Single-molecule characterization of the dynamics and regulation of endogenous transcription factor hubs

Thesis by

Shawn Yoshida

In Partial Fulfillment of the Requirements for the
degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2027
Defended May 27th, 2026

Shawn Yoshida
ORCID: 0000-0002-0866-2741

ACKNOWLEDGEMENTS

This thesis is dedicated to my family: my parents, E-Ling and Makoto; my best friend, Kohaku; and my partner in all, Yi-Ling.

First, I must thank my advisor, Prof. Shasha Chong, for her constant guidance and mentorship and for fostering a lab environment that allowed me to grow as an independent researcher. I feel very fortunate to have spent the last five years in her lab. I also owe thanks to my committee members, Profs. Mitch Guttman, Shu-ou Shan, and Lu Wei for their valuable feedback and advice over the years. I would also like to thank the many other people at Caltech and CWRU from whose teaching, advice, and mentorship I have greatly benefited, including Hoa Nguyen and Profs. Barbara Clemenson, Lydia Kisley, Brian McDermott, Michael Hinczewski, and Lior Pachter.

I also owe a great deal to lab members past and present: Qinyu Han for her valuable advice and constant friendship from day one, Victoria Walling for her endless patience in teaching me benchwork from the ground up, Dr. Shawn Irgen-Gioro for his mentorship across various domains of science, Jamie Takashima for endlessly helpful discussions on all things PTMs, Jonathan Banh for keeping the lab running and lively, and Dr. Barun Maity, Dr. Yanghao Zhong, Dr. Jiamin Guo, Dr. Donghun Lee, Dr. Leishu Lin, Michael Di Martino, and Anaya Gokarn, and for their friendship and for making our lab a wonderful place to work. I also must thank my BMB cohort, Cassie Zhang, Juliet Lee, Maria Carilli, Nikki Swarna, Nina Le, Allen Chen, Andres Orta, Tyler Brittain, and Jacob Parres-Gold, along with the many other friends who supported me through grad school including, Alex Spentzouris, Prem Rao, Jakob Myers, Kelvin Li, Sammie Lin, Marcus Eliot, Masami Hazu, Ashish Goel, Lynn Fang, Anthony Gutierrez, Kyogo Nagashima, Ruby Chen, Ziyao Luo, Jiaxi Wu, Nithin Varma, Albert Yang, Bonnie Nien, Sriram Srirangan, Martin Abt, and everyone else who colored in my life.

I also must thank Rebecca Fox, Aracely Sustaita, Andy Lo, Ann Mao, Michelle O'Camb, Sadie Rubalcava, Allen Lee, Joe Drew, and all the other professionals who kept everything running smoothly.

I am also ever so grateful to my family who helped make Los Angeles feel like home: Yunling Lou, Kevin Wheeler, Cameron Wheeler, Vivian Wheeler, Kek-Khee Loo, Masako Nakada, and Yi-Shen Loo. Thank you all for all the time, advice, and support you gave me over the years.

Finally, I owe everything to my family. To my parents, E-Ling Lou and Makoto Yoshida: thank you for your unending support no matter how many times I changed tack in life; I can only hope to begin repaying you for everything you did to give me this life. To Kohaku Yoshida: thank you for being my guidelight in the dark; although you are likely uninterested in reading this work, your impact upon it has been immeasurable. And to Yi-Ling Loo: thank you for your endless supply of love and patience and for being a constant reminder of what really matters in this life.

ABSTRACT

DNA-binding transcription factors (TFs) regulate eukaryotic gene expression through intrinsically disordered low-complexity domains (LCDs) that drive the formation of local, high-concentration hubs at target genomic loci via dynamic, multivalent LCD-LCD interactions. These hubs are critical to activate transcription of target genes; despite this, many of their physical properties and regulatory mechanisms remain unexplored at endogenous expression levels. This is largely due to the time-consuming nature of gene editing and the necessity for super-resolution microscopy methods that can resolve sub-diffraction-limit endogenous TF hubs compared to overexpression, which offers a faster alternative that also produces larger, micron-scale droplets that do not require specialized microscopy. However, a growing body of evidence suggests that the exquisite spatiotemporal coordination of endogenous TF hubs is poorly recapitulated by overexpression studies. Thus, here, we address these gaps by combining CRISPR/Cas9-mediated endogenous fluorescent labeling with single-molecule and super-resolution imaging methods including photoactivated localization microscopy (PALM), single-particle tracking (SPT), and stimulated emission depletion microscopy (STED), to characterize the biophysical properties and regulation mechanisms of our model oncogenic fusion TF, EWS::FLI1, at unprecedented resolution in Ewing sarcoma cells. We first establish that sample fixation with paraformaldehyde can differentially affect the appearance of certain LCD-driven droplets, artificially increasing their size and number in some cases and diminishing it in others. The spatial distribution of EWS::FLI1 does not appear to be sensitive to paraformaldehyde fixation. We then used PALM and showed that endogenous EWS::FLI1 forms sub-diffraction-limit hubs of mean diameter around 100 nm, and extracted the energetics and kinetics of hub formation, showing that they are not liquid-liquid phase-separated droplets, rather, they behave as supersaturation-driven nucleating clusters. We next showed that EWS::FLI1 hub formation is a neomorphic property of the fusion and is not directly conferred by either of its parental proteins. We also investigated numerous axes of EWS::FLI1 hub regulation using a combination of PALM, SPT, STED, and confocal microscopy. We showed that during mitosis, hubs dissolve and EWS::FLI1 remains weakly enriched on mitotic chromatin, that nascent RNA modulates EWS::FLI1-hub binding

kinetics without altering hub dimensions, that the post-translational modification O-GlcNAc can make hubs more fluid and reduce EWS::FLI1 retention in hubs, and that EWS::FLI1 preferentially associates with initiating serine-5 phosphorylated RNA Polymerase II. We also showed that treatment with small molecules LY2835219 or trabectedin dissolves and mislocalizes EWS::FLI1 hubs, respectively, suppressing transcription of EWS::FLI1 target genes, highlighting their therapeutic potential. Together, this work provides a physically grounded, single-molecule view into the formation, regulation, and function of oncogenic TF hubs, and establishes an imaging-based framework broadly applicable to the study of other proteins involved in transcriptional control and disease.

PUBLISHED CONTENT AND CONTRIBUTIONS

Irgen-Gioro, S[†], **Yoshida, S.**[†], Walling, V. & Chong, S. Fixation can change the appearance of phase separation in living cells. *Elife* 11, e79903 (2022). https://doi.org/10.7554/eLife.79903

S.R.Y. performed confocal microscopy for Figure 3C, single-particle tracking experiments for Figure 6, and wrote the corresponding sections of the manuscript.

Yoshida, S., Zhong, Y., Kumagai, A., Dunphy, W. G. & Chong, S. Dynamic regulation of endogenous transcription factor hubs at single-molecule resolution. 2026.01.25.701606 (2026). https://doi.org/10.64898/2026.01.25.701606

S.R.Y. performed all imaging experiments, physical modeling, and writing of the manuscript.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

DNA-binding transcription factors (TFs) serve as critical regulators of eukaryotic transcription and are typically composed of well-structured DNA-binding domains (DBDs) that bind to cognate sites throughout the genome and intrinsically disordered, low-complexity domains (LCDs) responsible for activating the transcription of target genes. These intrinsically disordered regions (IDRs) do not fold into stable three-dimensional conformations; rather, they rapidly sample large conformational spaces. The lack of a well-defined structure has historically made these regions resistant to mechanistic characterization through conventional structure-function relation-based approaches.

Work from the past few decades has revealed that TF LCDs engage in dynamic, highly multivalent, and selective LCD-LCD interactions. Homotypic LCD-LCD interactions of key TFs drive the formation of local, high-concentration TF hubs at target genomic loci, and heterotypic LCD-LCD interactions lead to the recruitment of other TFs[1,2], RNA Polymerase II (RNAP II)[3–7], and other transcriptional cofactors[3,4,8] to those hubs, leading to transcriptional activation of downstream target genes. Further, it has been shown that target genes require an optimal level of TF LCD-LCD interactions in order to be efficiently activated; insufficient levels fail to drive transcription while excessive levels actually serve to suppress it[9]. This "Goldilocks" behavior has been observed across a range of hub-forming transcription regulators[2,9], suggesting that it may be a more general feature of hub-mediated transcriptional control.

Under certain conditions, particularly the overexpression of TF LCD fragments, these hubs can grow to form macroscopic, phase-separated droplets likely through processes like liquid-liquid phase separation (LLPS) or other similar mechanisms[1]. However, whether endogenous TF hubs are bona fide LLPS droplets or a distinct class of biomolecular assembly remains unclear. This is primarily due to the fact that characterizing the physical underpinnings of endogenous TF hubs requires imaging the spatial distribution and dynamics of TFs at

endogenous expression levels, an undertaking complicated by the sub-diffraction-limit size of endogenous TF hubs[10]. Thus, to date, much of our mechanistic understanding of TF hubs and their transactivation function instead comes from overexpression systems which do not always faithfully represent endogenous TF behavior.

Identifying the physical processes underlying endogenous TF hub formation is valuable not only to enhance mechanistic understanding, but also because the specific physical mechanism driving hub formation can suggest biological or chemical perturbations that can dissolve or modulate the material properties of a hub, which has direct therapeutic implications. Whether a given TF hub is formed by bona fide LLPS[11], LLPS coupled to percolation[12,13], supersaturation-driven nucleation[10,12], or a host of other mechanisms can dictate its sensitivity to protein concentration, post-translational modification, or small molecules. The chemistry underlying these LCD-LCD interactions is multivalent and can include electrostatic forces, charge-block interactions, hydrophobic interactions, hydrogen bonds, π-π stacking, and cation-π interactions[14–17], leading to different flavors of hubs (i.e. those with different constituents) depending on the combination of forces driving their formation[18,19]. Precise and consistent terminology then becomes critical as such IDR-mediated assemblies have been referred to as “hubs”, “condensates”, “droplets”, “puncta”, and “clusters” throughout the literature. While many studies immediately link the term condensates with LLPS, out of respect for the sheer breadth of physical mechanisms that can drive the formation of local, high-concentration regions, here, we will treat all these terms, including condensate, as being mechanism-agnostic[13], with no particular term immediately implying any underlying physics. Further, we will attempt to retain the preferred terminology of the subfields, with “hubs” being used to preferentially refer to small TF assemblies.

TF LCDs are frequently mutated or dysregulated in cancers[20–24]. One of the most extensively studied examples is the fusion oncoprotein EWS::FLI1, a TF that causes Ewing sarcoma, an aggressive bone and soft tissue cancer that has no known targeted molecular therapies[20]. EWS::FLI1 is formed by a t(11;22)(q24;q12) chromosomal translocation that fuses the intrinsically disordered LCD of the FET-family RNA-binding protein EWSR1 with the well-

structured DBD of the ETS-family TF, FLI1[25]. EWS::FLI1 has been shown to exhibit numerous neomorphic behaviors not directly conferred by either parental protein, including an ability to remodel GGAA microsatellites into de novo enhancers, inactivate canonical ETS enhancers by displacing wild-type ETS TFs, and drive aberrant alternative splicing[26,27]. Recent work showed that EWS::FLI1 forms hubs at GGAA microsatellites throughout the genome, primarily driven by multivalent EWS LCD-LCD interactions[1,28,29]. These hubs have been shown to be critical for driving transcription of target genes, and disruption of these hubs by targeted mutagenesis of EWS::FLI1 has been shown to be able to interfere with oncogenic transformation[1]. The potential for hub disruption as a therapeutic vulnerability is an important finding, as TFs including EWS::FLI1 are extremely difficult to target with drugs owing to their lack of a large, well-defined binding pocket.

Despite the critical role of TF hubs in transcriptional control and human disease, many fundamental questions about the biophysical properties and regulation of TF hubs remain open. One key reason is due to methodological limitations, as the exquisite spatiotemporal organization and regulation of TF hubs are highly sensitive and demand to be studied in their endogenous contexts. This requires a combination of biological methods like CRISPR/Cas9-mediated genome editing to fluorescently label endogenous TFs without perturbing their expression level, and of single-molecule imaging methods like photoactivated localization microscopy (PALM) and single-particle tracking (SPT) that can resolve sub-diffraction-limit endogenous TF hubs and characterize the behavior of individual TF molecules rather than their ensemble behavior. Fixation also introduces an additional layer of complexity, as recent work has shown that many of the most commonly used fixatives in biology, including paraformaldehyde and alcohols, are not able to faithfully preserve the nuclear localization of TFs[30]. Using live-cell approaches when possible and performing proper controls when fixation is unavoidable is becoming increasingly critical when imaging TF hubs and other dynamic intracellular phenomena.

In this thesis, we first characterize how paraformaldehyde fixation artifacts can differentially affect different types of biomolecular condensates and establish a framework for

understanding how and when such artifacts arise. Then, we focus on EWS::FLI1 as our model system for TF hubs and employ an array of live-cell, single-molecule, and/or super-resolution imaging methods to characterize the biophysical properties and regulation strategies of endogenous EWS::FLI1 hubs. Here, we uncover the nucleation-driven process underlying EWS::FLI1 hub formation along with hub dimensions, energetics, kinetics, and material properties. We then examine the multiple layers of regulation governing endogenous EWS::FLI1 hub behavior, including by mitosis and chromatin compaction, nascent RNA, post-translational modification by O-GlcNAc, and association with distinct functional states of RNAP II. We also visualize how small-molecule-mediated dissolution and mislocalization of endogenous EWS::FLI1 hubs can suppress target gene transcription. We finally identify a novel, proximity-induced photophysical coupling of densely packed fluorophores commonly used in single-molecule imaging, and discuss its implications for future super-resolution studies.

References

1. Chong, S. et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science 361, eaar2555 (2018).
2. Cai, D. et al. Phase separation of YAP reorganizes genome topology for long-term YAP target gene expression. Nat. Cell Biol. 21, 1578–1589 (2019).
3. Lyons, H. et al. Functional partitioning of transcriptional regulators by patterned charge blocks. Cell 186, 327-345.e28 (2023).
4. Cho, W.-K. et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science 361, 412–415 (2018).
5. Guo, Y. E. et al. Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. Nature 572, 543–548 (2019).
6. Boehning, M. et al. RNA polymerase II clustering through carboxy-terminal domain phase separation. Nat. Struct. Mol. Biol. 25, 833–840 (2018).

7. Lyons, H. et al. RNA polymerase II partitioning is a shared feature of diverse oncofusion condensates. Cell 188, 3843-3862.e28 (2025).
8. Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. Science 361, eaar3958 (2018).
9. Chong, S. et al. Tuning levels of low-complexity domain interactions to modulate endogenous oncogenic transcription. Mol. Cell 82, 2084-2097.e5 (2022).
10. Yoshida, S., Zhong, Y., Kumagai, A., Dunphy, W. G. & Chong, S. Dynamic regulation of endogenous transcription factor hubs at single-molecule resolution. 2026.01.25.701606 Preprint at https://doi.org/10.64898/2026.01.25.701606 (2026).
11. Alberti, S., Gladfelter, A. & Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. Cell 176, 419–434 (2019).
12. Choi, J.-M., Holehouse, A. S. & Pappu, R. V. Physical principles underlying the complex biology of intracellular phase transitions. Annu. Rev. Biophys. 49, 107–133 (2020).
13. Mittag, T. & Pappu, R. V. A conceptual framework for understanding phase separation and addressing open questions and challenges. Mol. Cell 82, 2201–2214 (2022).
14. Kilgore, H. R. & Young, R. A. Learning the chemical grammar of biomolecular condensates. Nat. Chem. Biol. 18, 1298–1306 (2022).
15. Pappu, R. V., Wang, X., Vitalis, A. & Crick, S. L. A polymer physics perspective on driving forces and mechanisms for protein aggregation. Arch. Biochem. Biophys. 469, 132–141 (2008).
16. Wang, J. et al. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. Cell 174, 688–699 (2018).
17. Lin, Y., Currie, S. L. & Rosen, M. K. Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. J. Biol. Chem. 292, 19110–19120 (2017).
18. Ambadi Thody, S. et al. Small-molecule properties define partitioning into biomolecular condensates. Nat. Chem. 16, 1794–1802 (2024).

19. Han, Q. et al. Decoding Condensate Composition Reveals Tunable Interactomes and Functions of Disordered Proteins. 2025.10.27.684971 Preprint at https://doi.org/10.1101/2025.10.27.684971 (2025).
20. Kovar, H. Dr. Jekyll and Mr. Hyde: The Two Faces of the FUS/EWS/TAF15 Protein Family. Sarcoma 2011, 837474 (2011).
21. Linardic, C. M. PAX3-FOXO1 fusion gene in rhabdomyosarcoma. Cancer Lett. 270, 10–18 (2008).
22. Ahn, J. H. et al. Phase separation drives aberrant chromatin looping and cancer development. Nature 595, 591–595 (2021).
23. Browne, A. L. et al. Network analysis of SRC-1 reveals a novel transcription factor hub which regulates endocrine resistant breast cancer. Oncogene 37, 2008–2021 (2018).
24. Zhao, J. et al. Oncogenic transcription factors instruct promoter-enhancer hubs in individual triple negative breast cancer cells. Sci. Adv. 10, eadl4043 (2024).
25. Li, Y., Luo, H., Liu, T., Zacksenhaus, E. & Ben-David, Y. The ets transcription factor Fli-1 in development, cancer and disease. Oncogene 34, 2022–2031 (2015).
26. Gangwal, K. et al. Microsatellites as EWS/FLI response elements in Ewing's sarcoma. Proc. Natl. Acad. Sci. U. S. A. 105, 10149–10154 (2008).
27. Selvanathan, S. P. et al. Oncogenic fusion protein EWS-FLI1 is a network hub that regulates alternative splicing. Proc. Natl. Acad. Sci. U. S. A. 112, E1307-1316 (2015).
28. Johnson, K. M. et al. Role for the EWS domain of EWS/FLI in binding GGAA-microsatellites required for Ewing sarcoma anchorage independent growth. Proc. Natl. Acad. Sci. U. S. A. 114, 9870–9875 (2017).
29. Boulay, G. et al. Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. Cell 171, 163-178.e19 (2017).
30. Irgen-Gioro, S., Yoshida, S., Walling, V. & Chong, S. Fixation can change the appearance of phase separation in living cells. eLife 11, e79903 (2022).

Chapter 2

FIXATION CAN CHANGE THE APPEARANCE OF PHASE SEPARATION IN LIVING CELLS

Adapted from: Irgen-Gioro, S[†], **Yoshida, S.**[†], Walling, V. & Chong, S. Fixation can change the appearance of phase separation in living cells. *Elife* 11, e79903 (2022).

Abstract

Fixing cells with paraformaldehyde (PFA) is an essential step in numerous biological techniques, as it is thought to preserve a snapshot of biomolecular transactions in living cells. Fixed cell imaging techniques such as immunofluorescence have been widely used to detect liquid-liquid phase separation (LLPS) *in vivo*. In prior work, we compared images, before and after fixation, of cells expressing intrinsically disordered proteins that are able to undergo LLPS. Surprisingly, we found that PFA fixation can both enhance and diminish putative LLPS behaviors. For specific proteins, fixation can even cause their droplet-like puncta to artificially appear in cells that do not have any detectable puncta in the live condition. Fixing cells in the presence of glycine, a molecule that modulates fixation rates, can reverse the fixation effect from enhancing to diminishing LLPS appearance. We further established a kinetic model of fixation in the context of dynamic protein-protein interactions. Simulations based on the model suggest that protein localization in fixed cells depends on an intricate balance of protein-protein interaction dynamics, the overall rate of fixation, and notably, the difference between fixation rates of different proteins. Here, consistent with simulations, live-cell single-molecule imaging experiments showed that a fast overall rate of fixation relative to protein-protein interaction dynamics can minimize fixation artifacts. Our work reveals that PFA fixation changes the appearance of LLPS from living cells, presents a caveat in studying LLPS using fixation-based methods, and suggests a mechanism underlying the fixation artifact.

Introduction

Fixing cells to preserve a snapshot of biomolecular transactions *in vivo* is a widely used strategy in numerous techniques in biology and medicine. Due to its small size and high reactivity with a wide range of biological entities, paraformaldehyde (PFA) is one of the most commonly used fixatives to create covalent cross-linking between biomolecules, e.g., proteins and nucleic acids. PFA non-selectively "fixes" or cross-links molecules in proximity to enable characterization of biomolecular interactions formed in living cells. Examples of popular techniques that use PFA to fix cells include ChIP-sequencing[1], chromosome conformation capture (3C) based techniques[2], immunofluorescence[3], fluorescence in situ hybridization (FISH)[4], cross-linking mass spectrometry[5], super-resolution expansion microscopy[6], and super-resolution localization microscopies such as stochastic optical reconstruction microscopy (STORM)[7]. Although PFA fixation has been used to faithfully preserve live-cell conditions in many scenarios, a number of studies have uncovered situations in which fixation fails to cross-link DNA-protein interactions formed in living cells. By imaging different transcription factors (TFs) in live and fixed cells, Schmiedeberg et al. showed that TFs bound to DNA with fast dissociation dynamics (<5 sec residence times as determined by fluorescence recovery after photobleaching (FRAP)) are not cross-linked to DNA upon PFA fixation[8]. Using live-cell single-molecule imaging, Teves et al. showed that TFs stay bound to chromosomes during mitosis, and fixing cells can artificially deplete transiently bound TFs from mitotic chromosomes[9]. These studies exemplify the fact that fixation, with limited reaction rates, cannot provide an instantaneous snapshot and may miss or obfuscate biomolecular interactions that happen either at or faster than the timescale of fixation. What further complicates the result of cell fixation is that the reactivity and reaction rates of PFA are variable and dependent on its biomolecule substrates[10,11]. For example, the efficiency and rates at which PFA reacts with proteins can vary by orders of magnitude[12] and are dependent on their amino acid sequences[5,12,13] and tertiary structures[14].

Among the numerous biomolecular transactions investigated using fixed cell imaging is liquid-liquid phase separation (LLPS), a long-observed behavior of polymers in solution[15–17] that has recently generated much excitement in biological research communities due to its proposed roles in cellular organization and functions[18–21]. LLPS is driven by excessive levels

of transient, selective, and multivalent protein-protein interactions mediated by intrinsically disordered regions (IDRs) within the proteins of interest[22–24]. Whereas rigorous characterization of LLPS *in vivo* has been challenging and remains a question under active investigation[25], detection of discrete puncta that have a spherical shape, undergo fusion and fission, and dynamically exchange biomolecules with the surrounding according to FRAP is often considered evidence of putative LLPS in living cells. While such diverse measurements have been widely used for studying proteins under overexpression conditions, far fewer approaches are available to probe LLPS under physiological conditions. Detecting local high-concentration regions or puncta of an endogenously expressed protein using immunofluorescence of fixed cells has been used in many studies as evidence of LLPS[26–32]. Not only is the detection of puncta an inconclusive metric for establishing LLPS; whether a punctate distribution observed in fixed cells actually represents the live-cell scenario remains unclear, as fixation has only been assumed, but not directly shown, to faithfully preserve multivalent interactions and LLPS formed in living cells. This knowledge gap motivated us to image cells that overexpress various known IDR-containing proteins before and after fixation to evaluate the ability of PFA fixation to preserve LLPS behaviors. We found that, interestingly, fixation can significantly alter the appearance of droplet-like puncta in cells. Our quantitative image analysis suggests that depending on the LLPS-driving protein, fixing cells can either enhance or diminish the apparent LLPS behaviors *in vivo*. In certain cases, fixation can even cause droplet-like puncta to artificially appear in cells that have a homogeneous protein distribution and no detectable puncta in the live condition. Conversely, fixation can also cause droplet-like puncta in living cells to completely disappear. Combining experiments that modulate fixation rates, live-cell single-molecule imaging that quantifies protein binding dynamics, and simulations based on a kinetic model, we further demonstrated that protein localization in fixed cells depends on an intricate balance of protein-protein interaction dynamics, the overall rate of fixation, and the difference between protein fixation rates in and out of droplet-like puncta. Our work urges caution in the interpretation of previous claims of *in vivo* phase separation based solely on immunofluorescence imaging of fixed cells and serves to guide future judicious application of PFA fixation.

Results

As shown in the published version of this chapter[33], overexpression of the GFP-tagged IDRs of the FET (FUS, EWS, and TAF15) in live U2OS cells resulted in spherical puncta, and upon fixation with 4% paraformaldehyde (PFA), the number and size of puncta increased, quantified as outlined in the Methods. The effect of fixation on the appearance of droplets also was dependent on the fluorescent tag used, as PFA fixation of U2OS cells transfected with DsRed2-TAF15(IDR) resulted in enhanced LLPS with slightly larger droplets and the emergence of numerous small puncta in what was the dilute phase. In contrast, the same fixation method on U2OS cells transfected with Halo-TAF15(IDR) resulted in diminished LLPS with puncta becoming smaller and dimmer, and with many disappearing entirely.

