon, M.H. Van Regenmortel, and E. Weiss. 2013. Targeting endogenous nuclear antigens by electrotransfer of monoclonal antibodies in living cells. *MAbs.* 5:518–522. https://doi.org/10.4161/mabs.25084

Gorski, S.A., S.K. Snyder, S. John, I. Grummt, and T. Misteli. 2008. Modulation of RNA polymerase assembly dynamics in transcriptional regulation. *Mol. Cell.* 30:486–497. https://doi.org/10.1016/j.molcel.2008.04.021

Hager, G.L., J.G. McNally, and T. Misteli. 2009. Transcription dynamics. *Mol. Cell.* 35:741–753. https://doi.org/10.1016/j.molcel.2009.09.005

Hayashi-Takanaka, Y., K. Yamagata, T. Wakayama, T.J. Stasevich, T. Kainuma, T. Tsurimoto, M. Tachibana, Y. Shinkai, H. Kurumizaka, N. Nozaki, and H. Kimura. 2011. Tracking epigenetic histone modifications in single cells using Fab-based live endogenous modification labeling. *Nucleic Acids Res.* 39:6475–6488. https://doi.org/10.1093/nar/gkr343

Helmlinger, D., S. Hardy, G. Abou-Sleymane, A. Eberlin, A.B. Bowman, A. Gansmüller, S. Picaud, H.Y. Zoghbi, Y. Trottier, L. Tora, and D. Devys. 2006. Glutamine-expanded ataxin-7 alters TFTC/STAGA recruitment and chromatin structure leading to photoreceptor dysfunction. *PLoS Biol.* 4:e67. https://doi.org/10.1371/journal.pbio.0040067

Hernandez, N. 1993. TBP, a universal eukaryotic transcription factor? *Genes Dev.* 7(7b, 7B):1291–1308. https://doi.org/10.1101/gad.7.7b.1291

Hieda, M., H. Winstanley, P. Maini, F.J. Iborra, and P.R. Cook. 2005. Different populations of RNA polymerase II in living mammalian cells. *Chromosome Res.* 13:135–144.

Hnisz, D., K. Shrinivas, R.A. Young, A.K. Chakraborty, and P.A. Sharp. 2017. A phase separation model for transcriptional control. *Cell.* 169:13–23. https://doi.org/10.1016/j.cell.2017.02.007

Kimura, H. 2005. Histone dynamics in living cells revealed by photobleaching. *DNA Repair (Amst.).* 4:939–950. https://doi.org/10.1016/j.dnarep.2005.04.012

Kimura, H., Y. Tao, R.G. Roeder, and P.R. Cook. 1999. Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: Little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure. *Mol. Cell. Biol.* 19:5383–5392. https://doi.org/10.1128/MCB.19.8.5383

Kimura, H., K. Sugaya, and P.R. Cook. 2002. The transcription cycle of RNA polymerase II in living cells. *J. Cell Biol.* 159:777–782. https://doi.org/10.1083/jcb.200206019

Krah, S., C. Schröter, S. Zielonka, M. Empting, B. Valldorf, and H. Kolmar. 2016. Single-domain antibodies for biomedical applications. *Immunopharmacol. Immunotoxicol.* 38:21–28. https://doi.org/10.3109/08923973.2015.1102934

Lebedeva, L.A., E.N. Nabirochkina, M.M. Kurshakova, F. Robert, A.N. Krasnov, M.B. Evgen'ev, J.T. Kadonaga, S.G. Georgieva, and L. Tora. 2005. Occupancy of the Drosophila hsp70 promoter by a subset of basal transcription factors diminishes upon transcriptional activation. *Proc. Natl. Acad. Sci. USA.* 102:18087–18092. https://doi.org/10.1073/pnas.0509063102

Lescure, A., Y. Lutz, D. Eberhard, X. Jacq, A. Krol, I. Grummt, I. Davidson, P. Chambon, and L. Tora. 1994. The N-terminal domain of the human TATA-binding protein plays a role in transcription from TATA-containing RNA polymerase II and III promoters. *EMBO J.* 13:1166–1175.

Lukas, J., J. Bartek, and M. Strauss. 1994. Efficient transfer of antibodies into mammalian cells by electroporation. *J. Immunol. Methods.* 170:255–259. https://doi.org/10.1016/0022-1759(94)90400-6

Macháň, R., and T. Wohland. 2014. Recent applications of fluorescence correlation spectroscopy in live systems. *FEBS Lett.* 588:3571–3584. https://doi.org/10.1016/j.febslet.2014.03.056

Manders, E.M., H. Kimura, and P.R. Cook. 1999. Direct imaging of DNA in living cells reveals the dynamics of chromosome formation. *J. Cell Biol.* 144:813–821. https://doi.org/10.1083/jcb.144.5.813

Markaki, Y., M. Gunkel, L. Schermelleh, S. Beichmanis, J. Neumann, M. Heidemann, H. Leonhardt, D. Eick, C. Cremer, and T. Cremer. 2010. Functional nuclear organization of transcription and DNA replication: A topographical marriage between chromatin domains and the interchromatin compartment. *Cold Spring Harb. Symp. Quant. Biol.* 75:475–492. https://doi.org/10.1101/sqb.2010.75.042

Marschall, A.L., A. Frenzel, T. Schirrmann, M. Schüngel, and S. Dübel. 2011. Targeting antibodies to the cytoplasm. *MAbs.* 3:3–16. https://doi.org/10.4161/mabs.3.1.14110

Marschall, A.L., C. Zhang, A. Frenzel, T. Schirrmann, M. Hust, F. Perez, and S. Dübel. 2014. Delivery of antibodies to the cytosol: debunking the myths. *MAbs.* 6:943–956. https://doi.org/10.4161/mabs.29268

Miyazaki, K. 2011. MEGAWHOP cloning: A method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. *Methods Enzymol.* 498:399–406. https://doi.org/10.1016/B978-0-12-385120-8.00017-6

Muratoglu, S., S. Georgieva, G. Pápai, E. Scheer, I. Enünlü, O. Komonyi, I. Cserpán, L. Lebedeva, E. Nabirochkina, A. Udvardy, et al. 2003. Two different Drosophila ADA2 homologues are present in distinct GCN5 histone acetyltransferase-containing complexes. *Mol. Cell. Biol.* 23:306–321. https://doi.org/10.1128/MCB.23.1.306-321.2003

Nakade, S., T. Tsubota, Y. Sakane, S. Kume, N. Sakamoto, M. Obara, T. Daimon, H. Sezutsu, T. Yamamoto, T. Sakuma, and K.T. Suzuki. 2014. Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. *Nat. Commun.* 5:5560. https://doi.org/10.1038/ncomms6560

Natale, F., A. Rapp, W. Yu, A. Maiser, H. Harz, A. Scholl, S. Grulich, T. Anton, D. Hörl, W. Chen, et al. 2017. Identification of the elementary structural units of the DNA damage response. *Nat. Commun.* 8:15760. https://doi.org/10.1038/ncomms15760

Ollion, J., J. Cochennec, F. Loll, C. Escudé, and T. Boudier. 2013. TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. *Bioinformatics.* 29:1840–1841. https://doi.org/10.1093/bioinformatics/btt276

Ratz, M., I. Testa, S.W. Hell, and S. Jakobs. 2015. CRISPR/Cas9-mediated endogenous protein tagging for RESOLFT super-resolution microscopy of living human cells. *Sci. Rep.* 5:9592. https://doi.org/10.1038/srep09592

Renaud, E., P. Martineau, and L. Guglielmi. 2017. Solubility characterization and imaging of intrabodies using GFP-fusions. *Methods Mol. Biol.* 1575:165–174. https://doi.org/10.1007/978-1-4939-6857-2_9

Rinaldi, A.S., G. Freund, D. Desplancq, A.P. Sibler, M. Baltzinger, N. Rochel, Y. Mély, P. Didier, and E. Weiss. 2013. The use of fluorescent intrabodies to detect endogenous gankyrin in living cancer cells. *Exp. Cell Res.* 319:838–849. https://doi.org/10.1016/j.yexcr.2013.01.011

Röder, R., J. Helma, T. Preiß, J.O. Rädler, H. Leonhardt, and E. Wagner. 2017. Intracellular delivery of nanobodies for imaging of target proteins in live cells. *Pharm. Res.* 34:161–174. https://doi.org/10.1007/s11095-016-2052-8

Rothbauer, U., K. Zolghadr, S. Tillib, D. Nowak, L. Schermelleh, A. Gahl, N. Backmann, K. Conrath, S. Muyldermans, M.C. Cardoso, and H. Leonhardt. 2006. Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat. Methods.* 3:887–889. https://doi.org/10.1038/nmeth953

Schermelleh, L., P.M. Carlton, S. Haase, L. Shao, L. Winoto, P. Kner, B. Burke, M.C. Cardoso, D.A. Agard, M.G. Gustafsson, et al. 2008. Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science.* 320:1332–1336. https://doi.org/10.1126/science.1156947

Schneider, A.F.L., and C.P.R. Hackenberger. 2017. Fluorescent labelling in living cells. *Curr. Opin. Biotechnol.* 48:61–68. https://doi.org/10.1016/j.copbio.2017.03.012

Schnell, U., F. Dijk, K.A. Sjollema, and B.N. Giepmans. 2012. Immunolabeling artifacts and the need for live-cell imaging. *Nat. Methods.* 9:152–158. https://doi.org/10.1038/nmeth.1855

Siddiqui, M.S., M. François, M.F. Fenech, and W.R. Leifert. 2015. Persistent γH2AX: A promising molecular marker of DNA damage and aging. *Mutat. Res. Rev. Mutat. Res.* 766:1–19. https://doi.org/10.1016/j.mrrev.2015.07.001

Snapp, E.L., N. Altan, and J. Lippincott-Schwartz. 2003. Measuring protein mobility by photobleaching GFP chimeras in living cells. Curr. Protoc. Cell Biol. Chapter 21:Unit 21.

Soutoglou, E., M.A. Demény, E. Scheer, G. Fienga, P. Sassone-Corsi, and L. Tora. 2005. The nuclear import of TAF10 is regulated by one of its three histone fold domain-containing interaction partners. *Mol. Cell. Biol.* 25:4092–4104. https://doi.org/10.1128/MCB.25.10.4092-4104.2005

Tantale, K., F. Mueller, A. Kozulic-Pirher, A. Lesne, J.M. Victor, M.C. Robert, S. Capozi, R. Chouaib, V. Bäcker, J. Mateos-Langerak, et al. 2016. A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat. Commun.* 7:12248. https://doi.org/10.1038/ncomms12248

Teng, K.W., Y. Ishitsuka, P. Ren, Y. Youn, X. Deng, P. Ge, A.S. Belmont, and P.R. Selvin. 2016. Labeling proteins inside living cells using external fluorophores for microscopy. *Elife.* 5:e20378. https://doi.org/10.7554/eLife.20378

Teves, S.S., L. An, A.S. Hansen, L. Xie, X. Darzacq, and R. Tjian. 2016. A dynamic mode of mitotic bookmarking by transcription factors. *Elife.* 5:e22280. https://doi.org/10.7554/eLife.22280

Van Regenmortel, M.H. 2014. Specificity, polyspecificity, and heterospecificity of antibody-antigen recognition. *J. Mol. Recognit.* 27:627–639. https://doi.org/10.1002/jmr.2394

van Royen, M.E., A. Zotter, S.M. Ibrahim, B. Geverts, and A.B. Houtsmuller. 2011. Nuclear proteins: finding and binding target sites in chromatin. *Chromosome Res.* 19:83–98.

Vosnakis, N., M. Koch, E. Scheer, P. Kessler, Y. Mély, P. Didier, and L. Tora. 2017. Coactivators and general transcription factors have two distinct dynamic populations dependent on transcription. *EMBO J.* 36:2710–2725. https://doi.org/10.15252/embj.201696035

Wild, T., and P. Cramer. 2012. Biogenesis of multisubunit RNA polymerases. *Trends Biochem. Sci.* 37:99–105. https://doi.org/10.1016/j.tibs.2011.12.001

Ying, Q.L., J. Wray, J. Nichols, L. Batlle-Morera, B. Doble, J. Woodgett, P. Cohen, and A. Smith. 2008. The ground state of embryonic stem cell self-renewal. *Nature.* 453:519–523. https://doi.org/10.1038/nature06968

Zeder-Lutz, G., A. Benito, and M.H. Van Regenmortel. 1999. Active concentration measurements of recombinant biomolecules using biosensor technology. *J. Mol. Recognit.* 12:300–309. https://doi.org/10.1002/(SICI)1099-1352(199909/10)12:5<300::AID-JMR467>3.0.CO;2-N

Zhao, Z.W., R. Roy, J.C. Gebhardt, D.M. Suter, A.R. Chapman, and X.S. Xie. 2014. Spatial organization of RNA polymerase II inside a mammalian cell nucleus revealed by reflected light-sheet superresolution microscopy. *Proc. Natl. Acad. Sci. USA.* 111:681–686. https://doi.org/10.1073/pnas.1318496111

JCB

Conic et al., https://doi.org/10.1083/jcb.201709153

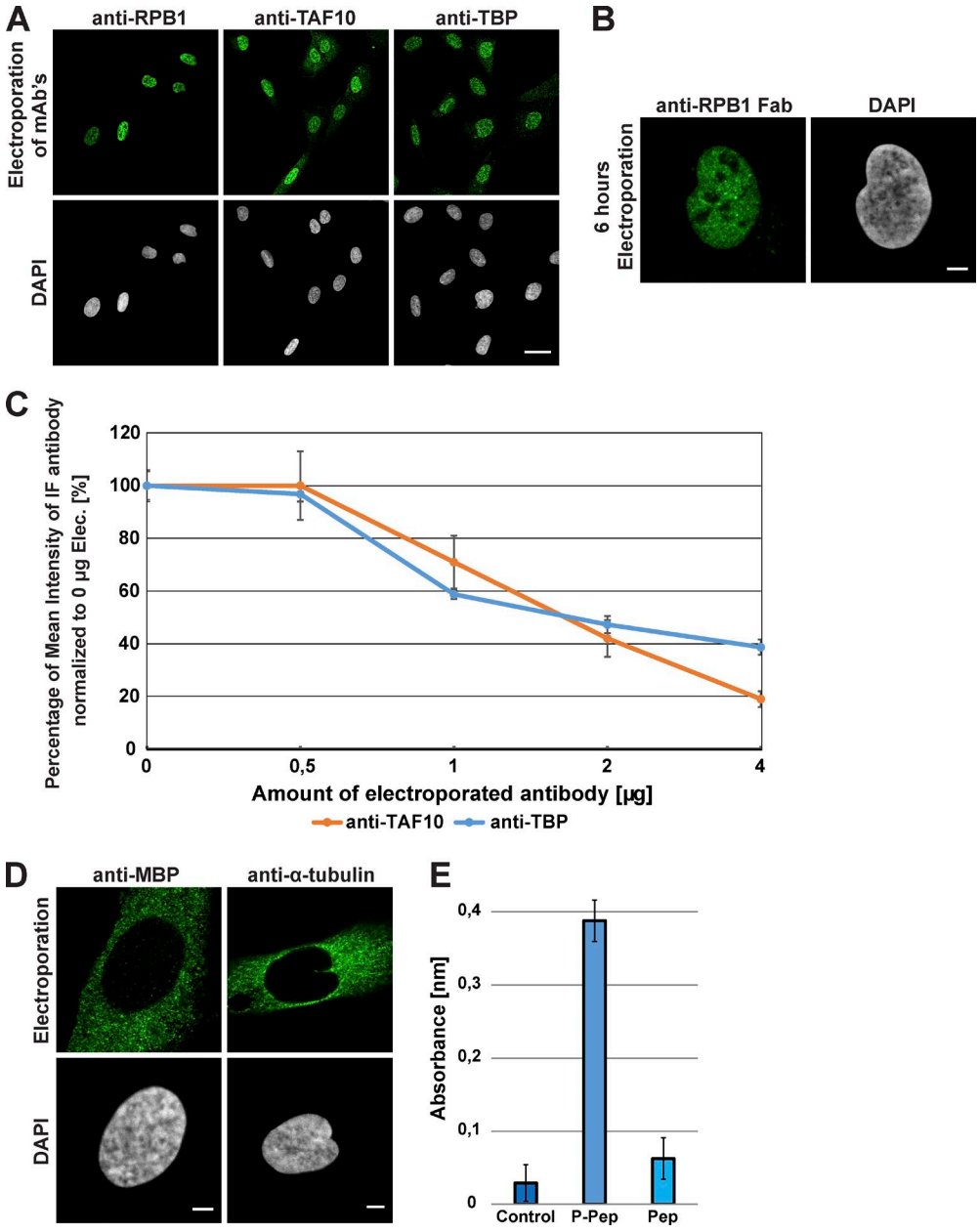


Figure S1. **Efficiency, localization, target binding, and affinity of different antibodies by using VANIMA. (A)** The Alexa Fluor 488–labeled mAbs binding specifically to the transcription factors RPB1, TAF10, and TBP were transduced in U2OS cells, and their localization in the cells was monitored 24 h after treatment (see also Fig. 2 A). Bar, 30 µm. **(B)** The Alexa Fluor 488–labeled Fab fragment against RPB1 was transduced in U2OS cells and monitored by confocal microscopy 6 h after electroporation. Bar, 5 µm. **(C)** Quantification of the competition assay shown in Fig. 2 C. Alexa Fluor 488–labeled antibodies against TAF10 (anti-TAF10) or TBP (anti-TBP) were transduced into U2OS cells in increasing amounts. To verify binding of the antibodies to their target, a competition assay was performed afterward by adding a constant amount (2 µg) of the same antibody, but Alexa Fluor 568–labeled as IF antibody, after fixation. The graph shows the mean fluorescence intensity of the nuclei labeled with the IF antibody for each condition. The anti-TAF10 and anti-TBP measurements were done 24 or 48 h after transduction, respectively. The percentage of mean intensity was normalized to the 0-µg transduction. Error bars represent the SD obtained with 10 recorded cells for each condition. **(D)** Antibodies against MBP or α-tubulin were transduced into U2OS cells, and their localization in the cells was monitored 24 h after electroporation. Bar, 5 µm. **(E)** Analysis of the binding specificity of the anti-γH2AX Fab by ELISA. The ELISA plate was coated with either the nonphosphorylated (Pep) or the phosphorylated peptide (P-Pep) corresponding to the C-terminal end of H2AX (see Materials and methods). After addition of 1 µg/ml anti-γH2AX Fab and subsequent washing, bound Fab was revealed with HRP-labeled anti-mouse immunoglobulins. The histogram shows the mean value of the absorbance measured in several wells after hydrolysis of the HRP substrate and error bars indicate the SD of the measurements. Control, no coating. All images in this figure are showing typical nuclei after fixation and counterstaining with DAPI as well as single z sections.

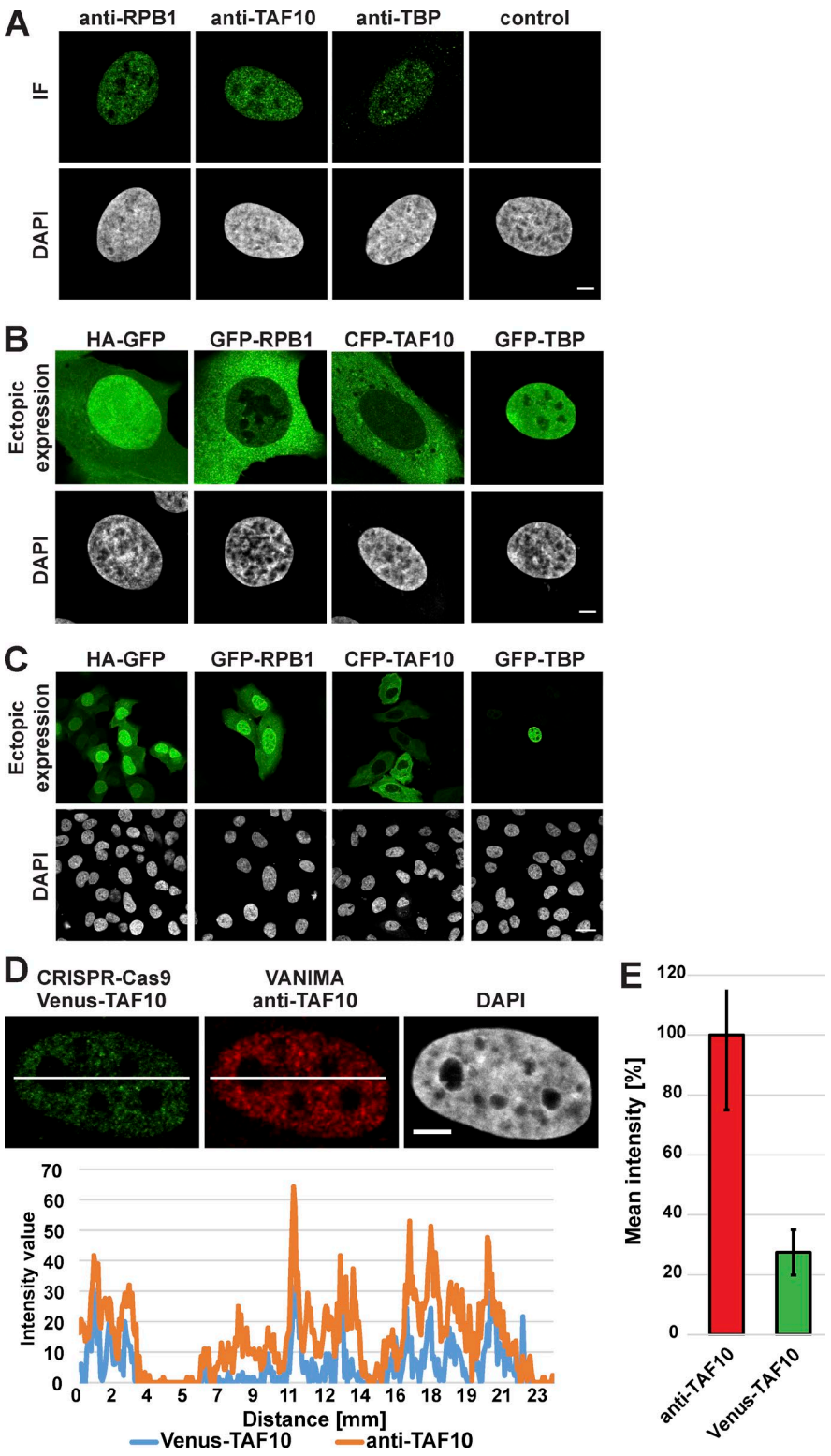
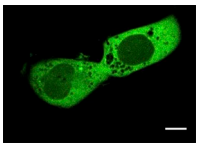
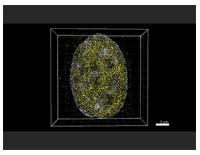


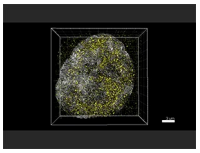
Figure S2. **Imaging transcription factors with classical IF or after genetic tagging with fluorescent tags by using confocal microscopy.** All images show single z sections of the nuclei. **(A)** The endogenous transcription factors were detected in fixed U2OS cells with the indicated antibodies by classical IF. Control: Cells treated only with Alexa Fluor 488–labeled secondary antibodies. Bar, 5 µm. **(B)** The indicated fluorescently tagged transcription factors were visualized 48 h after transfection with the plasmids expressing the indicated fusion proteins. Bar, 5 µm. **(C)** Same as in B, but cells were observed with lower magnification. Bar, 30 µm. **(D)** Detection of electroporated Alexa Fluor 568–labeled anti-TAF10 mAb (red) in CRISPR/Cas9-modified U2OS cells stably expressing Venus-TAF10 (green). The area taken for the intensity profile measurements is indicated by white lines in the nuclei corresponding to the Venus-TAF10 and anti-TAF10 images. An intensity profile of the Venus-TAF10 (blue) and anti-TAF10 Alexa Fluor 568 (orange) signal is shown under the confocal images. Bar, 5 µm. **(E)** Quantification of the mean intensity of nuclei either transduced with Alexa Fluor 488–labeled anti-TAF10 or expressing Venus-TAF10. The percentage of mean intensity was normalized to the anti-TAF10 transduced sample. Error bars represent the SD obtained with 10 recorded cells for each condition. All micrographs correspond to typical nuclei observed in each case after counterstaining with DAPI.



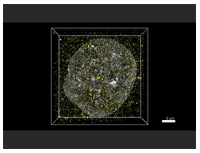
Video 1. **Transport of the labeled anti-RPB1 mAb from the cytoplasm into the nucleus of living cells (see Fig. 1 A).** U2OS cells were transduced with Alexa Fluor 488–labeled anti-RPB1 antibody and incubated for 6 h before starting image acquisition. Imaging was performed on a confocal microscope focusing on one single z plane and by taking one image every hour. Total time of analysis: 20 h. Bar, 15 µm.



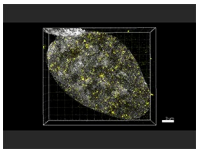
Video 2. **Analysis of a nucleus of U2OS cells transduced with the labeled anti-RPB1 mAb (yellow) by 3D-SIM super-resolution microscopy (see Fig. 4 A).** Images were taken 24 h after transduction and correspond to a full z stack of the whole nucleus. The video represents a typical nucleus recorded after fixation of the cells and subsequent counterstaining with DAPI (gray). Bar, 3 µm.



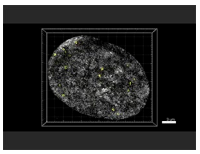
Video 3. **Analysis of a nucleus of U2OS cells transduced with the labeled anti-TAF10 mAb (yellow) and counterstaining with DAPI (gray) by 3D-SIM microscopy (see Fig. 4 A).** Images were taken as indicated in the legend of Video 2. Bar, 3 µm.



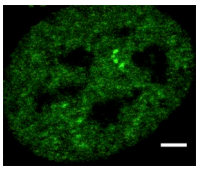
Video 4. **Analysis of a nucleus of U2OS cells transduced with the labeled anti-TBP mAb (yellow) and counterstaining with DAPI (gray) by 3D-SIM microscopy (see Fig. 4 A).** Images were taken as indicated in the legend of Video 2. Bar, 2 µm.



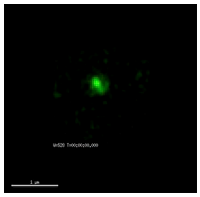
Video 5. **Analysis of U2OS nuclei after transduction with labeled anti-γH2AX Fab (yellow) in the presence of HU treatment by 3D-SIM microscopy (see Fig. 6 A).** The images correspond to a full z stack of a typical nucleus recorded in each case after cell fixation and counterstaining with DAPI (gray). Bar, 3 µm.



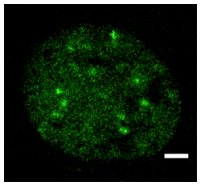
Video 6. **Analysis of U2OS nuclei after transduction with labeled anti-γH2AX Fab (yellow) in the absence of HU treatment by 3D-SIM microscopy (see Fig. 6 A).** The images correspond to a full z stack of a typical nucleus recorded in each case after cell fixation and counterstaining with DAPI (gray). Bar, 3 µm.



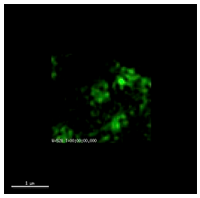
Video 7. **Live-cell imaging of RNA Pol II after transduction of labeled anti-RPB1 mAb (see Fig. 7 A).** 24 h after electroporation in the presence of Alexa Fluor 488–labeled anti-RPB1 mAb, the U2OS cells were analyzed by confocal microscopy. The nuclei were imaged by focusing on one z section and over a period of 2.5 h. The pictures were taken every 10 min. Bar, 3 µm.



Video 8. **3D-SIM live-cell imaging of one distinct RNA Pol II cluster after transduction of labeled anti-RPB1 mAb into U2OS cells (see Fig. 7 B).** The dynamics of a RNA Pol II cluster were analyzed 24 h after electroporation with Alexa Fluor 488–labeled anti-RPB1 mAb by using 3D-SIM microscopy. The video shown is a maximum-intensity projection of a 1-µm z stack. The nuclei were imaged over a period of 45 s, and pictures were taken every 4.1 s. Bar, 1 µm.



Video 9. **Confocal live-cell imaging of γH2AX foci (see Fig. 7 C).** 24 h after electroporation, U2OS cells transduced with Alexa Fluor 488–labeled anti-γH2AX Fab were analyzed by confocal microscopy after treatment with NCS. The nuclei were imaged by focusing on one z section over a period of 4 h. Pictures were taken every 10 min. Bar, 5 µm.



Video 10. **3D-SIM live-cell imaging of a γH2AX foci by using VANIMA (see Fig. 7 D).** U2OS cells transduced with Alexa Fluor 488–labeled anti-γH2AX Fab were treated with NCS and analyzed by live-cell imaging by using 3D-SIM microscopy. The first time point was acquired 10 min after drug treatment. The video shown corresponds to a maximum-intensity projection of a 1-µm z stack. The nuclei were imaged over a period of 45 s, and pictures were taken every 15 s. Bar, 1 µm.

Table S1. **Efficiency of anti-RPB1 mAb delivery**

Cell line	Efficiency	Viability
	%	%
U2OS	99	92
HFF-1	99	99
mES cells	97	56
S2 cells	94	71

The indicated cell lines were electroporated in the presence of the Alexa Fluor 488–labelled anti-RPB1 mAb (2 µg). The efficiency (%) was calculated by counting 100 cells and determined the percentage of cells showing a positive nuclear staining 24 h after transduction. The percentage of viability corresponds to the number of live cells after the electric treatment normalized to the number of living cells in the electroporation mixture. All counts were performed in the presence of Trypan blue. HFF-1, human foreskin fibroblast cells; mES, mouse embryonic stem.

Table S2. **Validated primer pairs used for the quantification of Pol II pre-mRNA as well as Pol I (*RN18S*) and Pol III (*RN7SK, RPPH1*)**

Gene	Forward primer (5′–3′)	Reverse primer (5′–3′)
CCT4	AGAGCACTGACTGATACCAACAGA	AGACACTAAAAGCAACTTGTGCTG
EEF2	CGACTCTTCACTGACCGTCTC	TGTGTGTAAGGTCACCTCTTTCTC
EIF3L	CTGGATGGTGAATTTCAGTCAGC	AACACTTAATACAAGACCCCAAGC
GAPDH	CTCACATATTCTGGAGGAGCCTC	TTACCAGAGTTAAAAGCAGCCCT
GMPS	GGAGAGAGGGCATAGACCTTGT	AGCATACACAGAATTAGGTCCTCC
GNB1	ATCTCCAGTGTGTCCGGTAAAC	ACCCAAGAAGTTAAGGCTGATGTC
MYBL2	CAGGTGGATGTGAAGGGCTATG	TGTGCCATACTTCTTAACCAGCT
PKM	CAAAGCTTTCCGTGGCTGTG	GAGCTGGATTCTAGTGTGGGAG
RBBP5	AGACAATGCTCCCAATGTGTC	AGGTTTACCTCTGGAAGGATCAG
RN18S	AAACGGCTACCACATCCAAG	GGCCTCGAAAGAGTCCTGTA
RN7SK	CGGTCAAGGGTATACGAGTAGC	TTGGAAGCTTGACTACCCTACG
RPL8	ACGATTGTACCCTCAGGCATG	CGCATTGTTTCTTACTGTGCTGA
RPPH1	GGCGGAGGAGAGTAGTCTGAAT	CGGAGCTTGGAACAGACTCA
RPS18	CCTTATCGGCCTTACTGTTTGAT	AAATATGCTGGAACTTTTCAGGG
SF3B2	CACCTGTATCTTTTGTTTCCGCTT	CAGTGAAGAGCTGAGGTGTCTC
TPM2	AAATGGGATGAGAAGGTACAGGAC	GGAGAAAACCATCGATGACCTAGA

Transcripts shown in Fig. 3 A.

4. Imaging of RNA Pol II recruitment dynamics in single living cells

The following section contains preliminary and unpublished data of experiments which were performed and designed by myself and Emmanouela Vlachou-Portari, a Master student who I supervised. The results shown in Figure 28C were obtained with the help of Nacho Molina.

The dynamics of RNA Pol II transcription is a highly studied field and numerous studies performed in the past attempted to measure the kinetics of RNA Pol II in living cells. Labeling of RNA Pol II was performed either by overexpressing a fluorescently-tagged version (Kimura et al., 2002) or by replacing the endogenous version of RPB1 with a α-amanitin resistant overexpressed version to get closer to the endogenous situation (Darzacq et al., 2007). These studies suggested that around 25% of all RNA Pol II molecules are immobile and therefore bound to chromatin whereas the rest is diffusing rapidly through the nucleus. In addition, it was indicated that transcription initiation is the rate limiting step of the transcription cycle whereas elongation is faster with an elongation rate of 4.3 kb/minute. However, all these examples missed a visible reference point on the genome to ensure that slowed down kinetics of RNA Pol II really correspond to a chromatin bound fraction and showed not only decreased dynamics due to other factors induced by the tightly packed nuclear environment. Additionally, even if some studies tried to follow endogenous RNA Pol II using the CRISPR/Cas9 technology to introduce a fluorescent tag into the endogenous locus of RPB1, information about the dynamics of real endogenous RNA Pol II is still highly limited (Cisse et al., 2013).

Therefore, our attempt to analyze the dynamics of endogenous RNA Pol II in single living cells was to set up an imaging system combining our endogenous labeling strategy using VANIMA with a specific cell line (U2OS 2-6-3) which allows the labeling of the genomic locus and specific transcriptional activation using a gene array (see Introduction section 2.3.2 for more information) (Janicki et al., 2004; Rafalska-Metcalf et al., 2010). The specific questions for these preliminary experiments and in accordance with **aim d)** of this thesis were:

- Does the combination of VANIMA with genetic labeling work and can we see and accumulation of RNA Pol II at the gene array?

- How fast is endogenous RNA Pol II recruited to the genomic locus after transcription is activated?
- How fast is the activator or other factors of the transcription machinery like GTFs recruited to the gene array and can we detect any delays in the recruitment time of different factors?

However, to answer all these questions, the first step was to test the combination of VANIMA and gene array labeling in the U2OS 2-6-3 cells. Therefore, a stepwise electroporation procedure was implemented to import the required vectors for gene array labeling and transcriptional activation as well as the fluorescently labeled anti-RPB1 antibody into the cells. The first electroporation included the two vectors, one coding for the LacI protein fused with the fluorescent marker mCherry to visualize the gene array through the binding to the lac-operon (LacO) and the other for the reverse Tetracycline-controlled transactivator (rtTA) which after addition of Doxycycline (Dox) will enter the nucleus and activate transcription of the gene controlled by a CMV promoter (Figure 27A). After an incubation of 24 hours, the cells were electroporated again with the labeled anti-RPB1 antibody for VANIMA staining of RNA Pol II. Another 24 hours later, the cells were either left untreated, or treated with Dox for 2.5 h to induce transcription. The cells were fixed and analyzed either by confocal microscopy (Figure 27B) or 3D-SIM microscopy (Figure 27C) to test if the induced transcription led to an accumulation of endogenous RNA Pol II at the gene array. After transcription activation the labeled gene array decondensed as it was already shown in previous studies (Janicki et al., 2004). Furthermore, VANIMA labeled endogenous RNA Pol II was co-localizing with the labeled gene array after transcription activation in contrast to the control cells without addition of Dox (Figure 27B and 27C). These results showed that VANIMA can be used to study the recruitment of RNA Pol II at the gene array.

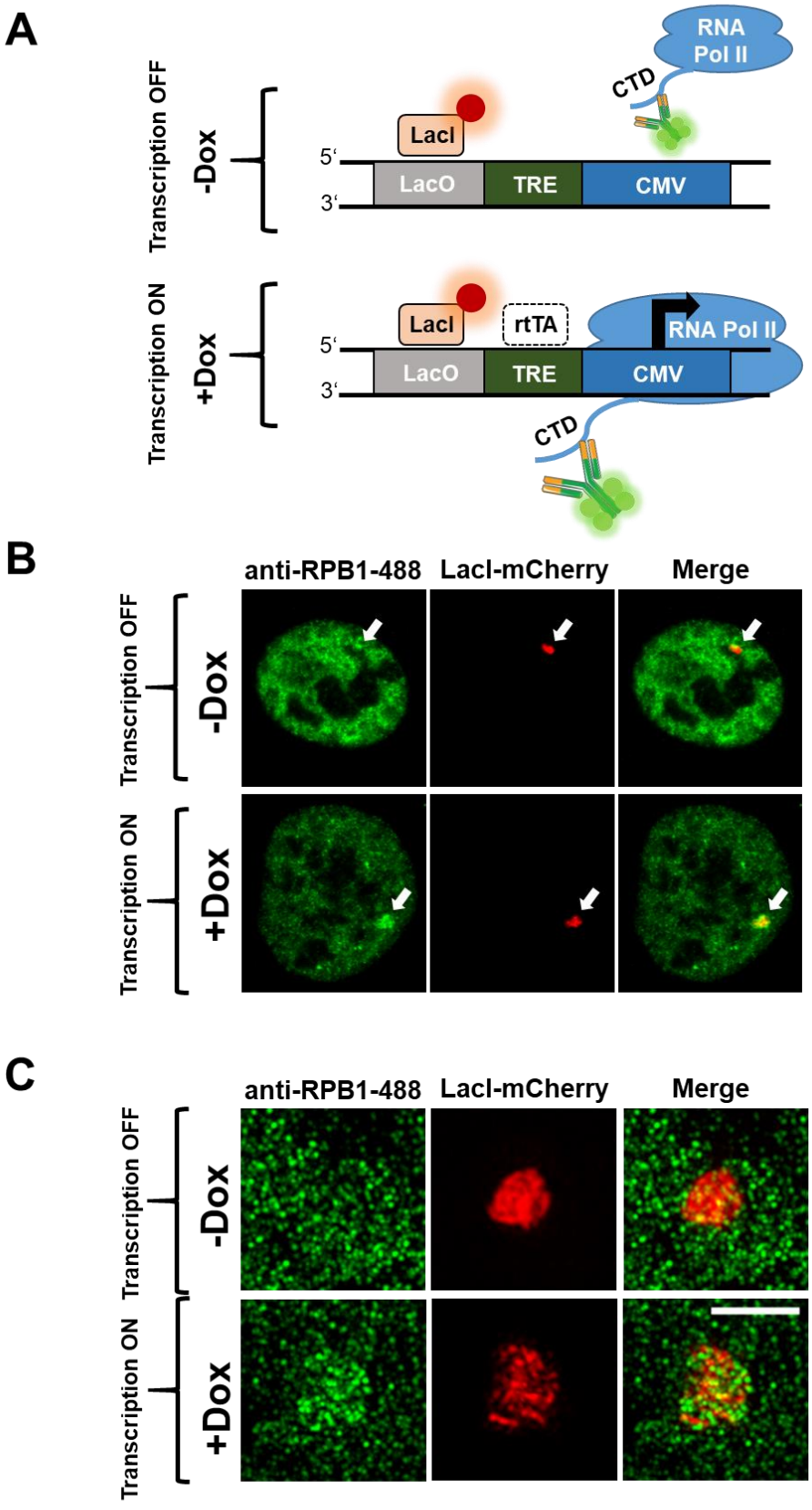


Figure 27: Legend on the next page →

Figure 27: Combination of VANIMA with an inducible and fluorescently labeled gene array. (A) *Schematic representation of the doxycycline inducible system including the lac-operator (LacO) the Tet-responsible element (TRE) and a CMV promoter as well as the elements bound to the gene array before and after transcription induction with doxycycline (+dox) or not (-dox). The mCherry-fused LacI protein (LacI) will bind and mark the gene array and the activator (rtTa) will only bind in the presence of doxycycline. The cells were electroporated with fluorescently labeled anti-RPB1 mAbs to label RNA Pol II (RNA Pol II) which will bind after transcription activation.* ***(B)*** *U2OS 2-6-3 cells were electroporated with two plasmids coding for the mCherry labeled LacI protein (red) and the activator (rtTa) as well as Alexa488 labeled anti-RPB1 mAbs (green). After 24 h, the cells were treated with doxycycline for 2.5 h, fixed and imaged using confocal microscopy. White arrows indicate the gene array and the accumulation of endogenous RNA Pol II which happens only after transcription activation. Scale bar: 5 µm.* ***(C)*** *Same treatment as in (B) but the cells were analyzed using 3D-SIM microscopy. The gene array visualized through the mCherry-labeled LacI and the VANIMA-labeled endogenous RNA Pol II are shown in red and green, respectively. Scale bar: 2 µm.*

However, for the live imaging experiments we changed to a different system in which we labeled the gene array through a fluorescently tagged transactivator instead of using the LacI-mCherry and the lac-operon. The new transactivator was a TetOFF-based activator fused to mCherry and the estrogen receptor (mCherry-tTA-ER) which stays cytoplasmic after expression but re-localizes into the nucleus after addition of tamoxifen to the cell medium to bind and activate transcription of the gene array (Figure 28A) (Rafalska-Metcalf et al., 2010). This system harbored several advantages as the localization of cells successfully electroporated with the transactivator plasmid could be defined easily under the microscope due to the cytoplasmic staining and it also allowed to measure not only the recruitment of RNA Pol II but also of the transcriptional activator to the gene array within the same acquisition. Therefore, as a first experiment we performed 2D confocal time-lapse imaging with electroporated cells for 3 h with images taken every 5 min and tamoxifen induction at time point zero to define the timespan in which transcription is activated and RNA Pol II is accumulating at the gene array in living cells (Figure 28B). The results showed that the activator is accumulating at the gene array around 5-10 min after tamoxifen addition which is in accordance with previous studies (Rafalska-Metcalf et al., 2010). The time-lapse also indicated that

RNA Pol II is starting to accumulate at the same time as the activator. In order to further define the recruitment of RNA Pol II and the activator to the gene array we quantified the fluorescent intensities fluctuations in both channels over time using several cells from the time-lapse. In more detail, the spot formed by the activator was tracked and the xy coordinates were determined. After the subtraction of the background fluorescence, the normalization of the fluorescence intensity of the spot to the maximum fluorescence and correction of bleaching-derived loss of fluorescence, the change in fluorescence intensity for the activator and RNA Pol II over time was obtained (Figure 28C). Only the quantification of two representative cells is shown. Interestingly, for both cases an accumulation of activator and RNA Pol II started at the same time with no delay visible indicating that RNA Pol II recruitment occurs very fast after the activator is binding. However, a very high cell to cell viability was observed concerning the time when accumulation of both molecules starts and how the fluctuations of the RNA Pol II signal is occurring. It is important to mention that these are very preliminary results and more acquisitions of many more cells and time-lapses with shorter timeframes within the time of RNA Pol II accumulation need to be performed to obtain statistically significant results concerning the recruitment of RNA Pol II after transcription activation. Nevertheless, these first results already showed that this new imaging system could be a promising tool to study endogenous RNA Pol II transcription in living cells.

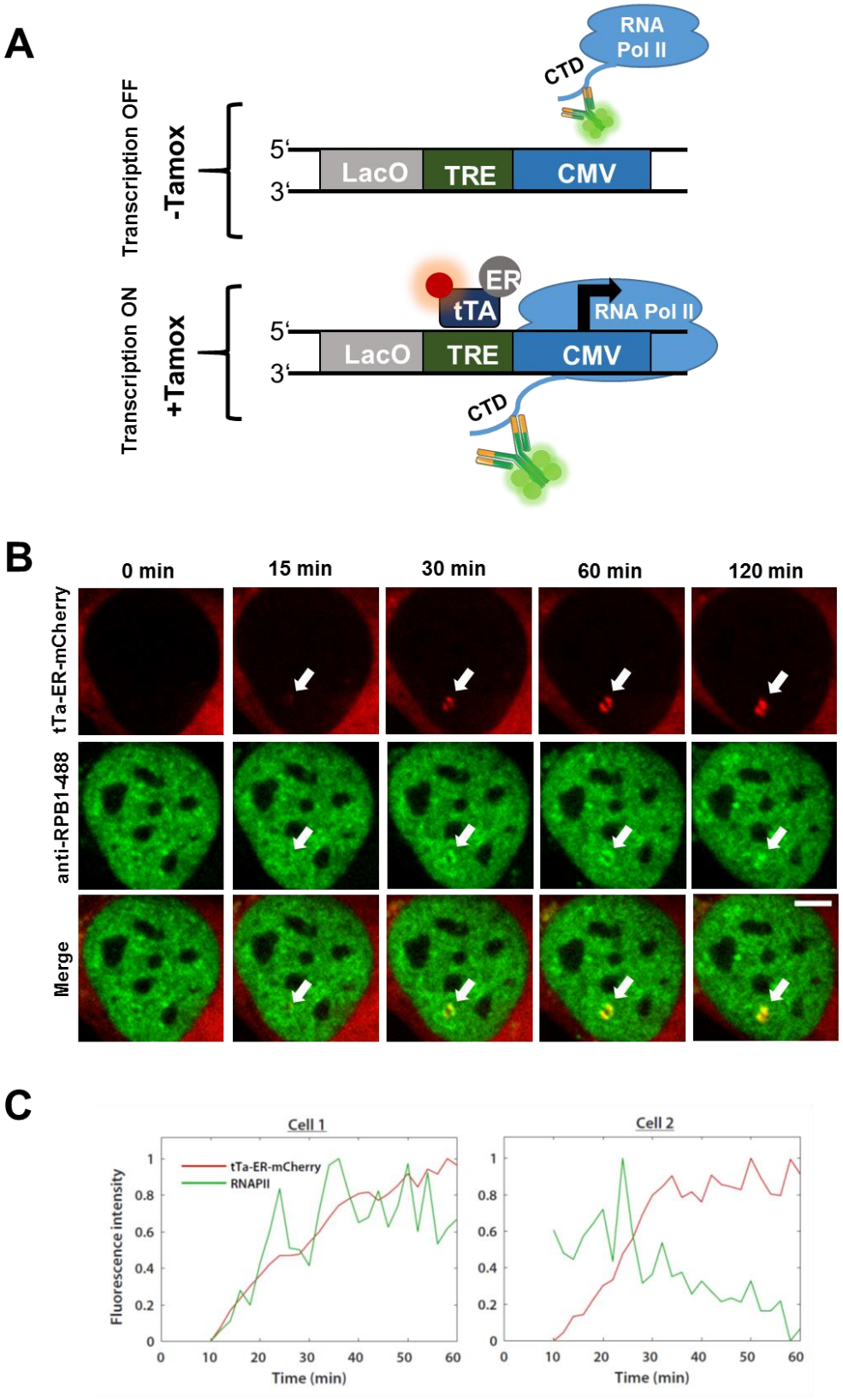


Figure 28: Legend on the next page →

Figure 28: **Tamoxifen inducible system and analysis of endogenous RNA Pol II recruitment in living cells. (A)** Schematic representation of the tamoxifen inducible system including the lac-operator (LacO) the Tet-responsible element (TRE) and a CMV promoter as well as the elements bound to the gene array before and after transcription induction with tamoxifen (+tamox) or not (-tamox). The mCherry-fused activator (tTa-ER) will bind and mark DNA only in presence of tamoxifen. The cells were electroporated with fluorescently labeled anti-RPB1 mAbs to label RNA Pol II (RNA Pol II) which will bind after transcription activation. **(B)** U2OS 2-6-3 cells were electroporated with the plasmid coding for the mCherry labeled tTa-ER activator and Alexa488 labeled anti-RPB1 mAbs. After 24 h, the cells were treated with tamoxifen and imaged using confocal time-lapse microscopy for 3 h with one image every 15 min (not all time points are shown). White arrows indicate the accumulation of the activator or RNA Pol II at the gene array. Scale bar: 5 µm. **(C)** Fluorescence fluctuations were measured at the gene array spot of the tTa-ER-mCherry activator and anti-RPB1-Alexa488 (RNAPII) for 1 h post-induction with tamoxifen. The maximum fluorescence intensity was set to 1. The transcriptional activator and the RNAPII intensity profiles are depicted in red and green, respectively. Two representative cells are shown to indicate the observed high cell to cell variability.

GENERAL DISCUSSION & PERSPECTIVES

Discussion and Perspectives

1. General discussion of the thesis project and summary of the results

The initial idea for this thesis project was to implement a new labeling strategy to study the assembly and dynamics of endogenous RNA Pol II transcription by using electroporation to transport fluorescently labeled antibodies into living cells. It was already shown that this delivery method harbors a very high delivery efficiency by keeping also a high cell viability (Freund et al., 2013). However, it is true that the project missed a specific scientific question about which aspects or proteins in detail will be analyzed using the new imaging approach and I would like to take the opportunity to explain why. The protocol of the new labeling technique needed to be developed and validated first to ensure that reliable results can be obtained concerning transcription dynamics in living cells. Additionally, many different antibodies were already available at the beginning against several targets of the transcription machinery and it was important to test first which of these antibodies could be used for the new technique and for imaging in general. For me, there are mainly two types of scientific projects. The first type of project in which a specific question is formulated at the beginning and afterwards already established techniques are used to answer the question. On the other side, there are the projects which have an idea for a new technique that can potentially revolutionize the field and give a new tool to answer questions that will be specified as soon as the new technique is validated. Therefore, the main aspect of this thesis project was to implement the new imaging approach and to test which antibodies can be used to be able to study the assembly and dynamics of specific factors of the transcription machinery. Thus, the main question for this project can be postulated as if it is possible to use electroporation to introduce antibodies or Fab fragments into living cells to label endogenous proteins or PTMs and answer important questions concerning transcription machinery assembly and dynamics?

Nevertheless, I achieved the goal to implement, validate and publish a labeling technique to track endogenous proteins and was also able to tackle some of the transcription-related aims of the project. In the first part of the results, I worked on the

transcription complex assembly aspect of my aims by joining a project to uncover the presence of a cytoplasmic TFIID sub-module consisting of TAF2-TAF8-TAF10 in the cytoplasm and nucleus of human cells. Furthermore, the second and third part of the results contained the main work of this thesis about intracellular antibody targeting of endogenous proteins or PTMs to either inhibit processes like replication using inhibiting antibodies or in form of the new technique VANIMA to use electroporated non-inhibiting immunoglobulins to visualize factors in living cells for the fluorescent imaging of transcription. In the last chapter of the results, to further develop the transcription dynamics aim of the thesis, the newly developed imaging technique was combined with other labeling strategies to obtain preliminary results about RNA Pol II recruitment dynamics in living cells. This serves now as a good starting point to use VANIMA to uncover new insights about transcription dynamics in living cells (see also chapter 3 of the *Discussion*).