Furthermore, to examine whether all phase-separating proteins show the fixation artifact, we compared live and fixed cell images of EGFP-tagged full-length FUS (FUS(FL)). Full-length FUS is reported to have a greater LLPS propensity *in vitro* than its IDR alone[34]. We found that EGFP-FUS(FL) overexpressed in live U2OS cells forms many small puncta throughout the nucleus, and we did not observe a significant change of this behavior after PFA fixation (**Figure 1A**). We also fused Halo-tagged TAF15(IDR) to FTH1, which forms a 24-mer[35,36], to make an artificial protein with a high LLPS propensity. We found that TAF15(IDR)-Halo-FTH1 overexpressed in live U2OS cells forms large droplet-like puncta, and the appearance of LLPS does not significantly change after PFA fixation (**Figure 1B**). In addition, we looked into a native IDR-containing protein, EWS::FLI1, an oncogenic TF causing Ewing sarcoma[37] and known to form local high-concentration hubs at target genes associated with GGAA microsatellites[22]. Although there is no convincing evidence that EWS::FLI1 undergoes LLPS under physiological conditions, the formation of its hubs is mediated by the homotypic multivalent interactions of EWS(IDR) within the protein. Excessive levels of such multivalent interactions often result in LLPS[24]. We previously Halo-tagged endogenous EWS::FLI1 in an Ewing sarcoma cell line A673 using CRISPR/Cas9-mediated genome editing[22]. Here, we compared live and fixed A673 cell images of endogenous EWS::FLI1-Halo and did not observe a significant difference in its distribution (**Figure 1C**). This result

suggests that PFA fixation does not change the intracellular distribution of all proteins that have LLPS potential.

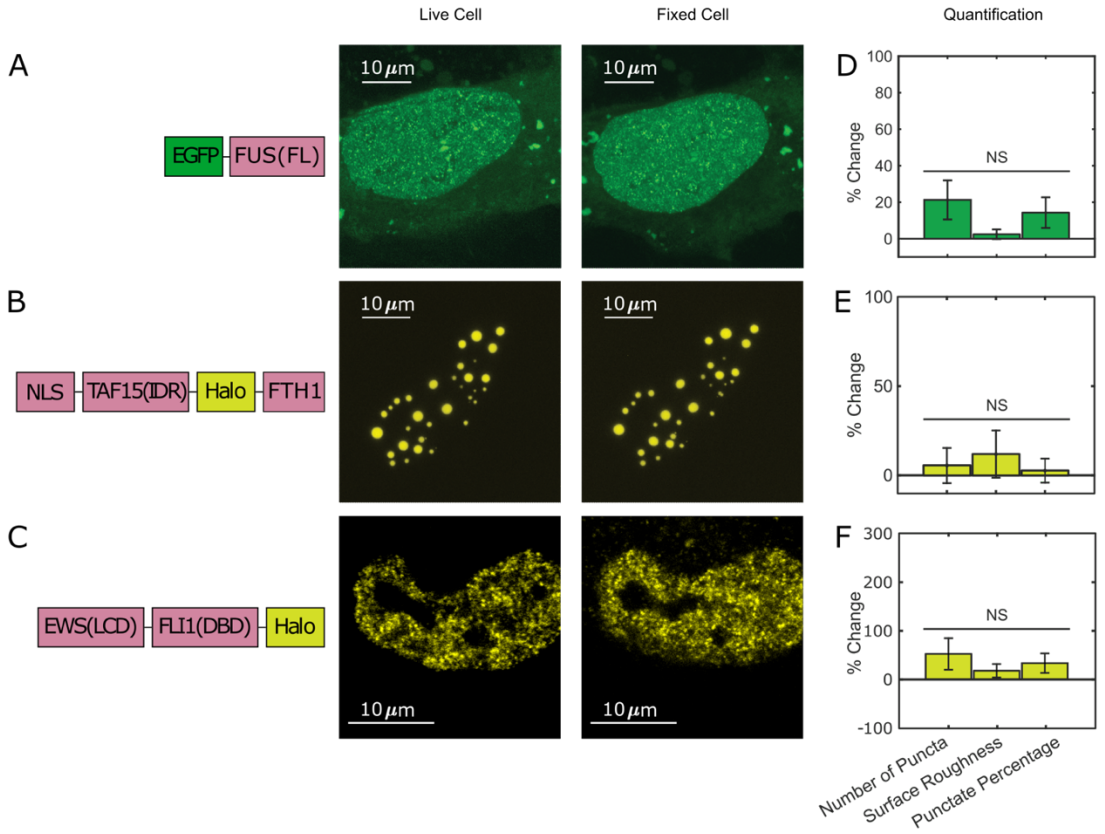

Figure 1. Not all puncta-forming proteins show the fixation artifact. U2OS cells expressing (A) EGFP-FUS(FL) and (B) TAF15(IDR)-Halo-FTH1, and (C) an A673 cell expressing endogenous EWS::FLI1-Halo are imaged using confocal fluorescence microscopy before and after 10 minutes of fixation with 4% PFA. Halo-tagged proteins are ligated with the JFX549 Halo ligand before imaging. Schematics of the protein constructs are shown on the left. Live and fixed cell images are compared. (D-F) Quantification of puncta parameters after fixation. The values are averaged from 21 (D), 16 (E), or 15 (F) cells measured in 1 (D), 4 (E), or 2(F) independent transfection and imaging sessions. Error bars represent standard errors. NS: not significant difference compared with 0 ($P < 0.05$, Wilcoxon signed-rank test). None of the examined proteins

show significant changes in their LLPS or hub appearance in the fixed cell image as compared to the live cell image.

In order to understand how fixation can either not affect, enhance, or diminish the appearance of LLPS depending on the puncta-forming protein, we added excess glycine, which competes with the protein-protein crosslinking under the fixative, to DsRed2-TAF15(IDR). After fixation, rather than enhancing the appearance of LLPS as before, it caused most puncta to disappear or collapse into donut-like structures. This indicates that the kinetics of fixation, rather than the presence of PFA alone, govern the appearance of LLPS after fixation.

To explain these results, we developed a four-state kinetic model of fixation in the context of protein-protein interaction dynamics, and simulations based on this model showed that fixation artifacts are a product of unequal fixation rates inside and outside of puncta. That is, when the in-puncta and out-of-puncta fixation rates are equal, there is no fixation artifact; however, when the in-puncta fixation rate is significantly faster, fixation causes the enhanced appearance of LLPS, and when the out-of-puncta fixation rate is significantly faster, fixation causes the diminished appearance of LLPS. Furthermore, our model predicts that a fast overall fixation rate relative to the dynamics of the underlying protein-protein interactions can reduce the fixation artifact.

A Fast Overall Fixation Rate Relative to Binding Dynamics Can Minimize Fixation Artifacts

As discussed above, our model suggests that when the overall fixation rate is fast compared with the dynamics of targeted protein-protein interactions, fixation artifacts can be minimized even with unequal fixation rates in and out of puncta. In order to test this prediction experimentally, we focused on Halo-TAF15(IDR), which exhibits significantly diminished LLPS behavior upon fixation, and TAF15(IDR)-Halo-FTH1, which does not exhibit a significant fixation artifact. The fact that fixation of both Halo-TAF15(IDR) and TAF15(IDR)-Halo-FTH1 are completed within 1~2 minutes suggests comparable overall fixation rates of the two proteins. Thus, our model predicts that TAF15(IDR)-Halo-FTH1 has more stable homotypic interactions than Halo-TAF15(IDR), resulting in a higher relative overall fixation rate of the former than the latter. To test this prediction, we performed live-

cell single-molecule imaging of Halo-TAF15(IDR) and TAF15(IDR)-Halo-FTH1 and measured their binding residence times (RTs) at respective droplet-like puncta. Using established single-particle tracking (SPT) analysis[22], we found the RTs of TAF15(IDR) and TAF15(IDR)-FTH1 to be 10.23 ± 1.10 and 64.15 ± 11.65 seconds, respectively (**Figure 2**). This result suggests significantly more stable binding of TAF15(IDR)-FTH1 than TAF15(IDR). Together, these imaging data are consistent with our model's prediction that a fast overall fixation rate relative to binding dynamics can minimize fixation artifacts.

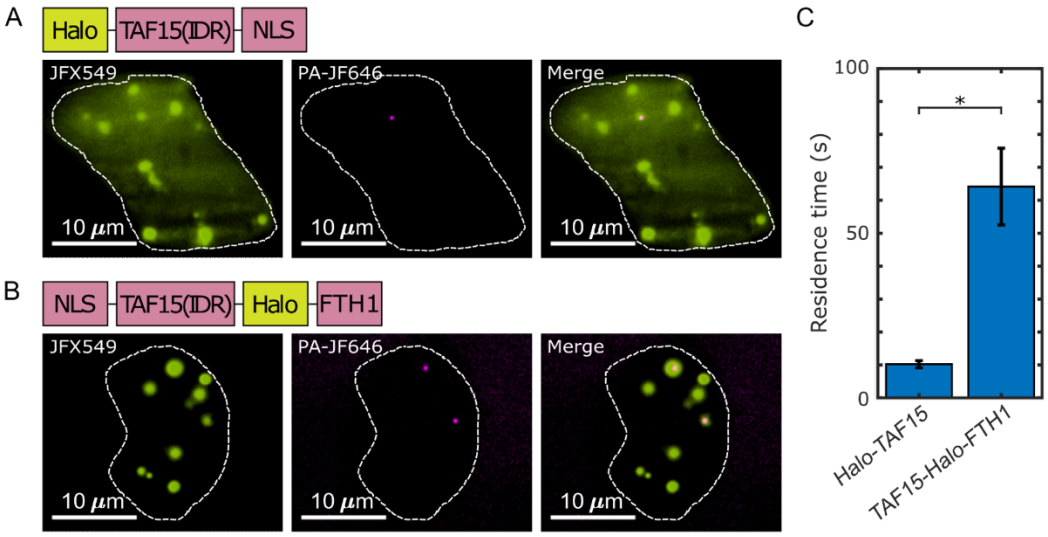

Figure 2. **The residence times of proteins in their droplet-like puncta vary**. Shown are individual frames from two-color single-molecule movies of (A) Halo-TAF15(IDR) and (B) TAF15(IDR)-Halo-FTH1. Each protein was labelled with a lower concentration of a photoactivatable dye for SPT (20 nM PA-JF646, magenta) and a higher concentration of non-photoactivatable dye for visualization of the droplet-like puncta (100 nM JFX549, yellow). A white dashed line outlines the nucleus. (C) The mean residence time of TAF15(IDR)-Halo-FTH1 in its puncta is significantly longer than that of Halo-TAF15(IDR) in its puncta. The value for each protein is averaged from 20 cells measured in 3 independent transfection and imaging sessions. Error bars represent standard errors. Asterisk indicates a significant difference between the two proteins ($P < 0.05$, Wilcoxon rank-sum test).

Discussion

Understanding situations in which PFA fixation can properly preserve live-cell conditions is essential in judicious applications of fixation-based biological techniques. Because approaches for rigorous determination of LLPS *in vivo* remain lacking[25] and detection of local high-concentration regions of an endogenously expressed protein in fixed cells via immunofluorescence has been widely used as evidence for LLPS[26–32], understanding how well fixation preserves LLPS behaviors is important for justifying the immunofluorescence-based diagnosis method and for studying the functional relevance of LLPS *in vivo*. In prior work, we imaged various LLPS systems in living cells before and after PFA fixation, quantified parameters that describe LLPS appearance in cells, and showed that fixation can either enhance or diminish the apparent LLPS behaviors *in vivo*. Lowering the PFA concentration and adding GA to PFA did not remove the fixation artifacts. For the first time, our work reveals an important caveat in using fixation-based methods to detect and characterize LLPS *in vivo* and suggests an advantage of using live-cell imaging to study LLPS systems over fixed-cell experiments. However, not all the proteins we examined have their puncta-forming or apparent LLPS behaviors in cells changed upon fixation. For example, PFA fixation faithfully preserves the appearance of EWS::FLI1 puncta in cells (**Figure 1c**). Nevertheless, our work points out a necessity to use live-cell imaging to confirm LLPS behaviors previously characterized with fixed-cell experiments. Live imaging techniques that allow estimation of protein diffusion coefficients within specific cellular compartments, e.g., SPT[38,39] and fluorescence correlation spectroscopy[40], can be useful alternative approaches for diagnosing LLPS *in vivo* without the potential artifact of fixation, as diffusion dynamics are recently shown to be affected by LLPS[41,41–45].

We note that fixation-induced changes of LLPS appearance may lead to potential misinterpretation of the functional relevance of LLPS in cellular processes. For example, recent work has uncovered that effective transcriptional activation requires an optimum of TF IDR-IDR interactions within TF hubs formed at target genes and that overly high levels of IDR-IDR interactions pushing the system toward LLPS can repress transcription[44,46]. Future characterization of the functionally optimal interaction level will require quantification of the sizes of hubs or droplet-like puncta while measuring transcription

activity. Because a fixation-induced increase or decrease in puncta sizes may lead to inaccurate determination of the functional optimum, scrutiny will be required in choosing between live-cell and fixed-cell imaging methods for quantifying LLPS appearance in these types of studies. Moreover, given that fixation can artificially generate intranuclear puncta of EGFP-EWS(IDR) that is homogenously distributed across the live cell nucleus, extra caution is required in interpreting immunofluorescence-detected intracellular puncta of an endogenously expressed protein as LLPS, as the same puncta-generating fixation artifact might happen to the protein even when it is not phase separating in living cells. To confirm puncta formation, counterpart live-cell imaging of the endogenous protein will be necessary, which requires engineering the cells, e.g., by CRISPR, to fluorescently tag the protein.

To understand the factors that can cause fixation-induced changes of LLPS appearance in the cell, we simulated the changes through kinetic modeling, which reveals that the dynamics of protein of interest (POI) binding to and dissociating from puncta, the absolute fixation rate of POI, and different fixation rates of POI in and out of puncta all play a role in inducing the fixation artifacts. Our kinetic model takes previous work studying fixation artifacts in the context of protein-DNA interactions[8,9,47] one step further by considering two fixed states of POI instead of one state, which are fixation in and out of puncta with different rates due to distinct local protein composition and concentrations. We then used live-cell single-molecule imaging experiments to demonstrate that as predicted by our model, a fast overall fixation rate of POI relative to its puncta-binding dynamics can minimize fixation artifacts.

We emphasize that because our four-state model makes no assumptions about any state being phase-separated, the logical implications of our model can extend beyond LLPS to other biomolecular transactions and cellular structures that have been found not well preserved by fixation or immunofluorescence, including localizations of cilia proteins[48], clustering of cell membrane receptors[49], splicing speckle formation[50], and chromatin organization and protein binding[9,51–55]. Our model can similarly extend beyond PFA to other fixatives. This is useful because different fixatives have been chosen for studying different types of structures. For example, PFA fixation is often preferable for preserving soluble proteins over dehydration fixatives such as methanol[56,57], yet methanol fixation can be preferable over PFA for

preserving proteins bound to mitotic chromatin[53,55]. Generally, our model predicts that fixation artifacts will occur whenever a protein can exist in multiple states that have different rates of fixation, and this artifact is most severe when the fixation is slower than the transition between states. For PFA fixation, because its rate is sensitively dependent on the amino acid sequence of POI, the structure of POI, and POI's cross-linked partners[12–14,58], POI in different states likely has different PFA fixation rates regardless of the type of interaction it undergoes.

One distinction between our study and previous studies is that we observe that PFA fixation can enhance apparent protein-protein interactions or LLPS behaviors in the cell, suggesting faster fixation for POI in the bound than dissociated state, whereas fixation has only been reported to diminish protein-DNA interactions, suggesting slower fixation for POI in the bound state[8,9,47]. We hypothesize that this is because fixing the bound state of an LLPS system (within puncta) is dominated by cross-linking reactions between IDRs enriched in puncta, which have reactive residues better exposed to solvent due to lack of well-defined tertiary structures and thereby likely cross-link faster than structured domains cross-linking to DNA[14]. It will be of future interest to measure fixation rates of different biomolecules including IDRs, structured proteins, and nucleic acids to prove the proposed chemical mechanism underlying fixation artifacts. Since our simulated results highlight the role of absolute fixation rates in the outcome of fixation, another future endeavor will be to design novel fixatives with significantly faster cross-linking rates than biomolecular interactions to eliminate fixation artifacts in the cell.

Materials and Methods

Cell Line and Sample Preparation. U2OS cells were grown in 1 g/L DMEM media (ThermoFisher, 10567014) supplemented with 10% FBS (Fisher Scientific, SH3039603) and 1% penicillin-streptomycin (ThermoFisher, 15140122). The cells were split onto an imaging plate (Mattek, P35G-1.5-14-C) and transfected with fluorescent protein constructs with Lipofectamine 3000 (Fisher Scientific, L3000001) according to manufacturer's instructions. One day after transfection, the culture media was changed to phenol-red-free DMEM (ThermoFisher, 11054001) with 10% FBS and 1% penicillin-streptomycin. For experiments

with additional glycine, glycine (Fisher Scientific, BP381-5) was added to the phenol red free media so that the final concentration was 50 mM (and 25 mM after the addition of 8% PFA, see below). It should be noted that normal DMEM media already contains 0.4 mM glycine. The knock-in A673 cell line expressing endogenous EWS::FLI1-Halo[22] was grown in 4.5 g/L DMEM media (ThermoFisher, 10566016) with 10% FBS (Fisher Scientific, SH3039603) and 1% penicillin-streptomycin (ThermoFisher, 15140122). The cells were similarly split onto an imaging plate (Mattek, P35G-1.5-14-C) and the culture media was changed to phenol-red-free DMEM (ThermoFisher, 31053028) just before imaging. The U2OS cell line used here was validated by whole-genome sequencing as described in[59]. The knock-in A673 cell line was generated by genome editing of the A673 cell line that was comprehensively authenticated by ATCC before distribution (ATCC, CRL-1598). By the time that the A673 cells had been genome-edited and used in this work, they had been cultured for 15 passages since purchased from ATCC. The genomic sequence of the locus encoding EWS::FLI1-Halo in the knock-in A673 cell line was confirmed by Sanger sequencing. Both U2OS and A673 cell lines were tested for mycoplasma using PCR-based assays in February 2022.

Fluorescence Microscopy. Confocal fluorescence microscopy was performed on Zeiss LSM 980 in the point-scanning mode with a 63x oil objective (Zeiss, 421782-9900-000). The pinhole was set to 1 airy unit for different emission wavelengths. The images displayed in the manuscript are maximum z-projections of z-stack images. A673 cell expressing endogenous EWS::FLI1-Halo were imaged in the Airyscan mode of the same Zeiss LSM 980 microscope. All postprocessing parameters in the Airyscan analysis module were kept constant to guarantee a fair comparison between the images taken before and after fixation. The culture dish contained 1 mL of phenol red free media, so that when 1 mL of 8% paraformaldehyde (PFA) (VWR, 100503-917) in PBS buffer was added to the dish, the final concentration of PFA was 4%. To achieve final PFA concentrations of 1%, 2%, and 8%, 1 mL of 2%, 4%, and 16% of PFA were diluted in PBS buffer and added to the culture dishes containing 1 mL of phenol red free media. A final concentration of 0% was achieved by following the same protocol only using 1 mL of PBS buffer in place of PFA. To achieve final

concentration of 4% PFA with 0.2% glutaraldehyde (GA) (Sigma-Aldrich, 340855-25ML), 1 mL of 8% PFA with 0.4% GA in PBS buffer was added to the culture dishes. After waiting 10 mins to allow PFA or PFA/GA fixation to complete, images of the same cells are taken again. For experiments with glycine, the final concentration of glycine after PFA addition was 25 mM and 8% PFA was used so that PFA was still in molar excess. Independent transfection and imaging sessions were performed on different days using different plates of cells.

LLPS Parameter Quantification. The three parameters we quantified were the number of puncta, surface roughness, and punctate percentage. The source code used to analyze the images is provided as a supplementary file "Puncta Quantification Processing Scripts.zip". To best compare the images of a cell before and after fixation, the two z-projection images were normalized so that the sum of the intensities within the nucleus is equal to 1. The border of the nucleus was manually drawn for each image. All analysis is done on normalized maximum z-projection images except for when calculating punctate percentage. We measured the number of puncta by quantifying the number of peaks within the nucleus. Specifically, the image was exported from MATLAB into ImageJ[60] using MIJ[61], and the "find maxima" processing function was used (**Figure 3**) with the same noise tolerance for both the live and fixed cell images.

To quantify the surface roughness of a cell nucleus image, the standard deviation of fluorescence intensities in the nucleus were compared before and after fixation (**Figure 8**). Utilizing this method of comparing images before and after fixation allows for quantification of change of nucleoplasm without peak fitting. The addition of structures such as puncta within a chosen patch will increase the standard deviation. Nuclei with puncta resulted in skewed (non-normal) distributions of intensities[62], leading to higher standard deviations.

The punctate percentage was determined with the first few steps identical to measuring the number of puncta as described above. The border of the nucleus was manually identified, the images were normalized, and preliminary peak locations were identified on maximum z-projection images using the "find maxima" function in ImageJ. The "find maxima" function

does not pick the perfect center of each punctum. Thus, to measure the full width at half maximum (FWHM) of a punctum, we made 36 different radial slices of the punctum crossing the preliminary punctum center pixel, extracted the intensity profile for each radial slice to calculate the punctum's FWHM, and selected the highest FWHM as the corresponding radial slice must have gone through the true center of the punctum. We then made a sum z-projection of the z-stack images, drew a circle with the maximum FWHM as its diameter centering the true central pixel of each punctum on the sum image, and integrated the fluorescence intensity across all circles. (**Figure 9**). The punctate percentage is calculated by dividing the in-circle total fluorescence intensity with the total fluorescence intensity integrated across the nucleus in the sum image.

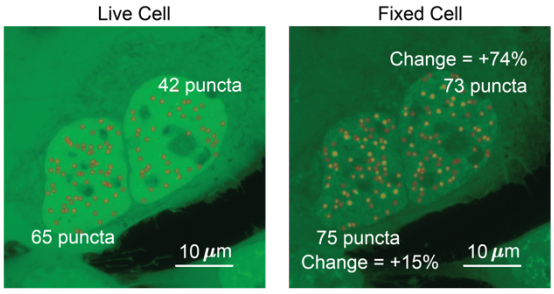

Figure 3. Determination of the number of puncta in the cell nucleus. Two cells expressing EGFP-TAF15(IDR) have the number of puncta before and after fixation compared. The cell on the left shows an increase of 10 puncta, a change of 15%. The cell on the right shows a change of 31 puncta, a change of 74%.

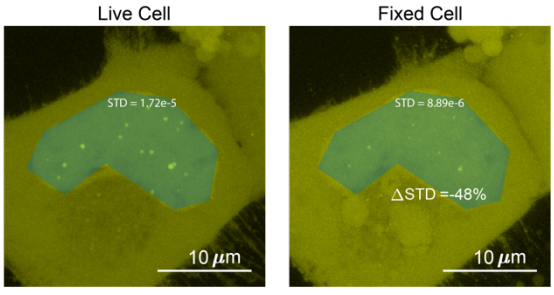

Figure 4. Determination of the surface roughness of a cell nucleus image. We drew a blue patch that covers the nucleus of a cell expressing Halo-TAF15(IDR) and compared the standard deviation of the pixel intensity within the blue patch before and after fixation. The change in standard deviation between the two images is -48%.

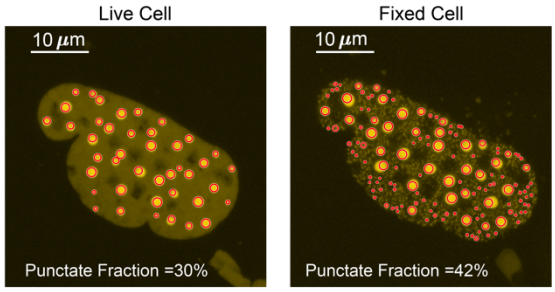

Figure 5. Determination of the punctate percentage. The punctate percentage of DsRed2-TAF15(IDR) is compared before and after fixation. The red circles represent the boundary within which the integrated fluorescence is considered "in puncta".

Single-particle Tracking (SPT). SPT of Halo-tagged TAF15(IDR) and TAF15(IDR)-FTH1 were performed on a Nikon Eclipse Ti2 TIRF microscope with a 100x/NA 1.49 oil-immersion objective (CFI SR HP Apochromat TIRF 100XAC Oil) under highly inclined and laminated optical sheet (HILO) illumination[64]. PA-JF646 was activated and excited under variable powers by 405 nm and 640 nm laser lines, respectively, while JFX549 was excited by a 561 nm laser line. The incubation chamber was held humidified at a 37°C with 5% CO_2 and the objective was also heated to 37°C.

Halo-tagged TAF15(IDR) and TAF15(IDR)-FTH1 were overexpressed in U2OS cells and stained with 100 nM JFX549 and 20 nM PA-JF646. Droplet-like puncta were visualized in the JFX549 channel, while individual molecules detected in the PA-JF646 channel were tracked in real-time. A low 405 nm activation power was used to ensure sufficiently sparse activation of PA-JF646-labeled proteins and allow for SPT. For SPT in the PA-JF646 channel, long camera exposure time (500 ms per frame, 2000 frames) blurred out faster

diffusing molecules and ensured that we only detect bound molecules. In the JFX549 channel, time-lapse images (500 ms per frame, one frame every 10 seconds) were taken to track the location of droplet-like puncta during the entire acquisition while limiting the effects of photobleaching.

The analysis was performed following Chong et al.[22] and is briefly described below. Single-molecule data from the PA-JF646 channel was analyzed using a SLIMfast[65], a GUI based on a MATLAB implementation of the MTT algorithm[66], and is available in the supplemental materials of Teves et al.[9]. SPT analysis was performed using the following parameters: localization error: 10^{-6}; deflation loops: 3; maximum number of competitors: 5; maximal expected diffusion constant ($\mu m^2/s$): 0.1.

Binary masks of the droplet-like puncta were generated from the JFX549 channel using custom-written Macros in ImageJ from Chong et al.[22]. Using custom-written MATLAB code also from Chong et al.[22], single-molecule trajectories were then sorted into in-puncta and out-of-puncta trajectories based on the fraction of time a molecule spent in the punctum, F. A trajectory with F > 50% was considered in-puncta and one with F < 5% was considered out-of-puncta. We only focused on the in-puncta trajectories.

Survival probability curves were then generated from the in-puncta trajectories and fit to the following two-component exponential model.

$$P(t) = Ae^{-k_1 t} + (1 - A)e^{-k_2 t},$$

$$1/k_1 = \tau_{ns}, 1/k_2 = \tau_s,$$

with τ_{ns} and τ_s as the specific and nonspecific residence times, respectively. Here, we only focused on the specific residence times.

In order to correct for photobleaching[59], the specific residence time of histone H2B (which is largely immobile on the chromatin) was measured via SPT as described above, except on

all trajectories rather than doing the in-puncta and out-of-puncta classification. We used PA-JF646-tagged H2B-Halo that was stably expressed in U2OS cells and imaged under illumination and acquisition parameters identical to those used to image Halo-tagged TAF15 and TAF15-FTH1. The corrected specific residence times of the Halo-tagged TAF15 and TAF15-FTH1 ($\tau_{\mathrm{corrected}}$) were computed based on the following model.

$$\tau_{\mathrm{corrected}} = 1/(1/\tau_{\mathrm{s}} - 1/\tau_{\mathrm{H2B}}),$$

with τ_{H2B} as the specific residence time of H2B.

Independent experiments were performed across at least three days for both Halo-tagged TAF15 and TAF15-FTH1. In each session, multiple movies of both constructs were taken along with three movies of Halo-tagged H2B to perform correction for that specific day. We reported the mean corrected residence times.

Statistical Analysis. Non-parametric tests used throughout because the data were often not normally distributed. Statistical significance of the LLPS parameters was calculated using the Wilcoxon signed-rank test and statistical significance of the residence times from SPT was using the Wilcoxon signed-rank test[67]. The Wilcoxon signed-rank test and Wilcoxon rank-sum test were performed using the built-in MATLAB functions *signrank* and *ranksum*, respectively.

References

1. Solomon, M. J. & Varshavsky, A. Formaldehyde-mediated DNA-protein crosslinking: a probe for in vivo chromatin structures. in Proceedings of the National Academy of Sciences vol. 82 6470–6474 (1985).
2. Quinodoz, S. A. *et al.* Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174**, 744-757 724 (2018).
3. Richter, K. N. *et al.* Glyoxal as an alternative fixative to formaldehyde in immunostaining and super-resolution microscopy. *The EMBO journal* **37**, 139–159 (2018).

4. Moter, A. & Göbel, U. B. Fluorescence in situ hybridization (FISH) for direct visualization of microorganisms. *Journal of microbiological methods* **41**, 85–112 (2000).
5. Sutherland, B. W., Toews, J. & Kast, J. Utility of formaldehyde cross-linking and mass spectrometry in the study of protein–protein interactions. *Journal of mass spectrometry* **43**, 699–715 (2008).
6. Chen, F., Tillberg, P. W. & Boyden, E. S. Expansion microscopy. *Science* **347**, 543–548 (2015).
7. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM. *Nature methods* **3**, 793–796 (2006).
8. Schmiedeberg, L., Skene, P., Deaton, A. & Bird, A. A temporal threshold for formaldehyde crosslinking and fixation. *PLoS One* **4**, 4636 (2009).
9. Teves, S. S. *et al.* A dynamic mode of mitotic bookmarking by transcription factors. *Elife* **5**, e22280 (2016).
10. Gavrilov, A., Razin, S. V. & Cavalli, G. In vivo formaldehyde cross-linking: it is time for black box analysis. *Briefings in functional genomics* **14**, 163–165 (2015).
11. Shishodia, S. *et al.* NMR analyses on N-hydroxymethylated nucleobases–implications for formaldehyde toxicity and nucleic acid demethylases. *Organic & biomolecular chemistry* **16**, 4021–4032 (2018).
12. Kamps, J. J., Hopkinson, R. J., Schofield, C. J. & Claridge, T. D. How formaldehyde reacts with amino acids. *Communications Chemistry* **2**, 1–14 (2019).
13. Metz, B. *et al.* Identification of formaldehyde-induced modifications in proteins: reactions with model peptides. *Journal of Biological Chemistry* **279**, 6235–6243 (2004).
14. Hoffman, E. A., Frey, B. L., Smith, L. M. & Auble, D. T. Formaldehyde crosslinking: a tool for the study of chromatin complexes. *Journal of Biological Chemistry* **290**, 26404–26411 (2015).
15. Gibbs, J. W. On the equilibrium of heterogeneous substances. (1879).
16. Graham, T. X. Liquid diffusion applied to analysis. *Philosophical transactions of the Royal Society of London* 183–224 (1861).
17. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid-liquid phase separation in biology. *Annual review of cell and developmental biology* **30**, 39–58 (2014).
18. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nature reviews Molecular cell biology* **18**, 285–298 (2017).

19. Boeynaems, S. *et al.* Protein phase separation: a new phase in cell biology. *Trends in cell biology* **28**, 420–435 (2018).
20. Mitrea, D. M. & Kriwacki, R. W. Phase separation in biology; functional organization of a higher order. *Cell Communication and Signaling* **14**, 1–20 (2016).
21. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, 4382 (2017).
22. Chong, S. *et al.* Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555 (2018).
23. Kato, M. & McKnight, S. L. A solid-state conceptualization of information transfer from gene to message to protein. *Annual review of biochemistry* **87**, 351–390 (2018).
24. Li, P. *et al.* Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336–340 (2012).
25. McSwiggen, D. T., Mir, M., Darzacq, X. & Tjian, R. Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev* **33**, 1619–1634 (2019).
26. Hayes, M. H. & Weeks, D. L. Amyloids assemble as part of recognizable structures during oogenesis in Xenopus. *Biology open* **5**, 801–806 (2016).
27. Kato, M. *et al.* Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
28. Lin, Y. *et al.* Toxic PR poly-dipeptides encoded by the C9orf72 repeat expansion target LC domain polymers. *Cell* **167**, 789-802 712 (2016).
29. Maharana, S. *et al.* RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* **360**, 918–921 (2018).
30. Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
31. Wang, J., Shi, C., Xu, Q. & Yin, H. SARS-CoV-2 nucleocapsid protein undergoes liquid–liquid phase separation into stress granules through its N-terminal intrinsically disordered region. *Cell discovery* **7**, 1–5 (2021).
32. Yang, P. *et al.* G3BP1 is a tunable switch that triggers phase separation to assemble stress granules. *Cell* **181**, 325-345 328 (2020).
33. Irgen-Gioro, S., Yoshida, S., Walling, V. & Chong, S. Fixation can change the appearance of phase separation in living cells. *Elife* **11**, e79903 (2022).
34. Wang, J. *et al.* A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**, 688–699 (2018).

35. Bracha, D. *et al.* Mapping local and global liquid phase behavior in living cells using photo-oligomerizable seeds. *Cell* **175**, 1467-1480 1413 (2018).
36. Bellapadrona, G. & Elbaum, M. Supramolecular protein assemblies in the nucleus of human cells. *Angewandte Chemie* **53**, 1534–1537 (2014).
37. Grünewald, T. *et al.* Ewing sarcoma. *Nature reviews Disease primers* **4**, 5 (2018).
38. Hansen, A. S. *et al.* Robust model-based analysis of single-particle tracking experiments with Spot-On. *eLife* **7**, 33125 (2018).
39. Heckert, A., Dahal, L., Tjian, R. & Darzacq, X. Recovering mixtures of fast-diffusing states from short single-particle trajectories. *Elife* **11**, e70169 (2022).
40. Lanzanò, L. *et al.* Measurement of nanoscale three-dimensional diffusion in the interior of living cells by STED-FCS. *Nature communications* **8**, 65 (2017).
41. Heltberg, M. L., Miné-Hattab, J., Taddei, A., Walczak, A. M. & Mora, T. Physical observables to determine the nature of membrane-less cellular sub-compartments. *eLife* **10**, 69181 (2021).
42. McSwiggen, D. T. *et al.* Evidence for DNA-mediated nuclear compartmentalization distinct from phase separation. *eLife* **8**, 47098 (2019).
43. Miné-Hattab, J. *et al.* Single molecule microscopy reveals key physical features of repair foci in living cells. *eLife* **10**, 60577 (2021).
44. Chong, S. *et al.* Tuning levels of low-complexity domain interactions to modulate endogenous oncogenic transcription. *Mol Cell* **82**, 2084-2097.e5 (2022).
45. Ladouceur, A. M. *et al.* Clusters of bacterial RNA polymerase are biomolecular condensates that assemble through liquid-liquid phase separation. in *Proceedings of the National Academy of Sciences of the United States of America* vol. 117 18540–18549 (2020).
46. Trojanowski, J. *et al.* Transcription activation is enhanced by multivalent interactions independent of phase separation. *Mol Cell* **82**, 1878-1893.e10 (2022).
47. Poorey, K. *et al.* Measuring chromatin interaction dynamics on the second time scale at single-copy genes. *Science* **342**, 369–372 (2013).
48. Hua, K. & Ferland, R. J. Fixation methods can differentially affect ciliary protein immunolabeling. *Cilia* **6**, 1–17 (2017).
49. Stanly, T. A. *et al.* Critical importance of appropriate fixation conditions for faithful imaging of receptor microclusters. *Biology Open* **5**, 1343–1350 (2016).
50. Neugebauer, K. M. & Roth, M. B. Distribution of pre-mRNA splicing factors at sites of RNA polymerase II transcription. *Genes and development* **11**, 1148–1159 (1997).

51. Zarębski, M. *et al.* Translocation of chromatin proteins to nucleoli-The influence of protein dynamics on post-fixation localization. *Cytometry. Part A : the journal of the International Society for Analytical Cytology* **99**, 1230–1239 (2021).
52. Lorber, D. & Volk, T. Evaluation of chromatin mesoscale organization. *APL bioengineering* **6**, 010902 (2022).
53. Lerner, J. *et al.* Human mutations affect the epigenetic/bookmarking function of HNF1B. *Nucleic acids research* **44**, 8097–8111 (2016).
54. Pallier, C. *et al.* Association of chromatin proteins high mobility group box (HMGB) 1 and HMGB2 with mitotic chromosomes. *Molecular biology of the cell* **14**, 3414–3426 (2003).
55. Kumar, S., Chaturvedi, N. K., Kumar, S. & Tyagi, R. K. Agonist-mediated docking of androgen receptor onto the mitotic chromatin platform discriminates intrinsic mode of action of prostate cancer drugs. *Biochimica et biophysica acta 1783* 59–73 (2008).
56. Stadler, C., Skogs, M., Brismar, H., Uhlén, M. & Lundberg, E. A single fixation protocol for proteome-wide immunofluorescence localization studies. *Journal of Proteomics* **73**, 1067–1078 (2010).
57. Schnell, U., Dijk, F., Sjollema, K. A. & Giepmans, B. N. Immunolabeling artifacts and the need for live-cell imaging. *Nature methods* **9**, 152–158 (2012).
58. Metz, B. *et al.* Identification of formaldehyde-induced modifications in proteins: reactions with insulin. *Bioconjugate chemistry* **17**, 815–822 (2006).
59. Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R. & Darzacq, X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *Elife* **6**, e25776 (2017).
60. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature methods* **9**, 676–682 (2012).
61. Sage, D., Prodanov, D., Tinevez, J.-Y. & Schindelin, J. MIJ: making interoperability between ImageJ and Matlab possible. in *ImageJ User & Developer Conference* (2012).
62. Jachowicz, J. W. *et al.* Xist spatially amplifies SHARP recruitment to balance chromosome-wide silencing and specificity to the X chromosome. (2021).
63. Hoops, S. *et al.* COPASI—a complex pathway simulator. *Bioinformatics* **22**, 3067–3074 (2006).
64. Tokunaga, M., Imamoto, N. & Sakata-Sogawa, K. Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nat Methods* **5**, 159–161 (2008).
65. Sergé, A., Bertaux, N., Rigneault, H. & Marguet, D. Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes. *Nat Methods* **5**, 687–694 (2008).

66. Normanno, D. *et al.* Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. *Nat Commun* **6**, 7357 (2015).
67. Gibbons, J. D. & Chakraborti, S. Nonparametric statistical inference. (2014).

Chapter 3

DYNAMIC REGULATION OF ENDOGENOUS TRANSCRIPTION FACTOR HUBS AT SINGLE-MOLECULE RESOLUTION

Adapted from: **Yoshida, S.**, Zhong, Y., Kumagai, A., Dunphy, W. G. & Chong, S. Dynamic regulation of endogenous transcription factor hubs at single-molecule resolution. 2026.01.25.701606

Preprint at https://doi.org/10.64898/2026.01.25.701606 (2026).

Abstract

Eukaryotic transcription factors (TFs) form local, high-concentration hubs at specific genomic loci through dynamic, multivalent protein-protein interactions mediated by their low-complexity domains. The hub formation behavior plays an essential role in TFs' transcriptional activation activities. Characterizing the dimensions, dynamics, and regulation of TF hubs requires high-resolution imaging of TFs in their native cellular environment, but much of such biophysical characterization remains missing. Here, we combined CRISPR/Cas9-mediated genome editing and advanced quantitative cell imaging, including single-molecule microscopy, to investigate the dynamic behaviors of the endogenous oncogenic fusion TF EWS::FLI1 in Ewing sarcoma cells. We found that endogenous EWS::FLI1 forms dynamic, sub-diffraction-limit hubs with mechanisms of dissolution that prevent the hubs from achieving macroscopic liquid-liquid phase separation. Hub formation is a neomorphic behavior of EWS::FLI1 that is not directly conferred by its parental proteins, EWSR1 and FLI1. We found that during mitosis, EWS::FLI1 hubs dissolve, but EWS::FLI1 molecules continue to dynamically bind and unbind mitotic chromosomes, revealing a role of EWS::FLI1 in mitotic bookmarking. Nascent RNA destabilizes EWS::FLI1 hubs on chromatin, but it does not affect the dimensions of the hubs. Finally, we visualized endogenous EWS::FLI1 hubs upon treatment with various compounds that were previously indicated to affect EWS::FLI1 function. We found that LY2835219 and trabectedin significantly alter the nuclear distribution of endogenous EWS::FLI1, disrupting and

mislocalizing EWS::FLI1 hubs, respectively. This finding highlights the therapeutic potential of both compounds for Ewing sarcoma. Together, our results reveal new insights into the assembly and regulation of endogenous EWS::FLI1 hubs at an unprecedented resolution. The methodology developed here will be useful for characterizing the functional hubs of many regular and pathological TFs in the future.

Introduction

DNA-binding transcription factors (TFs) are crucial regulators of eukaryotic transcription, comprising DNA-binding domains and transactivation domains that interact with other TFs and components of the transcriptional machinery to activate transcription. The transactivation domains usually contain intrinsically disordered low-complexity domains (LCDs) that do not fold into stable three-dimensional structures, thus are not amenable to investigation through the conventional structure-function framework. Recent work based on live-cell single-molecule imaging has revealed that TF LCDs undergo dynamic, multivalent, and selective protein-protein interactions, which drive the formation of high local concentration TF hubs at target genes. Under certain conditions, such as when LCDs are overexpressed, the hubs can develop into phase separated condensates. The multivalent interactions and hub formation behaviors of a TF enable the recruitment of more TFs[1,2], RNA polymerase II[3–7], and transcriptional cofactors to target genes[3,4,8], leading to transcriptional activation. Notably, endogenous target genes often require optimal levels of multivalent LCD-LCD interactions for transactivation; both low and excessive levels of multivalent interactions can repress transcription[9]. This “Goldilocks behavior” applies to a variety of hub-forming transcriptional regulators[10–12]. TF LCDs are also frequently mutated in diseases, including many types of cancers, and the pathological implications of aberrant TF hubs have been documented in an increasing number of reports[13–17]. We note that here we use the term “hub” to refer to local high-concentration TF clusters mediated by LCD-LCD interactions. Historically, similar clusters have been described using various other terms, most often "condensate", which was initially linked to liquid-liquid phase separation (LLPS) and now more broadly defined as membrane-less compartments that selectively concentrate

biomolecules, regardless of whether LLPS is involved[18]. We adopt this broader definition of condensates, preserve the original terminology used in the cited studies, and use "assembly" as an inclusive, mechanism-agnostic term to describe LCD-mediated biomolecular clusters.

Despite the critical role of TF hubs in transcriptional control and disease mechanisms, many questions about them remain unanswered due to limitations of widely used imaging and sample preparation methods. Specifically, most existing works have investigated TF hubs using conventional, diffraction-limited microscopy techniques, which are unable to precisely characterize the dynamics, spatial distribution, and dimensions of sub-diffraction objects, thereby limiting the ability to link the physical properties of small TF hubs to their functional contributions[19]. In addition, many studies imaged fluorescently labeled proteins overexpressed in cells. While valuable insights have been generated in such settings, a significant caveat is that multivalent interactions and hub formation behaviors are sensitively dependent on protein concentrations[20,21]. Studies that perform extensive characterization of endogenous TF hubs using high-resolution imaging techniques are highly desirable but remain very limited in number[2,22].

To fill in this gap, here we use single-molecule imaging methods, including photoactivated localization microscopy (PALM)[23] and single-particle tracking (SPT)[24,25], to characterize the dimensions, dynamics, and regulation of endogenous TF hubs in the cell. We use EWS::FLI1, an oncogenic fusion TF that is the hallmark of Ewing sarcoma, as the model TF in this study. It is encoded by a fusion gene produced by the t(11;22)(q24;q12) chromosomal translocation that fuses the intrinsically disordered LCD of EWSR1, an RNA-binding protein, with the well-structured DNA-binding domain of FLI1, a TF that plays roles in development and homeostasis[26]. We and others have previously reported that endogenous EWS::FLI1 forms local, high-concentration hubs at GGAA microsatellites via a combination of EWS::FLI1-DNA binding and homotypic, multivalent interactions of EWS LCD[1,27,28]. Yet, the dimensions of endogenous EWS::FLI1 hubs across the cell nucleus remain unquantified. It is also unknown whether the hub formation behavior is a property of either

parental protein of EWS::FLI1 (EWSR1 or FLI1) or a neomorphic property of the oncogenic fusion. Another unexplored question is how EWS::FLI1 hubs are regulated during mitosis, when some TFs are evicted from chromatin as it compacts, and others remain bound to specific genomic loci[29]. The behaviors of TF hubs during mitosis impact how the gene expression program is restored after cell division[30]. In addition, although RNA is known to play a role in nuclear organization[31] and modulate phase separation of specific transcription regulators, including MED1 and ENL[32,33], it remains unknown whether RNA contributes to regulating endogenous TF hubs. Also unexplored is whether any compounds can modulate endogenous pathological TF hubs, such as EWS::FLI1 hubs, despite emerging research identifying small molecules that target condensates involved in transcriptional dysregulation[34]. We address these open questions by imaging endogenous EWS::FLI1 hubs at single-molecule resolution in Ewing sarcoma cells. This work sheds light on both the mechanistic underpinnings of transcriptional control and the therapeutic targeting of pathological TFs.

Results

Endogenous EWS::FLI1 forms sub-diffraction-limit hubs that undergo dynamic turnover

To quantify the dimensions of endogenous EWS::FLI1 hubs, we performed PALM on our previously established knock-in Ewing sarcoma cell line A673, where we fused a HaloTag to the C-terminus of endogenous EWS::FLI1 via CRISPR/Cas9-mediated genome editing[1]. With this labeling, we visualized the dynamic behaviors of EWS::FLI1 in its native cellular context with single-molecule resolution. PALM, based on the single-molecule localization technique, offers a spatial resolution of ~20 nm with our imaging setup, significantly higher than diffraction-limited microscopy. A well-documented issue with traditional PALM is spurious clustering caused by repeated appearances of the same molecule across multiple frames, due to factors including frame discretization of a single emission event, variable photobleaching times, and photoblinking[35–41]. Among different correction methods, pair-correlation PALM (PC-PALM) is known to faithfully report true clustering, making it an excellent tool for probing the presence of clusters and their length scale[40,42,43]. We applied

PC-PALM to the raw single-molecule localizations of EWS::FLI1[40,44,45] to determine whether the apparent clustering of localizations represents true EWS::FLI1 clusters or artifacts. We observed statistically significant clustering of EWS::FLI1 with an average cluster correlation length of $\xi = 29.87 \pm 1.89\ nm$ (hereafter, mean $\pm$ SEM) (Figure 1a). The correlation length is known to provide an estimate of average cluster radius, though this is only an approximation and should not be interpreted as an exact measure[40] (Methods). Our measurement suggests that endogenous EWS::FLI1 hubs have an average diameter of ~60 nm (2ξ). In addition, while PC-PALM can robustly identify various molecule density fluctuations, including clustering, it cannot be used to generate a reconstructed image of the whole nucleus or to recover the distribution of cluster sizes across the nucleus. Nevertheless, PC-PALM confirms that endogenous EWS::FLI1 hubs are smaller than the diffraction limit of light and justifies the need for super-resolution microscopy for their characterization.

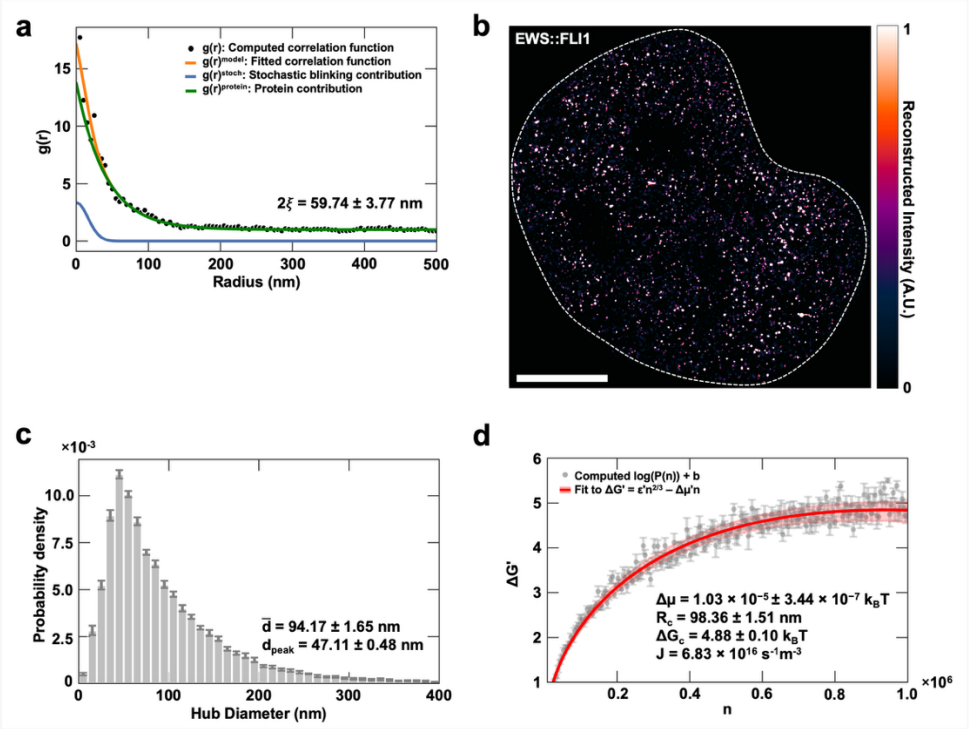

Figure 1. Super-resolution microscopy characterizes the dimensions of endogenous EWS::FLI1 hubs. (a) Pair-correlation analysis implemented as described previously[40,45]

characterizes the spatial distribution of EWS::FLI1 hubs. The computed pair-wise autocorrelation function of the raw PALM localizations (black dots) was fit to a model that accounts for the contributions of both true protein clustering and the multiple appearances of the same fluorophores due to stochastic blinking (orange curve). The protein correlation term and stochastic blinking correlation term are shown as green and blue curves, respectively. The correlation length of protein clusters, ξ, is proportional to the cluster size and 2ξ can be used as an estimate of average cluster diameter. Statistically significant clustering was observed with 2ξ = 59.74 ± 3.77 nm. (b) Photoblinking-corrected PALM image of a knock-in A673 cell expressing endogenous EWS::FLI1-Halo (labeled with the Halo ligand PA-JF549) reveals a punctate distribution of EWS::FLI1 in the nucleus, which forms numerous sub-diffraction-limit hubs. A white dashed contour outlines the cell nucleus. Scale bar is 5 μm. (c) The distribution of effective hub diameters generated by clustering the photoblinking-corrected localizations with DBSCAN[46]. (d) Plot of $-\log(P(n))$ vs n fitted with $\varepsilon' n^{2/3} - \Delta\mu' n$ yields values of ε' and $\Delta\mu'$ used to calculate the free energy barrier to nucleation and critical cluster size. Data is collected from 10 cells. SEM represents the uncertainty.