Altogether, the results demonstrated that:

- A TFIID sub-module consisting of TAF2-TAF8-TAF10 exists in the cytoplasm of living human cells which controls the import of TAF2 into the nucleus.
- Transduced anti-PCNA or anti-DNA Polα antibodies and Fabs can inhibit replication in various cancer cell types and induce cell death.
- These intracellular inhibiting antibodies are inducing a huge amount of DNA damage through replication stress.
- Non-inhibiting antibodies or Fabs can be fluorescently labeled and transduced into living cells to label and track endogenous proteins and PTMs.
- This new technique was called VANIMA for versatile antibody-based imaging approach (which also means “beautiful” and “fair” in elvish language) and was used to label and track RNA Pol II, different transcription factors and histone modifications in form of γH2AX.
- VANIMA can be combined with other labeling strategies like an inducible gene array to study the dynamics of RNA Pol II recruitment after transcription activation.

Importantly, a more detailed discussion can be found also in the *Results* section where specific results are discussed hand in hand with the current knowledge. The following more general discussion will be divided into two parts. The first part will discuss only the method VANIMA and its advantages as well as disadvantages.

Furthermore, I will give some ideas of how VANIMA could be improved even further and how it could be combined with different imaging techniques. The second chapter will concentrate on more specific scientific questions concerning transcription assembly and dynamics that can be tackled now after the imaging approach is implemented. This includes also a follow-up discussion about the preliminary results of chapter four of the *Results* section.

2. Intracellular antibodies and VANIMA: Past, present and future

Antibodies are known to be one of the most important tools in scientific research. Their ability to bind proteins or even protein modifications with high specificity and affinity is used in a plethora of different *in vitro* and *in vivo* techniques including ChIP, IF or Western blot analysis. However, their application in living cells was always difficult due to their big size of around 150 kDa and the resulting problems in the delivery into living cells through the cell membrane (Marschall et al., 2011). Several delivery methods were developed to overcome this obstacle and to transport antibodies into living cells including microinjection, osmotic lysis of pinocytic vesicles, protein transfection methods by using various cationic lipids/polymers or the loading with glass beads (Courtête et al., 2007; Hayashi-Takanaka et al., 2009; Röder et al., 2017). However, microinjection has the problem that only a subset of cells can be treated with the antibody within a short time and other methods like the loading with glass beads or the recently published technique using a bacterial toxin (streptolysin O) to open the membrane suffer from low delivery efficiencies or low viability of the cells due to toxicity of the treatments (Teng et al., 2016). Nevertheless, all these techniques showed that as soon as the antibodies or Fabs are inside the cells, they can be used to target and inhibit specific factors or, by using non-inhibiting antibodies, can give new insights about the behavior of proteins and protein modifications like specific histone modifications (Hayashi-Takanaka et al., 2011).

Another method to deliver molecules into living cells is electroporation. This technique, in which an electric shock is used to open pores in the membrane of living

cells for a short time, was already shown to be very effective to deliver proteins like antibodies (Berglund and Starkey, 1989; Chakrabarti et al., 1989; Marschall et al., 2014). Thus, with this technique it was, for example, possible to introduce inhibiting monoclonal antibodies against cyclin D1 into living cells to specifically inhibit the transition from G1 to S-phase of the cell cycle (Lukas et al., 1994). Furthermore, electroporation was able to conquer some of the problems of over delivery techniques as it was shown that transduced cells keep a high cell viability after the electric shock but also were harboring a high delivery efficiency of over 90% (Freund et al., 2013). This is why I decided to use electroporation as the delivery method for the new labeling technique VANIMA. However, what are the pros and cons of using VANIMA and how could it be improved to extend its application?

2.1 VANIMA: “beautiful” and “fair” but not perfect

No scientific technique is perfect, each method has its strengths and weaknesses which is the case for fluorescent labeling strategies as well. In the following section, I want to critically analyze and summarize the advantages and disadvantages of VANIMA but also compare the new technique with already existing labeling strategies.

The most obvious advantage of VANIMA is that it enables the labeling and tracking of endogenous proteins and PTMs in living cells in contrast to exogenously produced fusion proteins. Moreover, the delivery using electroporation is highly efficient in contrast to previous methods used for intracellular antibody delivery. Furthermore, tens of thousands of full length antibodies which can potentially be used for VANIMA have been already developed over the past decades and are commercially available. Another advantage is the possibility to electroporate several different antibodies or Fabs which can be labeled with different colors to perform multicolor and therefore multi-target imaging in living cells. This can be used to perform co-localization studies or tracking experiments with two to three different targets. Another advantage of VANIMA is that the technique was tested in several different cell lines and in all of them, the desired staining of the target protein could be detected.

On the other side, VANIMA also has some important disadvantages. One of the biggest issues for VANIMA is definitely the requirement of a validated antibody or Fab. This means that the antibody needs to be validated and tested before using it for

VANIMA to ensure that it can specifically bind to the target protein and that its epitope of the antigen is accessible in the living cells. Furthermore, it is important to test that the antibody does not affect the function of the target protein if the goal is to perform imaging and not inhibition of cellular processes. This problem can be tackled in a certain extent by using Fabs in contrast to full length antibodies which due to their smaller size can already reduce the inhibiting effects of steric hindrances. However, it is important to notice that none of the tested full length antibodies showed any inhibiting effect even if they are quite big and have bivalent binding properties.

Another drawback is the labeling efficiency of the antibodies or Fabs which can vary depending on which labeling strategy and antibody is used. In the current protocol, the labeling was performed using fluorophores conjugated with a succinimidyl-ester which react randomly with primary amines on the antibodies or Fabs. However, this labeling strategy can induce variations in the labeling efficiency depending on how many primary amines are present in the primary sequence of the antibodies and Fabs. Furthermore, it can happen that labeling sites are present in the target recognition sites of the antibodies or Fabs which can lower the affinity for the antigen or even destroys completely the interaction of the antibody with the target. Therefore, more site specific labeling strategies would be necessary to keep the integrity of the antibodies or Fabs and to ensure a reproducible labeling efficiency (see also the following section 2.2 of the *Discussion* for some suggestions).

Furthermore, it was not tested until now if VANIMA can be applied on tissues as well. There would be a need for other new delivery methods for antibodies into tissues as electroporation is probably not efficient enough to transport the labeled molecules into every cell of a tissue. However, there are already some new delivery methods which could improve the transport into tissues, for example nickel-immobilized polymers (Postupalenko et al., 2015).

Lastly, another disadvantage of VANIMA is that the electroporated antibody or Fab will be diluted over time by either degradation by the cell or simply by cell division. This hinders the possibility to perform very long acquisitions over several days as the signal will decline drastically around three days after electroporation when using cancer cell lines with a doubling rate of 24 h and can be even lower when cell lines with a faster doubling time are used. Therefore, acquisitions should be performed quickly after

electroporation, around 6 hours to a maximum of three days after transduction depending on the antibody or Fab and the cell line that was used.

In general, VANIMA can be seen as a combination of IF and the endogenous labeling using CRISPR/Cas9 knock-ins as the technique uses antibodies or Fabs to label endogenous proteins. However, VANIMA has some important advantages over the mentioned techniques. First, in contrast to IF, VANIMA can be applied on living cells and therefore no chemically fixation or permeabilization steps are needed to visualize the target protein. This erases the possibility of artifacts that could be induced through fixation-related processes.

On the other hand, VANIMA has also advantages over the knock-in of fluorescent tags into the endogenous locus of proteins. As mentioned before, thousands of antibodies are already existing which could be used to perform VANIMA and the protocol is easy and fast to perform in contrast with knock-in experiments using CRISPR/Cas9 which can take several months to obtain a homozygous knock-in clone. Especially for cancer cell lines which are highly used in scientific research it can be quite difficult as they often are hypertriploid. Additionally, the chosen fluorescent tag that was knocked-in cannot be easily changed afterwards, whereas for VANIMA the color of the labeling can be exchanged just by conjugating the antibodies the next time with a new dye. Moreover, VANIMA can be used to label several targets in the same cell which is very difficult when using the knock-in strategy as it means that several single knock-in need to be performed using the same clone over and over. However, the most important advantage of VANIMA over CRISPR/Cas9 knock-ins that VANIMA can label specific PTMs in the cells. In conclusion, VANIMA is a very powerful tool to perform imaging in living cells but there is still room for improvements which will be discussed in the following section.

2.2 VANIMA: How can it become even more “beautiful”?

Some of the drawbacks mentioned in the previous section could be avoided by further developing VANIMA. Therefore, in the following section, I would like to give some ideas of how VANIMA could be improved by, for example, using nanobodies in contrast to full length antibodies or by introducing more site specific labeling strategies.

One possibility for improvement would be to use either single-chain variable fragments (scFv; also called intrabodies or mintbodies) or single-domain antibody fragments (VHHs or nanobodies) for VANIMA. Both types of antibodies have the advantage that they are much smaller than full length antibodies or even Fabs (scFvs with around 28 kDa and VHHs with around 15 kDa). Therefore, they can freely diffuse into the nucleus, just like Fabs, but also decrease the overall size of the antibody-endogenous protein complex which could be advantageous for live tracking experiments. Mintbodies were already used successfully to image specific histone modifications during zebrafish development just by transgenic expression of the scFv inside the embryo (Sato et al., 2013). However, scFv often have a tendency to become insoluble and aggregate in the reducing environment of the cytosol and are therefore difficult to handle (Renaud et al., 2017). Thus, the better choice would be to use single-domain antibodies from camelids or sharks. These VHHs are functional antibodies consisting of only a single heavy chain (Hamers-Casterman et al., 1993). Therefore, they miss any form of disulfide bridges which could be reduced in the cytoplasm to induce denaturation of the molecule. In brief, after immunization of the camelid with the desired antigen or peptide, the VHH producing lymphocytes can be isolated from the serum and their mRNA can be isolated to produce a cDNA bank which can be cloned in phagemid vectors and transformed into *E. coli*. Furthermore, the antigen positive VHHs are selected using phage display and the resulting VHH can be produced in huge amounts in *E. coli* or the production vector for the nanobody can be transfected into living cells to produce the VHH directly inside the desired cell line (Harmsen and Haard, 2007; Schoonooghe et al., 2012). Another advantage of VHHs is that their expression vector can be genetically modified easily to add any tag as for example fluorescent proteins like GFP or other tags that could be used to add fluorescent labels as it will be discussed in the next paragraph. However it is important to mention that VHHs have also their problems. The ability to produce recombinant VHHs in living cells can lead to a high signal intensity but also inherits the drawbacks of exogenous overexpression-based systems. It is difficult to control the expression level of the VHH in the cell and this can result in a high background signal as well as a high amount of unbound VHHs which can falsify potentially the imaging data. The use of conventional antibodies or Fabs on the other side can be controlled specifically to ensure that only target bound antibodies are present in the cell. Therefore, the best strategy, in my opinion, would be to produce the VHHs beforehand in *E. coli* with a

purification tag (like His-tag) to be able to purify, fluorescently label and afterwards electroporate them into living cells. We are currently producing several new VHHs against different subunits of TFIID and specific PTMs of RNA Pol II in form of phosphorylation of Ser2 and Ser5 of the CTD to test them for VANIMA.

Thus, another problem of the current version of VANIMA is the aforementioned variations in the labeling efficiency when using random labeling strategies and full length antibodies or Fab fragments. Therefore, it would be advantageous to introduce a site specific labeling to ensure that the integrity of the antibody is kept and that always the same amount of dyes are attached. One possibility would be to use the thiol group of cysteine residues which are present at the hinge region of the antibodies. By reducing the disulfide bridges, it is possible to reveal these cysteine residues which can afterwards be covalently linked with a fluorophore conjugated with a maleimide-group. However, this technique would reduce the total number of dyes that can be put on the antibody as every antibody half would only have two cysteine residues to label at the hinge region. Furthermore, as for the labeling on primary amines it needs to be tested if other cysteine residues within the sequence of the antibody are not affected by the labeling and eventually alters the binding properties.

Another idea would be to use the aforementioned VHH technology to produce tagged nanobodies which can be labeled. One example would be a tag with several cysteine residues or amino acids with many primary amines which could be labeled after the production of the tagged VHH with fluorophores conjugated with maleimide- or succinimidyl-ester-groups. This would be also superior over using just a GFP-fused VHH as the fluorescent dyes have huge advantages over fluorescent proteins such as wider spectral range, greater photostability, smaller size and often a higher brightness (Toseland, 2013). However, it would need to be tested if such a highly labeled VHH is still soluble inside living cells or if the overall length of the peptide tag and the number of conjugated dyes need to be adapted. Interestingly, it is also possible to produce recombinant Fab fragments using a baculovirus expression system by first defining the genetic sequence of the Fab and afterwards introducing it into a baculovirus to infect insect cells and to produce the recombinant Fab. Like this, a genetic tag could be also added to the Fab sequence for site specific labeling without the long production procedure of a VHH (Etienne Weiss, personal communication).

Another tag that could be added to recombinant Fabs or VHHs are self-labelling protein domains. These tags are also known as SNAP-, Halo- or CLIP-tag depending on the chemistry that is used (Gautier et al., 2008; Los et al., 2008; Sun et al., 2011). They all belong to the class of protein-tags consisting of self-labelling enzymes that covalently link to substrates which can be fluorescently conjugated. Thus, the fluorescent labeling of the tagged protein can be started at any time within the cell just by adding the substrate to the medium. However, one problem is the high background that occurs due to the high concentration of non-bound fluorescent substrate inside the cell. Therefore, new protein-tags like the photoactive yellow protein (PYP)-tag were developed which are not only smaller than other protein-tags (14 kDa) but the fluorescent probes were designed so that fluorescence can be only detected when the dye is covalently linked to the PYP-tag (Schneider and Hackenberger, 2017). Another labeling strategy to add a fluorescently labeled peptide directly to the N- or C-terminus of a Fab or VHH would be by transpeptidation using a bacterial sortase enzyme derived from *Staphylococcus aureus*. These natural or recombinant enzymes can catalyze a ligation reaction between oligoglycine nucleophiles and a specific five amino acid substrate motif (LPXTG) (Antos et al., 2017). By adding one of these features to the VHH/Fab and the other one to a fluorescently labeled peptide, it would be possible to site-specifically link both molecules without risking to add any fluorescent dyes to an amino acid of the VHH or Fab. These examples show that many different other labeling strategies exist or are still in development which could improve the labeling efficiency and specificity of VANIMA.

2.3 Possibilities for VANIMA: Combination with different imaging techniques

It was already shown that VANIMA can be used with different imaging/microscopy techniques like confocal time-lapse imaging or 3D-SIM microscopy. However, there are several other imaging techniques which could profit from VANIMA to gain new insights about endogenous transcription dynamics. In this section, I would like to discuss some of the ideas I have to combine VANIMA with other imaging techniques like FRAP, FCS, SPT or Förster resonance energy transfer (FRET) and how good these techniques would synergize with VANIMA as well as what are the points of caution.

As already described in the *Introduction* (section 2.1) are FRAP, FCS and SPT three of the most important techniques to measure protein diffusion and dynamics in living cells. All three techniques can be combined with VANIMA. As it was developed to enable the labeling of endogenous proteins for tracking experiments it seems to be the perfect match for a combination with SPT. The endogenous labeling level results in a lower amount of fluorescent molecules inside the cells in contrast to overexpression based techniques. This can help to obtain more reliable results using SPT, as the number of fluorescent molecules can be controlled precisely and therefore single molecules can be more easily followed without the fear of too many other fluorescent particles crossing the detection volume. This advantage could also be useful for FCS measurements as they rely on very low concentrations of fluorescent molecules passing through the detection volume to generate reliable diffusion correlations (Magde et al., 1972). However, it is important to mention that controls need to be added to ensure that really the movement of the target protein is measured and not only free antibodies or Fabs are tracked inside the cells. This can be accomplished by either measuring the dynamics of free antibodies/Fabs in solution or by electroporating antibodies/Fabs which have no target in the cells to measure their diffusion and/or movements. Whereas FCS and SPT could be used together with VANIMA to measure dynamics of fast moving proteins or PTMs, FRAP could be employed to study endogenous chromatin bound and therefore slower processes in the nucleus which cannot be detected using FCS. However, it needs to be tested if the overall signal intensity of the endogenous VANIMA labeling is sufficient to perform reliable FRAP experiments. I think a combination of experiments including the three imaging techniques together with endogenous labeling using VANIMA could give new insights into the dynamics of transcription factors and nuclear PTMs in living cells. Specific examples with more defined biological questions will be also discussed in the following section of the *Discussion*.

FRET is an imaging technique in which two fluorophores are used, acting as energy donor and acceptor. The energy transfer occurs when the two fluorophores are in close proximity of around 10 nm. After the excitation of the donor fluorophore, there is a nonradiative transfer of energy to the acceptor inducing a fluorescent signal (Ha, 2001). This mechanism could be used in combination with VANIMA in different ways. It could be used to ensure that the fluorescent molecules that are detected are really

antibody bound targets by using two antibodies/Fabs labeled each with either the donor or acceptor fluorophores which bind to the target protein in close proximity and the resulting acceptor fluorescence would indicate target bound antibodies/Fabs. However, it is important to test if both antibodies together don't affect the target protein inside living cells and therefore it is probably better to use two Fabs or VHHs to decrease the overall size of the target bound molecules. Another application could be the use of two antibodies/Fabs against different targets which are meant to incorporate into a complex or just interact with each other inside the cells. Thus, the FRET signal that could be detected could indicate the co-localization of the two target proteins inside the living cells. It is important to mention that further experiments need to be performed to test if these ideas are really applicable.

3. Following transcription in living cells

As it was already mentioned in the *Introduction* of this thesis (section 2.1), had the development of new imaging/microscopy techniques a huge impact on the study of protein behavior and dynamics in living cells. These studies helped to either confirm data from original biochemical experiments or to obtain new insights into the dynamics of transcription related processes in the nucleus of living cells (Liu and Tjian, 2018). Therefore, in the following section of the *Discussion*, I would like to give some perspectives and ideas on which biological questions could be tackled now concerning transcription assembly and dynamics using the newly developed VANIMA technique.

3.1 Analysis of PIC assembly dynamics in single living cells

In chapter 4 of the *Results* section, I presented some preliminary results as a proof of principle in which we used VANIMA in combination with an activatable fluorescently labeled gene array to measure the recruitment of RNA Pol II to the gene array after transcription activation. Further experiments need to be performed to get statistically significant quantitative data about the recruitment dynamics of RNA Pol II to this gene array in human U2OS cells. In previous studies it was already shown that the different steps of RNA Pol II transcription can be visualized at a gene array in mouse cells

together with dynamic changes in histone modifications over time (Stasevich et al., 2014a; Stasevich et al., 2014b). They suggested rather slow recruitment dynamics for RNA Pol II to the gene array of around 2.3 min.

However, to my knowledge, no data exists until now concerning the recruitment dynamics of other GTFs to a specific gene or gene array in living cells. Therefore, I would be interested in using VANIMA in the combination with the gene array to measure the recruitment dynamics of different components of the PIC by utilizing differently labeled antibodies/Fabs against RNA Pol II, TFIID and other GTFs in single living cells. Just by measuring fluorescence intensity fluctuations at the gene array of the aforementioned VANIMA labeled endogenous factors and the activator by using confocal spinning disc microscopy or live 3D-SIM microscopy, it would be possible to study the recruitment dynamics of the sequential PIC assembly in single living cells. However, it is important to mention that 3D-SIM live imaging together with multicolor VANIMA labeling needs to be still implemented to test if the signal intensity of the endogenous labeling is high enough to visualize and follow the recruitment of the factors over time. Furthermore, depending on the time window in which the recruitment is happening, it needs to be tested if the 3D-SIM movies can be acquired fast enough to obtain reliable results. Nevertheless, the recruitment probably occurs in a time window of a few minutes which could be followed with live 3D-SIM if the exposure time of the single images is not too long. Further experiments could also include Fabs or VHHs against phosphorylated serine 5 of the CTD of RNA Pol II to be able to distinguish the dynamics of RNA Pol II and GTF recruitment and transcription initiation.

FRAP and FCS measurements could be performed as well to accompany the time-lapse experiments and to obtain specific diffusion coefficients of RNA Pol II and GTFs at the gene array. Furthermore, it would be very interesting to repeat the experiments using a single gene system to see if the recruitment dynamics are changing. However, the advantage of a gene array is that it can be visualized easily inside the nucleus due to its high amount of repeats. Nevertheless, new DNA labeling strategies like the ANCHOR 3 system try to overcome this problem to be able to visualize a single gene in living cells (Germier et al., 2017). Further experiments using different transcription inhibitors to block specifically initiation or elongation could be used as well to analyze changes in the recruitment or initiation of the factors (Bensaude, 2011). Furthermore, all these experiments could be easily repeated in different human cell lines including

primary cell lines to see if differences can be observed concerning recruitment dynamics between cancer and non-cancer cells. Altogether, these experiments could uncover the dynamics of the assembly of GTFs and RNA Pol II at a promoter in single living cells and test our current mostly biochemical knowledge about PIC assembly and therefore answer the question which assembly method is the more prevalent in single living cells: sequential or holo-RNA Pol II complex PIC assembly? Moreover, the results could also lead to the identification of specific targets for cancer therapy if different dynamics can be observed in cancer cell lines. Lastly, once established, this imaging procedure could be used as well to test and identify new molecules which have a direct effect on transcription and PIC assembly.

3.2 Analysis of RNA Pol II clusters and their implication in phase separation

In chapter 3 of the *Results* section, I showed results about RNA Pol II cluster dynamics and distribution in single cells. Previous studies suggested that RNA Pol II cluster are rapidly assembling and disassembling at active promoters and are mainly involved at the step of transcription initiation (Liu et al., 2014; Liu and Tjian, 2018). However, our results indicated that RNA Pol II clusters indeed disappear after transcription elongation inhibition with flavopiridol. Therefore, I would like to further investigate these RNA Pol II foci in living cells to be able to answer in which phase of transcription these clusters really are formed and also if they are connected with the formation of liquid-liquid phase separation droplets (Hyman et al., 2014). The theory of phase separation suggests that in addition to classical cellular compartments which are separated by membranes, high concentrations of molecules (including proteins and nucleic acids) can induce liquid-liquid phase separation to form a dense phase (or droplet) enriched with these molecules which serve as non-membrane containing compartments (Alberti, 2017). They are formed to separate and localize specific biochemical reactions in space. It was suggested that phase separated droplets are also formed by factors of the transcription machinery to increase the concentration of these factors at the gene and that real interactions are only occurring transiently at the promoter (Liu and Tjian, 2018).

First, I would like to repeat the experiments in 3D-SIM using different transcription inhibitors like triptolide to specifically inhibit transcription initiation. This would answer the question if the clusters are transcription elongation specific. The aforementioned gene array could also be used to include a reference point where RNA Pol II accumulation can be induced at any time. Furthermore, specific Fabs or VHHs against initiation or elongation specific CTD phosphorylation could be used to see if the clusters are appearing in these specific stages of transcription as well.

Previous co-localization studies showed that 3D-SIM can be also used to test if two molecules are in close proximity or not (Cerase et al., 2014). Therefore, antibodies or Fabs against other factors specific for transcription initiation (for example TFIID), elongation (for example TFIIS) or specific histone modifications indicating active transcription (for example H3K9ac or H3K27ac) could be used in two color 3D-SIM experiments to verify their presence or increased concentration within the RNA Pol II clusters by performing co-localization studies. However, a positive control for co-localization in form of an IF experiment with one primary antibody and two secondary antibodies conjugated with different dyes or fluorescent beads is necessary to define “real” co-localization. This is important as complete overlapping of two signals is difficult to detect even if two molecules are co-localizing due to the increased resolution of 3D-SIM and the fact that the microscope is always detecting the fluorescently labeled antibodies/Fabs and not the target proteins directly. However, as mentioned before, real interactions between the factors could also be highly transient if they are indeed forming phase separated droplets. Then, these co-localization experiments could also reveal if the different factors of the transcription machinery are at least enriched within the defined RNA Pol II clusters.

All these experiments would be repeated using confocal or 3D-SIM live imaging to see the dynamics of the clusters in living cells before and after transcription inhibition. Additionally, FCS measurements could be performed to calculate the concentration of RNA Pol II molecules in wild type or gene array bound RNA Pol II clusters. Moreover, to tackle the question if these RNA Pol II clusters are involved in liquid-liquid phase separation, the 3D-SIM and live imaging experiments could be repeated using a specific inhibitor for phase separation called 1,6-hexanediol (Lu et al., 2018). This molecule is routinely used to inhibit the formation of phase separated droplets in living cells and a decrease in the number of RNA Pol II clusters after treatment with the drug

could indicate that these accumulation of the polymerase are indeed involved in phase separation. All these experiments could give new insights into the behavior of endogenous RNA Pol II clusters in living cells and answer the longstanding debates if they are formed at the initiation step as part of transcription factories or during elongation as RNA Pol II trains on the gene.

CONCLUSION

Conclusion

During my thesis, I have explored the assembly of the general transcription factor TFIID and the use of intracellular antibody targeting both to inhibit important functions in form of replication or to label and track endogenous proteins in living cells using specific non-inhibiting antibodies.