To extract the true EWS::FLI1 localizations for image reconstruction, we applied statistical corrections for photoblinking to the raw single-molecule localizations. Briefly, the blinking behavior of PA-JF549 in A673 cells was characterized within the framework of a stochastic model, and the resulting rate parameters were used to filter out repeated localizations arising from blinking and frame-discretization artifacts from raw PALM datasets (Methods). The resulting PALM reconstruction shows the punctate distribution of endogenous EWS::FLI1 (Figure 1b). Density-based spatial clustering of applications with noise (DBSCAN) was applied to the corrected localizations to identify individual EWS::FLI1 hubs[46]. We then plotted the distribution of hub diameters and found that the hubs have a mean diameter of $94.17 \pm 1.65\ nm$ and an estimated mode diameter of $47.11 \pm 0.48\ nm$ (Figure 1c). This is

consistent with the PC-PALM result, confirming that endogenous EWS::FLI1 forms sub-diffraction hubs.

We next extracted parameters governing the energetics and kinetics of endogenous EWS::FLI1 hub formation by analyzing the hub size distribution through the lens of classical nucleation theory, a powerful theoretical framework for predicting nucleation kinetics. While nucleation classically describes the first step in the formation of a new thermodynamic phase, its use here does not imply that EWS::FLI1 undergoes LLPS; rather, we apply it as a general mathematical framework for describing the stochastic formation of clusters as established in the field[47]. In the classical, mean-field treatment of homogeneous nucleation, ΔG, the free energy change for the nucleation of a cluster of size n, can be compactly written as[48]

$$\Delta G = \varepsilon n^{2/3} - \Delta\mu n \quad (1)$$

The first term in Eq. 1 represents the contribution from the surface energy to ΔG, i.e. the energetic cost of maintaining the interface between the cluster and the surrounding environment, encapsulated by the prefactor, ε. The energetic cost depends on factors including the molecule size, the geometry of the interface, and the magnitude of interfacial energy, and it always increases with the cluster size. The second term in Eq. 1 represents the contribution from the bulk free energy to ΔG, i.e. the change in free energy due to the difference in chemical potential, $\Delta\mu$, between the inside of the cluster and the surrounding environment, where $\Delta\mu = \mu - \mu^* = k_B T \log(c/c^*)$. c/c^* is the ratio of the actual in-cluster concentration to the equilibrium or saturation concentration. While the concept of a saturation concentration is often associated with LLPS, here it instead denotes the concentration above which the formation of stable clusters becomes favorable, independent of macroscopic phase separation. In subsaturated systems ($\Delta\mu < 0$), both the surface and bulk energy contributions increase with cluster size, making any clusters that spontaneously form due to thermal fluctuations tend to dissipate. Conversely, in supersaturated systems ($\Delta\mu > 0$), the surface energy still increases with increasing cluster size, but the bulk energy decreases with increasing cluster size. As a result, while some clusters still tend to dissipate, others can overcome the free energy barrier and continue to grow larger. ε and $\Delta\mu$ together

set the free energy barrier to nucleation, ΔG_c, and the corresponding critical radius, R_c, which represents the minimum radius above which it is thermodynamically favorable for a cluster to continue growing rather than being driven to dissipation. ε, $\Delta\mu$, R_c, and ΔG_c, can all be determined from the distribution of cluster sizes, $P(n)$, with[49]

$$\Delta G = -k_B T \log P(n) + b \quad (2)$$

Here, b is a self-consistent term that ensures the normalization of $P(n)$. The derivation of Eq. (1), along with expressions for R_c and ΔG_c, is included in the Methods.

We fitted Eq. (2) against the EWS::FLI1 hub size distribution (Figure 1d), and remarkably, the shape of the distribution matches what would be expected of a supersaturated system in the framework of the classical nucleation theory. The difference in chemical potential ($\Delta\mu$) is $1.03 \times 10^{-5} k_B T$, whose positive sign suggests that at endogenous expression levels, EWS::FLI1 is supersaturated. The critical radius (R_c) is $98.36 \pm 1.51\, nm$, and the free energy barrier to nucleation (ΔG_c) is $4.88 \pm 0.10\, k_B T$. The relatively modest free energy barrier suggests that EWS::FLI1 hub nucleation is a frequent and reversible process.

Nucleation of a supersaturated state is typically transient and followed by cluster growth. As such, if we consider supersaturated nucleation of EWS::FLI1 hubs in isolation, an increasing number of hubs would overcome the free energy barrier and continue to grow beyond the critical radius, eventually becoming large LLPS droplets. However, the actual size distribution of EWS::FLI1 hubs is stable over time in the cell, suggesting that a balance is maintained between nucleation of new hubs and dissolution of hubs larger than the critical size. Using the Zeldovich form of the steady-state nucleation rate equation[50] along with the energetic parameters extracted earlier, we estimate that EWS::FLI1 hubs have a nucleation rate of $6.83 \times 10^{16}\ m^{-3} s^{-1}$ (Methods). This rate should match the bulk dissolution rate to maintain the observed steady-state hub size distribution. As such, approximating the volume of an A673 nucleus as $500\ \mu m^3$ from our imaging data, the dissolution rate is about 34 hubs per second. While the mechanism underlying hub dissolution remains unknown, this estimate sets the timescale for dissolution, allowing us to narrow down possible dissolution

pathways. Previously reported processes of regulating intrinsically disordered region (IDR)-mediated assemblies or condensates, including chaperone-mediated dissolution[51–53], phosphorylation-mediated dissolution[54–56], and competition between multivalent interactions and DNA-protein binding[57], can occur at rates near 34 hubs per second. In contrast, slower processes, including targeted proteasomal degradation, are unlikely to operate on this timescale.

EWS::FLI1 hub formation is a neomorphic behavior not shared by EWSR1 or FLI1

Many fusion oncoproteins resulting from chromosomal translocations can phase separate or form biomolecular assemblies via multivalent interactions[58–60]. EWS::FLI1 is one of such oncoproteins. It remains unexplored if the assembly formation ability of a fusion oncoprotein is directly conferred by any of the parental proteins or a neomorphic property of the fusion. To address this question for EWS::FLI1, we used PALM, along with the above-described photoblinking corrections, to visualize the spatial distributions of EWSR1 and FLI1 in the cell and compare them with the distribution of EWS::FLI1. It was crucial to image both parental proteins at native expression levels, given that multivalent interactions and hub formation behaviors, if present, are highly dependent on protein concentration. To fluorescently label EWSR1, we engineered the A673 line with CRISPR/Cas9-mediated genome editing to fuse a HaloTag to the N-terminus of endogenously expressed EWSR1. Halo-tagging endogenous EWSR1 did not affect its expression levels (Figure 2a). Because FLI1 is not expressed in A673, we could not image endogenous Halo-tagged FLI1 in this cell line[61]. The A673 line is also not ideal for studying exogenously expressed FLI1, whose interaction behaviors are likely affected by the presence of endogenous EWS::FLI1 that may compete with FLI1 for the same binding sites on chromatin[62]. We instead opted to express exogenous Halo-tagged FLI1 in U2OS cells with no endogenous EWS::FLI1 or FLI1[63,64]. To ensure that we measure the intracellular distribution of FLI1-Halo at expression levels comparable to those of endogenous EWS::FLI1-Halo in A673 cells, we performed PALM for FLI1-Halo in U2OS and EWS::FLI1-Halo in A673 cells until complete depletion of fluorescence, i.e., no more molecules could be detected, quantified the total number of single-

molecule localizations detected throughout the PALM movie as a measure of the protein's expression level, and selected only the U2OS cells expressing FLI1-Halo at levels similar to endogenous EWS::FLI1-Halo in A673 cells for downstream analysis of FLI1 distribution (Figure 2b).

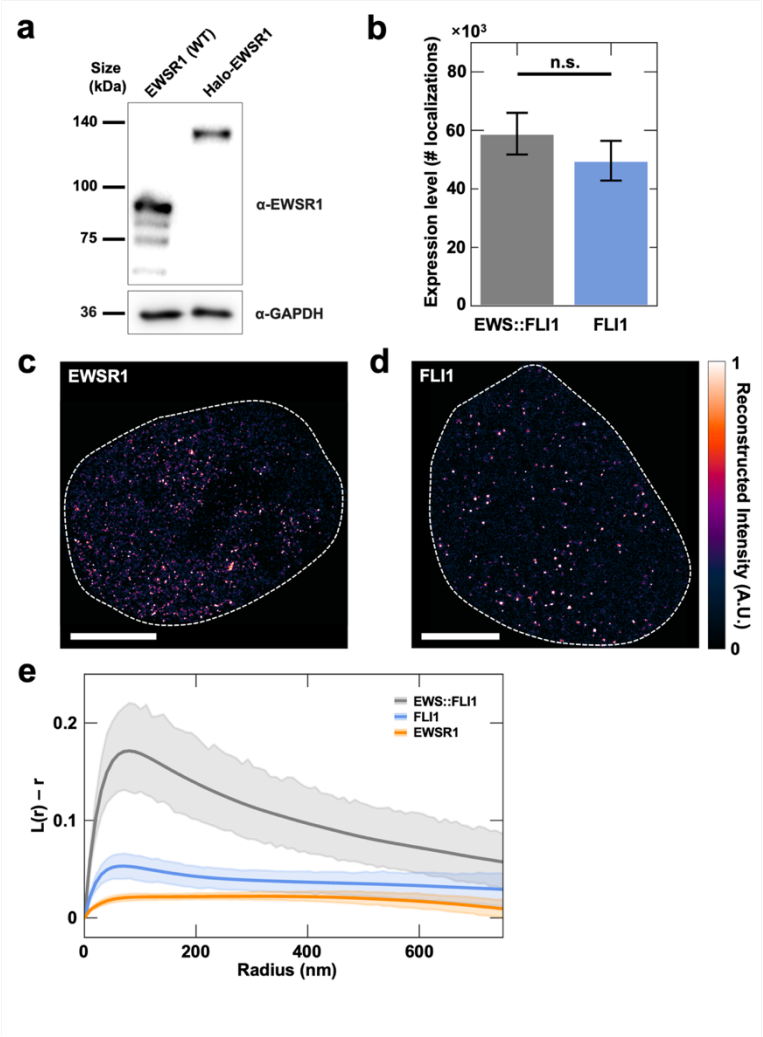

Figure 2. EWS::FLI1 has a stronger hub-forming ability than its parental proteins. (a) Western blot of EWSR1 and GAPDH from wild-type (WT) A673 and Halo-EWSR1 knock-in A673 cell lines. (b) Plot showing no statistically significant difference in the expression levels of endogenous EWS::FLI1-Halo in knock-in A673 cells and exogenous FLI1-Halo in U2OS cells ($p = 3.89 \times 10^{-1}$, Student's t-test). The

expression level is quantified by the total number of single-molecule localizations detected throughout the PALM movie of each cell. 10 cells were analyzed for each condition. Error bars represent SEM. Not statistically significant is abbreviated as n.s. (c-d) Photoblinking-corrected PALM images of a knock-in A673 cell expressing endogenous Halo-EWSR1 (c) and a U2OS cell transiently expressing FLI1-Halo at a level comparable to endogenous EWS::FLI1-Halo in A673 (d). The cells were stained with the Halo ligand PA-JF549. Both EWSR1 and FLI1 are less clustered than EWS::FLI1, though they still form puncta. A white dashed contour outlines the cell nucleus. Scale bars are 5 μm. (e) Ripley's L(r) - r curves show that over relevant length scales, EWS::FLI1 is more clustered than both EWSR1 and FLI1. 10 cells were analyzed for each condition. The semi-transparent regions represent 95% confidence intervals.

PALM revealed that both EWSR1 and FLI1 have punctate distributions in the cell nucleus when expressed at endogenous levels (Figure 2c and 2d). However, their puncta are significantly less pronounced than the hubs formed by endogenous EWS::FLI1. We quantified the protein distribution by computing Ripley's L function, $L(r)$, a spatial descriptive statistic that determines whether a set of points has a random, dispersed, or clustered distribution at a given length scale[65]. Specifically, we compared plots of $L(r) - r$ vs r for the localizations of EWS::FLI1, EWSR1, and FLI1 (Figure 2e), where randomly distributed localizations would yield a $L(r) - r$ value of 0, and the more punctate the localization distribution, the larger the magnitude of $L(r) - r$. As shown in Figure 2e, EWS::FLI1 has a significantly more punctate distribution than both its parental proteins, EWSR1 and FLI1. This result suggests that the hub-forming ability of EWS::FLI1 is neomorphic instead of being directly conferred by either of its parental proteins alone. We then applied PC-PALM and found that the approximate diameter of the hubs formed by EWSR1 and FLI1 were $2\xi = 31.65 \pm 2.03\ nm$ and $2\xi = 47.54 \pm 3.01\ nm$, respectively: significantly smaller than those formed by EWS::FLI1 (**Figure S1**). Although here we focus on a specific oncogenic fusion TF, EWS::FLI1, it is possible that some other aberrant fusion

proteins resulting from chromosomal translocation similarly have their intracellular distributions altered from their parental proteins and form neomorphic assemblies, and the assembly formation behaviors play a role in the oncogenesis of different cancer types beyond Ewing sarcoma.

EWS::FLI1 hubs dissolve during mitosis, and EWS::FLI1 dynamically interacts with mitotic chromatin

Since IDR-mediated biomolecular assemblies are involved in many cellular processes, their evolution during mitosis can impact many cellular functions. Despite this, the mitotic behaviors of only a handful of assembly-forming proteins have been investigated thus far; for example, mitotic hyperphosphorylation of the nucleolar proteins Ki-67 and NPM1 is reported to alter the charge blockiness of their IDRs, i.e. whether charged amino acids are arranged in large blocks or randomly dispersed, thereby modulating their condensation levels[66]. However, mitotic regulation of TF hubs remains unexplored. To fill this gap, we tracked the fate of endogenous EWS::FLI1 hubs throughout the cell cycle and compared the interaction and diffusion dynamics of EWS::FLI1 during mitosis and interphase. Since the association of many TFs with mitotic chromatin is known to be affected by cell fixation[67], here we imaged EWS::FLI1 hubs in live knock-in A673 cells that express endogenous EWS::FLI1-Halo. We treated the cells with the anti-mitotic agent nocodazole to synchronize them in the cell cycle. Meanwhile, we stained the cells with a DNA dye (Hoechst 33342) that reveals the morphology of chromatin or chromosomes, enabling identification of specific phases of the cell cycle. Using live-cell confocal fluorescence microscopy, we found that endogenous EWS::FLI1 hubs dissolve upon mitotic entry, with no visible hubs detected during prometaphase, metaphase, or anaphase. However, during telophase, the hubs began to reform, coinciding with chromatin decondensation (Figure 3a). Interestingly, despite the absence of visible EWS::FLI1 hubs through most of mitosis, EWS::FLI1 remained weakly enriched on mitotic chromatin, consistent with previous report by Teves et al. that many TFs bind to mitotic chromatin and serve as “mitotic bookmarks”, which maintain established transcriptional programs throughout the cell cycle and restart them in daughter cells after

mitosis[67]. Mitotic bookmarking by EWS::FLI1 is likely especially important in Ewing sarcoma cells, given its role in maintaining the oncogenic transformation phenotype.

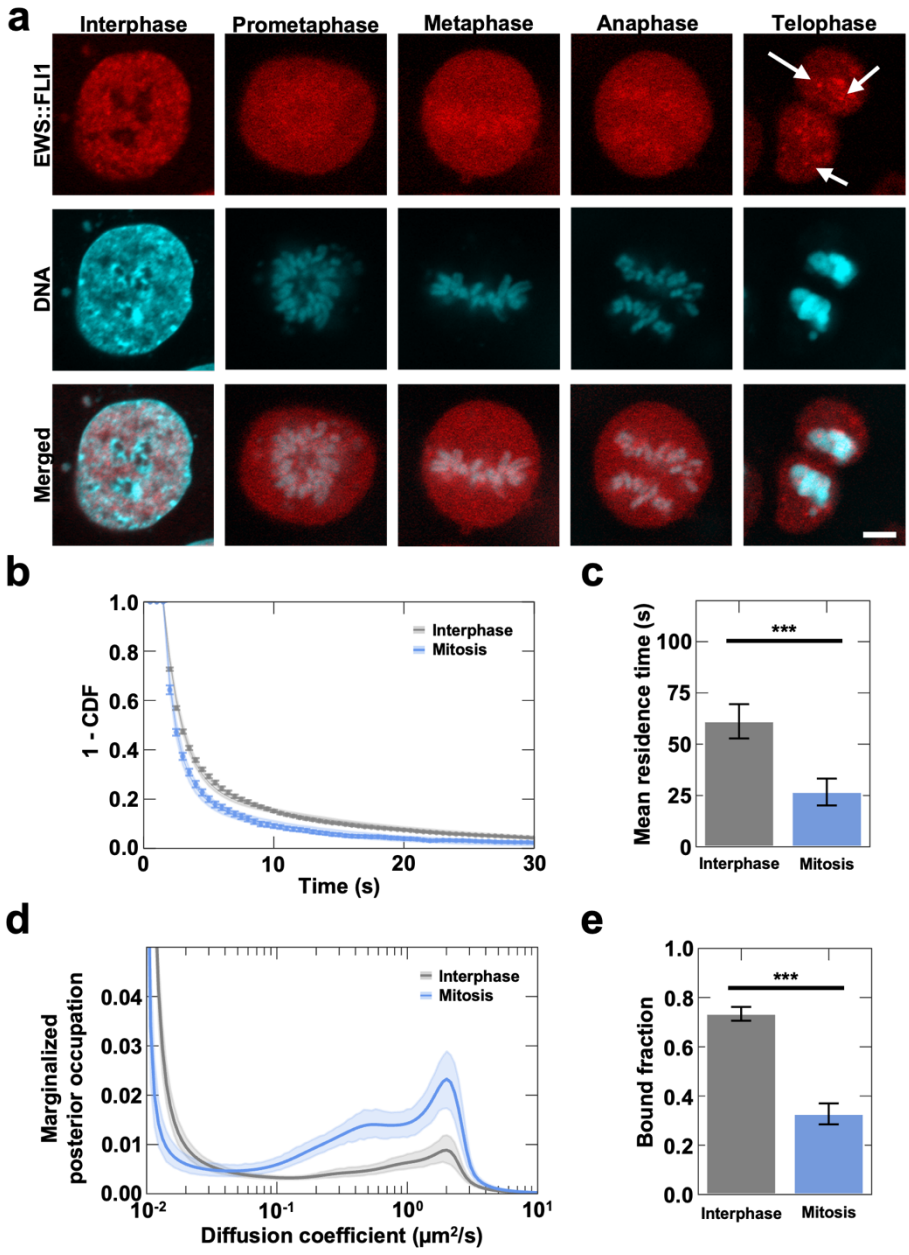

Figure 3. Interaction dynamics of EWS::FLI1 during mitosis. (a) Live-cell confocal images of the mitotic localization of endogenous EWS::FLI1-Halo in the knock-in A673 cell line. The cells were labeled with the Halo ligand JFX646 and the DNA stain Hoechst

33342. EWS::FLI1 hubs dissolve during mitosis and reform during telophase. Arrows highlight the reformed hubs. Scale bar is 5 μm. (b) Survival curves (scatter plots) were generated based on the binding residence times of individual EWS::FLI1 molecules on chromatin during interphase and mitosis, which were fit as described in the Methods (line plots). (c) The plot of the mean residence times shows that EWS::FLI1 binds more stably during interphase than during mitosis (***, $p = 1.46 \times 10^{-4}$, Mann-Whitney U test). For each condition, 21 cells were measured in three independent imaging sessions, each session performed on a different day. (d) EWS::FLI1 diffusion profiles show comparable diffusion coefficients during interphase and mitosis. Diffusion coefficients were extracted as the local maxima of the respective diffusion coefficient distributions. The semi-transparent regions represent 95% confidence intervals. (e) Plot of bound fractions shows that a greater fraction of EWS::FLI1 is bound during interphase than during mitosis (***, $p = 2.70 \times 10^{-10}$, Welch's t-test). EWS::FLI1 molecules were considered bound if they had a diffusion coefficient smaller than the threshold value $0.1\mu m^2/s$. For each condition in (d) and (e), 30 cells were measured in three independent imaging sessions, each session performed on a different day. Error bars represent SEM in (c) and (e).

We next used SPT to compare the chromatin binding dynamics of endogenous EWS::FLI1 during interphase and mitosis in the EWS::FLI1-Halo knock-in A673 cells[1]. Our SPT protocol has been described in detail previously[68]. Briefly, we performed single-molecule imaging with a relatively long exposure time (500 ms, "slow" SPT) such that quickly diffusing molecules are blurred out and bound molecules are clearly resolvable. We then fitted the survival curve generated using the binding residence times of all the detected molecules in a cell to a two-component model of dissociation that accounts for both specific and nonspecific binding (Figure 3b)[68]. This allowed us to extract the mean residence times of EWS::FLI1 specifically bound to chromatin. We found that EWS::FLI1 binds to chromatin with a mean residence time of $61.15 \pm 8.36\ s$ during interphase and $26.71 \pm$

$6.55\ s$ during mitosis (Figure 3c). The fact that EWS::FLI1 binds to mitotic chromatin more transiently coincides with the dissolution of EWS::FLI1 hubs during mitosis. This coincident timing is consistent with our previous finding that hub formation of EWS::FLI1 during interphase stabilizes its chromatin binding through multivalent LCD-LCD interactions in the hubs[1], further supporting the role of the multivalent interactions in regulating TF-chromatin binding.

We also compared the diffusion dynamics of endogenous EWS::FLI1 in the EWS::FLI1-Halo knock-in A673 cells during interphase and mitosis using SPT. To ensure the tracking of both bound and freely diffusing molecules in single-molecule images, we used a much shorter exposure time (5 ms, "fast" SPT) combined with stroboscopic illumination[25,69]. We then analyzed the resulting single-molecule trajectories using the state-array SPT (saSPT) algorithm, which extracts the diffusion coefficient distributions of a target protein without *a priori* knowledge of the number or identity of specific diffusing states of the protein[70] (Figure 3d). We found that while the characteristic diffusion coefficient of freely diffusing EWS::FLI1 varies minimally with the cell cycle, with its value being $1.96 \pm 0.02\ \mu m^2/s$ during interphase and $2.00 \pm 0.04\ \mu m^2/s$ during mitosis, the bound fraction of EWS::FLI1 significantly varies with the cell cycle, i.e., $73.41 \pm 2.80\%$ during interphase and $32.73 \pm 4.25\%$ during mitosis (Figure 3e). The decreased bound fraction during mitosis coincided with the dissolution of EWS::FLI1 hubs, the formation of which during interphase likely helps to retain EWS::FLI1 on chromatin.

Together, these results suggest that EWS::FLI1 hub assembly and disassembly are regulated throughout the cell cycle. The dissolution of hubs during mitosis coincides with decreased binding stability and bound fraction of EWS::FLI1 to chromatin. Nonetheless, during mitosis, a fraction of EWS::FLI1 remains bound to chromatin, suggesting it may function as a mitotic bookmark[67,71].

RNA modulates the hub assembling stability of EWS::FLI1 but not its spatial distribution

It has been increasingly appreciated that RNA can play a role in regulating biomolecular assemblies. For example, Henninger et al. recently proposed an RNA-mediated feedback mechanism for Mediator condensates, where condensate formation is promoted by low levels of RNA when genes start to be transcribed, and the condensates are dissolved upon accumulation of the transcribed RNA[33]. However, it remains unknown whether this model is applicable to endogenous TF hubs that often form at actively transcribed genes. To tackle this question, we investigated the role of RNA in the organization and dynamics of EWS::FLI1 hubs in the knock-in A673 cells that express endogenous EWS::FLI1-Halo[1] by imaging the hubs before and after we treated the cells with triptolide, a transcription initiation inhibitor that induces the proteasomal degradation of Rpb1, the largest subunit of RNA polymerase II[72]. Triptolide is known to globally inhibit transcription within minutes, but additional time is needed for all nascent RNAs to be degraded or diffuse away from their gene loci after transcriptional inhibition. To determine the time needed to deplete nascent RNA from the target genes of EWS::FLI1, we performed intron RNA fluorescence in situ hybridization (FISH)[70] to detect nascent RNA of two well-characterized target genes of EWS::FLI1, *ABHD6* and *CAV1*[73], at multiple timepoints after we treated the cells with 1 μM triptolide dissolved in DMSO. We found that the nascent RNA of both genes became undetectable two hours after triptolide treatment (Figure 4a). We then performed PALM to visualize the distribution of endogenous EWS::FLI1 after treating the cells with triptolide for two hours (Figure 4b). The $L(r) - r$ vs r plots showed that the spatial distribution of EWS::FLI1 in the nucleus after the triptolide treatment does not differ significantly from after the treatment of DMSO only for two hours (Figure 4c). By measuring the EWS::FLI1 hub dimensions after the triptolide treatment, we found that hubs have a mean diameter of $46.54 \pm 0.42\ nm$ and an estimated mode diameter of $90.03 \pm 2.25\ nm$ (Figure S2), which are not significantly different from those in cells treated with DMSO only (Figure 1c). These results suggest that RNA does not play a significant role in maintaining the assembly of EWS::FLI1 hubs.

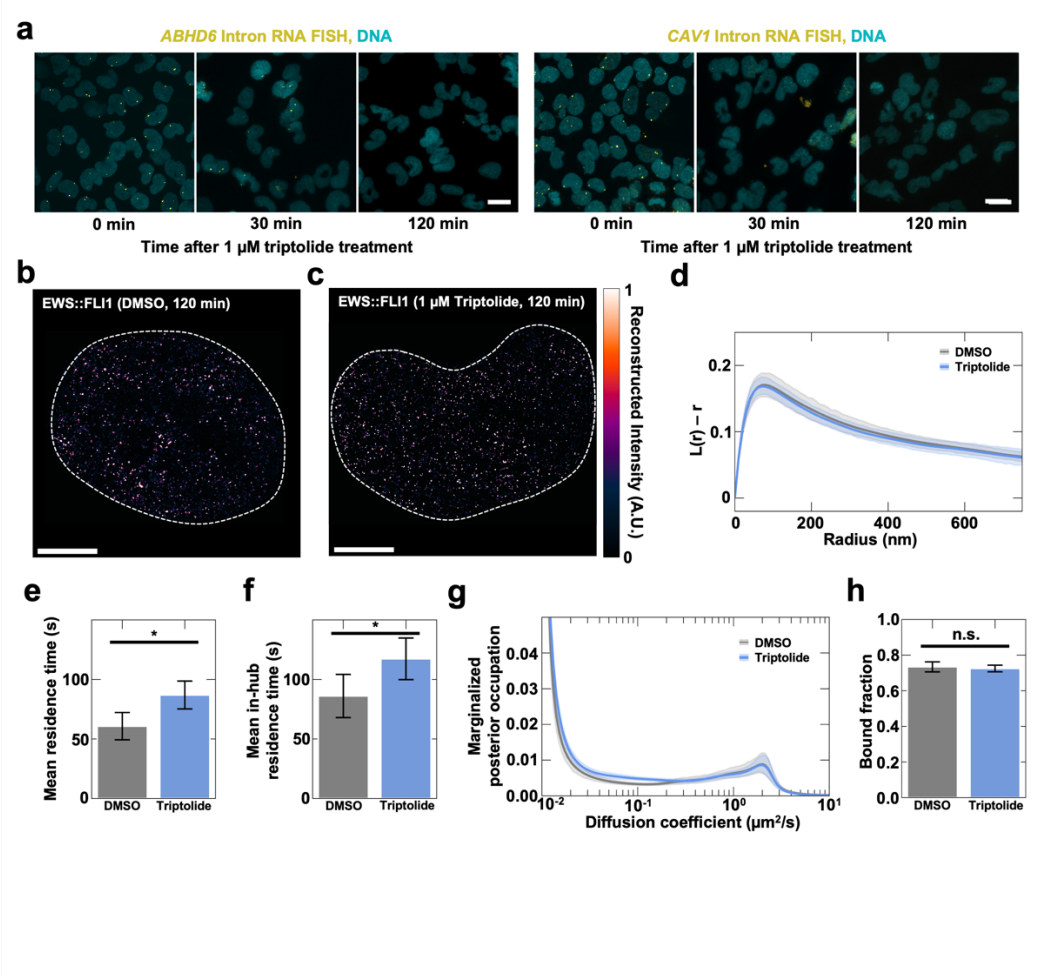

Figure 4. Transcriptional inhibition stabilizes EWS::FLI1 binding to hubs. (a) Confocal microscopy time-lapse of intron RNA FISH targeting *ABHD6* transcripts (FISH probes labeled with Quasar 670, yellow) or *CAV1* transcripts (FISH probes labeled with Quasar 570, yellow), and DNA (labeled with Hoechst 33342, cyan) in the knock-in A673 cells expressing endogenous EWS::FLI1-Halo after the treatment with 1 μM of triptolide. Scale bar is 20 μm. (b) Photoblinking-corrected PALM image of a knock-in A673 cell expressing endogenous EWS::FLI1-Halo (labeled with the Halo ligand PA-JF549) after transcriptional inhibition with triptolide. Scale bar is 5 μm. (c) Ripley's L(r) – r curves show that the punctate distribution of EWS::FLI1 does not differ significantly between the DMSO-treated and triptolide-treated cells. (d-e) Plots of the mean residence times of EWS::FLI1 show that it binds more stably across the chromatin

(*, $p = 2.06 \times 10^{-2}$, Mann-Whitney U test) (d) and within its hubs (*, $p = 2.21 \times 10^{-2}$, Mann-Whitney U test) (e) upon transcriptional inhibition. For each condition, 21 cells were measured in three independent imaging sessions, each session performed on a different day. (f) EWS::FLI1 diffusion profiles show comparable diffusion coefficients between the DMSO-treated and triptolide-treated cells. (g) Plot of bound fractions shows comparable bound fractions in DMSO-treated and triptolide-treated cells (n.s., $p = 1.96 \times 10^{-1}$, Mann-Whitney U test). For each condition in (f) and (g), 30 cells were measured in three independent imaging sessions, each session performed on a different day. The semi-transparent regions represent 95% confidence intervals in (c) and (f). Error bars represent SEM in (d), (e), and (g).

We next asked whether the dynamics of EWS::FLI1 hub assembly are dependent on the presence of RNA. To this end, we measured the mean chromatin-binding residence times of EWS::FLI1 after triptolide or DMSO-only treatment using slow SPT. We found that in cells treated with DMSO only, EWS::FLI1 binds to chromatin with a mean residence time of $60.77 \pm 11.44\ s$. The residence time increases to $86.93 \pm 11.71\ s$ in the cells treated with triptolide (Figure 4d). Similarly, the mean residence time of EWS::FLI1 in its hubs increases from $86.09 \pm 18.13\ s$ to $117.37 \pm 17.48\ s$ upon triptolide treatment (Figure 4e)[1]. This result suggests that the presence of RNA in EWS::FLI1 hubs facilitates dissociation of EWS::FLI1 from the hubs, i.e., increases its dissociation rate constant k_{off}. We then examined whether the diffusion dynamics of EWS::FLI1 depend on RNA using fast SPT combined with the saSPT analysis method. The characteristic diffusion coefficient of freely diffusing EWS::FLI1 in the triptolide-treated cells is $1.97 \pm 0.03\ \mu m^2/s$, nearly identical to that in the DMSO-only-treated cells (Figure 4f). The bound fraction of EWS::FLI1 also does not significantly change upon triptolide treatment, with $72.48 \pm 1.86\%$ in triptolide-treated cells and $73.87 \pm 2.66\%$ in DMSO-only-treated cells, respectively (Figure 4g). These results suggest that the presence of RNA does not affect the steady-state balance of the bound

and unbound EWS::FLI1 molecules in the cell nucleus. The bound fraction f_{bound} is directly linked to the dissociation constant K_d of EWS::FLI1 binding to chromatin via

$$f_{bound} = c/(c + K_d), \quad (3)$$

where c is a constant representing the intranuclear concentration of the binding sites of EWS::FLI1 on chromatin. Thus, K_d, the ratio of the dissociation and association rate constants (k_{off}/k_{on}), is also insensitive to the presence of RNA. Given this, and because having RNA in EWS::FLI1 hubs increases k_{off} of EWS::FLI1, the presence of RNA also proportionally increases its k_{on}.

Together, our results suggest that RNA plays a role in the assembly dynamics of EWS::FLI1 hubs but not their dimensions. Interestingly, nascent RNA in EWS::FLI1 hubs can increase both k_{on} and k_{off} of EWS::FLI1 binding to chromatin, such that their ratio is largely independent of the presence of RNA. The specific mechanism by which RNA affects the binding dynamics of EWS::FLI1 remains unclear.

Small molecules can disrupt the nuclear distribution of endogenous EWS::FLI1

Several small molecules have been reported in the literature to affect the functions of EWS::FLI1 and some are being explored as potential therapeutic candidates. However, how these compounds affect the nuclear organization of endogenous EWS::FLI1 in Ewing sarcoma cells, especially EWS::FLI1 hub formation, remains unknown. We sought to fill this gap by examining the spatial distribution of endogenous EWS::FLI1-Halo in the knock-in A673 cells after treating the cells with these compounds using live-cell confocal imaging.

The first compound we investigated was a CDK4/6 inhibitor, LY2835219, which is reported to disrupt the condensates formed by mCherry-labeled EWS::FLI1 overexpressed in U2OS cells, partly through activating lysosomal acidification[34]. The method that identified LY2835219, DropScan, screens condensate-modulating compounds using high-content imaging of fluorescently labeled condensates in U2OS cells formed via protein

overexpression. To examine the effect of LY2835219 on endogenous EWS::FLI1 hubs in Ewing sarcoma cells, we performed live-cell confocal imaging of endogenous EWS::FLI1-Halo in the knock-in A673 cells. The cells were imaged after 6 hours of treatment with 10 μM LY2835219 (stock solution in DMSO) or with DMSO only as a control (Figure 5a).

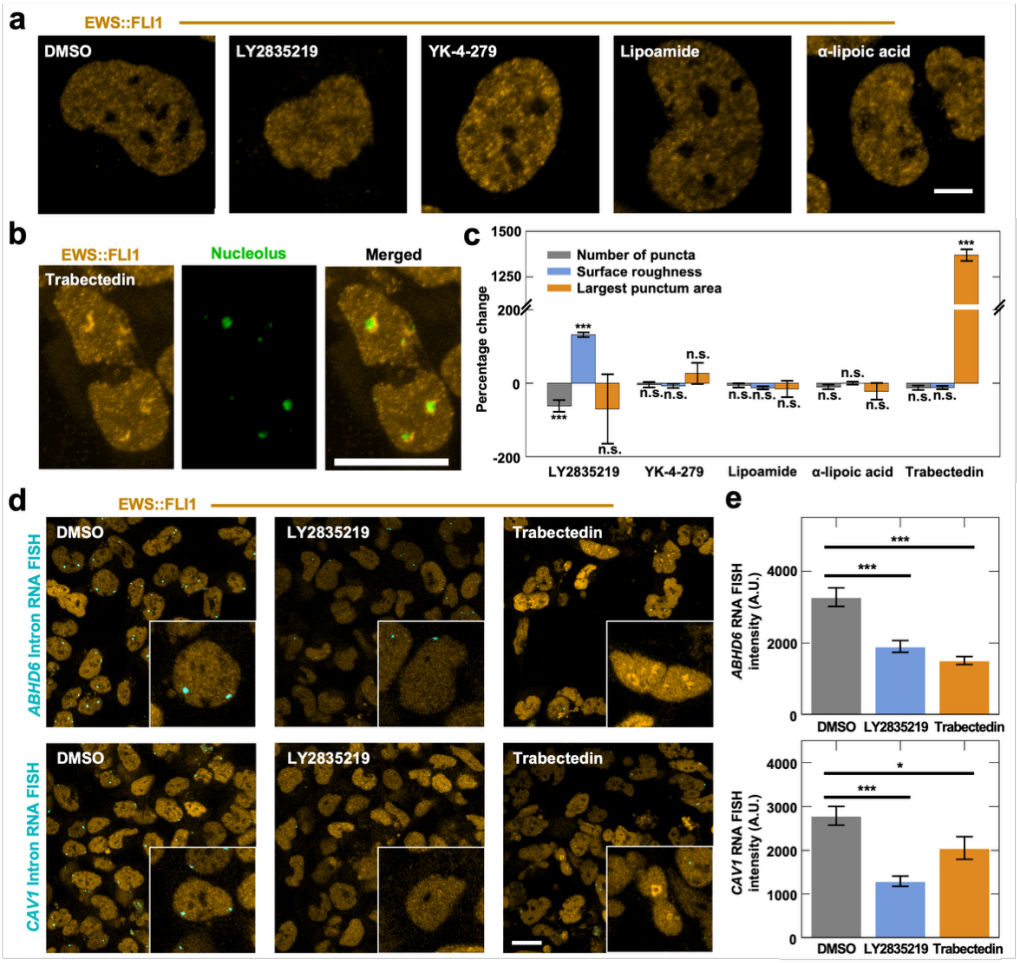

Figure 5: Small molecules can disrupt the nuclear organization of EWS::FLI1 and suppress the transcription of target genes. Live-cell confocal images of endogenous EWS::FLI1 (JFX549, yellow) in a knock-in A673 cell line after treatment with (a) DMSO only, 10 μM LY2835219 for 6 hrs, 10 μM YK-4-279 for 15 hrs, 10 μM lipoamide for 1 hr, and 10 μM α-lipoic acid for 1 hr. Scale bar is 5 μm. (b) Live-cell confocal images of endogenous EWS::FLI1 (JFX549, yellow) in a knock-in A673 cell line

transfected with mNeonGreen-NPM1 marking the nucleoli (green) and treated with 24 nM trabectedin for 1 hr followed by 23 hrs of culture in untreated media. Scale bar is 5 μm. (c) Plot of the percent change in the number of puncta, surface roughness, and largest punctum area per cell nucleus calculated as described previously (Irgen-Gioro et al., 2022) and in the Methods. Percent changes were calculated between cells treated with each of the chemicals and cells treated with an identical final concentration of pure DMSO as solvent controls. Treatment with LY2835219 reduces the number of puncta (***, $p = 5.82 \times 10^{-10}$, Welch's t-test) and increases the surface roughness (***, $p = 1.29 \times 10^{-6}$, Welch's t-test). Treatment with trabectedin increases the size of the largest punctum area (***, $p = 1.85 \times 10^{-6}$, Mann-Whitney U test). Data was acquired from n = 18 cells for each condition and error bars are SEM. (d) Confocal microscopy of intron RNA FISH targeting ABHD6 transcripts (Quasar 570, cyan) or CAV transcripts (Quasar 570, cyan), and EWS::FLI1 (JFX646, yellow) in the knock-in A673 cells after treatment with DMSO only, 10 μM LY2835219 for 6 hrs, or 24 nM trabectedin for 1 hr followed by 23 hrs of culture in untreated media. Scale bar is 20 μm. (e) Plot of the integrated intensity of RNA-FISH puncta per cell. Treatment with LY2835219 significantly suppressed ABHD6 transcription (***, $p = 1.57 \times 10^{-4}$, Mann-Whitney U test) and CAV1 transcription (***, $p = 5.81 \times 10^{-7}$, Mann-Whitney U test). Treatment with trabectedin similarly significantly suppressed ABHD6 transcription (***, $p = 2.61 \times 10^{-7}$, Mann-Whitney U test) and CAV1 transcription (*, $p = 3.72 \times 10^{-2}$, Mann-Whitney U test). Data was acquired from ABHD6: DMSO (n = 58), LY2835219 (n = 41), Trabectedin (n = 36) cells and CAV1: DMSO (n = 37), LY2835219 (n = 12), Trabectedin (n = 23) cells. Error bars are SEM.

While LY2835219 did not completely dissolve EWS::FLI1 hubs, it did reduce their number per cell nucleus. To quantify this difference, we calculated the percentage change in the number of puncta in the nucleus, surface roughness (standard deviation of pixel intensities across the nucleus), and area of the largest punctum in the nucleus in the cells with the

LY2835219 treatment compared to the cells with the DMSO-only treatment, following an established procedure[74] described in the Methods. We observed a statistically significant decrease in the number of puncta (Figure 5c), suggesting that LY2835219 disrupts endogenous EWS::FLI1 hubs in the native cellular context. Endogenous EWS::FLI1 hubs observed in our experiment, consistent with many previous studies[1,9,74–77], are located in the cell nucleus, which is different from the puncta of overexpressed mCherry-EWS::FLI1 that exhibited both nuclear and cytoplasmic localizations in DropScan[34]. Regardless, our result supports the finding from DropScan that LY2835219 can disrupt EWS::FLI1 puncta. Furthermore, LY2835219 treatment was shown to significantly downregulate transcription of 20 out of 32 EWS::FLI1 target genes[34]. Wang et al. proposed that LY2835219 dissolves the condensates of overexpressed EWS::FLI1 by lysosomal activity rather than a CDK4/6 inhibition-dependent mechanism, as they ruled out the latter by showing that other selective CDK4/6 inhibitors failed to dissolve the condensates[34]. However, since many EWS::FLI1 condensates were cytoplasmic in this previous study, it remains unclear whether endogenous EWS::FLI1 hubs in the nucleus would be similarly insensitive to CDK4/6 inhibitors or be regulated by a lysosome-mediated process. Together, our findings from live-cell imaging of endogenous EWS::FLI1, combined with these previous results, highlight the therapeutic potential of LY2835219 and warrant further investigation.

The second compound we investigated was YK-4-279, a well-characterized inhibitor of EWS::FLI1 that disrupts EWS::FLI1 interactions with RNA helicase A, leading to altered splicing and reducing transcription of EWS::FLI1 target genes, such as *UBE2C*[78]. We asked whether the effect of YK-4-279 involves disrupting the multivalent LCD-LCD interactions of EWS::FLI1 that enable its hub formation. We imaged the knock-in A673 cells expressing endogenous EWS::FLI1-Halo after treating the cells with 10 μM YK-4-279 (stock solution in DMSO) or DMSO only for 15 hours (Figure 5a). We found that the YK-4-279 treatment did not significantly change the number of puncta, surface roughness, or area of the largest punctum in the cells compared with the DMSO-only treatment (Figure 5c). This result suggests that YK-4-279 does not disrupt endogenous EWS::FLI1 hubs.

We next investigated lipoamide and α-lipoic acid, which are reported to partition into FUS droplets in vitro. The compounds also accumulate in and dissolve cytoplasmic stress granules formed upon oxidative stress, which were marked by fluorescently labeled FUS and contain all the FET family proteins (FUS, EWSR1, and TAF15)[79]. FUS is a homolog of EWSR1, and EWS::FLI1 contains the LCD of EWSR1. We thus hypothesized that lipoamide and α-lipoic acid can partition into and disrupt EWS::FLI1 hubs similarly. We imaged the knock-in A673 cells expressing EWS::FLI1-Halo after treating them with 10 μM lipoamide or α-lipoic acid (stock solutions in DMSO) or DMSO alone for 1 hour (Figure 5a). However, contrary to our hypothesis, neither lipoamide nor α-lipoic acid affected the number of puncta, surface roughness, or area of the largest punctum in the cells compared with the DMSO-only treatment (Figure 5c). The finding that neither lipoamide nor α-lipoic acid disrupts endogenous EWS::FLI1 hubs, despite the sequence similarities between the LCDs of EWS::FLI1 and FET family proteins, suggests that there are differences in the molecular mechanisms driving the formation of EWS::FLI1 hubs and cytoplasmic stress granules. This also indicates that the mechanisms by which lipoamide and α-lipoic acid mediate stress granule dissolution are independent of the LCD of EWSR1 that is present in the stress granules.

Finally, we investigated trabectedin, a DNA minor groove-binding compound that has been shown to suppress the expression of EWS::FLI1 target genes, including *NR0B1* and *CCND1*, through indirect, epigenetic mechanisms, e.g., evicting the SWI/SNF chromatin-remodeling complex from the genes, rather than through direct disruption of protein structure or interactions[80]. Previous immunofluorescence data also showed redistribution of EWS::FLI1 into the nucleolus upon trabectedin treatment, suggesting that the mislocalization of EWS::FLI1 away from its target genes may also contribute to the suppression of its transcriptional activation function. However, these experiments were performed in fixed cells[80], and fixation can artificially change intracellular distribution of specific proteins[74]. Here, to examine the localization of endogenous EWS::FLI1 in live cells upon trabectedin treatment, we performed live-cell confocal microscopy on the knock-in A673 cells expressing EWS::FLI1-Halo after treating the cells with 24 nM trabectedin (stock solution

in DMSO) or DMSO alone for 1 hour. Prior to either treatment, we incubated the cells in untreated media for 23 hours, a condition identical to the previous study[80]. However, contrary to the previous results, live-cell imaging showed that upon the trabectedin treatment, EWS::FLI1 accumulates at the periphery of the nucleolus instead of entering its interior. This localization was further confirmed by two-color imaging of endogenous EWS::FLI1-Halo and overexpressed NPM1 labeled with mNeonGreen, where the nucleolar protein NPM1 marks the location of the nucleolus (Figure 5b). Further analysis of EWS::FLI1 distribution showed that while the trabectedin treatment did not change the number of puncta and the surface roughness in the cells compared with the DMSO-only treatment, trabectedin increased the area of the largest punctum significantly (Figure 5c). According to the cell images, the largest puncta are often located on the periphery of the nucleolus. This result provides a revised view of trabectedin's mode of action, i.e., instead of redistributing EWS::FLI1 into the nucleolus, a compartment inactive for RNA Polymerase II transcription, trabectedin enriches EWS::FLI1 at the nucleolar periphery, which is likely spatially distinct from EWS::FLI1 target genes. Our result also suggests that trabectedin does not dissolve endogenous EWS::FLI1 hubs.

To test whether the dissolution of EWS::FLI1 hubs by LY2835219 and the mislocalization of EWS::FLI1 to the periphery of the nucleolus by trabectedin could affect EWS::FLI1-mediated transcription, we performed intron RNA FISH on target genes ABHD6 and CAV1 after drug treatment (Figure 5d). Treatment with either LY2835219 or trabectedin resulted in a clear reduction in both ABHD6 and CAV1 RNA FISH puncta intensity compared to DMSO controls (Figure 5e). This shows that both LY2835219 and trabectedin can suppress transcription of EWS::FLI1 target genes and suggests that both the dissolution of EWS::FLI1 hubs and redistribution of EWS::FLI1 to the periphery of the nucleolus are associated with transcriptional repression.

Together, this set of microscopy data provides a useful resource showing how the distribution of endogenous EWS::FLI1 in live Ewing sarcoma cells can be affected by treatment with various compounds with previously reported therapeutic effects or potential. These results,

together with previous publications showing functional perturbation of EWS::FLI1 by some of the examined compounds, suggest that the transcriptional activation function of EWS::FLI1 can be inhibited through diverse mechanisms, which may and may not involve disruption of its hubs.

Discussion

An important caveat in the characterization of IDR-mediated biomolecular assemblies arises from their strong dependence on protein concentration. Numerous studies have shown that overexpressed IDR-containing proteins can form large, liquid-droplet-like puncta in the cell, which may or may not form under endogenous concentrations[20,21]. The formation of excessive condensates in such overexpression systems can distort or obscure the true properties, functions, and regulatory mechanisms of the IDR of interest. Despite this, few assembly-forming transcription regulators have been studied at strictly endogenous concentrations. Here, we fill this gap by systematically investigating the dimensions, energetics, dynamics, and regulation of endogenous EWS::FLI1 hubs in Ewing sarcoma cells using quantitative live-cell imaging methods with single-molecule resolution (Figure 6).

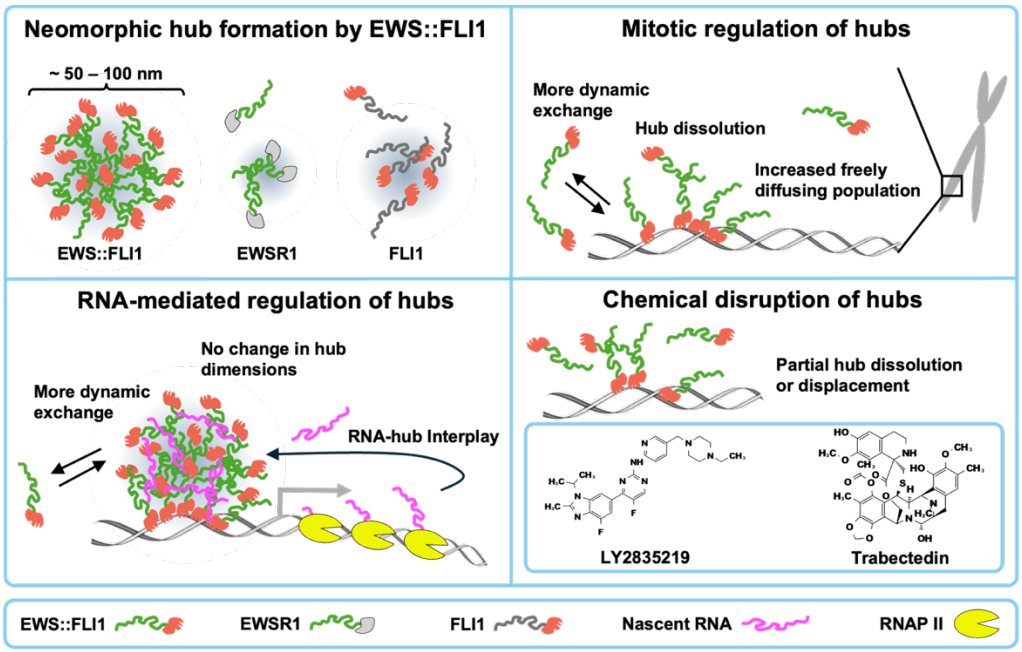

Figure 6. A summary of the formation, regulation, and perturbation of endogenous EWS::FLI1 hubs characterized in this work. 1) EWS::FLI1, not EWSR1 and FLI1, undergoes hub formation at native expression levels. 2) EWS::FLI1 hubs dissolve during mitosis. EWS::FLI1 retains its ability to associate with mitotic chromatin, albeit with faster binding kinetics. 3) The presence of RNA increases the in-hub binding kinetics of EWS::FLI1 without altering the hub dimensions. 4) Small-molecules LY2835219 and Trabectedin can dissolve EWS::FLI1 hubs and mislocalize EWS::FLI1 to the nuclear periphery, respectively.