We were able to show that a TFIID sub-module consisting of TAF2-TAF8-TAF10 exist *in vivo* in the cytoplasm and nucleus of the cells. These results support the view of the presence of stable partial TFIID complexes which could have potentially important functions on their own but are surely necessary for the cytoplasmic-nuclear transport of certain subunits and for the stepwise holo-TFIID assembly in the nucleus.

Furthermore, we showed that intracellular antibodies against important replication factors like PCNA and DNA Pol α can induce a huge amount of replication stress followed by high DNA damage and cell death of various cancer cell lines. These results showed that intracellular antibodies could be used as a potential novel approach for cancer treatment by inducing replication stress.

In my own publication, I have developed a strategy that is simple to implement for visualizing target antigens in their native form in single living cells without causing any toxicity in the treated cells. This approach includes the highly efficient delivery of fluorescently labeled antibodies or Fabs into living cells by electroporation and was named versatile antibody-based imaging approach (VANIMA). It can be used for live- and single-cell super-resolution detection of a large variety of factors and PTMs. Moreover, by using VANIMA, dynamic processes of fundamental biological mechanisms can be visualized in nonfixed cells at high resolution. The results suggested that larger endogenous RNA Pol II cluster are present in the nucleus of living cells and may contain several transcribing RNA Pol II assemblies or RNA Pol II “trains” possibly organized in transcription-related compartments and/or other control regions. Moreover, we also show that large RNA Pol II cluster are constantly forming, dynamically associating and dissociating.

Furthermore, by coupling VANIMA with genetic labeling, controlled transcription activation and confocal live-imaging, we could observe the recruitment of endogenous RNA Pol II to a gene array after transcription activation and compare it to activator

binding, which serves as a good starting point to further analyze RNA Pol II dynamics in living cells. In the future, by using VANIMA coupled to live confocal or 3D-SIM imaging and/or the aforementioned genome-labeling technology, it will become possible to investigate, characterize, and dissect the function and dynamics of RNA Pol II transcription in single living cells. Hence, it would be possible for example to uncover the dynamics of RNA Pol II PIC assembly directly at the promoter to challenge the current predominant biochemical knowledge and answer the question of how transcription is really initiated in living cells.

MATERIAL & METHODS

Material & Methods

1. Antibody-based Imaging Approach to Visualize Endogenous Proteins and Posttranslational Modifications in Living Metazoan Cell Types (S. Conic *et al*.; Bioprotocol, under review)

The following section, in contrast to a classical Material & Methods chapter, contains a step by step protocol to guide other scientists who want to apply VANIMA staining in their laboratory. All materials, reagents and equipment which were used to perform VANIMA are listed. Furthermore, the procedure chapter contains the whole VANIMA protocol from the first validation test using IF to verify if the desired antibody is suitable for VANIMA, over the purification of the antibody, the digestion protocol to obtain Fab fragments, the fluorescent labeling of the antibodies or Fabs until the electroporation procedure for proteins into living cells. Moreover, it includes important tips and notes of caution for the VANIMA procedure and the following data analysis. Lastly, the recipes for all the buffers used to perform VANIMA are also included.

This protocol was submitted on the 12th of July 2018 to *Bioprotocol* and is currently under review.

Author's contributions

Sascha Conic – Designed the protocol. Performed all the experiments except for Figure 1. Wrote the manuscript.

Dominique Desplancq – Performed the experiment shown in Figure 1. Assisted in the writing of the manuscript.

László Tora – Designed the protocol.

Etienne Weiss – Designed the protocol.

Antibody-based Imaging Approach to Visualize Endogenous Proteins and Posttranslational Modifications in Living Metazoan Cell Types

Sascha Conic[1, 2, 3, 4, *], Dominique Desplancq[4, 5], László Tora[1, 2, 3, 4] and Etienne Weiss[4, 5, *]

[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire, 67404 Illkirch, France; [2]Centre National de la Recherche Scientifique, UMR7104, 67404, Illkirch, France; [3]Institut National de la Santé et de la Recherche Médicale, U964, 67404, Illkirch, France; [4]Université de Strasbourg, 67404, Illkirch, France; [5]Biotechnology and Cell Signaling, UMR 7242, 67404 Illkirch, France

*For correspondence: conic@igbmc.fr; etienne.weiss@unistra.fr

[Abstract] The spatiotemporal localization of different intracellular factors in real-time and their detection in live cells are important parameters to understand dynamic protein-based processes. Therefore, there is a demand to perform live-cell imaging and to measure endogenous protein dynamics in single cells. However, fluorescent labeling of endogenous protein in living cells without overexpression of fusion proteins or genetic tagging has not been routinely possible. Here we describe a versatile antibody-based imaging approach (VANIMA) to be able to precisely locate and track endogenous proteins in living cells. The labeling is achieved by the efficient and harmless delivery of fluorescent dye-conjugated antibodies or antibody fragments (Fabs) into living cells and the specific binding of these antibodies to the target protein inside of the cell. Our protocol describes step by step the procedure from testing of the suitability of the desired antibody, over the digestion of the antibody to Fabs until the labeling and the delivery by electroporation of the antibody or Fab into the cells. VANIMA can be adapted to any monoclonal antibody, self-produced or commercial, and many different metazoan cell lines. Additionally, our method is simple to implement and can be used not only to visualize and track endogenous factors, but also to specifically label posttranslational modifications, which cannot be achieved by any other labeling technique so far.

Keywords: Antibodies, Fab fragments, Live-imaging, Antibody delivery, Single cells, endogenous proteins, posttranslational modifications

[Background] The fluorescent labeling of proteins to follow in real time their spatiotemporal localization in living cells was mainly achieved until now by using transgenic or overexpression-based approaches. However, the labeling of specific endogenous proteins or even posttranslational modifications in living cells is not yet routinely possible. Imaging of cellular structures and processes is typically performed by either immunofluorescence (IF) labeling on fixed cells or by exogenously overexpressing fluorescent fusion proteins in living cells. Although these well-established techniques showed to be very powerful to locate or follow proteins inside the cells, they inherit also some important drawbacks. In IF, the cells need to be chemically fixed and permeabilized to be able to incubate them with specific primary and secondary antibodies. Despite many variables and potential artifacts (Schnell *et al*., 2012; Teves *et al*., 2016) like fixation-related protein denaturation or permeabilization efficiency, IF is still often used to visualize target proteins in fixed cells or tissues. Otherwise, imaging of proteins in living cells is mainly achieved through the exogenously expression of fluorescent fusion proteins (Ellenberg *et al*., 1999; Betzig *et al*., 2006; Schneider and Hackenberger, 2017) or by knock-in of a fluorescent tag into the endogenous locus using the CRISPR/Cas9 technology (Ratz *et al*., 2015). Although fluorescent fusion proteins have been proven to be very powerful, they often do not behave as their endogenous counterparts due to their increased levels when exogenously overexpressed (Burgess *et al*., 2012). On the other hand, endogenous fusion proteins containing knocked in tags are difficult to obtain as knock-in efficiencies are often very low. Consequently, there is a need for new and easy to implement imaging approaches to visualize endogenous target proteins in single living cells. Previous studies and methods, like FabLEM or the expression of mintbodies, showed that intracellular labeling of proteins with fluorescently labeled antibody fragments can give new insights into the dynamics of histone modifications (Hayashi-Takanaka *et al*., 2009; Hayashi-Takanaka *et al*., 2011; Sato *et al*., 2013). However, these techniques suffer from lower delivery efficiencies into living cells, or potential poor solubility of the intracellular expressed mintbodies. Recently, another method achieved fluorescent labeling of endogenous proteins by using a bacterial toxin called streptolysin O, which creates pores in the membrane of cells and allows for the delivery of fluorescent probes into living cells (Teng *et al*., 2016). However, this method requires additional steps to

be able to reseal the membrane after treatment which can be quite harmful for the cells and can decrease cell viability. In contrast, our versatile antibody-based imaging approach (VANIMA) uses fluorescent dye-conjugated antibodies or Fabs, which are delivered into the cells by electroporation (Freund *et al*., 2013; Brees and Fransen, 2014). The antibody labeling reaction is highly efficient and can result in up to 5-7 fluorescent dyes per molecule of antibody depending on the antibody and the labeling kit used. The transduction of the antibodies has a very high delivery efficiency and viability of the cells is above 90% in human cancer cell lines such as U2OS. Afterwards, the transduced antibodies will bind to the endogenous target protein inside the cell and for nuclear targets they will be transported with the target protein into the nucleus. Otherwise, for faster delivery into the nucleus of the cells, the antibodies can be digested to produce Fabs which can freely diffuse into the nucleus to find and bind their target. Thus, even proteins with posttranslational modifications in the nucleus can be visualized specifically using fluorescently-labeled Fabs against the target. Considering that there are several thousands of commercially-available antibodies that specifically recognize intracellular target proteins with high affinity, VANIMA can be used to uncover the dynamical behavior of a plethora of targets in living cells (Conic *et al*., 2018). Additionally, the method is easy to implement in any laboratory and can also be used to perform multicolor imaging with different targets just by labeling two different antibodies with different dyes or by combining it with an already established endogenous knock-in clone. Finally, VANIMA can also be used with identified inhibiting antibodies to disrupt protein functions inside living cells.

Materials and Reagents

1. 15 ml conical Falcon tubes (Corning, Falcon, catalog number: 352095)
2. 1.5 ml Eppendorf tubes (Sigma-Aldrich, Eppendorf, catalog number: Z66505-100EA)
3. 0.5 ml Eppendorf tubes (Sigma-Aldrich, Eppendorf, catalog number: Z666491-100EA)
4. Falcon 12-well clear flat bottom cell culture plate (Corning, catalog number: 351143)

5. µ-slide 8-well glass bottom: No. 1.5H (170 µm +/- 5 µm) (Ibidi, catalog number: 80827)
6. 18 mm high precision cover glasses (Marienfeld, catalog number: 117580)
7. Microscope slides ground edges plain (VWR, catalog number: 631-1552)
8. Poly-Prep chromatography column (Bio-Rad, catalog number: 731-1550)
9. DiaEasy dialyzer (3 ml) MWCO 6-8 kDa (Biovision, catalog number: K1013-25)
10. DiaEasy dialyzer (800 µl) MWCO 6-8 kDa (Biovision, catalog number: K1019-25)
11. Amicon Ultra-4 centrifugal filter units 10 kDa (Merck-Millipore, catalog number: UFC801024)
12. Amicon Ultra-0.5 centrifugal filter units 10 kDa (Merck-Millipore, catalog number: UFC501096)
13. Countess™ cell counting chamber slides (Thermo Fisher, catalog number: C10312)
14. Sterile individually packaged 5 ml pipettes (Sigma-Aldrich, catalog number: SIAL1487)
15. Sterile individually packaged 10 ml pipettes (Sigma-Aldrich, catalog number: SIAL1488)
16. U2OS osteosarcoma cells [American Type Culture Collection (ATCC, catalog number: HTB-96)]
17. Neon transfection 10 µl kit (including the Neon 10 µl tips) (Invitrogen, catalog number: MPK1096)
18. AlexaFluor-488 antibody labeling kit (Invitrogen, catalog number: A20181)
19. Dulbecco's Modified Eagle Medium (DMEM) (Thermo Scientific, Gibco, catalog number: 10567-014)
20. Heat-inactivated fetal calf serum (FCS) (Gibco, catalog number: 15750-037)
21. Gentamicin (Gibco, catalog number: 15750-037)
22. 16% Paraformaldehyde (16% PFA) (Electron Microscopy Sciences, catalog number: 50-980-487)
23. Phosphate Buffered Saline (PBS) (GE Healthcare, catalog number: SH30013.03)
24. Triton X-100 (Sigma-Aldrich, catalog number: X100-100ML)
25. Vectashield antifade mounting medium with DAPI (Vector-Laboratories, catalog number: H-1200-10)

26. Protein G Sepharose FastFlow (GE Healthcare, catalog number: GE17-0618-01)
27. Protein A Sepharose FastFlow (GE Healthcare, catalog number: GE17-5280-01)
28. Papain-coated magnetic beads (Spherotech, catalog number: PAPM-40-2)
29. Tris-HCl (Sigma-Aldrich, catalog number: 10812846001)
30. Glycine (Sigma-Aldrich, catalog number: G8898)
31. Sodium bicarbonate (Sigma-Aldrich, catalog number: S5761-1KG)
32. Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) (Sigma-Aldrich; catalog number: C4706)
33. Sodium dodecyl sulfate (SDS) (Euromedex; catalog number: EU0660)
34. Acrylamide/Bis-acrylamide 40% solution (Euromedex; catalog number: EU0077-B)
35. Tetramethylethylendiamin (TEMED) (Serva; catalog number: 35930.01)
36. Ammonium persulfate (APS)(Sigma-Aldrich; catalog number: A3678)
37. U2OS growth medium (see Recipes)
38. 4% Paraformaldehyde (4% PFA) (see Recipes)
39. 10x Phosphate Buffered Saline (10x PBS) (see Recipes)
40. 1x Phosphate Buffered Saline (PBS)
41. 10% Triton X-100 (see Recipes)
42. 0.1% Triton X-100 (see Recipes)
43. 0.02% Triton X-100 (see Recipes)
44. 1 M sodium bicarbonate pH 8.2 (see Recipes)
45. 0.1 M sodium bicarbonate (see Recipes)
46. 1 M Tris-HCl pH 8.2 (see Recipes)
47. 0.1 M glycine-HCl pH 2.7 (see Recipes)
48. 2.5% Trypsin (see Recipes)

Equipment

1. Pipetman P2 pipette (Gilson, catalog number: F144801)
2. Pipetman P20 pipette (Gilson, catalog number: F123600)
3. Pipetman P200 pipette (Gilson, catalog number: F123601)

4. Pipetman P1000 pipette (Gilson, catalog number: F123602)
5. Jewelers forceps, Dumont No. 5 (Sigma-Aldrich, Dumont, catalog number: 6521)
6. Pipette boy (Corning, Falcon, catalog number: 357469)
7. Water bath (Julabo; model: ED (v.2))
8. Magnetic tube rack (Diagenode, catalog number: B04000001)
9. Fume hood Hera Safe KS (Thermo Scientific, catalog number: 51023175)
10. Cell culture incubator with CO_2 supply (Sanyo, catalog number: MCO-19AIC)
11. SP8UV confocal microscope (Leica)
12. Eppendorf centrifuge 5804 R (Eppendorf, model: 5804 R, catalog number: 805000620)
13. Beckman Coulter Allegra centrifuge (Beckman, catalog number: 21R)
14. Nanodrop 2000 spectrophotometer (Thermo Scientific, model: NanoDrop™ 2000, catalog number: ND2000)
15. Countess Cell Counter (Thermo-Fisher, catalog number: AMQAX1000)
16. Neon Transfection System (Invitrogen, catalog number: MPK5000S)

Software

1. Fiji/Image J (https://fiji.sc/)

Procedure

Notes:

a. This protocol was optimized for U2OS cells but can be adapted to any adherent metazoan cell line.

b. Use 500 µl of buffer for every wash-step if 12-well plates are used, otherwise the volume needs to be adjusted.

A. Validation of antibody for VANIMA by immunofluorescence

1. Seed around 10^5 cells of U2OS cells in a 12-well plate containing growth

medium and a glass coverslip.

2. Let cells re-attach to the coverslip surface overnight.
3. Pre-warm 4% PFA, diluted in PBS to 37 °C in a water bath.
4. Remove growth medium from the cells and rinse them twice with PBS.
5. Fix the cells by adding the pre-warmed 4% PFA for 5 min at RT.
6. Rinse the cells two times with PBS.
7. Permeabilize the cells by incubating them in PBS containing 0.1% Triton X-100 for 20 min at RT.
8. Rinse the cells twice with PBS.
9. Incubate the cells with different dilutions of your antibody in PBS supplemented with 10% FCS for 1 h at RT (see also Note 1).
10. Wash the cells twice with PBS supplemented with 0.02% Triton X-100 for 5 min at RT.
11. Wash the cells once with PBS for 5 min at RT.
12. Incubate the cells with the corresponding fluorescently-labeled secondary antibody (like anti-mouse-IgG-Alexa488) in a dilution of 1/3,000 in 10% FCS in PBS for 1 h at RT.
13. Repeat the wash-steps as mentioned in Steps A10 and A11.
14. Mount the coverslip on a microscope slide using 6-10 µl of Vectashield mounting medium containing DAPI.
15. Observe your cells under the microscope to define if the antibody shows the expected staining of the target (see Note 2).

Note: Procedure B can be skipped if the antibody for VANIMA is already pure or a commercially available antibody in PBS (see Notes of Procedure D).

B. Purification of mouse monoclonal antibodies for electroporation (see Figure 1)

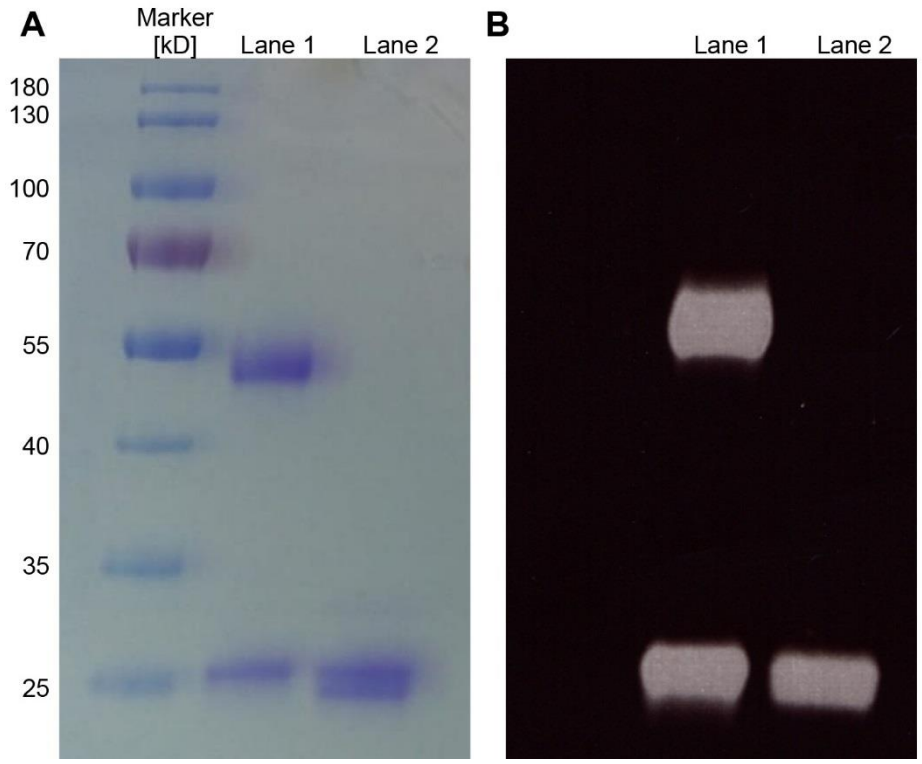


Figure 1. Analysis of antibodies and Fabs by SDS-PAGE. *A. Purified aliquots of antibodies (Lane 1) or Fabs (Lane 2) were chemically-labeled with Alexa488 and subsequently analyzed by SDS-PAGE and Coomassie staining. B. Corresponds to the same gel shown in A analyzed under UV illumination before staining.*

1. Transfer 1 ml of Protein G Sepharose FastFlow beads into a 15 ml tube.
2. Centrifuge the bead solution at 277 *x g* for 3 min at 4 °C to pellet the beads.
3. Remove the storage solution and add 5 ml of PBS. Resuspend the beads and centrifuge them again as mentioned before.
4. Repeat this step 4 times to equilibrate the beads in PBS and to remove all the storage solution.
5. Remove all PBS from the beads and add the solution containing the antibodies to the beads.
6. Incubate the beads for 2 h at 4 °C under constant shaking.
7. Centrifuge the beads for 5 min at 277 *x g* at 4°C.
8. Remove the supernatant and keep it on ice. This is the flow through (FT) which shouldn't contain any antibodies anymore.
9. Add 2 ml of PBS to the beads, resuspend them and transfer them to a Poly-

Prep chromatography column.

10. Add a total of 20 ml of PBS to wash the beads and to remove all unspecific bound proteins.
11. Prepare ten 1.5 ml Eppendorf tubes with 70 µl of 1 M Tris-HCl pH 8.2 for fractionation and neutralization.
12. After all the PBS passed through the column, start the elution of the antibody from the beads by adding stepwise 10 ml of 0.1 M glycine-HCl pH 2.7 in 1 ml steps and collect the fractions in the prepared Eppendorf tubes containing the neutralization buffer.
13. Analyze an aliquot of the every elution fraction by SDS-PAGE using a 12% SDS-acrylamide gel. The following samples can be:

    1) The input antibody solution

    2) The flow-through (FT)

    3) All ten fractions collected

14. Perform a Coomassie staining after the electrophoresis and pool all the fractions containing the purified antibodies.
15. Dialyze the pooled fractions against a total of 4 L of PBS in two steps using DiaEasy dialyzer tubes. The first step overnight and the second for 4 h with 2 L of PBS each at 4 °C.
16. Measure the concentration of the dialyzed antibody by 280 nm absorption using a Nanodrop spectrophotometer. Concentrate the purified antibodies using the 4 ml Amicon filter units with a cutoff of 10 kDa by centrifugation at 4,000 *x g* until the concentration is 1 mg/ml or higher.

Note: This protocol is optimized for an antibody input of 200 µg which corresponds to antibody samples that are commercially available.

C. Digestion of monoclonal antibodies to Fab fragments (see Note 3 and Figure 1)

1. Prepare 200 µg of monoclonal antibodies in PBS (1 mg/ml) (see Procedure B).
2. Add 1.2 µl of 0.17 M TCEP (1 mM final concentration) in 200 µl of antibody solution.
3. Transfer 100 µl of magnetic Papain coated bead solution per digestion into 0.5 ml Eppendorf tubes.

4. Use an Eppendorf magnetic stand to fix the magnetic beads and remove the storage solution.
5. Wash the beads 3 x with 300 µl of PBS by resuspending them in the buffer and afterwards removing the washing buffer again with the help of the magnet.
6. Remove all PBS and add the antibody solution with TCEP to the beads.
7. Incubate for 3 h at 37 °C under shaking.
8. Remove and collect the supernatant (S1) from the beads using the magnet (this includes the Fab's).
9. Wash the beads again 3 x with PBS and store them at 4 °C in PBS to be able to reuse them.
10. For the purification of the Fab fragments, transfer 100 µl of Protein A Sepharose beads into a 0.5 ml Eppendorf tube.
11. Equilibrate the beads by washing them 4 x with 300 µl of PBS. Centrifuge them for each wash-step for 3 min at 277 *x g*.
12. Remove all PBS and add supernatant (S1) from the digestion to the beads.
13. Incubate for 30 min at 4 °C under shaking.
14. Centrifuge for 5 min at 277 *x g* and collect the supernatant (S2).
15. Wash beads with 300 µl of PBS and centrifuge again.
16. Collect wash step and pool with supernatant (S2).
17. Concentrate the fraction S2 using an Amicon filter unit with a cutoff of 10 kDa (0.5 ml or 4 ml tubes) to about 100 µl volume (5 min at 14,000 *x g*). This is now fraction S3.
18. Determine concentration of the Fab by measuring the absorption at 280 nm using the Nanodrop spectrophotometer.
19. Perform a SDS-PAGE using a 4-15% acrylamide gradient gel and the following samples that need to be boiled during 5 min before loading:
    Input antibody solution

    Supernatant S1

    Supernatant S2

    Supernatant S3

    Filtrate from the Amicon concentration step

20. The Fab fragments are now purified and ready for labeling.

D. Fluorescent labeling of monoclonal antibodies or Fab fragments for VANIMA (see Note 4)

1. Prepare 1 L of 0.1 M sodium bicarbonate pH 8.2 (see Recipes).
2. Dialyze 100 µg of antibody/Fab solution in a volume of 100 µl (1 mg/ml) against 1 L of 0.1 M sodium bicarbonate for 4 h at 4 °C using DiaEasy dialyzing tubes (see Note 8).
3. Mix the antibody/Fab solution with fluorescent dyes as written in the manufacturer's protocol and incubate at RT for 1 h in the dark. Mix the solution every 15 min by inverting the tube. The labeling kits used to label the antibodies/Fabs are the Alexa Fluor Monoclonal Antibody labeling kits from Invitrogen (see Materials and Reagents).
4. Remove non-bound dyes by purifying the labeling mix using the gel filtration columns supplied in the labeling kit (see Note 9).
5. Concentrate the labeled antibody/Fab by using 0.5 ml Amicon filter units with a cut-off of 10 kDa. Centrifuge for 10 min at 14,000 *x g* to concentrate the solution to a volume of approximately 50 µl.
6. Measure the concentration of the labeled antibody using a Nanodrop spectrophotometer and the Protein and Labels mode.
7. Labeling efficiency can be calculated by measuring the absorption at 280 nm and at the dye specific wavelength. The dye/antibody labeling ratio can then be calculated using the formula mentioned in the protocol of Invitrogen (see Note 10).