Using PALM, we quantified the dimensions of sub-diffraction-limit EWS::FLI1 hubs and found that their size distribution is consistent with that of a supersaturated system predicted by the classical theory of homogeneous nucleation. While preceding studies have predicted and observed supersaturated (and subsaturated) protein clusters *in vitro*[81–86] or in the cell[87,88], this is the first time that the energetics and kinetics of endogenous TF hub formation have been characterized in the framework of the classical nucleation theory. Our finding raises the question of why the EWS::FLI1 hub size distribution appears to be held at a stationary nonequilibrium steady state without progressing to larger, phase-separated droplets, implying the existence of a hub dissolution mechanism. Although the identity of the dissolution pathway remains unknown, our estimated hub nucleation rate (34.15 hubs per second) sets the approximate timescale of the dissolution process. It will be of future interest to investigate the possible TF hub dissolution pathways that can occur at rates in this scale, such as chaperone-mediated dissolution[51–53], phosphorylation-mediated dissolution[54–56], and competition between multivalent interactions and DNA-protein binding[57]. In addition, our results suggest that endogenous EWS::FLI1 hubs are supersaturated clusters[8] rather than *bona fide* LLPS droplets and are likely different from transcription condensates formed at super-enhancers reported in previous work[8,89,90].

EWS::FLI1 exhibits multiple neomorphic behaviors that its parental proteins (EWSR1 and FLI1) lack, including remodeling GGAA microsatellites into *de novo* enhancers[61,62],

inactivating canonical ETS enhancers by displacing wild-type ETS TFs[62], and inducing aberrant alternative splicing[91]. Here, using PALM to quantify protein distributions in the cell with sub-diffraction resolution, we showed that hub formation is another neomorphic behavior of EWS::FLI1. Specifically, we found that while both parental proteins formed puncta, neither exhibited the same degree of clustering as EWS::FLI1. This result supports our previous conclusion that hub formation behavior plays an important role in the oncogenic functions of EWS::FLI1, underscoring the potential of targeting this behavior in therapeutic development. Many oncogenic fusion proteins resulting from chromosomal translocations have been reported to undergo LLPS when overexpressed in the cell[92]. In the future, it will be interesting to investigate whether LLPS or related assembly formation is a neomorphic behavior of fusion proteins other than EWS::FLI1 and how these neomorphic assemblies may contribute to the oncogenesis of different cancers. The imaging and analysis approach we develop here will provide helpful tools for this purpose.

We then investigated whether EWS::FLI1 hubs are regulated during mitosis or by RNA using various live-cell imaging approaches. We found that during mitosis, EWS::FLI1 hubs dissolve, but EWS::FLI1 remains capable of binding to and dissociating from chromatin, albeit more dynamically and with a smaller bound fraction than during interphase. This suggest that similar to several reported TFs[67], EWS::FLI1 can also serve as a mitotic bookmark that helps to restart transcriptional programs in daughter cells after mitosis. It will be of future interest to investigate whether and how mitotic bookmarking by EWS::FLI1 may contribute to the tumorigenesis of Ewing sarcoma. In addition, since a burst in protein phosphorylation occurs as cells go into and progress through mitosis[93,94], our observed dissolution of EWS::FLI1 hubs during mitosis indicates a potential role of phosphorylation in regulating the hubs and warrants future investigation of this regulation mechanism.

Contrary to mitosis, when we remove nascent RNA from the cells, EWS::FLI1 binds to chromatin more stably with the hub dimensions unchanged. This finding highlights similarities and differences between EWS::FLI1 hubs and other transcription-related condensates, such as Mediator condensates, where RNA plays an essential role in condensate

formation. Specifically, varying RNA concentrations strongly influences both the dimensions and dynamics of Mediator condensates, with transcriptional inhibition increasing both condensate size and lifetime. In contrast, EWS::FLI1 hubs are predominantly maintained by protein-DNA and protein-protein interactions, with RNA contributing only to regulating EWS::FLI1 binding kinetics. One possible explanation is that nascent RNA synthesized at EWS::FLI1 target genes may dynamically interact with EWS::FLI1 hubs directly through aromatic interactions[95] and/or indirectly through RNA-binding proteins such as RNA helicase A[91,96], allowing RNA to serve as a multivalent scaffold[97–99] that helps concentrate EWS::FLI1 at the gene loci, thus increasing the likelihood of EWS::FLI1 binding to chromatin (increasing k_{on}). Meanwhile, RNA might increase k_{off} by destabilizing existing EWS::FLI1-chromatin interactions through charge-based repulsion or steric hindrance[33]. The differential roles of RNA in the regulation of Mediator condensates and EWS::FLI1 hubs could be due to differences in the chemistry underlying the molecular interactions that drive the formation of the respective assemblies. The IDR of the Mediator subunit MED1 contains alternating blocks of charged amino acids that drive condensate formation[3,100]. RNA, a highly-charged polyanion, also interacts with the MED1 IDR primarily via electrostatic interactions, thus strongly coupling the formation and dissolution of MED1 condensates to local RNA concentration[33]. On the other hand, EWS::FLI1 hub formation is driven mainly by interactions between aromatic residues, particularly tyrosines in the EWS LCD[1]. These interactions are less sensitive to changes in bulk electrostatics [33,101], thus resulting in a lower impact of local RNA concentration on EWS::FLI1 hub assembly. Our observations are also related to previous work showing that an increase in the RNA-protein ratio can reduce the ability of EWSR1 and FUS, members of the FET family of RNA-binding proteins[102], to form condensates. Notably, this effect was previously observed when the proteins were artificially driven to phase separation via overexpression. In contrast, our measurements of EWS::FLI1 under endogenous expression levels suggest that RNA regulates only the binding dynamics of EWS::FLI1, without altering its hub size.

Finally, we showed that in live Ewing sarcoma cells, LY2835219 partially dissolves endogenous EWS::FLI1 hubs and trabectedin redistributes EWS::FLI1 to the periphery of

the nucleolus, away from its target genes. This is a useful finding given that recent work indicates that assemblies of disease-causing TFs are potential alternative therapeutic targets, owing to the notoriously "undruggable" nature of most TFs, which is due to their largely disordered nature and lack of well-defined binding pockets. For EWS::FLI1 specifically, our earlier work shows that its hub formation mediated by multivalent LCD-LCD interactions plays an essential role in oncogenic transcription in Ewing sarcoma [1,9], suggesting disruption or mislocalization EWS::FLI1 hubs as a potential therapeutic strategy. Here, our finding that LY2835219 and trabectedin can alter the nuclear distribution of endogenous EWS::FLI1 further supports the previously reported therapeutic potential of the compounds.

Overall, our study combined CRISPR/Cas9-mediated genome editing and quantitative live-cell and single-molecule imaging techniques to characterize the formation and regulation of endogenous TF hubs at an unprecedented resolution. We quantified the dimensions of EWS::FLI1 hubs along with the energetics and kinetics of hub nucleation. We uncovered the neomorphic nature of EWS::FLI1 hubs and dynamic hub regulation during mitosis and by RNA. We also showed that endogenous EWS::FLI1 hubs in Ewing sarcoma cells can be disrupted by small molecules, including LY2835219 and trabectedin, suggesting a route for therapeutic intervention. While our study focuses on a single TF EWS::FLI1, the findings we made here, particularly those related to the formation and regulation of TF hubs, may apply to diverse TFs, including regular and pathological ones. Future studies applying similar approaches to a broad range of TFs will inform the generalizability of our findings.

Methods

Cell Culture:

The knock-in A673 cell line expressing endogenous EWS::FLI1-Halo from Chong et al.[1] was grown in 4.5 g/L high-glucose DMEM (Thermo Fisher, 10566016) with 10% FBS (Fisher Scientific, SH3039603) and 1% penicillin-streptomycin (Thermo Fisher, 15140122). For live-cell imaging, the phenol-red-free version of high-glucose DMEM (Thermo Fisher, 31053028) was used instead. The human U2OS osteosarcoma cell line was cultured in 1 g/L

low-glucose DMEM (Thermo Fisher, 10567014) with 10% FBS and 1% penicillin-streptomycin. Cells were synchronized using nocodazole (Sigma Aldrich, M1404-2MG). All cell lines were cultured in a humidified incubator at 37°C and 5% CO_2.

Single-molecule imaging:

All single-molecule imaging was performed on 25 mm #1.5H coverslips (Marienfeld Superior, 0117650). Coverslips were cleaned by immersing in 1 M KOH and sonicating for 1 hr, rinsing thrice with double-distilled water, immersing in pure ethanol and sonicating for 1 hr, then immersing in 20% ethanol diluted in double-distilled water for storage.

Single-molecule imaging was performed on a Nikon Eclipse Ti2 N-STORM microscope setup with a x100/NA 1.49 oil-immersion objective (Nikon, CFI SR HP Apochromat TIRF 100XAC Oil, MRD01995). Samples were illuminated under a highly inclined and laminated optical sheet (HILO)[103] with a Nikon LUN-F solid-state laser unit using the following lasers: 405 nm (photoactivation of PA-JF549 and PA-JF646[104] and excitation of Hoechst 33342), 561 nm (excitation of JFX549 and PA-JF549), and 640 nm (excitation of JFX646[105] and PA-JF646). For single-channel experiments, fluorescence from JFX549 and PA-JF549 was filtered through a 593/40 nm single-band bandpass filter (Semrock, FF01-593/40-25) and fluorescence from JFX646 and PA-JF646 was filtered through a 676/37 nm single-band bandpass filter (Semrock, FF01-676/37-25) and detected on an iXon Ultra 897 EMCCD (Andor). For two-channel experiments, the emission fluorescence was first split using a 635 nm single-edge dichroic beamsplitter (Semrock, Di02-R635-25x36) before being similarly filtered through the appropriate single-band bandpass filters and detected on iXon Ultra 897 EMCCDs. For live-cell imaging, a stage incubator was used to maintain humidity and 37°C with 5% CO_2. Fixed-cell imaging was performed at room temperature. All hardware was controlled through the NIS-Elements software (Nikon).

Photoactivated localization microscopy (PALM):

Cells were cultured on cleaned coverslips then stained with 200 nM PA-JF549 following the protocol described in Yoshida et al.[68]. Cells were then fixed using 4% paraformaldehyde (VWR, 100503-917) supplemented with 0.2% glutaraldehyde (Sigma Aldrich, 340855-25ML) for 15 minutes, then rinsed thrice with 1X PBS for 5 min (Thermo Fisher, 18912014). Coverslips were then transferred into Attofluor Cell Chambers (Thermo Fisher, A7816). 0.1 μm yellow-green Invitrogen Fluospheres (Thermo Fisher, F8803) were diluted in DPBS with calcium and magnesium (Thermo Fisher, 14-040-141) and added to the coverslips for 10 minutes at room temperature to allow the beads to adsorb to the coverslip.

The samples were imaged on the single-molecule microscope described above under continuous, simultaneous photoactivation (405 nm) and excitation (561 nm). The 405 nm laser intensity was gradually increased over the course of the acquisition to maintain an approximately constant number of photoactivated molecules, and the 561 nm laser was held constant. A 50 ms integration time was used and cells were imaged until all single molecules were bleached (typically at least 24000 frames). Images of the fiducial marker positions were obtained every 100 frames.

NIS-Elements was used to identify single-molecule localizations in each frame, with the following settings: minimum height: 1500, maximum height: 65535, CCD baseline: 100, minimum width: 200 nm, maximum width: 600 nm, initial fit width: 300 nm, maximum axial ratio: 1.2, maximum displacement: 0. Fiducial markers were identified under identical settings except for minimum height: 3000. Drift correction was performed based on the fiducial marker positions, and the localizations were exported as a .csv file.

Even when using well-behaved fluorophores, single-molecules typically appear in multiple frames due to a combination of variable photobleaching times and photoblinking[106]. More specifically, rather than photobleaching permanently after a fixed amount of time, fluorophores can enter a temporary dark state from which they recover and reappear, often “photoblinking” a number of times before permanently bleaching. This is especially problematic in the characterization of TF hubs because it becomes difficult to distinguish between the contributions of true TF clustering and blinking. Thus, to minimize these

artifacts, endogenous EWS::FLI1-Halo was labeled with the photoactivatable ligand PA-JF549 which features minimal blinking[107], and the PALM acquisitions were performed with low intensity photoactivation to minimize the probability of multiple molecules within a diffraction-limited spot fluorescing simultaneously. Even with these strategies, it is still necessary to correct for multiple appearances of the same protein, motivating pair-correlation PALM (PC-PALM) and our photoblinking correction.

Raw localizations were analyzed using a custom python implementation of PC-PALM[40,44,45]. While PC-PALM provides an excellent way of discriminating between localization clusters due to true protein clusters and due to photoblinking (particularly for emitters with high blinking rates or long temporary dark states), it cannot identify which localizations correspond to a single protein, i.e. it cannot recover the set of true protein localizations. Further, it is particularly sensitive to edge effects and variable localization densities, thus only a subset of cellular localizations (corresponding to a representative, relatively homogeneous subregion of the nucleus) can be analyzed. In order to generate a reconstructed image with true localizations, we had to directly filter out the localizations due to photoblinking.

Often, the time between reappearances of a single emitter is much shorter than the amount of time before another emitter is excited in the same position within what is expected with localization error and previous work has shown that identifying spatiotemporal clusters can be an effective way to identify photoblinking artifacts[36]. Existing strategies often choose a threshold cutoff time for either the time until photobleaching, t_{bleach}, or the time between appearances, t_{off}. For the t_{bleach} strategy, all localizations within the threshold time of an initial localization are assumed to have come from the same molecule, and their positions are averaged and they are treated as one molecule. For the t_{off} strategy, after localization disappears, a reappearance that takes place within the threshold time elapsing is assumed to have come from the same molecule, and the localization are similarly averaged. The specific threshold values used, whether to use the t_{off} or t_{bleach} strategy, and whether either of these

strategies is suitable at all, are all dependent on the fluorophore used. Thus, we first needed to characterize the blinking behavior of PA-JF549.

Many groups have investigated the blinking properties of fluorophores by purifying fluorescent proteins and tethering them to coverslips[40]. However, the blinking properties of fluorophores is also well known to be highly environment dependent[108]. Thus, we chose to examine the blinking of PA-JF549 in fixed, A673 cells stably expressing Halo-NLS, which does not significantly cluster. Other than staining concentration, the samples were prepared and imaged under identical conditions as our PALM experiments. The cells were stained with 0.5 nM of PA-JF549, making it highly unlikely that any Halo-NLS bound by a ligand would appear in the same location as another ligand-bound molecule. This ensured that any localizations in a particular region would all be due to a single PA-JF549 molecule. We then performed PALM imaging across ten cells, and in each cells the localizations were clustered using DBSCAN ($\varepsilon = 10$ nm, minimum points $= 1$). We confirmed that PA-JF549 features extremely low blinking, appearing on average 1.12 times, on par with an existing estimate from the literature[107]. Next, we adopted a strategy similar to those of Annibale et al.[36] and Lee et al.[38] for determining the appropriate t_{off} threshold value (hereafter, t^*_{off}) using our Halo-NLS data. Given the total number of localizations within a region n_{total}, the true number of proteins n_{true}, and the mean number of blinks $\overline{N}_{blinks}$, $\frac{n_{total}}{1+\overline{N}_{blinks}}$ becomes an excellent estimator of n_{true} for increasing n_{total} as long as $\overline{N}_{blinks} \ll n_{total}$. Thus, we can determine an appropriate value of t^*_{off} by estimating n_{true} across all Halo-NLS cells using n_{total} and the observed value of $\overline{N}_{blinks}$, then filtering blinking using different values of t^*_{off} and totaling the number of "true" localizations, n^*_{true}, until $n^*_{true} \approx n_{true}$. We did this and found that $t_{off} = 0.05$ ms, that is tolerating a single empty frame between blinking events, was sufficient to recover $n^*_{true} = 19055$ of $n_{true} = 19079$ total Halo-NLS localizations, corresponding to a relative error of 0.13%. We then applied this threshold value to our raw PALM localizations.

To visualize whether there was a significant difference in the spatial distribution of proteins in different PALM datasets, we computed Ripley $L(r) - r$ for each cell over values of r

from 0 to 750 nm in Python using the `RipleysKEstimator` function from the `astropy.stats` package[109]. The area of each nucleus was estimated by generating an alpha shape around the localizations ($\alpha = 2$) using the `alphashape` package. While spatial descriptive statistics provides an excellent visualization of which proteins are more clustered than others and over what length scales, a great deal of information on the single cluster level is lost. Thus, we identified individual hubs by clustering the raw localizations using DBSCAN ($\varepsilon = 10$ nm, minimum points $= 3$), confirming by eye the suitability of these parameters. For each hub, the diameter was calculated as the maximum pair-wise distance between localizations and the "mode" diameter was extracted by approximating the hub diameter distribution as lognormal.

Our analysis investigating the hub distribution through the lens of homogeneous nucleation was based on the framework established by Naranayan et al.[88]. Measurements of both critical radius R_c and free energy barrier to nucleation ΔG_c are insensitive to constant multiplicative factors in hub "size" n[88], thus $n = \left(\frac{R\ (nm)}{1\ nm}\right)^3$ was calculated for each hub with diameter, d, larger than our localization precision, ~20 nm. Here, n is proportional to the number of EWS::FLI1 molecules in each hub. The offset factor is primarily due to the fact that every EWS::FLI1-Halo molecule does not end up being detected; literature measurements of the effective labeling efficiency of PA-JF549 place it at around 21%[107]. Eqs. (1) and (2) are valid only for subcritical hubs, thus only hubs with $n \leq 1 \times 10^6$ were analyzed. This condition ensured that only subcritical hubs, accounting for over 85% of hubs in each cell, were included. Then $P(n)$ was generated using bins of size ~ 5000 and then $-\log P(n)$ was calculated and fit as $-\log P(n) = \varepsilon' n^{2/3} - \Delta\mu' n - b$. Fit values of ε' and $\Delta\mu'$ were used to calculate R_c and ΔG_c as described in Supplementary Note A. Estimation of the nucleation rate, J, is described in Supplementary Note B.

Slow single-particle tracking (SPT):

Slow SPT was performed following Yoshida et al.[68]. Briefly, cells were cultured on cleaned coverslips then for slow SPT during mitosis, cells were stained with 20 nM PA-JF646 and

DNA was labeled with Hoechst 33342. For slow SPT after transcriptional inhibition, cells were stained with 20 nM PA-JF646 and 100 nM JFX549. This enabled the simultaneous visualization of individual EWS::FLI1 molecules in the PA-JF646 channel against the real-time distribution of DNA or EWS::FLI1 hubs. Coverslips were then transferred into Attofluor Cell Chambers and DMEM was exchanged for phenol-red-free DMEM.

The samples were then imaged on the single-molecule microscope described above. Single molecules were excited using the 640 nm laser during the exposure time and they were photoactivated using the 405 nm laser during the transition time between exposures. 500 ms exposure times were used such that mobile particles were blurred out. The 405 nm laser intensity was gradually increased over the course of the acquisition to maintain an approximately constant number of photoactivated molecules, and the 640 nm laser was held constant. For simultaneous single-molecule imaging with DNA, one frame of the Hoechst 33342 channel was taken before and after acquisition. For simultaneous single-molecule imaging with EWS::FLI1 hubs, every 10 seconds, one frame of the bulk EWS::FLI1 channel was taken. One the timescale of the acquisition, DNA is not that mobile, thus before and after images are sufficient for a timelapse. EWS::FLI1 hubs remain mobile, thus the need for more frequent timelapse images. For each movie, 2000 frames were acquired, and movies were split into single-molecule channels (PA-JF646) and bulk channels (Hoechst 33342 or JFX549).

Single-molecules were localized and assembled into trajectories using SLIMfast[110], a MATLAB-based GUI implementation of the MTT algorithm[111]. The following parameters were used: localization error: 10^{-6}, deflation loops: 3, maximum number of competitors: 5, maximum expected diffusion coefficient: $0.1\ \mu m^2/s$. Using the bulk channels, real-time binary masks of DNA/EWS::FLI1 hub position were generated with custom ImageJ macros from Chong et al.[1]. By changing the threshold for EWS::FLI1, masks of both EWS::FLI1 hubs and the whole nucleus were created. Then, using different custom ImageJ macros also from [], trajectories were sorted into those overlapping with DNA/EWS::FLI1 hub masks and those not. The trajectories overlapping with masks represent stable binding to

DNA/EWS::FLI1 hubs. Survival curves defined as $1 - CDF$, i.e. the complement of the cumulative distribution function, were generated from the overlapping trajectories and fit to the following two-component exponential model.

$$P(t) = Ae^{-k_{obs,ns}t} + (1 - A)e^{-k_{obs,s}t}$$

Here, A is the fraction of non-specifically bound molecules, $k_{obs,ns}$ is the observed dissociation rate of nonspecifically bound molecules, and $k_{obs,s}$ is the observed dissociation rate of specifically bound molecules. The observed dissociate rate of specifically bound molecules is the sum of the photobleaching rate, k_{pb}, and the true specific dissociation rate, $k_{true,s}$, since both dissociation and photobleaching can result in a trajectory ending. The photobleaching rate is obtained by doing slow SPT under identical settings on an A673 cell line stably expressing H2B-Halo described in Chong et al.[1] and extracting the $k_{obs,s}$ for H2B, which rarely dissociates on the timescales measured by slow SPT. The mean residence time, $\bar{\tau}_s$, is then given as

$$\bar{\tau}_s = \frac{1}{k_{true,s}} = \frac{1}{k_{obs,s} - k_{pb}}$$

Fast single-particle tracking (SPT):

Cells were cultured on cleaned coverslips then stained with 20 nM PA-JF646. Coverslips were then transferred into Attofluor Cell Chambers and DMEM was exchanged for phenol-red-free DMEM.

The samples were then imaged on the single-molecule microscope described above under stroboscopic illumination based on Chong et al.[9]. Cameras were synchronized at the beginning of each acquisition with 0.3 μs shift speeds and +4 vertical clock voltage amplitudes enabling 5 ms exposure times with 158.01 μs transition times. Single molecules were excited using the 640 nm laser during only the first 1 ms of each exposure. Single molecules were photoactivated using the 405 nm laser during the transition times. The 405 nm laser intensity was gradually increased over the course of the acquisition to maintain an approximately constant number of photoactivated molecules, and the 640 nm laser was held constant. For each movie, 20000 frames were acquired.

Single molecules were localized and assembled into trajectories using SLIMfast[110], using the following parameters: localization error: 10^{-5}, deflation loops: 2, maximum number of competitors: 3, maximum expected diffusion coefficient: $3.5\,\mu m^2/s$. The resulting trajectories were then analyzed using the state array SPT (saSPT) python package developed in Heckert et al.[70]. The state array was defined over one hundred logarithmically spaced diffusion coefficients from 10^{-2} to $10^{1}\,\mu m^2/s$, inclusive, and twenty linearly spaced localization errors from 0 to $0.07\,\mu m$, inclusive. The posterior occupations were then marginalized over localization error to obtain the diffusion coefficient distribution. Bound fractions were computed as the sum of the occupation values of diffusion coefficients under $0.1\,\mu m^2/s$ corresponding to stably bound molecules.

Live-cell confocal microscopy:

Cells were cultured on 35 mm MatTek glass-bottom dishes (MatTek, P35G-1.5-20-C) and stained with 200 nM JFX549 or JFX646. For two-color imaging during mitosis, DNA was labeled with Hoechst 33342. For two-color imaging with the nucleolus, cells were transfected with the mNeonGreen-NPM1 construct 24 hrs before imaging using Lipofectamine 3000 (Fisher Scientific, L3000001) following the manufacturer's protocol. DMEM was then replaced with phenol-red-free DMEM.

Confocal imaging was performed on a Nikon Eclipse Ti2 AX R point-scanning confocal microscope setup with a x60/NA 1.42 (Nikon, CFI Plan Apochromat Lambda D 60X Oil, MRD71670). Samples were illuminated with a Nikon LUA-S4 laser unit using the following lasers: 405 nm (excitation of Hoechst 33342, (VWR, PI62249)), 488 nm (excitation of mNeonGreen), 561 nm (excitation of JFX549[105] and Quasar 570), and 640 nm (excitation of JFX646[105] and Quasar 670). For multichannel images, emission filters and detection sequence were chosen such that there was no bleed-through between the channels. Z-stacks were acquired with slice intervals of 0.5 µm and a pinhole size of 1 Airy unit. A stage incubator was used to maintain humidity and 37°C with 5% CO_2.

RNA fluorescence in situ hybridization (RNA FISH):

Knock-in A673 cells were plated on 18 mm circular No 1.5 coverslips (Electron Microscopy Sciences, 72222-01) cleaned by soaking with 70% ethanol. Cells were then stained with 200 nM JFX549 or JFX646 chosen such that the labeled EWS::FLI1-Halo is spectrally distinct from the RNA FISH probe. Intron RNA FISH was performed following the protocol for adherent cells published online by Stellaris using FISH probes targeting the introns of *ABHD6* and *CAV1* from Chong et al.[9] designed using the online Stellaris Probe Designer software. The *ABHD6* probe was conjugated with Quasar 670 and *CAV1* probe was conjugated with Quasar 570. Samples were then imaged at room temperature on the confocal system described above.