Notes:

a. *Your antibody or Fab fragments need to be diluted in sterile PBS with no other ingredients prior to electroporation as preservatives like sodium azide or traces of BSA will lower the viability of your cells after electroporation significantly.*
b. *The Neon transfection system and the corresponding Neon transduction kits are used for antibody transduction.*
c. *This protocol describes electroporation using the 10 µl Neon tips but 100 µl Neon tips can be used as well. The number of cells and antibodies/Fabs need to be adjusted accordingly.*
d. *All buffer and solutions need to be filtered and sterile for electroporation to avoid*

contamination of the antibodies and cells.

e. *Pre-warm 2.5% trypsin and the growth culture medium without antibiotics to 37 °C prior to the experiment.*

E. Electroporation procedure for monoclonal antibodies or Fab fragments (see also Video 1)

Video 1. Electroporation procedure

1. Transfer the volume corresponding to 1-5 µg of antibody/Fab solution into a 1.5 ml Eppendorf tube and keep them in the dark on ice.
2. Add antibiotics free growth medium into the cell culture plate (12-well plate with glass coverslip) or live-imaging chamber slides (µ-slides from Ibidi) and store them in the incubator at 37 °C.
3. Wash the U2OS cells twice with PBS.
4. Detach the cells from the surface by incubating them with 2.5% trypsin for 4 min and dilute them in antibiotic-free growth medium.
5. Count the cells and transfer the volume of cell suspension needed for a final cell number of 8 x 10^5 cells (see Note 6) to a 15 ml tube and pellet the cells by centrifuging them at 200 *x g* for 5 min. Every electroporation uses 10^5 cells which means that with this pellet one can perform 8 transductions in total.
6. Remove the growth medium and resuspend the cell pellet in 4 ml of PBS and centrifuge them again.
7. Remove the PBS and resuspend the pellet in 80 µl of resuspension buffer (R-buffer) (see Note 7).
8. Mix 10 µl of the resuspended cells with the antibody/Fab solution and pipette this mix with the Neon pipette using the 10 µl Neon tips. Be careful that no bubble can be seen in the Neon tip as this will induce an electric short-circuit which will kill the cells during the electroporation.
9. Put the Neon pipette into the pipette station and electroporate the cells using the following parameters:

   Voltage: 1550 V

   Number of pulses: 3

Time width of pulse: 10 msec

10. Pipette the cells after the electric pulse directly into the cell culture plate prepared in Step E2.
11. Repeat Steps E8-E10 until all transduction were performed.
12. Incubate the cells in the incubator at 37 °C with 5% CO_2.
    a. For Fab transductions, the cells can be observed or fixed as early as 6 h post electroporation.
    b. For full-length antibody transductions, the incubation time can vary depending on the neo-synthesis rate of the target protein in the cell and its localization (see Note 3).
13. The same protocol as in Procedure A can be followed if the cells need to be fixed except that for the electroporation of labeled antibodies/Fabs Steps E6-E12 can be skipped.
14. For live-cell imaging the samples can be observed under the microscope after the incubation time (see Figure 2).

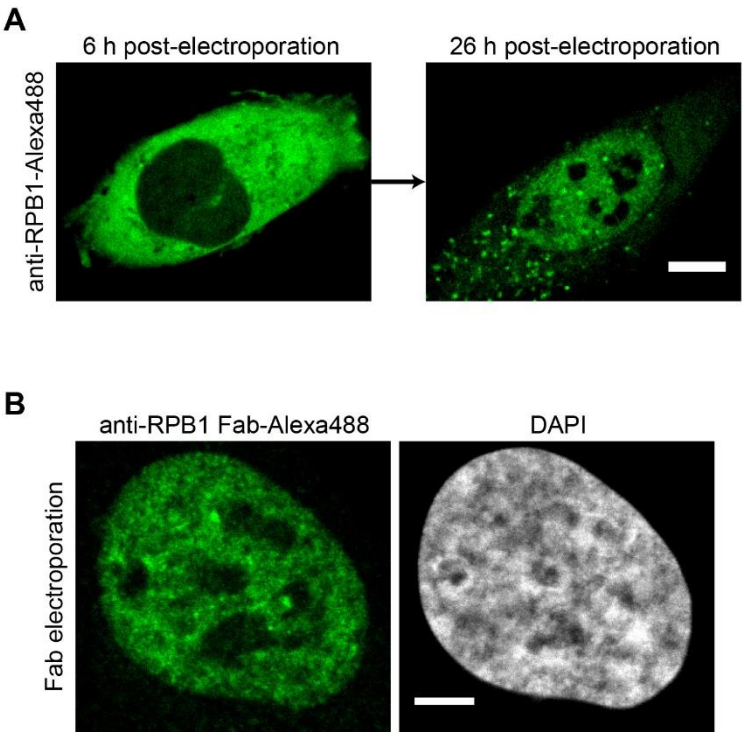


Figure 2. Imaging results after the electroporation of anti-RPB1-Alexa488 antibodies or Fabs. *A. U2OS cells were electroporated with anti-RPB1-Alexa488 labeled antibodies and the same cell was imaged by confocal microscopy 6 h and 26 h post-electroporation. The transport of the labeled antibody from the cytoplasm into the nucleus can be detected. Scale bar = 10 µm. B. Electroporation as in A but this time anti-RPB1-Alexa488 Fab fragments were transduced. The cells were fixed and imaged 6 h post-electroporation by confocal microscopy and a specific nuclear staining for RPB1 can be observed. Scale bar = 5 µm.*

Data analysis

1. The microscopes used for the analysis were either confocal, spinning disk or 3D-SIM super resolution microscopes for live-imaging or fixed cell acquisition. For each electroporation experiment, we perform at least three replicates. All samples for one data set were acquired on the same day with the same microscope settings. The use of higher laser power or longer exposure times for these samples is completely normal as labeling of endogenous proteins always results in lower signal intensities. The laser power and exposure times can vary from 5-20% or 40-200 msec respectively depending on the labeling efficiency, target protein as well as the microscope used.
2. Microscopic analysis of electroporated cells should be performed within 3 days after transduction. Afterwards the antibody/Fab will be diluted and signal intensity will decrease significantly due to degradation of the antibody/Fab or by dilution due to cell division.
3. Bright spots eventually present in the cytoplasm of cells after electroporation may correspond to antibodies that were structurally altered due to dye conjugation and they thus tend to aggregate in the cytoplasm. A possibility to avoid the formation of these aggregates is to electroporate a lesser amount of antibodies/Fabs or to repeat the labeling experiment using less dye.
4. Image processing can be performed using classical Fiji/ImageJ software. Different plugins can be used to calculate protein distribution or dynamics using 3D-SIM images. For images of the same data set, the image processing should be performed the same way.
5. It is recommended to perform validation experiments for cell viability and function every time a new antibody/Fab is used for electroporation. Apoptosis and proliferation assays should be performed to test for antibody toxicity. Also, specific validation tests concerning the mechanism involving the target protein inside the cell should be performed to ensure that the antibody is not inhibiting any important function of the protein (*e.g.,* measuring the amount for nascent RNA transcription for a target involved in transcription).

Notes

1. To verify if the desired antibody is suitable for VANIMA it is recommended to test it first in classical immunofluorescence (IF). Procedure A of this protocol can be skipped if the antibody was already characterized by IF. Depending on the antibody source (hybridoma supernatant, commercial antibody, *etc*.). It is important to test several dilutions or concentration of the antibody for the first test.
2. The expected staining for IF and VANIMA depends on the target antigen that needs to be analyzed. If the antibody recognizes a nuclear target and if its use leads to a strong unspecific signal in the cytoplasm, then this antibody is likely not suitable for VANIMA.
3. Full-length antibodies are too big to be able to diffuse freely into the nucleus. They need to bind to the newly synthesized target protein in the cytoplasm and get piggybacked with it into the nucleus. Therefore, the time for the transport of the antibody into the nucleus depends highly on the turnover of the target protein and antibodies against posttranslational modifications in the nucleus will never reach their target and will remain cytoplasmic. In contrast, Fab fragments can diffuse freely into the nucleus and are particularly adapted to target posttranscriptional modifications or proteins with a low cellular turnover.
4. To test if the antibody or Fab is functional inside the cell, it is recommended to electroporate them once before they get fluorescently labeled. They can be visualized after fixation with a fluorescently labeled secondary antibody.
5. To test a new antibody or Fab it is also important to verify how much antibody is needed to be electroporated to bind and label all of the target protein in the cell. Therefore, for the first electroporation, a titration experiment with different amounts of antibodies or Fabs in a range between 1-5 µg is recommended. Antibody amounts higher than 10 µg should be avoided as these high amounts of protein electroporated will start to get toxic for the cell.
6. A total of 8 electroporation experiments can be performed with a pellet of 8 x 10^5 cells (10^5 cells per transduction). Even if a lower number of electroporation are planned it is still better to keep a higher cell number as lower number will result in very small pellets and inaccuracies which can lower cell viability

drastically.

7. As soon as the cells are resuspended in R-buffer, it is important to proceed quickly to the electroporation. Cell viability could decrease significantly if the cells remain longer than 15 min in R-buffer (please contact us if you're interested in the recipe of the R-buffer).
8. The dialysis against the labeling buffer (0.1 M sodium bicarbonate) is an important step as it will raise the pH of the antibody solution over a pH of 8 which will increase labeling efficiency. The dialysis also showed to give better labeling efficiencies (3-4 dyes per molecule more) than the pH raising step described in the Invitrogen labeling protocol.
9. Washing of the gel filtration column with PBS prior to loading of the labeling reaction is recommended as it will remove any traces of NaN_3 that is present in the storage buffer of the gel filtration beads.
10. To label antibodies or Fabs, the labeling kits are using N-hydroxysuccinimide ester fluorophores that react with the amine group at the tip of the side chain of lysines. This works fine with antibodies that do not harbor lysine residues in their binding site (paratope). If the binding capacity of the labeled antibodies or Fab (that can be easily tested by IF) is affected by this technique, we propose to set up a site-directed labeling. This labeling procedure consists in the preparation of (Fab')2 fragments, which can be specifically labeled at the typical cysteine residues in the C-terminal of the Fab' (hinge region) with maleimide-activated fluorophores upon mild reduction. Whilst this method preserves theoretically the antibody-binding site from any deleterious chemical alteration, it allows the addition of a maximum of two to three fluorophore molecules only per antibody or Fab.

Recipes

1. U2OS growth medium

   Supplement 450 ml of DMEM medium with 50 ml of FCS (10% FCS) and 40 µg/ml gentamicin

   Store at 4 °C

2. 4% PFA

   Dilute 10 ml of 16% PFA with 30 ml of PBS

   Store at -20 °C for a maximum of 1 month

3. 10x PBS
   a. Dissolve the whole content of one bottle in 1 L of filtered and sterile dH_2O
   b. Autoclave the solution to get sterile 10x PBS
   c. Store at RT
4. Triton X-100 solutions

   10% Triton X-100

   Dilute 1 ml of Triton X-100 in 9 ml of PBS

   Store at RT in the dark for several months

   0.1% Triton X-100

   Dilute 0.5 ml of 10% Triton X-100 in 50 ml of PBS

   Store at RT in the dark for several months

   0.02% Triton X-100

   Dilute 0.1 ml of 10% Triton X-100 in 50 ml of PBS

   Store at RT in the dark for several months

5. 1 M Tris-HCl pH 8.2
   a. Weigh 15.8 g of Tris-HCl and dissolve it in 100 ml of sterile dH_2O
   b. Adjust the pH to 8.2
   c. Filter sterilize the solution using a 0.22 µm filter
   d. Store at RT for several months
6. 0.1 M glycine-HCl pH 2.7
   a. Weigh 3.8 g of glycine and dissolve it in 500 ml of sterile dH_2O
   b. Adjust the pH to 2.7
   c. Filter sterilize the solution using a 0.22 µm filter
   d. Store at RT for several months

7. Sodium bicarbonate buffer

   1 M sodium bicarbonate

   a. Dissolve 42 g of sodium bicarbonate in 500 ml of sterile dH_2O
   b. Filter the solution through a 0.22 µm filter
   c. Store at 4 °C for a maximum of 2 weeks

   0.1 M sodium bicarbonate

   a. Dilute 100 ml of 1 M sodium bicarbonate with 900 ml of sterile dH_2O
   b. Filter the solution through a 0.22 µm filter
   c. Store at 4 °C for a maximum of 2 weeks

8. 2.5% Trypsin
   a. Weigh 1 g of Trypsin and dissolve it in 40 ml of PBS
   b. Filter the solution through a 0.22 µm filter to sterilize it
   c. Store at 4 °C for a maximum of one month

Acknowledgments

This work was supported by funds from CNRS, INSERM, University of Strasbourg, Ligue Régionale contre le Cancer (CCIRGE-BFC) (to EW), by the European Research Council (ERC) Advanced grant (ERC-2013-340551, Birtoaction) (to LT) and a grant ANR-10-LABX-0030-INRT, a French State fund managed by the Agence Nationale de la Recherche under the frame program Investissements d'Avenir ANR-10-IDEX-0002-02.

The present protocol is based on the following published research papers: Freund *et al.*, 2013; Desplancq *et al.*, 2016 and Conic *et al.*, 2018.

The authors declare no conflict of interest.

References

1. Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J. and Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science* 313(5793): 1642-1645.
2. Brees, C. and Fransen, M. (2014). A cost-effective approach to microporate mammalian cells with the neon transfection system. *Anal Biochem* 466: 49-50.
3. Burgess, A., Lorca, T. and Castro, A. (2012). Quantitative live imaging of endogenous DNA replication in mammalian cells. *PLoS One* 7(9): e45726.
4. Conic, S., Desplancq, D., Ferrand, A., Fischer, V., Heyer, V., Reina San Martin, B., Pontabry, J., Oulad-Abdelghani, M., Babu, N. K., Wright, G. D., Molina, N., Weiss, E. and Tora, L. (2018). Imaging of native transcription factors and histone phosphorylation at high resolution in live cells. *J Cell Biol* 217(4): 1537-1552.
5. Desplancq, D., Freund, G., Conic, S., Sibler, A. P., Didier, P., Stoessel, A., Oulad-Abdelghani, M., Vigneron, M., Wagner, J., Mely, Y., Chatton, B., Tora, L. and Weiss, E. (2016). Targeting the replisome with transduced monoclonal antibodies triggers lethal DNA replication stress in cancer cells. *Exp Cell Res* 342(2): 145-158.
6. Ellenberg, J., Lippincott-Schwartz, J. and Presley, J. F. (1999). Dual-colour imaging with GFP variants. *Trends Cell Biol* 9(2): 52-56.
7. Freund, G., Sibler, A. P., Desplancq, D., Oulad-Abdelghani, M., Vigneron, M., Gannon, J., Van Regenmortel, M. H. and Weiss, E. (2013). Targeting endogenous nuclear antigens by electrotransfer of monoclonal antibodies in living cells. *MAbs* 5(4): 518-522.
8. Hayashi-Takanaka, Y., Yamagata, K., Nozaki, N. and Kimura, H. (2009). Visualizing histone modifications in living cells: spatiotemporal dynamics of H3 phosphorylation during interphase. *J Cell Biol* 187(6): 781-790.
9. Hayashi-Takanaka, Y., Yamagata, K., Wakayama, T., Stasevich, T. J., Kainuma, T., Tsurimoto, T., Tachibana, M., Shinkai, Y., Kurumizaka, H., Nozaki, N. and Kimura, H. (2011). Tracking epigenetic histone modifications in single

cells using fab-based live endogenous modification labeling. *Nucleic Acids Res* 39(15): 6475-6488.

10. Ratz, M., Testa, I., Hell, S. W. and Jakobs, S. (2015). CRISPR/Cas9-mediated endogenous protein tagging for RESOLFT super-resolution microscopy of living human cells. *Sci Rep* 5: 9592.
11. Sato, Y., Mukai, M., Ueda, J., Muraki, M., Stasevich, T. J., Horikoshi, N., Kujirai, T., Kita, H., Kimura, T., Hira, S., Okada, Y., Hayashi-Takanaka, Y., Obuse, C., Kurumizaka, H., Kawahara, A., Yamagata, K., Nozaki, N. and Kimura, H. (2013). Genetically encoded system to track histone modification *in vivo*. *Sci Rep* 3: 2436.
12. Schneider, A. F. L. and Hackenberger, C. P. R. (2017). Fluorescent labelling in living cells. *Curr Opin Biotechnol* 48: 61-68.
13. Schnell, U., Dijk, F., Sjollema, K. A. and Giepmans, B. N. (2012). Immunolabeling artifacts and the need for live-cell imaging. *Nat Methods* 9(2): 152-158.
14. Teng, K. W., Ishitsuka, Y., Ren, P., Youn, Y., Deng, X., Ge, P., Lee, S. H., Belmont, A. S. and Selvin, P. R. (2016). Labeling proteins inside living cells using external fluorophores for microscopy. *Elife* 5: e20378.
15. Teves, S. S., An, L., Hansen, A. S., Xie, L., Darzacq, X. and Tjian, R. (2016). A dynamic mode of mitotic bookmarking by transcription factors. *Elife* 5: e22280.

Please download Video 1 via the following link,

http://os.bio-protocol.org/doc/upprotocol/p2193/Abstract2193_20180712044303406/Video%201%20Elec%20procedure.mp4

BIBLIOGRAPHY

Bibliography

Abbe, E. (1873). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. Archiv f. mikrosk. Anatomie *9*, 413-418.

Adachi, N., and Lieber, M.R. (2002). Bidirectional gene organization: a common architectural feature of the human genome. Cell *109*, 807-809.

Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nature reviews. Genetics *13*, 720-731.

Ahn, S.H., Kim, M., and Buratowski, S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. Molecular cell *13*, 67-76.

Akhtar, M.S., Heidemann, M., Tietjen, J., Zhang, D., Chapman, R.D., Eick, D., and Ansari, A.Z. (2009). TFIIH kinase places bivalent marks on the carboxyl-terminal domain of RNA polymerase II. Molecular cell *34*, 387-393.

Alberti, S. (2017). Phase separation in biology. Current biology : CB *27*, R1097-R1102.

Alekseev, S., Nagy, Z., Sandoz, J., Weiss, A., Egly, J.-M., Le May, N., and Coin, F. (2017). Transcription without XPB Establishes a Unified Helicase-Independent Mechanism of Promoter Opening in Eukaryotic Gene Expression. Molecular cell *65*, 504-514.e4.

ALLFREY, V.G., FAULKNER, R., and MIRSKY, A.E. (1964). ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. Proceedings of the National Academy of Sciences of the United States of America *51*, 786-794.

Allis, C.D., Berger, S.L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., and Shiekhattar, R., et al. (2007). New nomenclature for chromatin-modifying enzymes. Cell *131*, 633-636.

Ame, J.C., Schreiber, V., Fraulob, V., Dolle, P., Murcia, G. de, and Niedergang, C.P. (2001). A bidirectional promoter connects the poly(ADP-ribose) polymerase 2 (PARP-2) gene to the gene for RNase P RNA. structure and expression of the mouse PARP-2 gene. The Journal of biological chemistry *276*, 11092-11099.

Anandapadamanaban, M., Andresen, C., Helander, S., Ohyama, Y., Siponen, M.I., Lundström, P., Kokubo, T., Ikura, M., Moche, M., and Sunnerhagen, M. (2013). High-resolution structure of TBP with TAF1 reveals anchoring patterns in transcriptional regulation. Nature structural & molecular biology *20*, 1008-1014.

Andegeko, Y., Moyal, L., Mittelman, L., Tsarfaty, I., Shiloh, Y., and Rotman, G. (2001). Nuclear retention of ATM at sites of DNA double strand breaks. The Journal of biological chemistry *276*, 38224-38230.

Andel III, F. (1999). Three-Dimensional Structure of the Human TFIID-IIA-IIB Complex. Science *286*, 2153-2156.

Antos, J.M., Ingram, J., Fang, T., Pishesha, N., Truttmann, M.C., and Ploegh, H.L. (2017). Site-Specific Protein Labeling via Sortase-Mediated Transpeptidation. Current protocols in protein science *89*, 15.3.1-15.3.19.

Autour, A., C Y Jeng, S., D Cawte, A., Abdolahzadeh, A., Galli, A., Panchapakesan, S.S.S., Rueda, D., Ryckelynck, M., and Unrau, P.J. (2018). Fluorogenic RNA Mango aptamers for imaging small non-coding RNAs in mammalian cells. Nature communications *9*, 656.

Axelrod, D., Koppel, D.E., Schlessinger, J., Elson, E., and Webb, W.W. (1976). Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. Biophysical Journal *16*, 1055-1069.

Baek, H.J., Kang, Y.K., and Roeder, R.G. (2006). Human Mediator enhances basal transcription by facilitating recruitment of transcription factor IIB during preinitiation complex assembly. The Journal of biological chemistry *281*, 15172-15181.

Bagby, S., Mal, T.K., Liu, D., Raddatz, E., Nakatani, Y., and Ikura, M. (2000). TFIIA-TAF regulatory interplay: NMR evidence for overlapping binding sites on TBP. FEBS letters *468*, 149-154.

Bajic, V.B., Tan, S.L., Christoffels, A., Schönbach, C., Lipovich, L., Yang, L., Hofmann, O., Kruger, A., Hide, W., and Kai, C., et al. (2006). Mice and men: their promoter properties. PLoS genetics *2*, e54.

Baker, S.P., and Grant, P.A. (2007). The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. Oncogene *26*, 5329-5340.

Bakkenist, C.J., and Kastan, M.B. (2003). DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. Nature *421*, 499-506.

Balasubramanian, R., Pray-Grant, M.G., Selleck, W., Grant, P.A., and Tan, S. (2002). Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation. The Journal of biological chemistry *277*, 7989-7995.

Ball, G., Demmerle, J., Kaufmann, R., Davis, I., Dobbie, I.M., and Schermelleh, L. (2015). SIMcheck: a Toolbox for Successful Super-resolution Structured Illumination Microscopy. Scientific reports *5*, 15915.

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299-308.

Bannister, A.J., Schneider, R., and Kouzarides, T. (2002). Histone methylation: dynamic or static? Cell *109*, 801-806.

Baptista, T., Grünberg, S., Minoungou, N., Koster, M.J.E., Timmers, H.T.M., Hahn, S., Devys, D., and Tora, L. (2017). SAGA Is a General Cofactor for RNA Polymerase II Transcription. Molecular cell *68*, 130-143.e5.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

Bates, M., Huang, B., Dempsey, G.T., and Zhuang, X. (2007). Multicolor super-resolution imaging with photo-switchable fluorescent probes. Science (New York, N.Y.) *317*, 1749-1753.

Benke, A., Olivier, N., Gunzenhäuser, J., and Manley, S. (2012). Multicolor single molecule tracking of stochastically active synthetic dyes. Nano letters *12*, 2619-2624.

Benoist, C., and Chambon, P. (1981). In vivo sequence requirements of the SV40 early promotor region. Nature *290*, 304-310.

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

Berglund, D.L., and Starkey, J.R. (1989). Isolation of viable tumor cells following introduction of labelled antibody to an intracellular oncogene product using electroporation. Journal of immunological methods *125*, 79-87.

Bermejo, R., Lai, M.S., and Foiani, M. (2012). Preventing replication stress to maintain genome stability: resolving conflicts between replication and transcription. Molecular cell *45*, 710-718.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., and Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Bertolotti, A., Lutz, Y., Heard, D.J., Chambon, P., and Tora, L. (1996). hTAF(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. The EMBO Journal *15*, 5022-5031.

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., and Long, R.M. (1998). Localization of ASH1 mRNA particles in living yeast. Molecular cell *2*, 437-445.

Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacino, J.S., Davidson, M.W., Lippincott-Schwartz, J., and Hess, H.F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. Science (New York, N.Y.) *313*, 1642-1645.

Bieniossek, C., Papai, G., Schaffitzel, C., Garzoni, F., Chaillet, M., Scheer, E., Papadopoulos, P., Tora, L., Schultz, P., and Berger, I. (2013). The architecture of human general transcription factor TFIID core complex. Nature *493*, 699-702.

Biggar, S.R., and Crabtree, G.R. (1999). Continuous and widespread roles for the Swi-Snf complex in transcription. The EMBO Journal *18*, 2254-2264.

Bleichenbacher, M., Tan, S., and Richmond, T.J. (2003). Novel Interactions Between the Components of Human and Yeast TFIIA/TBP/DNA Complexes. Journal of molecular biology *332*, 783-793.

Blosser, T.R., Yang, J.G., Stone, M.D., Narlikar, G.J., and Zhuang, X. (2009). Dynamics of nucleosome remodelling by individual ACF complexes. Nature *462*, 1022-1027.

Blythe, S.A., and Wieschaus, E.F. (2015). Zygotic Genome Activation Triggers the DNA Replication Checkpoint at the Midblastula Transition. Cell *160*, 1169-1181.

Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A.S., Yu, T.-K., Marie-Nelly, H., McSwiggen, D.T., Kokic, G., Dailey, G.M., and Cramer, P., et al. (2018). RNA polymerase II clustering through CTD phase separation.

Boeing, S., Rigault, C., Heidemann, M., Eick, D., and Meisterernst, M. (2010). RNA polymerase II C-terminal heptarepeat domain Ser-7 phosphorylation is established in a mediator-dependent fashion. The Journal of biological chemistry *285*, 188-196.

Boireau, S., Maiuri, P., Basyuk, E., La Mata, M. de, Knezevich, A., Pradet-Balade, B., Bäcker, V., Kornblihtt, A., Marcello, A., and Bertrand, E. (2007). The transcriptional cycle of HIV-1 in real-time and live cells. The Journal of cell biology *179*, 291-304.

Bonnet, J., Wang, C.-Y., Baptista, T., Vincent, S.D., Hsiao, W.-C., Stierle, M., Kao, C.-F., Tora, L., and Devys, D. (2014). The SAGA coactivator complex acts on the whole transcribed genome and is required for RNA polymerase II transcription. Genes & Development *28*, 1999-2012.

Bornfleth, Satzler, Eils, and Cremer (1998). High-precision distance measurements and volume-conserving segmentation of objects near and below the resolution limit in three-dimensional confocal fluorescence microscopy. Journal of microscopy *189*, 118-136.

Brand, M. (1999). Three-Dimensional Structures of the TAFII-Containing Complexes TFIID and TFTC. Science *286*, 2151-2153.

Brown, J.M., Green, J., das Neves, R.P., Wallace, H.A.C., Smith, A.J.H., Hughes, J., Gray, N., Taylor, S., Wood, W.G., and Higgs, D.R., et al. (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. The Journal of cell biology *182*, 1083-1097.

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell *84*, 843-851.

Burke, T.W., and Kadonaga, J.T. (1996). Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes & Development *10*, 711-724.

Burke, T.W., and Kadonaga, J.T. (1997). The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila. Genes & Development *11*, 3020-3031.

Burley, S.K., and Roeder, R.G. (1996). Biochemistry and structural biology of transcription factor IID (TFIID). Annual review of biochemistry *65*, 769-799.

Burma, S., Chen, B.P., Murphy, M., Kurimasa, A., and Chen, D.J. (2001). ATM phosphorylates histone H2AX in response to DNA double-strand breaks. The Journal of biological chemistry *276*, 42462-42467.

Burri, O., Laroche, T., Guiet, R., and Seitz, A. (2017). Correlative SIM-STORM Microscopy. Methods in molecular biology (Clifton, N.J.) *1663*, 95-103.

Burton, Z.F., Killeen, M., Sopta, M., Ortolan, L.G., and Greenblatt, J. (1988). RAP30/74: a general initiation factor that binds to RNA polymerase II. Molecular and Cellular Biology *8*, 1602-1613.

Burton, Z.F., Ortolan, L.G., and Greenblatt, J. (1986). Proteins that bind to RNA polymerase II are required for accurate initiation of transcription at the adenovirus 2 major late promoter. The EMBO Journal *5*, 2923-2930.

Bushnell, D.A., and Kornberg, R.D. (2003). Complete, 12-subunit RNA polymerase II at 4.1-Å resolution: Implications for the initiation of transcription. Proceedings of the National Academy of Sciences of the United States of America *100*, 6969-6973.

Bushnell, D.A., Westover, K.D., Davis, R.E., and Kornberg, R.D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. Science (New York, N.Y.) *303*, 983-988.

Čabart, P., Újvári, A., Pal, M., and Luse, D.S. (2011). Transcription factor TFIIF is not required for initiation by RNA polymerase II, but it is essential to stabilize transcription factor TFIIB in early elongation complexes. Proceedings of the National Academy of Sciences of the United States of America *108*, 15786-15791.

Calvo, O., and Manley, J.L. (2003). Strange bedfellows: polyadenylation factors at the promoter. Genes & Development *17*, 1321-1327.

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C.A.M., Taylor, M.S., Engström, P.G., and Frith, M.C., et al. (2006). Genome-wide analysis of mammalian promoter architecture and evolution. Nature genetics *38*, 626-635.

Cerase, A., Smeets, D., Tang, Y.A., Gdula, M., Kraus, F., Spivakov, M., Moindrot, B., Leleu, M., Tattermusch, A., and Demmerle, J., et al. (2014). Spatial separation of Xist RNA and polycomb proteins revealed by superresolution microscopy. Proceedings of the National Academy of Sciences of the United States of America *111*, 2235-2240.

Chagin, V.O., Casas-Delucchi, C.S., Reinhart, M., Schermelleh, L., Markaki, Y., Maiser, A., Bolius, J.J., Bensimon, A., Fillies, M., and Domaing, P., et al. (2016). 4D Visualization of replication foci in mammalian cells corresponding to individual replicons. Nature communications *7*, 11231.

Chakrabarti, R., Wylie, D.E., and Schuster, S.M. (1989). Transfer of monoclonal antibodies into mammalian cells by electroporation. The Journal of biological chemistry *264*, 15494-15500.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. (1994). Green fluorescent protein as a marker for gene expression. Science (New York, N.Y.) *263*, 802-805.

Chalkley, G.E., and Verrijzer, C.P. (1999). DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the initiator. The EMBO Journal *18*, 4835-4845.

Chang, Y., Wang, Y., Cui, Y., and Ge, B. (2016). Investigation of non-linear imaging in high-resolution transmission electron microscopy. Microscopy (Oxford, England) *65*, 465-472.

Chapman, R.D., Heidemann, M., Albert, T.K., Mailhammer, R., Flatley, A., Meisterernst, M., Kremmer, E., and Eick, D. (2007). Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. Science (New York, N.Y.) *318*, 1780-1782.

Chen, J., Zhang, Z., Li, L., Chen, B.-C., Revyakin, A., Hajj, B., Legant, W., Dahan, M., Lionnet, T., and Betzig, E., et al. (2014). Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. Cell *156*, 1274-1285.

Cheng, B., and Price, D.H. (2007). Properties of RNA polymerase II elongation complexes before and after the P-TEFb-mediated transition into productive elongation. The Journal of biological chemistry *282*, 21901-21912.

Chesnut, J.D., Stephens, J.H., and Dahmus, M.E. (1992). The interaction of RNA polymerase II with the adenovirus-2 major late promoter is precluded by phosphorylation of the C-terminal domain of subunit IIa. The Journal of biological chemistry *267*, 10500-10506.

Cheung, A.C.M., and Cramer, P. (2011). Structural basis of RNA polymerase II backtracking, arrest and reactivation. Nature *471*, 249-253.

Ching, R.W., Ahmed, K., Boutros, P.C., Penn, L.Z., and Bazett-Jones, D.P. (2013). Identifying gene locus associations with promyelocytic leukemia nuclear bodies using immuno-TRAP. The Journal of cell biology *201*, 325-335.

Cho, E.-J., Kobor, M.S., Kim, M., Greenblatt, J., and Buratowski, S. (2001). Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. Genes & Development *15*, 3319-3329.

Cho, W.-K., Jayanth, N., English, B.P., Inoue, T., Andrews, J.O., Conway, W., Grimm, J.B., Spille, J.-H., Lavis, L.D., and Lionnet, T., et al. (2016). RNA Polymerase II cluster dynamics predict mRNA output in living cells. eLife *5*.

Chowdhury, D., Keogh, M.-C., Ishii, H., Peterson, C.L., Buratowski, S., and Lieberman, J. (2005). gamma-H2AX dephosphorylation by protein phosphatase 2A facilitates DNA double-strand break repair. Molecular cell *20*, 801-809.

Chowdhury, D., Xu, X., Zhong, X., Ahmed, F., Zhong, J., Liao, J., Dykxhoorn, D.M., Weinstock, D.M., Pfeifer, G.P., and Lieberman, J. (2008). A PP4-phosphatase complex dephosphorylates γ-H2AX generated during DNA replication. Molecular cell *31*, 33-46.

Ciccia, A., and Elledge, S.J. (2010). The DNA damage response: making it safe to play with knives. Molecular cell *40*, 179-204.

Cimprich, K.A., and Cortez, D. (2008). ATR: an essential regulator of genome integrity. Nature reviews. Molecular cell biology *9*, 616-627.

Cisse, I.I., Izeddin, I., Causse, S.Z., Boudarene, L., Senecal, A., Muresan, L., Dugast-Darzacq, C., Hajj, B., Dahan, M., and Darzacq, X. (2013). Real-time dynamics of RNA polymerase II clustering in live human cells. Science (New York, N.Y.) *341*, 664-667.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annual review of biochemistry *78*, 273-304.

Cler, E., Papai, G., Schultz, P., and Davidson, I. (2009). Recent advances in understanding the structure and function of general transcription factor TFIID. Cellular and molecular life sciences : CMLS *66*, 2123-2134.

Coin, F., Oksenych, V., and Egly, J.-M. (2007). Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair. Molecular cell *26*, 245-256.

Conaway, J.W., Florens, L., Sato, S., Tomomori-Sato, C., Parmely, T.J., Yao, T., Swanson, S.K., Banks, C.A.S., Washburn, M.P., and Conaway, R.C. (2005). The mammalian Mediator complex. FEBS letters *579*, 904-908.

Conaway, R.C., and Conaway, J.W. (1993). General initiation factors for RNA polymerase II. Annual review of biochemistry *62*, 161-190.

Conaway, R.C., and Conaway, J.W. (2013). The Mediator complex and transcription elongation. Biochimica et biophysica acta *1829*, 69-75.

Cook, P.R. (1999). The organization of replication and transcription. Science (New York, N.Y.) *284*, 1790-1795.

Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C., and Chambon, P. (1980). Promoter sequences of eukaryotic protein-coding genes. Science (New York, N.Y.) *209*, 1406-1414.

Corden, J.L., Cadena, D.L., Ahearn, J.M., and Dahmus, M.E. (1985). A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. Proceedings of the National Academy of Sciences of the United States of America *82*, 7934-7938.

Courey, A.J., and Tjian, R. (1988). Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell *55*, 887-898.

Courtête, J., Sibler, A.-P., Zeder-Lutz, G., Dalkara, D., Oulad-Abdelghani, M., Zuber, G., and Weiss, E. (2007). Suppression of cervical carcinoma cell growth by intracytoplasmic codelivery of anti-oncoprotein E6 antibody and small interfering RNA. Molecular cancer therapeutics *6*, 1728-1735.

Cramer, P. (2004). Structure and function of RNA polymerase II. Advances in protein chemistry *67*, 1-42.

Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science (New York, N.Y.) *292*, 1863-1876.

Danino, Y.M., Even, D., Ideses, D., and Juven-Gershon, T. (2015). The core promoter: At the heart of gene expression. Biochimica et biophysica acta *1849*, 1116-1131.

Darzacq, X., Shav-Tal, Y., Turris, V. de, Brody, Y., Shenoy, S.M., Phair, R.D., and Singer, R.H. (2007). In vivo dynamics of RNA polymerase II transcription. Nature structural & molecular biology *14*, 796-806.

Davidson, I. (2003). The genetics of TBP and TBP-related factors. Trends in biochemical sciences *28*, 391-398.

Dean, K.M., and Palmer, A.E. (2014). Advances in fluorescence labeling strategies for dynamic cellular imaging. Nature chemical biology *10*, 512-523.

Deng, W., and Roberts, S.G.E. (2005). A core promoter element downstream of the TATA box that is recognized by TFIIB. Genes & Development *19*, 2418-2423.

Dichtl, B., Blank, D., Ohnacker, M., Friedlein, A., Roeder, D., Langen, H., and Keller, W. (2002). A role for SSU72 in balancing RNA polymerase II transcription elongation and termination. Molecular cell *10*, 1139-1150.

Dikstein, R., Zhou, S., and Tjian, R. (1996). Human TAFII 105 is a cell type-specific TFIID subunit related to hTAFII130. Cell *87*, 137-146.

Dimitrova, D.S. (2011). DNA replication initiation patterns and spatial dynamics of the human ribosomal RNA gene loci. Journal of cell science *124*, 2743-2752.

Donner, A.J., Ebmeier, C.C., Taatjes, D.J., and Espinosa, J.M. (2010). CDK8 is a positive regulator of transcriptional elongation within the serum response network. Nature structural & molecular biology *17*, 194-201.

Donnert, G., Keller, J., Wurm, C.A., Rizzoli, S.O., Westphal, V., Schönle, A., Jahn, R., Jakobs, S., Eggeling, C., and Hell, S.W. (2007). Two-color far-field fluorescence nanoscopy. Biophysical Journal *92*, L67-9.

Doshi, R., Chen, B.R., Vibat, C.R.T., Huang, N., Lee, C.-W., and Chang, G. (2014). In vitro nanobody discovery for integral membrane protein targets. Scientific reports *4*, 6760.

Dürr, H., and Hopfner, K.-P. (2006). Structure-function analysis of SWI2/SNF2 enzymes. Methods in enzymology *409*, 375-388.

Dyba, M., and Hell, S.W. (2002). Focal spots of size lambda/23 open up far-field fluorescence microscopy at 33 nm axial resolution. Phys. Rev. Lett. *88*, 163901.

Dyba, M., Jakobs, S., and Hell, S.W. (2003). Immunofluorescence stimulated emission depletion microscopy. Nature biotechnology *21*, 1303-1304.

Egloff, S., O'Reilly, D., Chapman, R.D., Taylor, A., Tanzhaus, K., Pitts, L., Eick, D., and Murphy, S. (2007). Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. Science (New York, N.Y.) *318*, 1777-1779.

Egner, A., Geisler, C., Middendorff, C. von, Bock, H., Wenzel, D., Medda, R., Andresen, M., Stiel, A.C., Jakobs, S., and Eggeling, C., et al. (2007). Fluorescence nanoscopy in whole cells by asynchronous localization of photoswitching emitters. Biophysical Journal *93*, 3285-3290.

El Kaderi, B., Medler, S., Raghunayakula, S., and Ansari, A. (2009). Gene looping is conferred by activator-dependent interaction of transcription initiation and termination machineries. The Journal of biological chemistry *284*, 25015-25025.

Ernst, J., and Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. Nature biotechnology *28*, 817-825.

Ernst, J., and Kellis, M. (2013). Interplay between chromatin state, regulator binding, and regulatory motifs in six human cell types. Genome research *23*, 1142-1154.

Evans, R., Fairley, J.A., and Roberts, S.G.E. (2001). Activator-mediated disruption of sequence-specific DNA contacts by the general transcription factor TFIIB. Genes & Development *15*, 2945-2949.

Eychenne, T., Novikova, E., Barrault, M.-B., Alibert, O., Boschiero, C., Peixeiro, N., Cornu, D., Redeker, V., Kuras, L., and Nicolas, P., et al. (2016). Functional interplay between Mediator and TFIIB in preinitiation complex assembly in relation to promoter architecture. Genes & Development *30*, 2119-2132.

Famulok, M., Hartig, J.S., and Mayer, G. (2007). Functional aptamers and aptazymes in biotechnology, diagnostics, and therapy. Chemical reviews *107*, 3715-3743.

Feaver, W.J., Gileadi, O., Li, Y., and Kornberg, R.D. (1991). CTD kinase associated with yeast RNA polymerase II initiation factor b. Cell *67*, 1223-1230.

Femino, A.M., Fay, F.S., Fogarty, K., and Singer, R.H. (1998). Visualization of single RNA transcripts in situ. Science (New York, N.Y.) *280*, 585-590.

Fernandez-Capetillo, O., Chen, H.-T., Celeste, A., Ward, I., Romanienko, P.J., Morales, J.C., Naka, K., Xia, Z., Camerini-Otero, R.D., and Motoyama, N., et al. (2002). DNA damage-induced G2-M checkpoint activation by histone H2AX and 53BP1. Nature cell biology *4*, 993-997.

Fishburn, J., Tomko, E., Galburt, E., and Hahn, S. (2015). Double-stranded DNA translocase activity of transcription factor TFIIH and the mechanism of RNA polymerase II open complex formation. Proceedings of the National Academy of Sciences of the United States of America *112*, 3961-3966.

FitzGerald, P.C., Sturgill, D., Shyakhtenko, A., Oliver, B., and Vinson, C. (2006). Comparative genomics of Drosophila and human core promoters. Genome biology *7*, R53.

Fölling, J., Bossi, M., Bock, H., Medda, R., Wurm, C.A., Hein, B., Jakobs, S., Eggeling, C., and Hell, S.W. (2008). Fluorescence nanoscopy by ground-state depletion and single-molecule return. Nature methods *5*, 943-945.

Freund, G., Sibler, A.-P., Desplancq, D., Oulad-Abdelghani, M., Vigneron, M., Gannon, J., van Regenmortel, M.H., and Weiss, E. (2013). Targeting endogenous nuclear antigens by electrotransfer of monoclonal antibodies in living cells. mAbs *5*, 518-522.

Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T., and Peterlin, B.M. (2004). Dynamics of Human Immunodeficiency Virus Transcription: P-TEFb Phosphorylates RD and Dissociates Negative Effectors from the Transactivation Response Element. Molecular and Cellular Biology *24*, 787-795.

Gaillard, H., Herrera-Moyano, E., and Aguilera, A. (2013). Transcription-associated genome instability. Chemical reviews *113*, 8638-8661.

Ganem, C., Devaux, F., Torchet, C., Jacq, C., Quevillon-Cheruel, S., Labesse, G., Facca, C., and Faye, G. (2003). Ssu72 is a phosphatase essential for transcription

termination of snoRNAs and specific mRNAs in yeast. The EMBO Journal *22*, 1588-1598.

Gangaraju, V.K., and Bartholomew, B. (2007). Mechanisms of ATP dependent chromatin remodeling. Mutation research *618*, 3-17.

Garbett, K.A., Tripathi, M.K., Cencki, B., Layer, J.H., and Weil, P.A. (2007). Yeast TFIID serves as a coactivator for Rap1p by direct protein-protein interaction. Molecular and Cellular Biology *27*, 297-311.

García-Muse, T., and Aguilera, A. (2016). Transcription-replication conflicts: how they occur and how they are resolved. Nature reviews. Molecular cell biology *17*, 553-563.

Gautier, A., Juillerat, A., Heinis, C., Corrêa, I.R., Kindermann, M., Beaufils, F., and Johnsson, K. (2008). An engineered protein tag for multiprotein labeling in living cells. Chemistry & biology *15*, 128-136.

Gebhardt, J.C.M., Suter, D.M., Roy, R., Zhao, Z.W., Chapman, A.R., Basu, S., Maniatis, T., and Xie, X.S. (2013). Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. Nature methods *10*, 421-426.

Germier, T., Kocanova, S., Walther, N., Bancaud, A., Shaban, H.A., Sellou, H., Politi, A.Z., Ellenberg, J., Gallardo, F., and Bystricky, K. (2017). Real-Time Imaging of a Single Gene Reveals Transcription-Initiated Local Confinement. Biophysical Journal *113*, 1383-1394.

Gershenzon, N.I., Trifonov, E.N., and Ioshikhes, I.P. (2006). The features of Drosophila core promoters revealed by statistical analysis. BMC genomics *7*, 161.

Ghosh, A., Shuman, S., and Lima, C.D. (2008). The structure of Fcp1, an essential RNA polymerase II CTD phosphatase. Molecular cell *32*, 478-490.

Ghosh, R.N., and Webb, W.W. (1994). Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules. Biophysical Journal *66*, 1301-1318.

Gibbons, B.J., Brignole, E.J., Azubel, M., Murakami, K., Voss, N.R., Bushnell, D.A., Asturias, F.J., and Kornberg, R.D. (2012). Subunit architecture of general transcription factor TFIIH. Proceedings of the National Academy of Sciences of the United States of America *109*, 1949-1954.

Giglia-Mari, G., Zotter, A., and Vermeulen, W. (2011). DNA damage response. Cold Spring Harbor perspectives in biology *3*, a000745.

Glover-Cutter, K., Larochelle, S., Erickson, B., Zhang, C., Shokat, K., Fisher, R.P., and Bentley, D.L. (2009). TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. Molecular and Cellular Biology *29*, 5455-5464.

Godin, A.G., Lounis, B., and Cognet, L. (2014). Super-resolution microscopy approaches for live cell imaging. Biophysical Journal *107*, 1777-1784.

Goldknopf, I.L., and Busch, H. (1977). Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24. Proceedings of the National Academy of Sciences of the United States of America *74*, 864-868.

Goodrich, J.A., and Tjian, R. (2010). Unexpected Roles for Core Promoter Recognition Factors in Cell-type Specific Transcription and Gene Regulation. Nature reviews. Genetics *11*, 549-558.

Govind, C.K., Zhang, F., Qiu, H., Hofmeyer, K., and Hinnebusch, A.G. (2007). Gcn5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions. Molecular cell *25*, 31-42.

Gowrishankar, J., Leela, J.K., and Anupama, K. (2013). R-loops in bacterial transcription: Their causes and consequences. Transcription *4*, 153-157.

Gregoretti, I.V., Lee, Y.-M., and Goodson, H.V. (2004). Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis. Journal of molecular biology *338*, 17-31.

Grob, P., Cruse, M.J., Inouye, C., Peris, M., Penczek, P.A., Tjian, R., and Nogales, E. (2006). Cryo-electron microscopy studies of human TFIID: conformational breathing in the integration of gene regulatory cues. Structure (London, England : 1993) *14*, 511-520.

Grünberg, S., and Hahn, S. (2013). Structural insights into transcription initiation by RNA polymerase II. Trends in biochemical sciences *38*.

Grünberg, S., Warfield, L., and Hahn, S. (2012). Architecture of the RNA polymerase II preinitiation complex and mechanism of ATP-dependent promoter opening. Nature structural & molecular biology *19*, 788-796.

Gustafsson, M.G. (2000). Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. Journal of microscopy *198*, 82-87.

Gustafsson, M.G.L., Shao, L., Carlton, P.M., Wang, C.J.R., Golubovskaya, I.N., Cande, W.Z., Agard, D.A., and Sedat, J.W. (2008). Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. Biophysical Journal *94*, 4957-4970.

Ha, T. (2001). Single-molecule fluorescence resonance energy transfer. Methods (San Diego, Calif.) *25*, 78-86.

Hager, G.L., McNally, J.G., and Misteli, T. (2009). Transcription dynamics. Molecular cell *35*, 741-753.

Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. Nature structural & molecular biology *11*, 394-403.

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Songa, E.B., Bendahman, N., and Hamers, R. (1993). Naturally occurring antibodies devoid of light chains. Nature *363*, 446-448.

Hamperl, S., and Cimprich, K.A. (2016). Conflict resolution in the genome: how transcription and replication make it work. Cell *167*, 1455-1467.

Hanasoge, S., and Ljungman, M. (2007). H2AX phosphorylation after UV irradiation is triggered by DNA repair intermediates and is mediated by the ATR kinase. Carcinogenesis *28*, 2298-2304.

Hansen, S.K., Takada, S., Jacobson, R.H., Lis, J.T., and Tjian, R. (1997). Transcription properties of a cell type-specific TATA-binding protein, TRF. Cell *91*, 71-83.

Hantsche, M., and Cramer, P. (2017). Conserved RNA polymerase II initiation complex structure. Current opinion in structural biology *47*, 17-22.

Hargreaves, D.C., and Crabtree, G.R. (2011). ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. Cell Research *21*, 396-420.

Harmsen, M.M., and Haard, H.J. de (2007). Properties, production, and applications of camelid single-domain antibody fragments. Applied microbiology and biotechnology *77*, 13-22.

Hayashi-Takanaka, Y., Yamagata, K., Nozaki, N., and Kimura, H. (2009). Visualizing histone modifications in living cells: spatiotemporal dynamics of H3 phosphorylation during interphase. The Journal of cell biology *187*, 781-790.

Hayashi-Takanaka, Y., Yamagata, K., Wakayama, T., Stasevich, T.J., Kainuma, T., Tsurimoto, T., Tachibana, M., Shinkai, Y., Kurumizaka, H., and Nozaki, N., et al. (2011). Tracking epigenetic histone modifications in single cells using Fab-based live endogenous modification labeling. Nucleic acids research *39*, 6475-6488.

He, X., Khan, A.U., Cheng, H., Pappas, D.L., Hampsey, M., and Moore, C.L. (2003). Functional interactions between the transcription and mRNA 3′ end processing machineries mediated by Ssu72 and Sub1. Genes & Development *17*, 1030-1042.

He, Y., Fang, J., Taatjes, D.J., and Nogales, E. (2013). Structural visualization of key steps in human transcription initiation. Nature *495*, 481-486.

Heilemann, M., Herten, D.P., Heintzmann, R., Cremer, C., Muller, C., Tinnefeld, P., Weston, K.D., Wolfrum, J., and Sauer, M. (2002). High-resolution colocalization of single dye molecules by fluorescence lifetime imaging microscopy. Analytical chemistry *74*, 3511-3517.

Heilemann, M., van de Linde, S., Schüttpelz, M., Kasper, R., Seefeldt, B., Mukherjee, A., Tinnefeld, P., and Sauer, M. (2008). Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. Angewandte Chemie (International ed. in English) *47*, 6172-6176.

Heintzmann, R., and Cremer, C.G. (1999). Laterally modulated excitation microscopy: improvement of resolution by using a diffraction grating. In Optical Biopsies and Microscopic Techniques III, I.J. Bigio, H. Schneckenburger, J. Slavik, K. Svanberg and P.M. Viallet, eds. (SPIE), pp. 185–196.

Hell, S.W., and Wichmann, J. (1994). Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. Optics letters *19*, 780-782.

Helmrich, A., Ballarino, M., Nudler, E., and Tora, L. (2013). Transcription-replication encounters, consequences and genomic instability. Nature structural & molecular biology *20*, 412-418.

Heo, K., Kim, H., Choi, S.H., Choi, J., Kim, K., Gu, J., Lieber, M.R., Yang, A.S., and An, W. (2008). FACT-mediated exchange of histone variant H2AX regulated by phosphorylation of H2AX and ADP-ribosylation of Spt16. Molecular cell *30*, 86-97.