CRISPR/Cas9-mediated genome editing:

We created the knock-in A673 cell line expressing endogenous Halo-EWSR1 using published procedures[112], except that we exploited the HaloTag for fluorescence-activated cell sorting (FACS). We first co-transfected A673 using a repair plasmid together with the Cas9 plasmid (3:1 mass ratio for repair plasmid to Cas9 plasmid). For the repair vector, we modified a pUC57 plasmid to incorporate coding sequences for TEV protease recognition epitope (EDLYFQS), HaloTag and FLAG-tag flanked by ~750 bp of genomic homology sequence (homology arm) of EWSR1 N-terminus on either side. The TEV linker sequence

was between HaloTag and EWSR1. We introduced synonymous mutations in the homology sequence, where necessary, to prevent the Cas9-sgRNA complex from cutting the repair vector.

We designed two sgRNAs (CGGGTGAGTATGGTGGAACT and GGAGAGAAAATGGCGTCCAC) using the Zhang lab CRISPR design tool (http://tools.genome-engineering.org), cloned them into the Cas9 plasmid, and transfected A673 cells with the Cas9-sgRNA plasmid and the repair vector. 24 hours after transfection, we pooled cells transfected with each sgRNA individually and FACS-sorted for YFP (mVenus)- positive, successfully transfected cells. To generate the HaloTag knock-in, we grew the YFP-positive cells for 19 days, labeled them with 500 nM HaloTag TMR ligand (Promega, G8251), FACS-sorted for HaloTag-positive cells, and plated one cell per well into 96-well plates. Clones were then expanded and screened by genomic PCR using a genomic primer external to the homology sequence and an internal HaloTag primer. Successfully edited clones were further verified with western blot. EWSR1 shares the same N-terminus with EWS::FLI1. We chose a clone with HaloTag knock-in at EWSR1 and normal expression of EWS::FLI1 for further studies.

Stable cell line construction:

The A673 cell line stably expressing Halo-NLS was generated using PiggyBac transposition and drug selection. Briefly, the cDNA encoding Halo-NLS was cloned into a PiggyBac vector co-expressing a G418 resistance gene, and this vector was co-transfected together with a SuperPiggyBac transposase vector into the A673 cell line using Fugene 6 (Promega, E2691) according to the manufacturer's instructions. 48 hours after transfection, we started with selection by adding 0.5 mg/ml G418. Untransfected A673 cells were treated with G418 in parallel, and selection was judged to be complete once no untransfected cells remained (~ 1 week). The cells were further selected for 2 weeks under 0.25 mg/ml G418 before stocks were frozen.

Western blot:

Cells were washed with ice-cold PBS and then lysed in mild lysis buffer containing protease inhibitor cocktail (Millipore Sigma, 11836170001), 50 mM Tris (pH = 7.5), 150 mM NaCl, 0.5 % Triton X-100. Total cell lysates were incubated on ice for 30 min and then centrifuged at 15,000g at 4°C for 20 min. Equal amounts of total protein were separated via 4-15% SDS-PAGE and transferred to nitrocellulose membranes. The membranes were blocked with TBS containing 0.1% Tween-20 and 5% milk and then incubated with primary antibodies overnight at 4°C. The next day, membranes were washed, incubated with the appropriate horseradish peroxidase-conjugated secondary antibodies, and were developed using horseradish peroxidase-based chemiluminescent substrate (Thermo Scientific, 34580). The following primary antibodies were used for immunoblotting: EWSR1 antibody (Abcam, ab133288), GAPDH antibody (Cell Signaling Technology, 2118S), horseradish peroxidase-labeled goat anti-rabbit secondary antibodies (Invitrogen, A21245).

Statistical analyses:

For normally distributed data with equal variance, Student's t-tests were used. For normally distributed data with unequal variance, Welch's t-tests were used. For skew data, Mann-Whitney U test was used. All statistical tests were performed in Python using the `scipy.stats` package. Significance is abbreviated in figures according to the following key: n.s. is not statistically significant, * is $p < 0.05$, ** is $p < 0.01$, and *** is $p < 0.001$.

Author Contributions

S.C. and S.Y. conceived the study, designed the experiments, and interpreted the results. S.C. constructed all knock-in cell lines. S.Y. performed all imaging experiments, analysis, and physical modeling. Y.Z. performed western blots. W.D. and A.K. provided reagents and advised on cell cycle experimental design and interpretation. S.Y. prepared the original draft and S.C. and S.Y. reviewed and edited the manuscript. S.C. supervised the work and acquired funding.

References

1. Chong, S. *et al.* Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555 (2018).
2. Cai, D. *et al.* Phase separation of YAP reorganizes genome topology for long-term YAP target gene expression. *Nat. Cell Biol.* **21**, 1578–1589 (2019).
3. Lyons, H. *et al.* Functional partitioning of transcriptional regulators by patterned charge blocks. *Cell* **186**, 327-345.e28 (2023).
4. Cho, W.-K. *et al.* Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).
5. Guo, Y. E. *et al.* Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* **572**, 543–548 (2019).
6. Boehning, M. *et al.* RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.* **25**, 833–840 (2018).
7. Lyons, H. *et al.* RNA polymerase II partitioning is a shared feature of diverse oncofusion condensates. *Cell* **188**, 3843-3862.e28 (2025).
8. Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
9. Chong, S. *et al.* Tuning levels of low-complexity domain interactions to modulate endogenous oncogenic transcription. *Mol. Cell* **82**, 2084-2097.e5 (2022).
10. Song, L. *et al.* Hotspot mutations in the structured ENL YEATS domain link aberrant transcriptional condensates and cancer. *Mol. Cell* **82**, 4080-4098.e12 (2022).
11. Chen, L. *et al.* Hormone-induced enhancer assembly requires an optimal level of hormone receptor multivalent interactions. *Mol. Cell* **83**, 3438-3456.e12 (2023).
12. Trojanowski, J. *et al.* Transcription activation is enhanced by multivalent interactions independent of phase separation. *Mol. Cell* **82**, 1878-1893.e10 (2022).
13. Kovar, H. Dr. Jekyll and Mr. Hyde: The Two Faces of the FUS/EWS/TAF15 Protein Family. *Sarcoma* **2011**, 837474 (2011).
14. Linardic, C. M. PAX3-FOXO1 fusion gene in rhabdomyosarcoma. *Cancer Lett.* **270**, 10–18 (2008).
15. Ahn, J. H. *et al.* Phase separation drives aberrant chromatin looping and cancer development. *Nature* **595**, 591–595 (2021).

16. Browne, A. L. *et al.* Network analysis of SRC-1 reveals a novel transcription factor hub which regulates endocrine resistant breast cancer. *Oncogene* **37**, 2008–2021 (2018).
17. Zhao, J. *et al.* Oncogenic transcription factors instruct promoter-enhancer hubs in individual triple negative breast cancer cells. *Sci. Adv.* **10**, eadl4043 (2024).
18. Mittag, T. & Pappu, R. V. A conceptual framework for understanding phase separation and addressing open questions and challenges. *Mol. Cell* **82**, 2201–2214 (2022).
19. Yoshida, S., Zhong, Y., Banh, J., Guo, J. & Chong, S. Elucidating the Roles of Intrinsically Disordered Proteins in Eukaryotic Transcriptional Regulation with Single-Molecule Imaging. *J. Mol. Biol.* **438**, 169343 (2026).
20. McSwiggen, D. T., Mir, M., Darzacq, X. & Tjian, R. Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev.* **33**, 1619–1634 (2019).
21. Alberti, S., Gladfelter, A. & Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419–434 (2019).
22. Hao, S. *et al.* YAP condensates are highly organized hubs. *iScience* **27**, 109927 (2024).
23. Betzig, E. *et al.* Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science* **313**, 1642–1645 (2006).
24. Manley, S., Gillette, J. M. & Lippincott-Schwartz, J. Single-Particle Tracking Photoactivated Localization Microscopy for Mapping Single-Molecule Dynamics. *Methods Enzymol.* **475**, 109–120 (2010).
25. Elf, J., Li, G.-W. & Xie, X. S. Probing transcription factor dynamics at the single-molecule level in a living cell. *Science* **316**, 1191–1194 (2007).
26. Li, Y., Luo, H., Liu, T., Zacksenhaus, E. & Ben-David, Y. The ets transcription factor Fli-1 in development, cancer and disease. *Oncogene* **34**, 2022–2031 (2015).

27. Johnson, K. M. *et al.* Role for the EWS domain of EWS/FLI in binding GGAA-microsatellites required for Ewing sarcoma anchorage independent growth. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 9870–9875 (2017).
28. Boulay, G. *et al.* Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. *Cell* **171**, 163-178.e19 (2017).
29. Lodhi, N., Ji, Y. & Tulin, A. Mitotic bookmarking: maintaining post-mitotic reprogramming of transcription reactivation. *Curr. Mol. Biol. Rep.* **2**, 10–16 (2016).
30. Ito, K. & Zaret, K. S. Maintaining Transcriptional Specificity through Mitosis. *Annu. Rev. Genomics Hum. Genet.* **23**, 53–71 (2022).
31. Bhat, P., Honson, D. & Guttman, M. Nuclear compartmentalization as a mechanism of quantitative control of gene expression. *Nat. Rev. Mol. Cell Biol.* **22**, 653–670 (2021).
32. Budinich, K. A. *et al.* RNA reinforces condensate nucleation on chromatin to amplify oncogenic transcription. *BioRxiv Prepr. Serv. Biol.* 2025.07.29.665963 (2025) doi:10.1101/2025.07.29.665963.
33. Henninger, J. E. *et al.* RNA-Mediated Feedback Control of Transcriptional Condensates. *Cell* **184**, 207-225.e24 (2021).
34. Wang, Y. *et al.* Dissolution of oncofusion transcription factor condensates for cancer therapy. *Nat. Chem. Biol.* **19**, 1223–1234 (2023).
35. Platzer, R. *et al.* Unscrambling fluorophore blinking for comprehensive cluster detection via photoactivated localization microscopy. *Nat. Commun.* **11**, 4993 (2020).
36. Annibale, P., Vanni, S., Scarselli, M., Rothlisberger, U. & Radenovic, A. Identification of clustering artifacts in photoactivated localization microscopy. *Nat. Methods* **8**, 527–528 (2011).
37. Annibale, P., Vanni, S., Scarselli, M., Rothlisberger, U. & Radenovic, A. Quantitative photo activated localization microscopy: unraveling the effects of photoblinking. *PloS One* **6**, e22678 (2011).
38. Lee, S.-H., Shin, J. Y., Lee, A. & Bustamante, C. Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM). *Proc. Natl. Acad. Sci. U. S. A.* **109**, 17436–17441 (2012).

39. Veatch, S. L. *et al.* Correlation functions quantify super-resolution images and estimate apparent clustering due to over-counting. *PloS One* **7**, e31457 (2012).
40. Sengupta, P. *et al.* Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. *Nat. Methods* **8**, 969–975 (2011).
41. Annibale, P., Scarselli, M., Kodiyan, A. & Radenovic, A. Photoactivatable Fluorescent Protein mEos2 Displays Repeated Photoactivation after a Long-Lived Dark State in the Red Photoconverted Form. *J. Phys. Chem. Lett.* **1**, 1506–1510 (2010).
42. Jensen, L. G. *et al.* Correction of multiple-blinking artifacts in photoactivated localization microscopy. *Nat. Methods* **19**, 594–602 (2022).
43. Veatch, S. L. *et al.* Critical fluctuations in plasma membrane vesicles. *ACS Chem. Biol.* **3**, 287–293 (2008).
44. Sengupta, P. & Lippincott-Schwartz, J. Quantitative analysis of photoactivated localization microscopy (PALM) datasets using pair-correlation analysis. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **34**, 396–405 (2012).
45. Sengupta, P., Jovanovic-Talisman, T. & Lippincott-Schwartz, J. Quantifying spatial organization in point-localization superresolution images using pair correlation analysis. *Nat. Protoc.* **8**, 345–354 (2013).
46. Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* 226–231 (AAAI Press, Portland, Oregon, 1996).
47. Zandi, R., van der Schoot, P., Reguera, D., Kegel, W. & Reiss, H. Classical nucleation theory of virus capsids. *Biophys. J.* **90**, 1939–1948 (2006).
48. Clouet, E. Modeling of Nucleation Processes. Preprint at https://doi.org/10.48550/arXiv.1001.4131 (2010).
49. Slezov, V. V. *Kinetics of First-Order Phase Transitions*. (Wiley, 2009). doi:10.1002/9783527627769.
50. Becker, R. & Döring, W. Kinetische Behandlung der Keimbildung in übersättigten Dämpfen. *Ann. Phys.* **416**, 719–752 (1935).

51. Gu, J. *et al.* Hsp40 proteins phase separate to chaperone the assembly and maintenance of membraneless organelles. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 31123–31133 (2020).
52. Yoo, H., Bard, J. A. M., Pilipenko, E. V. & Drummond, D. A. Chaperones directly and efficiently disperse stress-triggered biomolecular condensates. *Mol. Cell* **82**, 741-755.e11 (2022).
53. Boczek, E. E. *et al.* HspB8 prevents aberrant phase transitions of FUS by chaperoning its folded RNA-binding domain. *eLife* **10**, e69377 (2021).
54. Lu, Y. *et al.* Phase separation of TAZ compartmentalizes the transcription machinery to promote gene expression. *Nat. Cell Biol.* **22**, 453–464 (2020).
55. Rai, A. K., Chen, J.-X., Selbach, M. & Pelkmans, L. Kinase-controlled phase transition of membraneless organelles in mitosis. *Nature* **559**, 211–216 (2018).
56. Reineke, L. C. *et al.* Casein Kinase 2 Is Linked to Stress Granule Dynamics through Phosphorylation of the Stress Granule Nucleating Protein G3BP1. *Mol. Cell. Biol.* **37**, e00596-16 (2017).
57. Strom, A. R. *et al.* Interplay of condensation and chromatin binding underlies BRD4 targeting. *Mol. Biol. Cell* **35**, ar88 (2024).
58. Owen, I. *et al.* The oncogenic transcription factor FUS-CHOP can undergo nuclear liquid-liquid phase separation. *J. Cell Sci.* **134**, jcs258578 (2021).
59. Chandra, B. *et al.* Phase Separation Mediates NUP98 Fusion Oncoprotein Leukemic Transformation. *Cancer Discov.* **12**, 1152–1169 (2022).
60. Terlecki-Zaniewicz, S. *et al.* Biomolecular condensation of NUP98 fusion proteins drives leukemogenic gene expression. *Nat. Struct. Mol. Biol.* **28**, 190–201 (2021).
61. Gangwal, K. *et al.* Microsatellites as EWS/FLI response elements in Ewing's sarcoma. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 10149–10154 (2008).
62. Riggi, N. *et al.* EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. *Cancer Cell* **26**, 668–681 (2014).
63. DepMap Public 25Q3. (2025).

64. Arafeh, R., Shibue, T., Dempster, J. M., Hahn, W. C. & Vazquez, F. The present and future of the Cancer Dependency Map. *Nat. Rev. Cancer* **25**, 59–73 (2025).
65. Ripley, B. D. *Spatial Statistics*. (Wiley, 1981).
66. Yamazaki, H., Takagi, M., Kosako, H., Hirano, T. & Yoshimura, S. H. Cell cycle-specific phase separation regulated by protein charge blockiness. *Nat. Cell Biol.* **24**, 625–632 (2022).
67. Teves, S. S. *et al.* A dynamic mode of mitotic bookmarking by transcription factors. *eLife* **5**, e22280 (2016).
68. Yoshida, S. R. & Chong, S. Single-Molecule Measurement of Protein Interaction Dynamics within Biomolecular Condensates. *J. Vis. Exp. JoVE* https://doi.org/10.3791/66169 (2024) doi:10.3791/66169.
69. Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R. & Darzacq, X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* **6**, e25776 (2017).
70. Heckert, A., Dahal, L., Tjian, R. & Darzacq, X. Recovering mixtures of fast-diffusing states from short single-particle trajectories. *eLife* **11**, e70169 (2022).
71. Caravaca, J. M. *et al.* Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor to mitotic chromosomes. *Genes Dev.* **27**, 251–260 (2013).
72. Bensaude, O. Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? *Transcription* **2**, 103–108 (2011).
73. Guillon, N. *et al.* The oncogenic EWS-FLI1 protein binds in vivo GGAA microsatellite sequences with potential transcriptional activation function. *PloS One* **4**, e4932 (2009).
74. Irgen-Gioro, S., Yoshida, S., Walling, V. & Chong, S. Fixation can change the appearance of phase separation in living cells. *eLife* **11**, e79903 (2022).
75. Park, H. *et al.* Chromosomal localization of Ewing sarcoma EWSR1/FLI1 protein promotes the induction of aneuploidy. *J. Biol. Chem.* **296**, 100164 (2021).
76. Kitagawa, T. *et al.* AT-hook DNA-binding motif-containing protein one knockdown downregulates EWS-FLI1 transcriptional activity in Ewing's sarcoma cells. 2022.05.16.492174 Preprint at https://doi.org/10.1101/2022.05.16.492174 (2022).

77. Selvanathan, S. & Toretsky, J. An EWS::FLI1 monoclonal antibody with multimodal utility and specificity for FLI1. 2023.11.13.566886 Preprint at https://doi.org/10.1101/2023.11.13.566886 (2023).
78. Zöllner, S. K. *et al.* Inhibition of the oncogenic fusion protein EWS-FLI1 causes G2-M cell cycle arrest and enhanced vincristine sensitivity in Ewing's sarcoma. *Sci. Signal.* **10**, eaam8429 (2017).
79. Uechi, H. *et al.* Small-molecule dissolution of stress granules by redox modulation benefits ALS models. *Nat. Chem. Biol.* **21**, 1577–1588 (2025).
80. Harlow, M. L. *et al.* Trabectedin Inhibits EWS-FLI1 and Evicts SWI/SNF from Chromatin in a Schedule-dependent Manner. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **25**, 3417–3429 (2019).
81. Ciryam, P., Tartaglia, G. G., Morimoto, R. I., Dobson, C. M. & Vendruscolo, M. Widespread aggregation and neurodegenerative diseases are associated with supersaturated proteins. *Cell Rep.* **5**, 781–790 (2013).
82. Ciryam, P., Kundra, R., Morimoto, R. I., Dobson, C. M. & Vendruscolo, M. Supersaturation is a major driving force for protein aggregation in neurodegenerative diseases. *Trends Pharmacol. Sci.* **36**, 72–77 (2015).
83. Pappu, R. V., Wang, X., Vitalis, A. & Crick, S. L. A polymer physics perspective on driving forces and mechanisms for protein aggregation. *Arch. Biochem. Biophys.* **469**, 132–141 (2008).
84. Vitalis, A., Wang, X. & Pappu, R. V. Atomistic simulations of the effects of polyglutamine chain length and solvent quality on conformational equilibria and spontaneous homodimerization. *J. Mol. Biol.* **384**, 279–297 (2008).
85. Kar, M. *et al.* Phase-separating RNA-binding proteins form heterogeneous distributions of clusters in subsaturated solutions. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2202222119 (2022).
86. Kar, M. *et al.* Solutes unmask differences in clustering versus phase separation of FET proteins. *Nat. Commun.* **15**, 4408 (2024).
87. Shimobayashi, S. F., Ronceray, P., Sanders, D. W., Haataja, M. P. & Brangwynne, C. P. Nucleation landscape of biomolecular condensates. *Nature* **599**, 503–506 (2021).

88. Narayanan, A. *et al.* A first order phase transition mechanism underlies protein aggregation in mammalian cells. *eLife* **8**, e39695 (2019).
89. Du, M. *et al.* Direct observation of a condensate effect on super-enhancer controlled gene bursting. *Cell* **187**, 331-344.e17 (2024).
90. Wang, S., Wang, Z. & Zang, C. Genomic clustering tendency of transcription factors reflects phase-separated transcriptional condensates at super-enhancers. *Nucleic Acids Res.* **53**, gkaf015 (2025).
91. Selvanathan, S. P. *et al.* Oncogenic fusion protein EWS-FLI1 is a network hub that regulates alternative splicing. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E1307-1316 (2015).
92. Tripathi, S. *et al.* Defining the condensate landscape of fusion oncoproteins. *Nat. Commun.* **14**, 6008 (2023).
93. Burgess, A., Vuong, J., Rogers, S., Malumbres, M. & O'Donoghue, S. I. SnapShot: Phosphoregulation of Mitosis. *Cell* **169**, 1358-1358.e1 (2017).
94. Moura, M. & Conde, C. Phosphatases in Mitosis: Roles and Regulation. *Biomolecules* **9**, 55 (2019).
95. Corley, M., Burns, M. C. & Yeo, G. W. How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Mol. Cell* **78**, 9–29 (2020).
96. Toretsky, J. A. *et al.* Oncoprotein EWS-FLI1 activity is enhanced by RNA helicase A. *Cancer Res.* **66**, 5574–5581 (2006).
97. Wadsworth, G. M. *et al.* RNA-driven phase transitions in biomolecular condensates. *Mol. Cell* **84**, 3692–3705 (2024).
98. Sanchez-Burgos, I., Herriott, L., Collepardo-Guevara, R. & Espinosa, J. R. Surfactants or scaffolds? RNAs of varying lengths control the thermodynamic stability of condensates differently. *Biophys. J.* **122**, 2973–2987 (2023).
99. Alshareedah, I. *et al.* Interplay between Short-Range Attraction and Long-Range Repulsion Controls Reentrant Liquid Condensation of Ribonucleoprotein–RNA Complexes. *J. Am. Chem. Soc.* **141**, 14593–14602 (2019).
100. De La Cruz, N. *et al.* Disorder-mediated interactions target proteins to specific condensates. *Mol. Cell* **84**, 3497-3512.e9 (2024).

101. Hazra, M. K. & Levy, Y. Cross-Talk of Cation−π Interactions with Electrostatic and Aromatic Interactions: A Salt-Dependent Trade-off in Biomolecular Condensates. *J. Phys. Chem. Lett.* **14**, 8460–8469 (2023).

102. Maharana, S. *et al.* RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* **360**, 918–921 (2018).

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

104. Grimm, J. B. *et al.* Bright photoactivatable fluorophores for single-molecule imaging. *Nat. Methods* **13**, 985–988 (2016).

105. Grimm, J. B. *et al.* A General Method to Improve Fluorophores Using Deuterated Auxochromes. *JACS Au* **1**, 690–696 (2021).

106. Baumgart, F. *et al.* Varying label density allows artifact-free analysis of membrane-protein nanoclusters. *Nat. Methods* **13**, 661–664 (2016).

107. Thevathasan, J. V. *et al.* Nuclear pores as versatile reference standards for quantitative superresolution microscopy. *Nat. Methods* **16**, 1045–1053 (2019).

108. Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat. Methods* **3**, 891–893 (2006).

109. Collaboration, T. A. *et al.* The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package. *Astrophys. J.* **935**, 167 (2022).

110. Normanno, D. *et al.* Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. *Nat. Commun.* **6**, 7357 (2015).

111. Sergé, A., Bertaux, N., Rigneault, H. & Marguet, D. Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes. *Nat. Methods* **5**, 687–694 (2008).

112. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).

Chapter 4

O-GLCNACYLATION MODULATES EWS::FLI1 HUB BIOPHYSICAL PROPERTIES AND ATTENUATES TRANSCRIPTIONAL ACTIVITY

Abstract

Eukaryotic transcription factors (TFs) typically undergo dynamic, multivalent interactions mediated by their low-complexity domains that drive the formation of local, high-concentration TF hubs at target genes; however, the role of post-translational modifications in the regulation of the biophysical properties and transactivation function of endogenous TF hubs in vivo remains poorly understood. Here, we investigated the role of O-linked β-N-acetylglucosaminylation (O-GlcNAcylation) in the regulation of oncogenic fusion protein EWS::FLI1, the TF that causes Ewing sarcoma. While O-GlcNAcylation induced by O-GlcNAcase inhibition robustly increased EWS::FLI1 glycosylation, super-resolution microscopy showed that excessive O-GlcNAcylation did not alter the nanoscale spatial distribution of endogenous EWS::FLI1. Despite this, single-particle tracking showed that, consistent with in vitro studies, increased O-GlcNAcylation resulted in a greater fraction of freely diffusing EWS::FLI1 and resulted in greater EWS::FLI1 mobility inside of hubs, suggesting a decrease in multivalent interaction strength. Further, using a dual-luciferase reporter assay, we found that O-GlcNAcylation suppressed EWS::FLI1-mediated transactivation at a synthetic GGAA microsatellite-driven gene. Together, these findings support a model where elevated O-GlcNAcylation can subtly regulate the EWS::FLI1 LCD-LCD multivalent interaction strength without significantly altering its spatial distribution, resulting in poorer retention of EWS::FLI1 in hubs and increased hub fluidity, in turn suppressing EWS::FLI1 transactivation function. More broadly, our results suggest that the transcriptional output of genes driven by EWS::FLI1 hubs is not only regulated by their size, but by the dynamics of their constituent molecules.