Hess, S.T., Girirajan, T.P.K., and Mason, M.D. (2006). Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. Biophysical Journal *91*, 4258-4272.

Hirose, Y., and Ohkuma, Y. (2007). Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the integrated events of eucaryotic gene expression. Journal of biochemistry *141*, 601-608.

Holstege, F.C., van der Vliet, P.C., and Timmers, H.T. (1996). Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. The EMBO Journal *15*, 1666-1677.

Hsin, J.-P., and Manley, J.L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes & Development *26*, 2119-2137.

Hsin, J.-P., Sheth, A., and Manley, J.L. (2011). RNAP II CTD Phosphorylated on Threonine 4 Is Required for Histone mRNA 3′ end Processing. Science (New York, N.Y.) *334*, 683-686.

Hu, Z., and Tee, W.-W. (2017). Enhancers and chromatin structures: regulatory hubs in gene expression and diseases. Bioscience reports *37*.

Huang, B., Jones, S.A., Brandenburg, B., and Zhuang, X. (2008). Whole-cell 3D STORM reveals interactions between cellular structures with nanometer-scale resolution. Nature methods *5*, 1047-1052.

Huisinga, K.L., and Pugh, B.F. (2004). A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae. Molecular cell *13*, 573-585.

Hyman, A.A., Weber, C.A., and Jülicher, F. (2014). Liquid-liquid phase separation in biology. Annual review of cell and developmental biology *30*, 39-58.

Imbalzano, A.N., Zaret, K.S., and Kingston, R.E. (1994). Transcription factor (TF) IIB and TFIIA can independently increase the affinity of the TATA-binding protein for DNA. The Journal of biological chemistry *269*, 8280-8286.

Jackson, D.A., Hassan, A.B., Errington, R.J., and Cook, P.R. (1993). Visualization of focal sites of transcription within human nuclei. The EMBO Journal *12*, 1059-1065.

Janicki, S.M., Tsukamoto, T., Salghetti, S.E., Tansey, W.P., Sachidanandam, R., Prasanth, K.V., Ried, T., Shav-Tal, Y., Bertrand, E., and Singer, R.H., et al. (2004). From silencing to gene expression: real-time analysis in single cells. Cell *116*, 683-698.

Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., and Smale, S.T. (1994). DNA sequence requirements for transcriptional initiator activity in mammalian cells. Molecular and Cellular Biology *14*, 116-127.

Jiang, Y.W., Veschambre, P., Erdjument-Bromage, H., Tempst, P., Conaway, J.W., Conaway, R.C., and Kornberg, R.D. (1998). Mammalian mediator of transcriptional regulation and its possible role as an end-point of signal transduction pathways. Proceedings of the National Academy of Sciences of the United States of America *95*, 8538-8543.

Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.-A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature *503*, 290-294.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science (New York, N.Y.) *337*, 816-821.

Jishage, M., Malik, S., Wagner, U., Uberheide, B., Ishihama, Y., Hu, X., Chait, B.T., Gnatt, A., Ren, B., and Roeder, R.G. (2012). Transcriptional regulation by Pol II(G) involving mediator and competitive interactions of Gdown1 and TFIIF with Pol II. Molecular cell *45*, 51-63.

Johnson, K.M., and Carey, M. (2003). Assembly of a mediator/TFIID/TFIIA complex bypasses the need for an activator. Current biology : CB *13*, 772-777.

Johnson, K.M., Wang, J., Smallwood, A., Arayata, C., and Carey, M. (2002). TFIID and human mediator coactivator complexes assemble cooperatively on promoter DNA. Genes & Development *16*, 1852-1863.

Jones, S.A., Shim, S.-H., He, J., and Zhuang, X. (2011). Fast, three-dimensional super-resolution imaging of live cells. Nature methods *8*, 499-508.

Juven-Gershon, T., Hsu, J.-Y., and Kadonaga, J.T. (2006). Perspectives on the RNA polymerase II core promoter. Biochemical Society transactions *34*, 1047-1050.

Juven-Gershon, T., and Kadonaga, J.T. (2009). Regulation of Gene Expression via the Core Promoter and the Basal Transcriptional Machinery. Developmental biology *339*, 225-229.

Kadonaga, J.T. (2012). Perspectives on the RNA polymerase II core promoter. Wiley interdisciplinary reviews. Developmental biology *1*, 40-51.

Kang, J.J., Auble, D.T., Ranish, J.A., and Hahn, S. (1995). Analysis of the yeast transcription factor TFIIA: distinct functional regions and a polymerase II-specific role in basal and activated transcription. Molecular and Cellular Biology *15*, 1234-1243.

Kaplan, C.D., Laprade, L., and Winston, F. (2003). Transcription elongation factors repress transcription initiation from cryptic sites. Science (New York, N.Y.) *301*, 1096-1099.

Kelly, W.G., Dahmus, M.E., and Hart, G.W. (1993). RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc. The Journal of biological chemistry *268*, 10416-10424.

Kim, B., Nesvizhskii, A.I., Rani, P.G., Hahn, S., Aebersold, R., and Ranish, J.A. (2007). The transcription elongation factor TFIIS is a component of RNA polymerase II preinitiation complexes. Proceedings of the National Academy of Sciences of the United States of America *104*, 16068-16073.

Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D., and Ren, B. (2005). A high-resolution map of active promoters in the human genome. Nature *436*, 876-880.

Kim, Y.J., Björklund, S., Li, Y., Sayre, M.H., and Kornberg, R.D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell *77*, 599-608.

Kimura, H., and Cook, P.R. (2001). Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. The Journal of cell biology *153*, 1341-1353.

Kimura, H., Sugaya, K., and Cook, P.R. (2002). The transcription cycle of RNA polymerase II in living cells. The Journal of cell biology *159*, 777-782.

Kimura, H., Takizawa, N., Allemand, E., Hori, T., Iborra, F.J., Nozaki, N., Muraki, M., Hagiwara, M., Krainer, A.R., and Fukagawa, T., et al. (2006). A novel histone exchange factor, protein phosphatase 2Cgamma, mediates the exchange and dephosphorylation of H2A-H2B. The Journal of cell biology *175*, 389-400.

Kinner, A., Wu, W., Staudt, C., and Iliakis, G. (2008). Gamma-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. Nucleic acids research *36*, 5678-5694.

Kitayner, M., Rozenberg, H., Rohs, R., Suad, O., Rabinovich, D., Honig, B., and Shakked, Z. (2010). Diversity in DNA recognition by p53 revealed by crystal structures with Hoogsteen base pairs. Nature structural & molecular biology *17*, 423-429.

Klar, T.A., and Hell, S.W. (1999). Subdiffraction resolution in far-field fluorescence microscopy. Optics letters *24*, 954-956.

Klar, T.A., Jakobs, S., Dyba, M., Egner, A., and Hell, S.W. (2000). Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. Proceedings of the National Academy of Sciences of the United States of America *97*, 8206-8210.

Klein, T., Löschberger, A., Proppert, S., Wolter, S., van de Linde, S., and Sauer, M. (2011). Live-cell dSTORM with SNAP-tag fusion proteins. Nature methods *8*, 7-9.

Kleinsmith, L.J., ALLFREY, V.G., and MIRSKY, A.E. (1966). Phosphorylation of nuclear protein early in the course of gene activation in lymphocytes. Science (New York, N.Y.) *154*, 780-781.

Koleske, A.J., and Young, R.A. (1994). An RNA polymerase II holoenzyme responsive to activators. Nature *368*, 466-469.

Kostrewa, D., Zeller, M.E., Armache, K.-J., Seizl, M., Leike, K., Thomm, M., and Cramer, P. (2009). RNA polymerase II-TFIIB structure and mechanism of transcription initiation. Nature *462*, 323-330.

Kotani, T., Miyake, T., Tsukihashi, Y., Hinnebusch, A.G., Nakatani, Y., Kawaichi, M., and Kokubo, T. (1998). Identification of highly conserved amino-terminal segments of dTAFII230 and yTAFII145 that are functionally interchangeable for inhibiting TBP-DNA interactions in vitro and in promoting yeast cell growth in vivo. The Journal of biological chemistry *273*, 32254-32264.

Koutelou, E., Hirsch, C.L., and Dent, S.Y.R. (2010). Multiple faces of the SAGA complex. Current opinion in cell biology *22*, 374-382.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Krah, S., Schröter, C., Zielonka, S., Empting, M., Valldorf, B., and Kolmar, H. (2016). Single-domain antibodies for biomedical applications. Immunopharmacology and immunotoxicology *38*, 21-28.

Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C., and Hampsey, M. (2004). Ssu72 Is an RNA polymerase II CTD phosphatase. Molecular cell *14*, 387-394.

Kuehner, J.N., Pearson, E.L., and Moore, C. (2011). Unravelling the means to an end: RNA polymerase II transcription termination. Nature reviews. Molecular cell biology *12*, 283-294.

Kuldell, N.H., and Buratowski, S. (1997). Genetic analysis of the large subunit of yeast transcription factor IIE reveals two regions with distinct functions. Molecular and Cellular Biology *17*, 5288-5298.

Kulkarni, M.M., and Arnosti, D.N. (2003). Information display by transcriptional enhancers. Development (Cambridge, England) *130*, 6569-6575.

Kutach, A.K., and Kadonaga, J.T. (2000). The Downstream Promoter Element DPE Appears To Be as Widely Used as the TATA Box in Drosophila Core Promoters. Molecular and Cellular Biology *20*, 4754-4764.

Kwon, I., Kato, M., Xiang, S., Wu, L., Theodoropoulos, P., Mirzaei, H., Han, T., Xie, S., Corden, J.L., and McKnight, S.L. (2013). Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. Cell *155*, 1049-1060.

Lagrange, T., Kapanidis, A.N., Tang, H., Reinberg, D., and Ebright, R.H. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes & Development *12*, 34-44.

Latt, S.A., and Wohlleb, J.C. (1975). Optical studies of the interaction of 33258 Hoechst with DNA, chromatin, and metaphase chromosomes. Chromosoma *52*, 297-316.

Le Treut, G., Képès, F., and Orland, H. (2016). Phase Behavior of DNA in the Presence of DNA-Binding Proteins. Biophysical Journal *110*, 51-62.

Lee, K.K., and Workman, J.L. (2007). Histone acetyltransferase complexes: one size doesn't fit all. Nature reviews. Molecular cell biology *8*, 284-295.

Legant, W.R., Shao, L., Grimm, J.B., Brown, T.A., Milkie, D.E., Avants, B.B., Lavis, L.D., and Betzig, E. (2016). High-density three-dimensional localization microscopy across large volumes. Nature methods *13*, 359-365.

Lenhard, B., Sandelin, A., and Carninci, P. (2012). Metazoan promoters: emerging characteristics and insights into transcriptional regulation. Nature reviews. Genetics *13*, 233-245.

Levine, M., Cattoglio, C., and Tjian, R. (2014). Looping back to leap forward: transcription enters a new era. Cell *157*, 13-25.

Liao, S.M., Zhang, J., Jeffery, D.A., Koleske, A.J., Thompson, C.M., Chao, D.M., Viljoen, M., van Vuuren, H.J., and Young, R.A. (1995). A kinase-cyclin pair in the RNA polymerase II holoenzyme. Nature *374*, 193-196.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., and Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science (New York, N.Y.) *326*, 289-293.

Lim, C.Y., Santoso, B., Boulay, T., Dong, E., Ohler, U., and Kadonaga, J.T. (2004). The MTE, a new core promoter element for transcription by RNA polymerase II. Genes & Development *18*, 1606-1617.

Limoli, C.L., Giedzinski, E., Bonner, W.M., and Cleaver, J.E. (2002). UV-induced replication arrest in the xeroderma pigmentosum variant leads to DNA double-strand breaks, gamma -H2AX formation, and Mre11 relocalization. Proceedings of the National Academy of Sciences of the United States of America *99*, 233-238.

Lin, Y.C., Choi, W.S., and Gralla, J.D. (2005). TFIIH XPB mutants suggest a unified bacterial-like mechanism for promoter opening but not escape. Nature structural & molecular biology *12*, 603-607.

Lippincott-Schwartz, J., and Patterson, G.H. (2008). Fluorescent proteins for photoactivation experiments. Methods in cell biology *85*, 45-61.

Lis, J.T., Mason, P., Peng, J., Price, D.H., and Werner, J. (2000). P-TEFb kinase recruitment and function at heat shock loci. Genes & Development *14*, 792-803.

Liu, W.-L., Coleman, R.A., Grob, P., King, D.S., Florens, L., Washburn, M.P., Geles, K.G., Yang, J.L., Ramey, V., and Nogales, E., et al. (2008). Structural changes in TAF4b-TFIID correlate with promoter selectivity. Molecular cell *29*, 81-91.

Liu, W.-L., Coleman, R.A., Ma, E., Grob, P., Yang, J.L., Zhang, Y., Dailey, G., Nogales, E., and Tjian, R. (2009). Structures of three distinct activator-TFIID complexes. Genes & Development *23*, 1510-1521.

Liu, X., Bushnell, D.A., and Kornberg, R.D. (2012). RNA Polymerase II Transcription: Structure and Mechanism. Biochimica et biophysica acta *1829*, 2-8.

Liu, X., Bushnell, D.A., Wang, D., Calero, G., and Kornberg, R.D. (2010). Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. Science (New York, N.Y.) *327*, 206-209.

Liu, Z., Lavis, L.D., and Betzig, E. (2015). Imaging live-cell dynamics and structure at the single-molecule level. Molecular cell *58*, 644-659.

Liu, Z., Legant, W.R., Chen, B.-C., Li, L., Grimm, J.B., Lavis, L.D., Betzig, E., and Tjian, R. (2014). 3D imaging of Sox2 enhancer clusters in embryonic stem cells. eLife *3*, e04236.

Liu, Z., and Tjian, R. (2018). Visualizing transcription factor dynamics in living cells. The Journal of cell biology *217*, 1181-1191.

Los, G.V., Encell, L.P., McDougall, M.G., Hartzell, D.D., Karassina, N., Zimprich, C., Wood, M.G., Learish, R., Ohana, R.F., and Urh, M., et al. (2008). HaloTag: a novel protein labeling technology for cell imaging and protein analysis. ACS chemical biology *3*, 373-382.

Louder, R.K., He, Y., López-Blanco, J.R., Fang, J., Chacón, P., and Nogales, E. (2016). Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. Nature *531*, 604-609.

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

Lu, H., Zawel, L., Fisher, L., Egly, J.M., and Reinberg, D. (1992). Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. Nature *358*, 641-645.

Lukas, J., Bartek, J., and Strauss, M. (1994). Efficient transfer of antibodies into mammalian cells by electroporation. Journal of immunological methods *170*, 255-259.

Lukyanov, K.A., Chudakov, D.M., Lukyanov, S., and Verkhusha, V.V. (2005). Innovation: Photoactivatable fluorescent proteins. Nature reviews. Molecular cell biology *6*, 885-891.

Magde, D., Elson, E., and Webb, W.W. (1972). Thermodynamic Fluctuations in a Reacting System—Measurement by Fluorescence Correlation Spectroscopy. Phys. Rev. Lett. *29*, 705-708.

Magnani, L., Eeckhoute, J., and Lupien, M. (2011). Pioneer factors: directing transcriptional regulators within the chromatin environment. Trends in genetics : TIG *27*, 465-474.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science (New York, N.Y.) *339*, 823-826.

Malik, S., Barrero, M.J., and Jones, T. (2007). Identification of a regulator of transcription elongation as an accessory factor for the human Mediator coactivator. Proceedings of the National Academy of Sciences of the United States of America *104*, 6182-6187.

Marfella, C.G.A., and Imbalzano, A.N. (2007). The Chd family of chromatin remodelers. Mutation research *618*, 30-40.

Markaki, Y., Gunkel, M., Schermelleh, L., Beichmanis, S., Neumann, J., Heidemann, M., Leonhardt, H., Eick, D., Cremer, C., and Cremer, T. (2010). Functional nuclear organization of transcription and DNA replication: a topographical marriage between chromatin domains and the interchromatin compartment. Cold Spring Harbor symposia on quantitative biology *75*, 475-492.

Marschall, A.L.J., Frenzel, A., Schirrmann, T., Schüngel, M., and Dübel, S. (2011). Targeting antibodies to the cytoplasm. mAbs *3*, 3-16.

Marschall, A.L.J., Zhang, C., Frenzel, A., Schirrmann, T., Hust, M., Perez, F., and Dübel, S. (2014). Delivery of antibodies to the cytosol: debunking the myths. mAbs *6*, 943-956.

Marshall, N.F., Peng, J., Xie, Z., and Price, D.H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. The Journal of biological chemistry *271*, 27176-27183.

Marshall, N.F., and Price, D.H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. The Journal of biological chemistry *270*, 12335-12338.

Martinez, E., Palhan, V.B., Tjernberg, A., Lymar, E.S., Gamper, A.M., Kundu, T.K., Chait, B.T., and Roeder, R.G. (2001). Human STAGA Complex Is a Chromatin-Acetylating Transcription Coactivator That Interacts with Pre-mRNA Splicing and DNA Damage-Binding Factors In Vivo. Molecular and Cellular Biology *21*, 6782-6795.

Maston, G.A., Zhu, L.J., Chamberlain, L., Lin, L., Fang, M., and Green, M.R. (2012). Non-canonical TAF complexes regulate active promoters in human embryonic stem cells. eLife *1*, e00068.

Maxon, M.E., Goodrich, J.A., and Tjian, R. (1994). Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance. Genes & Development *8*, 515-524.

Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Söding, J., and Cramer, P. (2010). Uniform transitions of the general RNA polymerase II transcription complex. Nature structural & molecular biology *17*, 1272-1278.

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

McNally, J.G., Müller, W.G., Walker, D., Wolford, R., and Hager, G.L. (2000). The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. Science (New York, N.Y.) *287*, 1262-1265.

Mirkin, E.V., and Mirkin, S.M. (2005). Mechanisms of transcription-replication collisions in bacteria. Molecular and Cellular Biology *25*, 888-895.

Mischo, H.E., Gómez-González, B., Grzechnik, P., Rondón, A.G., Wei, W., Steinmetz, L., Aguilera, A., and Proudfoot, N.J. (2011). Yeast Sen1 helicase protects the genome from transcription-associated instability. Molecular cell *41*, 21-32.

Misteli, T. (2001). Protein dynamics: implications for nuclear architecture and gene expression. Science (New York, N.Y.) *291*, 843-847.

Mohibullah, N., and Hahn, S. (2008). Site-specific cross-linking of TBP in vivo and in vitro reveals a direct functional interaction with the SAGA subunit Spt3. Genes & Development *22*, 2994-3006.

Moyle-Heyrman, G., Viswanathan, R., Widom, J., and Auble, D.T. (2012). Two-step mechanism for modifier of transcription 1 (Mot1) enzyme-catalyzed displacement of TATA-binding protein (TBP) from DNA. The Journal of biological chemistry *287*, 9002-9012.

Mukundan, B., and Ansari, A. (2011). Novel role for mediator complex subunit Srb5/Med18 in termination of transcription. The Journal of biological chemistry *286*, 37053-37057.

Müller, F., Zaucker, A., and Tora, L. (2010). Developmental regulation of transcription initiation: more than just changing the actors. Current opinion in genetics & development *20*, 533-540.

Murakami, K., Gibbons, B.J., Davis, R.E., Nagai, S., Liu, X., Robinson, P.J.J., Wu, T., Kaplan, C.D., and Kornberg, R.D. (2012). Tfb6, a previously unidentified subunit of the general transcription factor TFIIH, facilitates dissociation of Ssl2 helicase after transcription initiation. Proceedings of the National Academy of Sciences of the United States of America *109*, 4816-4821.

Murakami, K., Mattei, P.-J., Davis, R.E., Jin, H., Kaplan, C.D., and Kornberg, R.D. (2015). Uncoupling Promoter Opening from Start-Site Scanning. Molecular cell *59*, 133-138.

MURRAY, K. (1964). THE OCCURRENCE OF EPSILON-N-METHYL LYSINE IN HISTONES. Biochemistry *3*, 10-15.

Myers, L.C., and Kornberg, R.D. (2000). Mediator of transcriptional regulation. Annual review of biochemistry *69*, 729-749.

Nag, A., Narsinh, K., and Martinson, H.G. (2007). The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. Nature structural & molecular biology *14*, 662-669.

Nagy, Z., and Tora, L. (2007). Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. Oncogene *26*, 5341-5357.

Newhart, A., and Janicki, S.M. (2014). Seeing is believing: visualizing transcriptional dynamics in single cells. Journal of cellular physiology *229*, 259-265.

Ni, Z., Saunders, A., Fuda, N.J., Yao, J., Suarez, J.-R., Webb, W.W., and Lis, J.T. (2008). P-TEFb is critical for the maturation of RNA polymerase II into productive elongation in vivo. Molecular and Cellular Biology *28*, 1161-1170.

Ohkuma, Y., Sumimoto, H., Horikoshi, M., and Roeder, R.G. (1990). Factors involved in specific transcription by mammalian RNA polymerase II: purification and characterization of general transcription factor TFIIE. Proceedings of the National Academy of Sciences of the United States of America *87*, 9163-9167.

Ohler, U., Liao, G.-c., Niemann, H., and Rubin, G.M. (2002). Computational analysis of core promoters in the Drosophila genome. Genome biology *3*, RESEARCH0087.

Ohler, U., and Niemann, H. (2001). Identification and analysis of eukaryotic promoters: recent computational approaches. Trends in genetics : TIG *17*, 56-60.

Orekhova, A.S., and Rubtsov, P.M. (2013). Bidirectional promoters in the transcription of mammalian genomes. Biochemistry. Biokhimiia *78*, 335-341.

Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., and Reik, W., et al. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nature genetics *36*, 1065-1071.

Ossipow, V., Tassan, J.P., Nigg, E.A., and Schibler, U. (1995). A mammalian RNA polymerase II holoenzyme containing all components required for promoter-specific transcription initiation. Cell *83*, 137-146.

Ozer, J., Mitsouras, K., Zerby, D., Carey, M., and Lieberman, P.M. (1998). Transcription factor IIA derepresses TATA-binding protein (TBP)-associated factor inhibition of TBP-DNA binding. The Journal of biological chemistry *273*, 14293-14300.

Paige, J.S., Wu, K.Y., and Jaffrey, S.R. (2011). RNA mimics of green fluorescent protein. Science (New York, N.Y.) *333*, 642-646.

Pal, M., Ponticelli, A.S., and Luse, D.S. (2005). The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. Molecular cell *19*, 101-110.

Pan, G., and Greenblatt, J. (1994). Initiation of transcription by RNA polymerase II is limited by melting of the promoter DNA in the region immediately upstream of the initiation site. The Journal of biological chemistry *269*, 30101-30104.

Papai, G., Tripathi, M.K., Ruhlmann, C., Layer, J.H., Weil, P.A., and Schultz, P. (2010). TFIIA and the transactivator Rap1 cooperate to commit TFIID for transcription initiation. Nature *465*, 956-960.

Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., and Franzmann, T.M., et al. (2015). A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell *162*, 1066-1077.

Patterson, G.H., and Lippincott-Schwartz, J. (2002). A photoactivatable GFP for selective photolabeling of proteins and cells. Science (New York, N.Y.) *297*, 1873-1877.

Paulsen, R.D., and Cimprich, K.A. (2007). The ATR pathway: fine-tuning the fork. DNA repair *6*, 953-966.

Pavri, R., Zhu, B., Li, G., Trojer, P., Mandal, S., Shilatifard, A., and Reinberg, D. (2006). Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. Cell *125*, 703-717.

Peterson, C.L., and Herskowitz, I. (1992). Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. Cell *68*, 573-583.

Peterson, M.G., Inostroza, J., Maxon, M.E., Flores, O., Admon, A., Reinberg, D., and Tjian, R. (1991). Structure and functional properties of human general transcription factor IIE. Nature *354*, 369-373.

Petty, E., and Pillus, L. (2013). Balancing chromatin remodeling and histone modifications in transcription. Trends in genetics : TIG *29*, 621-629.

Phair, R.D., Gorski, S.A., and Misteli, T. (2004). Measurement of dynamic protein binding to chromatin in vivo, using photobleaching microscopy. Methods in enzymology *375*, 393-414.

Plaschka, C., Hantsche, M., Dienemann, C., Burzinski, C., Plitzko, J., and Cramer, P. (2016). Transcription initiation complex structures elucidate DNA opening. Nature *533*, 353-358.

Plaschka, C., Larivière, L., Wenzeck, L., Seizl, M., Hemann, M., Tegunov, D., Petrotchenko, E.V., Borchers, C.H., Baumeister, W., and Herzog, F., et al. (2015). Architecture of the RNA polymerase II-Mediator core initiation complex. Nature *518*, 376-380.

Ponjavic, J., Lenhard, B., Kai, C., Kawai, J., Carninci, P., Hayashizaki, Y., and Sandelin, A. (2006). Transcriptional and structural impact of TATA-initiation site spacing in mammalian core promoters. Genome biology *7*, R78.

Portugal, J., and Waring, M.J. (1988). Assignment of DNA binding sites for 4',6-diamidine-2-phenylindole and bisbenzimide (Hoechst 33258). A comparative footprinting study. Biochimica et biophysica acta *949*, 158-168.

Postupalenko, V., Desplancq, D., Orlov, I., Arntz, Y., Spehner, D., Mely, Y., Klaholz, B.P., Schultz, P., Weiss, E., and Zuber, G. (2015). Protein Delivery System Containing a Nickel-Immobilized Polymer for Multimerization of Affinity-Purified His-Tagged Proteins Enhances Cytosolic Transfer. Angewandte Chemie (International ed. in English) *54*, 10583-10586.

Powell, S.K., MacAlpine, H.K., Prinz, J.A., Li, Y., Belsky, J.A., and MacAlpine, D.M. (2015). Dynamic loading and redistribution of the Mcm2-7 helicase complex through the cell cycle. The EMBO Journal *34*, 531-543.

Prado, F., and Aguilera, A. (2005). Impairment of replication fork progression mediates RNA polII transcription-associated recombination. The EMBO Journal *24*, 1267-1276.

Price, D.H., Sluder, A.E., and Greenleaf, A.L. (1989). Dynamic interaction between a Drosophila transcription factor and RNA polymerase II. Molecular and Cellular Biology *9*, 1465-1475.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173-1183.

Rada-Iglesias, A., Bajpai, R., Prescott, S., Brugmann, S.A., Swigut, T., and Wysocka, J. (2012). Epigenomic annotation of enhancers predicts transcriptional regulators of human neural crest. Cell stem cell *11*, 633-648.

Rafalska-Metcalf, I.U., Powers, S.L., Joo, L.M., LeRoy, G., and Janicki, S.M. (2010). Single cell analysis of transcriptional activation dynamics. PloS one *5*, e10272.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. Cell *141*, 432-445.

Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. Nature methods *5*, 877-879.

Rayleigh (1903). On the Theory of Optical Images, with special reference to the Microscope. Journal of the Royal Microscopical Society *23*, 474-482.

Renaud, E., Martineau, P., and Guglielmi, L. (2017). Solubility Characterization and Imaging of Intrabodies Using GFP-Fusions. Methods in molecular biology (Clifton, N.J.) *1575*, 165-174.

Renner, D.B., Yamaguchi, Y., Wada, T., Handa, H., and Price, D.H. (2001). A highly purified RNA polymerase II elongation control system. The Journal of biological chemistry *276*, 42601-42609.

Robinson, P.J., Trnka, M.J., Bushnell, D.A., Davis, R., Mattei, P.-J., Burlingame, A.L., and Kornberg, R.D. (2016). Structure of a Complete Mediator-RNA Polymerase II Pre-Initiation Complex. Cell *166*, 1411-1422.e16.

Röder, R., Helma, J., Preiß, T., Rädler, J.O., Leonhardt, H., and Wagner, E. (2017). Intracellular Delivery of Nanobodies for Imaging of Target Proteins in Live Cells. Pharmaceutical research *34*, 161-174.

Rodríguez-Navarro, S. (2009). Insights into SAGA function during gene expression. EMBO reports *10*, 843-850.

Rohs, R., Jin, X., West, S.M., Joshi, R., Honig, B., and Mann, R.S. (2010). Origins of specificity in protein-DNA recognition. Annual review of biochemistry *79*, 233-269.

Rossignol, M., Kolb-Cheynel, I., and Egly, J.M. (1997). Substrate specificity of the cdk-activating kinase (CAK) is altered upon association with TFIIH. The EMBO Journal *16*, 1628-1637.

Rothbauer, U., Zolghadr, K., Tillib, S., Nowak, D., Schermelleh, L., Gahl, A., Backmann, N., Conrath, K., Muyldermans, S., and Cardoso, M.C., et al. (2006). Targeting and tracing antigens in live cells with fluorescent nanobodies. Nature methods *3*, 887-889.

Rust, M.J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nature methods *3*, 793-795.

Sainsbury, S., Bernecky, C., and Cramer, P. (2015). Structural basis of transcription initiation by RNA polymerase II. Nature reviews. Molecular cell biology *16*, 129-143.

Sainsbury, S., Niesser, J., and Cramer, P. (2013). Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. Nature *493*, 437-440.

Sanders, S.L., Jennings, J., Canutescu, A., Link, A.J., and Weil, P.A. (2002). Proteomics of the Eukaryotic Transcription Machinery: Identification of Proteins Associated with Components of Yeast TFIID by Multidimensional Mass Spectrometry. Molecular and Cellular Biology *22*, 4723-4738.

Saponaro, M., Kantidakis, T., Mitter, R., Kelly, G.P., Heron, M., Williams, H., Söding, J., Stewart, A., and Svejstrup, J.Q. (2014). RECQL5 controls transcript elongation and suppresses genome instability associated with transcription stress. Cell *157*, 1037-1049.

Sarge, K.D., and Park-Sarge, O.-K. (2005). Gene bookmarking: keeping the pages open. Trends in biochemical sciences *30*, 605-610.

Sato, Y., Mukai, M., Ueda, J., Muraki, M., Stasevich, T.J., Horikoshi, N., Kujirai, T., Kita, H., Kimura, T., and Hira, S., et al. (2013). Genetically encoded system to track histone modification in vivo. Scientific reports *3*, 2436.

Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J.H., Chambon, P., and Egly, J.M. (1993). DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. Science (New York, N.Y.) *260*, 58-63.

Schermelleh, L., Carlton, P.M., Haase, S., Shao, L., Winoto, L., Kner, P., Burke, B., Cardoso, M.C., Agard, D.A., and Gustafsson, M.G.L., et al. (2008). Subdiffraction

multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. Science (New York, N.Y.) *320*, 1332-1336.

Schermelleh, L., Heintzmann, R., and Leonhardt, H. (2010). A guide to super-resolution fluorescence microscopy. The Journal of cell biology *190*, 165-175.

Schmidt, R., Wurm, C.A., Jakobs, S., Engelhardt, J., Egner, A., and Hell, S.W. (2008). Spherical nanosized focal spot unravels the interior of cells. Nature methods *5*, 539-544.

Schneider, A.F.L., and Hackenberger, C.P.R. (2017). Fluorescent labelling in living cells. Current opinion in biotechnology *48*, 61-68.

Schnell, U., Dijk, F., Sjollema, K.A., and Giepmans, B.N.G. (2012). Immunolabeling artifacts and the need for live-cell imaging. Nature methods *9*, 152-158.

Schoonooghe, S., Laoui, D., van Ginderachter, J.A., Devoogdt, N., Lahoutte, T., Baetselier, P. de, and Raes, G. (2012). Novel applications of nanobodies for in vivo bio-imaging of inflamed tissues in inflammatory diseases and cancer. Immunobiology *217*, 1266-1272.

Schüller, R., Forné, I., Straub, T., Schreieck, A., Texier, Y., Shah, N., Decker, T.-M., Cramer, P., Imhof, A., and Eick, D. (2016). Heptad-Specific Phosphorylation of RNA Polymerase II CTD. Molecular cell *61*, 305-314.

Schwabish, M.A., and Struhl, K. (2004). Evidence for Eviction and Rapid Deposition of Histones upon Transcriptional Elongation by RNA Polymerase II. Molecular and Cellular Biology *24*, 10111-10117.

Serizawa, H., Mäkelä, T.P., Conaway, J.W., Conaway, R.C., Weinberg, R.A., and Young, R.A. (1995). Association of Cdk-activating kinase subunits with transcription factor TFIIH. Nature *374*, 280-282.

Shandilya, J., and Roberts, S.G.E. (2012). The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. Biochimica et biophysica acta *1819*, 391-400.

Shaner, N.C., Patterson, G.H., and Davidson, M.W. (2007). Advances in fluorescent protein technology. Journal of cell science *120*, 4247-4260.

Sherwood, R.I., Hashimoto, T., O'Donnell, C.W., Lewis, S., Barkal, A.A., van Hoff, J.P., Karun, V., Jaakkola, T., and Gifford, D.K. (2014). Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. Nature biotechnology *32*, 171-178.

SHIMOMURA, O., JOHNSON, F.H., and SAIGA, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. Journal of cellular and comparative physiology *59*, 223-239.

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

Shroff, H., Galbraith, C.G., Galbraith, J.A., and Betzig, E. (2008). Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. Nature methods *5*, 417-423.

Shroff, H., Galbraith, C.G., Galbraith, J.A., White, H., Gillette, J., Olenych, S., Davidson, M.W., and Betzig, E. (2007). Dual-color superresolution imaging of genetically expressed probes within individual adhesion complexes. Proceedings of the National Academy of Sciences of the United States of America *104*, 20308-20313.

Shu, X., Shaner, N.C., Yarbrough, C.A., Tsien, R.Y., and Remington, S.J. (2006). Novel chromophores and buried charges control color in mFruits. Biochemistry *45*, 9639-9647.

Sikorski, T.W., and Buratowski, S. (2009). The basal initiation machinery: beyond the general transcription factors. Current opinion in cell biology *21*, 344-351.

Sims, R.J., Belotserkovskaya, R., and Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. Genes & Development *18*, 2437-2468.

Sims, R.J., and Reinberg, D. (2006). Histone H3 Lys 4 methylation: caught in a bind? Genes & Development *20*, 2779-2786.

Sims, R.J., Rojas, L.A., Beck, D.B., Bonasio, R., Schüller, R., Drury, W.J., Eick, D., and Reinberg, D. (2011). The C-terminal domain of RNA polymerase II is modified by site-specific methylation. Science (New York, N.Y.) *332*, 99-103.

Singh, B.N., and Hampsey, M. (2007). A transcription-independent role for TFIIB in gene looping. Molecular cell *27*, 806-816.

Skourti-Stathaki, K., and Proudfoot, N.J. (2014). A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes & Development *28*, 1384-1396.

Slattery, M., Zhou, T., Yang, L., Dantas Machado, A.C., Gordân, R., and Rohs, R. (2014). Absence of a simple code: how transcription factors read the genome. Trends in biochemical sciences *39*, 381-399.

Smale, S.T., and Baltimore, D. (1989). The "initiator" as a transcription control element. Cell *57*, 103-113.

Smale, S.T., and Kadonaga, J.T. (2003). The RNA polymerase II core promoter. Annual review of biochemistry *72*, 449-479.

Soutourina, J., Wydau, S., Ambroise, Y., Boschiero, C., and Werner, M. (2011). Direct interaction of RNA polymerase II and mediator required for transcription in vivo. Science (New York, N.Y.) *331*, 1451-1454.

Sprague, B.L., Pego, R.L., Stavreva, D.A., and McNally, J.G. (2004). Analysis of binding reactions by fluorescence recovery after photobleaching. Biophysical Journal *86*, 3473-3495.

Srivatsan, A., Tehranchi, A., MacAlpine, D.M., and Wang, J.D. (2010). Co-orientation of replication and transcription preserves genome integrity. PLoS genetics *6*, e1000810.

Stamatoyannopoulos, J.A. (2010). Illuminating eukaryotic transcription start sites. Nature methods *7*, 501-503.

Stargell, L.A., Ogg, R.C., Adkins, J.N., Robinson, M.M., and Lumb, K.J. (2001). Transcriptional activity of the TFIIA four-helix bundle in vivo. Proteins *43*, 227-232.

Stasevich, T.J., Hayashi-Takanaka, Y., Sato, Y., Maehara, K., Ohkawa, Y., Sakata-Sogawa, K., Tokunaga, M., Nagase, T., Nozaki, N., and McNally, J.G., et al. (2014a). Regulation of RNA polymerase II activation by histone acetylation in single living cells. Nature *516*, 272-275.

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

Stasevich, T.J., Sato, Y., Nozaki, N., and Kimura, H. (2014b). Quantifying histone and RNA polymerase II post-translational modification dynamics in mother and daughter cells. Methods (San Diego, Calif.) *70*, 77-88.

Stella, S., Cascio, D., and Johnson, R.C. (2010). The shape of the DNA minor groove directs binding by the DNA-bending protein Fis. Genes & Development *24*, 814-826.

Stenoien, D.L., Patel, K., Mancini, M.G., Dutertre, M., Smith, C.L., O'Malley, B.W., and Mancini, M.A. (2001). FRAP reveals that mobility of oestrogen receptor-alpha is ligand- and proteasome-dependent. Nature cell biology *3*, 15-23.

Stiff, T., O'Driscoll, M., Rief, N., Iwabuchi, K., Löbrich, M., and Jeggo, P.A. (2004). ATM and DNA-PK function redundantly to phosphorylate H2AX after exposure to ionizing radiation. Cancer research *64*, 2390-2396.

Stirling, P.C., Chan, Y.A., Minaker, S.W., Aristizabal, M.J., Barrett, I., Sipahimalani, P., Kobor, M.S., and Hieter, P. (2012). R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. Genes & Development *26*, 163-175.

Strohner, R., Wachsmuth, M., Dachauer, K., Mazurkiewicz, J., Hochstatter, J., Rippe, K., and Längst, G. (2005). A 'loop recapture' mechanism for ACF-dependent nucleosome remodeling. Nature structural & molecular biology *12*, 683-690.

Suh, H., Ficarro, S.B., Kang, U.-B., Chun, Y., Marto, J.A., and Buratowski, S. (2016). Direct Analysis of Phosphorylation Sites on the Rpb1 C-Terminal Domain of RNA Polymerase II. Molecular cell *61*, 297-304.

Sun, X., Zhang, A., Baker, B., Sun, L., Howard, A., Buswell, J., Maurel, D., Masharina, A., Johnsson, K., and Noren, C.J., et al. (2011). Development of SNAP-tag fluorogenic probes for wash-free fluorescence imaging. Chembiochem : a European journal of chemical biology *12*, 2217-2226.

Sun, X., Zhang, Y., Cho, H., Rickert, P., Lees, E., Lane, W., and Reinberg, D. (1998). NAT, a human complex containing Srb polypeptides that functions as a negative regulator of activated transcription. Molecular cell *2*, 213-222.

Sun, Z.W., Tessmer, A., and Hampsey, M. (1996). Functional interaction between TFIIB and the Rpb9 (Ssu73) subunit of RNA polymerase II in Saccharomyces cerevisiae. Nucleic acids research *24*, 2560-2566.

Swygert, S.G., and Peterson, C.L. (2014). Chromatin dynamics: interplay between remodeling enzymes and histone modifications. Biochimica et biophysica acta *1839*, 728-736.

Taatjes, D.J., Näär, A.M., Andel, F., Nogales, E., and Tjian, R. (2002). Structure, function, and activator-induced conformations of the CRSP coactivator. Science (New York, N.Y.) *295*, 1058-1062.

Takahashi, H., Parmely, T.J., Sato, S., Tomomori-Sato, C., Banks, C.A.S., Kong, S.E., Szutorisz, H., Swanson, S.K., Martin-Brown, S., and Washburn, M.P., et al. (2011). Human Mediator Subunit Med26 Functions As A Docking Site For Transcription Elongation Factors. Cell *146*, 92-104.

Tee, W.-W., and Reinberg, D. (2014). Chromatin features and the epigenetic regulation of pluripotency states in ESCs. Development (Cambridge, England) *141*, 2376-2390.

Teng, K.W., Ishitsuka, Y., Ren, P., Youn, Y., Deng, X., Ge, P., Lee, S.H., Belmont, A.S., and Selvin, P.R. (2016). Labeling proteins inside living cells using external fluorophores for microscopy. eLife *5*.

Thanos, D., and Maniatis, T. (1995). Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. Cell *83*, 1091-1100.

Theisen, J.W.M., Lim, C.Y., and Kadonaga, J.T. (2010). Three key subregions contribute to the function of the downstream RNA polymerase II core promoter. Molecular and Cellular Biology *30*, 3471-3479.

Thomas, M.C., and Chiang, C.-M. (2006). The general transcription machinery and general cofactors. Critical reviews in biochemistry and molecular biology *41*, 105-178.

Tora, L. (2002). A unified nomenclature for TATA box binding protein (TBP)-associated factors (TAFs) involved in RNA polymerase II transcription. Genes & Development *16*, 673-675.

Toseland, C.P. (2013). Fluorescent labeling and modification of proteins. Journal of chemical biology *6*, 85-95.

Tsien, R.Y. (1998). The green fluorescent protein. Annual review of biochemistry *67*, 509-544.

Tuduri, S., Crabbé, L., Conti, C., Tourrière, H., Holtgreve-Grez, H., Jauch, A., Pantesco, V., Vos, J. de, Thomas, A., and Theillet, C., et al. (2009). Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. Nature cell biology *11*, 1315-1324.

Vasiljeva, L., and Buratowski, S. (2006). Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. Molecular cell *21*, 239-248.

Vosnakis, N., Koch, M., Scheer, E., Kessler, P., Mély, Y., Didier, P., and Tora, L. (2017). Coactivators and general transcription factors have two distinct dynamic populations dependent on transcription. The EMBO Journal *36*, 2710-2725.

Voss, T.C., Schiltz, R.L., Sung, M.-H., Yen, P.M., Stamatoyannopoulos, J.A., Biddie, S.C., Johnson, T.A., Miranda, T.B., John, S., and Hager, G.L. (2011). Dynamic exchange at regulatory elements during chromatin remodeling underlies assisted loading mechanism. Cell *146*, 544-554.

Wang, H., Wang, M., Wang, H., Böcker, W., and Iliakis, G. (2005). Complex H2AX phosphorylation patterns by multiple kinases including ATM and DNA-PK in human cells exposed to ionizing radiation and treated with kinase inhibitors. Journal of cellular physiology *202*, 492-502.

Wang, Q., Sawyer, I.A., Sung, M.-H., Sturgill, D., Shevtsov, S.P., Pegoraro, G., Hakim, O., Baek, S., Hager, G.L., and Dundr, M. (2016). Cajal bodies are linked to genome conformation. Nature communications *7*, 10966.

Ward, I.M., and Chen, J. (2001). Histone H2AX is phosphorylated in an ATR-dependent manner in response to replicational stress. The Journal of biological chemistry *276*, 47759-47762.

Wegel, E., Göhler, A., Lagerholm, B.C., Wainman, A., Uphoff, S., Kaufmann, R., and Dobbie, I.M. (2016). Imaging cellular structures in super-resolution with SIM, STED and Localisation Microscopy: A practical comparison. Scientific reports *6*, 27290.

Wei, X., Samarabandu, J., Devdhar, R.S., Siegel, A.J., Acharya, R., and Berezney, R. (1998). Segregation of transcription and replication sites into higher order domains. Science (New York, N.Y.) *281*, 1502-1506.

Wen, Y., and Shatkin, A.J. (1999). Transcription elongation factor hSPT5 stimulates mRNA capping. Genes & Development *13*, 1774-1779.

West, M.L., and Corden, J.L. (1995). Construction and Analysis of Yeast RNA Polymerase II Ctd Deletion and Substitution Mutations. Genetics *140*, 1223-1233.

Westphal, V., Rizzoli, S.O., Lauterbach, M.A., Kamin, D., Jahn, R., and Hell, S.W. (2008). Video-rate far-field optical nanoscopy dissects synaptic vesicle movement. Science (New York, N.Y.) *320*, 246-249.

White, M.D., Angiolini, J.F., Alvarez, Y.D., Kaur, G., Zhao, Z.W., Mocskos, E., Bruno, L., Bissiere, S., Levi, V., and Plachta, N. (2016). Long-Lived Binding of Sox2 to DNA Predicts Cell Fate in the Four-Cell Mouse Embryo. Cell *165*, 75-87.

Wright, K.J., Marr, M.T., and Tjian, R. (2006). TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. Proceedings of the National Academy of Sciences of the United States of America *103*, 12347-12352.

Wu, C.-H., Yamaguchi, Y., Benjamin, L.R., Horvat-Gordon, M., Washinsky, J., Enerly, E., Larsson, J., Lambertsson, A., Handa, H., and Gilmour, D. (2003). NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in Drosophila. Genes & Development *17*, 1402-1414.

Wyce, A., Xiao, T., Whelan, K.A., Kosman, C., Walter, W., Eick, D., Hughes, T.R., Krogan, N.J., Strahl, B.D., and Berger, S.L. (2007). H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex. Molecular cell *27*, 275-288.

Xie, L., Torigoe, S.E., Xiao, J., Mai, D.H., Li, L., Davis, F.P., Dong, P., Marie-Nelly, H., Grimm, J., and Lavis, L., et al. (2017). A dynamic interplay of enhancer elements regulates Klf4 expression in naïve pluripotency. Genes & Development *31*, 1795-1808.

Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., and Handa, H. (2006). P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. Molecular cell *21*, 227-237.

Yamaguchi, Y., Wada, T., Watanabe, D., Takagi, T., Hasegawa, J., and Handa, H. (1999). Structure and function of the human transcription elongation factor DSIF. The Journal of biological chemistry *274*, 8085-8092.

Yan, Q., Moreland, R.J., Conaway, J.W., and Conaway, R.C. (1999). Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. The Journal of biological chemistry *274*, 35668-35675.

Yang, C., Bolotin, E., Jiang, T., Sladek, F.M., and Martinez, E. (2007). Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. Gene *389*, 52-65.

Yankulov, K.Y., and Bentley, D.L. (1997). Regulation of CDK7 substrate specificity by MAT1 and TFIIH. The EMBO Journal *16*, 1638-1646.

Yao, J., Munson, K.M., Webb, W.W., and Lis, J.T. (2006). Dynamics of heat shock factor association with native gene loci in living cells. Nature *442*, 1050-1053.

Yao, T., Song, L., Jin, J., Cai, Y., Takahashi, H., Swanson, S.K., Washburn, M.P., Florens, L., Conaway, R.C., and Cohen, R.E., et al. (2008). Distinct modes of regulation of the Uch37 deubiquitinating enzyme in the proteasome and in the Ino80 chromatin-remodeling complex. Molecular cell *31*, 909-917.

Yokomori, K., Verrijzer, C.P., and Tjian, R. (1998). An interplay between TATA box-binding protein and transcription factors IIE and IIA modulates DNA binding and transcription. Molecular and Cellular Biology *95*, 6722-6727.

Yudkovsky, N., Ranish, J.A., and Hahn, S. (2000). A transcription reinitiation intermediate that is stabilized by activator. Nature *408*, 225-229.

Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. Genes & Development *25*, 2227-2241.

Zentner, G.E., and Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. Nature structural & molecular biology *20*, 259-266.

Zhang, S., Tian, D., Tran, N.H., Choi, K.P., and Zhang, L. (2014). Profiling the transcription factor regulatory networks of human cell types. Nucleic acids research *42*, 12380-12387.

Zhang, Y., Smith, C.L., Saha, A., Grill, S.W., Mihardja, S., Smith, S.B., Cairns, B.R., Peterson, C.L., and Bustamante, C. (2006). DNA translocation and loop formation mechanism of chromatin remodeling by SWI/SNF and RSC. Molecular cell *24*, 559-568.

Zhao, Z.W., Roy, R., Gebhardt, J.C.M., Suter, D.M., Chapman, A.R., and Xie, X.S. (2014). Spatial organization of RNA polymerase II inside a mammalian cell nucleus revealed by reflected light-sheet superresolution microscopy. Proceedings of the National Academy of Sciences of the United States of America *111*, 681-686.

Sascha CONIC

Analysis of transcription factor and histone modification dynamics in the nucleus of single living cells using a novel antibody-based imaging approach

Résumé

Dans les cellules des eucaryotes, la transcription des gènes est contrôlée par une pléthore de complexes protéiniques. Cependant, la plupart de nos connaissances fondamentales sur la régulation de la transcription viennent des expériences biochimiques ou des expériences d'immunofluorescences utilisant des cellules fixées. Par conséquent, beaucoup d'efforts ont été consacré récemment pour obtenir des informations sur les mouvements dynamiques ou sur l'assemblage des facteurs de transcription directement dans des cellules vivantes. Nous avons développé une stratégie de marquage, appelé « versatile antibody-based imaging approach » (VANIMA), dans laquelle des anticorps marqués avec un fluorochrome sont introduit dans des cellules vivantes pour visualiser spécifiquement des protéines endogènes ou des modifications post-traductionnelle. Nous avons pu montrer que VANIMA peut être utilisé pour étudier des processus dynamique des mécanismes fondamental de la biologie y compris les facteurs de la machinerie de transcription ainsi que les modifications des histones dans des cellules vivantes de cancer humaine en utilisant la microscopie conventionnelle ou à super-résolution. Dans l'avenir VANIMA va servir comme un outil valable pour révéler les dynamiques des processus endogènes en biologie y compris la transcription directement dans des cellules vivantes individuelles.

Mots-clés : livraison d'anticorps, imagerie sur cellules vivantes, cellules individuelles, protéine endogènes, modifications post-traductionnelles, transcription par ARN Polymérase II

Résumé en anglais

In eukaryotic cells, gene transcription is controlled by a plethora of protein complexes. However, most of our basic knowledge about transcription regulation originate from biochemical experiments or immunofluorescence experiments using fixed cells. Consequently, many efforts have been devoted recently to obtain information about the dynamic movements or assembly of transcription factors directly from living cells. Therefore, we developed a labeling strategy, named versatile antibody-based imaging approach (VANIMA), in which fluorescently labeled antibodies are introduced into living cells to image specific endogenous proteins or posttranslational modifications. We were able to show that VANIMA can be used to study dynamical processes of fundamental biological mechanisms including factors of the transcription machinery as well as histone modifications in living human cancer cells using conventional or super-resolution microscopy. Hence, in the future VANIMA will serve as a valuable tool to uncover the dynamics of endogenous biological processes including transcription directly in single living cells.

Keywords: Antibody delivery, Live-imaging, Single cells, Endogenous proteins, Posttranslational modifications, RNA Polymerase II transcription.