Introduction

DNA-binding transcription factors (TFs) are key regulators of eukaryotic gene expression typically comprising well-structured DNA-binding domains (DBDs) that can bind target

genes and intrinsically disordered low-complexity domains (LCDs) responsible for transactivation. These LCDs generally do not form stable three-dimensional conformations and undergo dynamic, multivalent interactions with other core components of the transcriptional machinery. Recent work has shown that homotypic LCD-LCD interactions can drive the formation of local, high-concentration transcription factor hubs at target genomic loci, leading to the recruitment of TFs[1,2], RNA polymerase II[3–7], and other transcriptional coregulators[3,5,8], which in turn drive the expression of target genes.

A growing body of work suggests that post-translational modifications (PTMs) are critical regulators of biomolecular condensates including TF hubs[9,10] including O-linked β-N-acetylglucosaminylation (O-GlcNAcylation), a reversible post-translational modification of serines and threonines that is cycled on by O-GlcNAc transferase (OGT) and cycled off by O-GlcNAcase (OGA)[11,12]. O-GlcNAcylation has been shown to be enriched in many biomolecular condensates[13,14] and is heavily implicated as an important regulator of their physical properties including their dynamics and phase behavior[15–22].

One well-studied example is the oncogenic fusion TF EWS::FLI1, which is pathognomonic of Ewing sarcoma and which has been shown to be O-GlcNAcylated in vivo[23]. It is the result of a t(11;22)(q24;q12) chromosomal translocation that fuses the intrinsically disordered LCD of the RNA-binding protein EWSR1 with the well-structured DBD of the TF FLI1[24]. Previous work has shown that EWS::FLI1 forms transcriptionally active hubs at GGAA microsatellites throughout the genome that drive an aberrant, oncogenic transcriptional program necessary for oncogenesis[1,25,26]. Recent in vitro studies have shown that O-GlcNAcylation of the EWS LCD weakens LCD-LCD interactions, reducing phase separation propensity and increasing condensate fluidity[22]; however, it is still not known whether O-GlcNAcylation similarly regulates endogenous EWS::FLI1 under physiological conditions and if any resulting biophysical changes have functional consequences.

Here, we address these open questions using super-resolution and single-molecule imaging approaches along with functional assays to make nanoscale measurements of how O-

GlcNAcylation regulates the spatial distribution, temporal dynamics, and transactivation function of EWS::FLI1 and EWS::FLI1 hubs.

Results

O-GlcNAcylation does not alter EWS(LCD)-driven condensate formation in cells

In order to investigate how O-GlcNAcylation affects EWS::FLI1 hub formation and function, we focused on investigating how it can modulate EWS(LCD) alone, the intrinsically disordered low-complexity domain responsible for EWS::FLI1 transactivation function. Previous in vitro work has shown that O-GlcNAcylation weakens EWS LCD-LCD interactions and causes condensates to become more liquid like[22]. To determine if O-GlcNAcylation can similarly modulate EWS(LCD) condensates in vivo, we transfected U2OS cells with EWS(LCD)-DsRed2, treated cells with either O-GlcNAc transferase (OGT) inhibitor OSMI-1 (hereafter OSMI), O-GlcNAcase (OGA) inhibitor Thiamet-G (TMG), or DMSO as a control, then imaged cells with confocal microscopy (Figure 1a). Quantification of the surface roughness (standard deviation of pixel intensities), number of puncta per cell, and punctate percentage revealed no statistically significant change due to OSMI or TMG treatment relative to the DMSO control. This suggests that the modulation of O-GlcNAcylation level does not substantially alter the morphology of EWS(LCD)-driven condensates in cells.

O-GlcNAcylation of endogenous EWS::FLI1 does not alter its nuclear distribution

Next, we investigated how O-GlcNAcylation affects the spatial organization of endogenous EWS::FLI1 using our previously established knock-in A673 cell line which contains a HaloTag fused to the C-terminus of endogenous EWS::FLI1 inserted using CRISPR/Cas9-mediated genome editing[1]. As shown in previous work, endogenous EWS::FLI1 forms hub with diameters on the order of ~100 nm, far below the resolution of conventional confocal microscopes[28]. Thus, we used photoactivated localization microscopy (PALM)[29] together with photoblinking corrections optimized for the study of biomolecular condensates[28] to

examine the effects of OSMI and TMG on endogenous EWS::FLI1 hubs with ~20 nm resolution (Figures 2a – c).

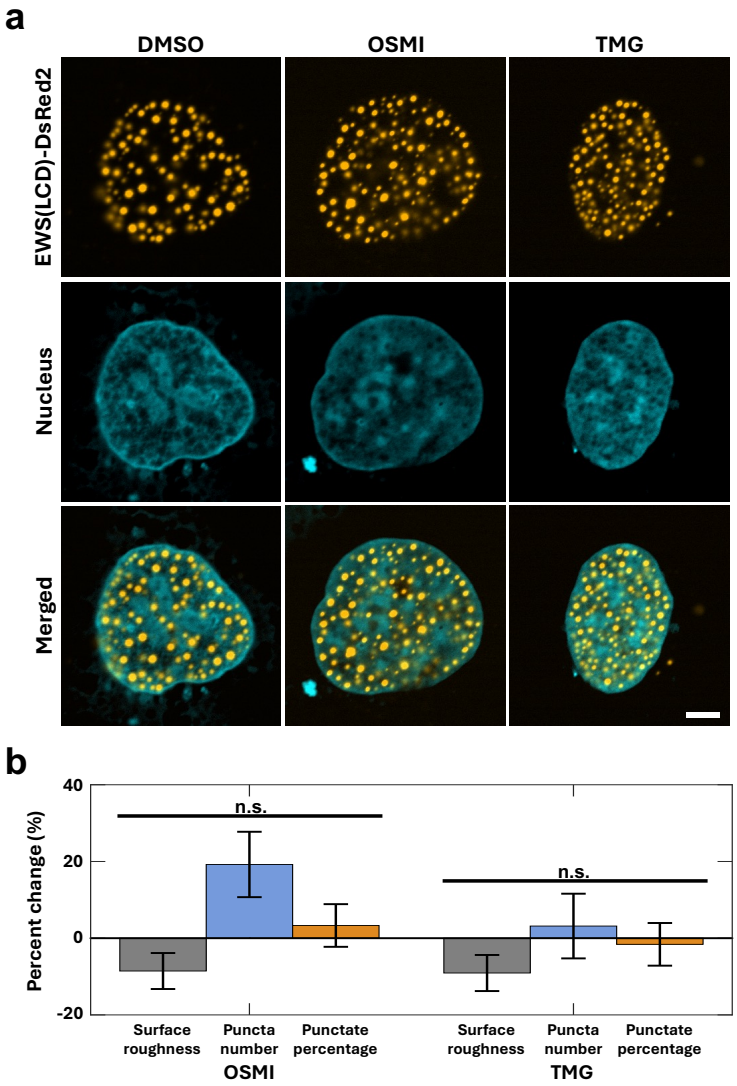

Figure 1: Increased O-GlcNAcylation does not alter the phase behavior of EWS(LCD) in vivo. (a) Dual-color confocal images of U2OS cells transfected with EWS(LCD)-DsRed2 treated with DMSO, 50 μM OSMI, or 200 μM TMG for 1 hr. DNA was labeled with Hoechst 33342 (cyan) and scale bars are 5 μm. (b) Quantification of the percent change of surface roughness, puncta number, and punctate percentage showed no statistically significant change between OSMI- or TMG-treated cells

compared to the DMSO vehicle (n.s., (left to right, $p =$ 0.20, 0.15, 0.56, 0.13, 0.90, 0.54), Mann-Whitney U test). For each condition, n = 60 cells were imaged, and extraction of parameters was performed as described in Irgen-Gioro et al.[27] and the Methods.

O-GlcNAcylation of endogenous EWS::FLI1 does not alter its nuclear distribution

Next, we investigated how O-GlcNAcylation affects the spatial organization of endogenous EWS::FLI1 using our previously established knock-in A673 cell line which contains a HaloTag fused to the C-terminus of endogenous EWS::FLI1 inserted using CRISPR/Cas9-mediated genome editing[1]. As shown in previous work, endogenous EWS::FLI1 forms hub with diameters on the order of ~100 nm, far below the resolution of conventional confocal microscopes[28]. Thus, we used photoactivated localization microscopy (PALM)[29] together with photoblinking corrections optimized for the study of biomolecular condensates[28] to examine the effects of OSMI and TMG on endogenous EWS::FLI1 hubs with ~20 nm resolution (Figures 2a – c).

We then quantified the spatial distribution of EWS::FLI1 in the PALM images using Ripley's L function[30], specifically L(r) – r, to characterize the degree of spatial clustering (Figure 2d). Here, a value of zero corresponds to complete spatial randomness, and the greater the values of L(r) – r, the greater the degree of clustering over the length-scale of r. As evident from Figure 2d, treatments with OSMI or TMG are both unable to affect the hub formation behavior of EWS::FLI1.

O-GlcNAcylation of endogenous EWS::FLI1 increases its mobility in hubs

We next examined the effects of O-GlcNAcylation on the dynamics of EWS::FLI1 hubs, namely the kinetics of exchange in and out of hubs along with their material properties. While it may seem unintuitive to expect changes in dynamics given that we observed that the sizes of EWS(LCD)-driven condensates and endogenous EWS::FLI1 hubs were insensitive to O-

GlcNAcylation level, it has previously been found that in certain cases, the structure and dynamics of EWS::FLI1 hubs can be, to an extent, divorced[28]. For example, the presence or absence of RNA can modulate the dynamics of EWS::FLI1 exchange without altering hub dimensions.

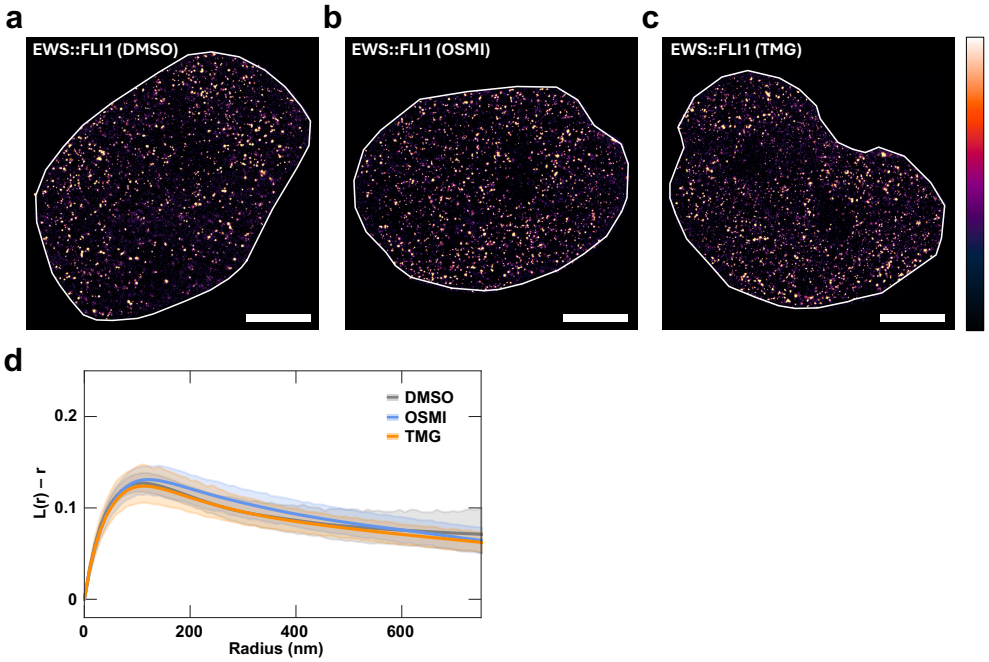

Figure 2: O-GlcNAcylation does not alter the distribution of EWS::FLI1 hubs in cells. Super-resolution microscopy image of endogenous EWS::FLI1-Halo stained with PA-JF549 in A673 cells after treatment with (a) DMSO, (b) 50 μM OSMI, or (c) 200 μM TMG for 1 hr. A white, dashed contour outlines the nucleus, and scale bars are 5 μm. (d) Ripley's L(r) – r curves show that the spatial clustering of EWS::FLI1 is not significantly affected by inhibition of OGA or OGT on the relevant length scales. For each condition, n = 10 cells were analyzed, and transparent regions are 95% confidence intervals.

Given previous reports that O-GlcNAcylation can increase EWS(LCD) condensate fluidity in vitro[22], we transfected U2OS cells with EWS(LCD)-DsRed2 and treated them with DMSO, OSMI, or TMG and performed fluorescence recovery after photobleaching (FRAP) to test whether this effect was replicable in vivo (Figure 3a). The recovery kinetics of EWS(LCD) in cells treated with DMSO and TMG were not statistically significant, with nearly identical recovery half-times ($p = 0.82$, Mann-Whitney U test) and immobile fractions ($p = 0.79$, Mann-Whitney U test). These measurements suggest that in live cells and under overexpression, EWS(LCD)-driven condensates are not sensitive to O-GlcNAcylation in the same way as they are in vitro. However, it is also plausible that the presence of the tetramerizing fluorescent protein DsRed2 masked the increase in fluidity and/or that the effect is masked by the complex intracellular environment, and FRAP, an ensemble-averaged method, lacks the sensitivity to detect subtle changes in fluidity.

To address this, we next used single-particle tracking (SPT) to measure the diffusion behavior of endogenous EWS::FLI1 with and without excessive glycosylation in the EWS::FLI1-Halo knock-in A673 cell line[1]. The SPT protocol has been described previously[28], but briefly, in order to ensure that quickly diffusing molecules are well resolved and sampled frequently enough to be tracked, cells were illuminated using stroboscopic illumination[32,33] and acquired using 5 ms exposure times. We then analyzed the resulting trajectories using a state-array SPT (saSPT)[34], a variational Bayesian method that can be used to extract a diffusion coefficient distribution without *a priori* knowledge of the number or identity of the underlying diffusing populations (Figure 3b). Using $D \geq 1\,\mu m^2/s$ as the cutoff for what is considered freely diffusing, we found that excessive O-GlcNAcylation appears to lead to an increase in freely diffusing EWS::FLI1 (Figure 3c). This suggests that O-GlcNAcylation of EWS::FLI1 weakens its homotypic interactions, resulting in reduced ability to retain EWS::FLI1 in hubs. This is consistent with in vitro reports that O-GlcNAcylation can result in more liquid-like EWS(LCD) droplets, given that EWS LCD-LCD multivalent interactions are primary drivers of EWS::FLI1 hub formation.

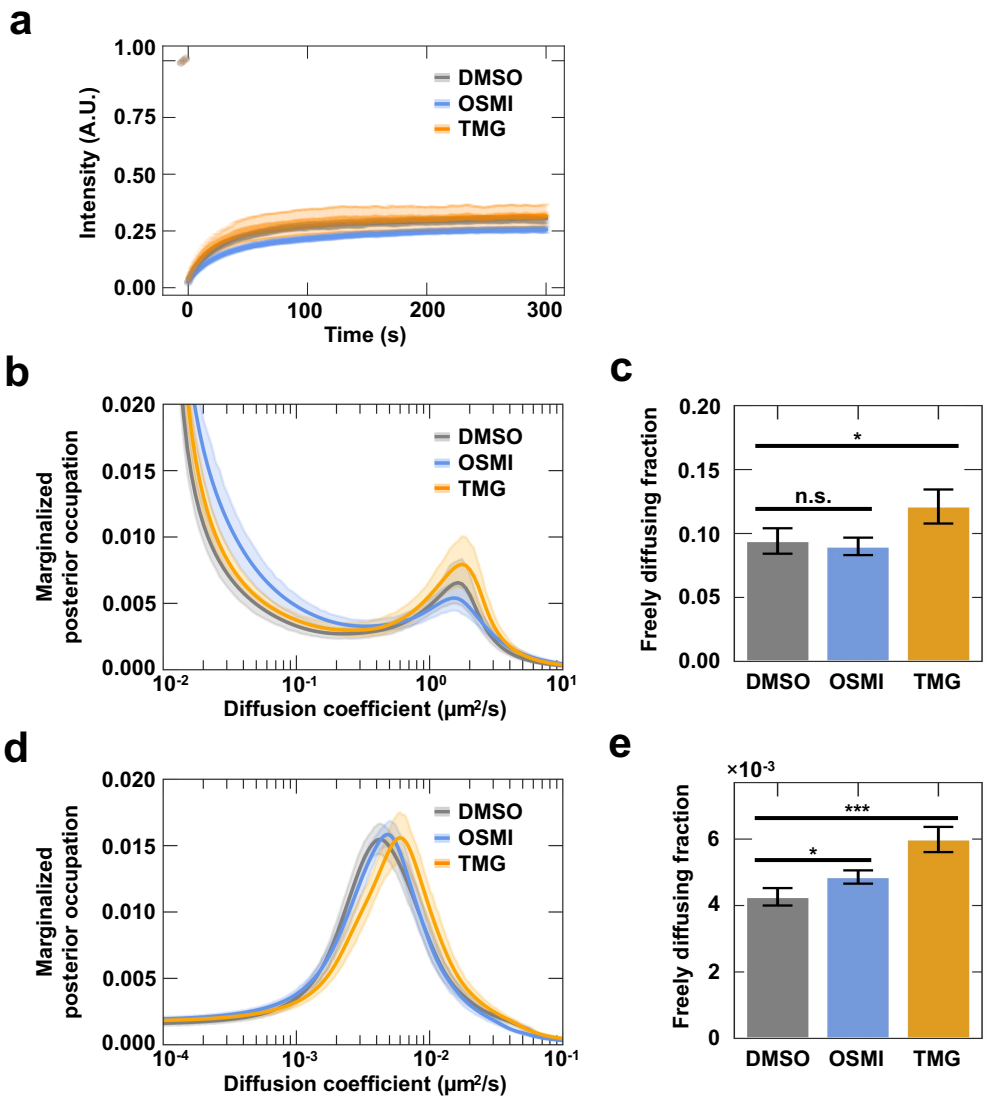

Figure 3: O-GlcNAcylation increases the mobility of EWS::FLI1 inside hubs. (a) Averaged FRAP curves for U2OS cells transiently expressing EWS(LCD)-DsRed2 show no statistically significant difference in the recovery dynamics of DMSO- and TMG-treated cells for n = 8 cells per condition. (b) Diffusion coefficient profiles acquired from fast SPT show differences in EWS::FLI1 diffusion behavior between cells treated with DMSO, OSMI, and TMG. The marginalized posterior occupation can be though of as approximating a probability mass function. (c) Plots of the freely diffusing fraction show that treatment with TMG increases population of freely diffusing EWS::FLI1 (*, $p = 3.1\times10^{-2}$, Mann-Whitney U test). EWS::FLI1 diffusing at D ≥

$1\,\mu m^2/s$ were considered freely diffusing for n = 30 cells per condition. (d) Diffusion coefficient profiles acquired from 25 Hz SPT show differences in the mobility of in-hub EWS::FLI1. (e) Plot of the characteristic diffusion coefficients shows that treatment with TMG increases the mobility of in-hub EWS::FLI1 (***, $p = 4.8 \times 10^{-4}$, Mann-Whitney U test). The characteristic diffusion coefficient was calculated as the mode for n = 35 cells per condition. Shaded regions are 95% confidence intervals and error bars are SEM.

Next, we directly probed the mobility of EWS::FLI1 inside hubs at different O-GlcNAcylation levels with 25 Hz SPT. The SPT protocol is described in more detail in the methods, but briefly, rather than using 5 ms exposure times with stroboscopic illumination, which is very good at capturing highly dynamic trajectories at the cost of quickly photobleaching bound molecules and producing more inaccurate and imprecise measurements of slower diffusive behavior, we opted to use 40 ms exposure times without stroboscopic illumination, which allowed us to accurately measure the diffusion coefficient of EWS::FLI1 bound in hubs. Once again using saSPT, we generated diffusion coefficient distributions for EWS::FLI1 after treatment with DMSO, OSMI, and TMG only over a range of diffusion coefficients corresponding to hub-bound molecules (Figure 3d). We quantified the modes of the distributions and found that TMG resulted in a statistically significant increase in the characteristic diffusion coefficient of bound EWS::FLI1, which suggests that highly O-GlcNAcylated EWS::FLI1 molecules retained in hubs are able to diffuse more freely, implying an increase in hub fluidity (Figure 3e). Together, these SPT experiments suggest that increased O-GlcNAcylation of EWS::FLI1 slightly weakens homotypic interactions, leading to an increase in freely diffusing EWS::FLI1 resulting from poorer retention in hubs, and an increase in EWS::FLI1 in-hub mobility resulting from increased hub fluidity. *O-GlcNAcylation of endogenous EWS::FLI1 can suppress its transactivation ability*

Finally, we examined the functional consequences of excessive O-GlcNAcylation on EWS::FLI1 transactivation function using a dual-luciferase reporter assay driven by a synthetic GGAA microsatellite promoter in HEK293T cells. Here, cells were co-transfected with an EWS::FLI1 expression construct together with a firefly luciferase reporter driven by a synthetic GGAA microsatellite-containing promoter; thus, increased transactivation function would result in increased luminescence, and vice versa. Reporter activity was normalized to Renilla luciferase, a control used to account for differences in transfection efficiency and cell number across conditions, and compared across treatment conditions (Figure 4). As expected, EWS::FLI1 robustly activated the reporter relative to negative controls, including empty vector and the FLI1 DNA-binding domain alone, confirming that transcriptional activation requires the EWS LCD, and the DMSO treatment did not significantly alter reporter activity. In contrast, increasing O-GlcNAcylation with TMG significantly decreased EWS::FLI1-driven transcriptional activation, showing that elevated O-GlcNAcylation can suppress the transactivation function of EWS::FLI1 at GGAA microsatellite regulatory elements. Together with the single-molecule measurements, this suggests that weakening of EWS::FLI1 homotypic interactions and increased hub fluidity by excessive O-GlcNAcylation are sufficient to impair EWS::FLI1 transactivation even in the absence of detectable changes in hub size and distribution.

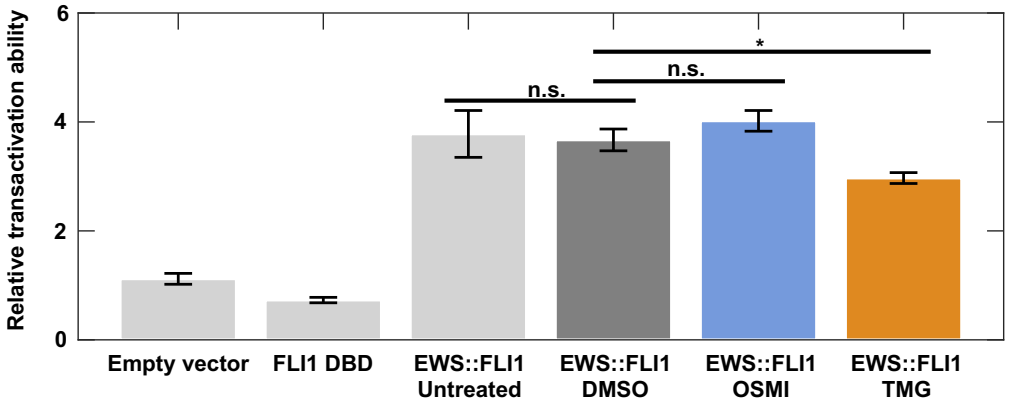

Figure 4: Increased O-GlcNAcylation suppresses the transcriptional activity of EWS::FLI1. Dual Luciferase reporter assay shows that EWS::FLI1 transactivation

function is suppressed upon treatment with TMG (*, $p = 1.1\times10^{-2}$, Student's t-test). Error bars are SEM.

Discussion

The LCD of EWS::FLI1 has been shown to be central to its ability to spatially reorganize and execute the aberrant transcriptional program characteristic of Ewing sarcoma cells through multivalent interactions that drive hub formation. Post-translational modifications like O-GlcNAcylation have long been implicated as having the potential to modulate multivalent LCD-LCD interactions, and while previous studies have investigated the effect of O-GlcNAcylation on the EWS LCD alone using in vitro systems, the effect of O-GlcNAcylation on full-length EWS::FLI1 under more physiological conditions has not been characterized thus far[22]. Here, we addressed this gap using a variety of single-molecule methods to probe how O-GlcNAcylation modulates the spatial organization, dynamics, and transactivation function of endogenous EWS::FLI1 in an Ewing sarcoma cell line.

A previous in vitro study showed that excessive O-GlcNAcylation of the EWS LCD weakens its homotypic interactions, leading to decreased phase-separation propensity and increased condensate fluidity[22]. Surprisingly, despite observing robust increases in EWS::FLI1 O-GlcNAcylation levels upon treatment with TMG, we were unable to detect any significant changes in the morphological properties of overexpressed EWS(LCD)-driven condensates in U2OS using confocal microscopy. These findings suggest that in complex cellular conditions, the effect of O-GlcNAcylation on the phase behavior of EWS(LCD) is limited. Further, we were unable to detect any nanoscale changes in the nuclear distribution of endogenous EWS::FLI1 hubs using PALM. Together, these results suggest that in the complex cellular environment, increased O-GlcNAcylation does not substantially alter the steady-state organization of EWS(LCD)-driven condensates or of endogenous EWS::FLI1 hubs.

Importantly, the absence of detectable structural changes does not imply an absence of biophysical or functional effects. As observed previously, the absence of RNA in EWS::FLI1

hubs led to altered binding dynamics in the absence of structural changes[28]. Here, using SPT, we found that increased O-GlcNAcylation resulted in a larger freely diffusing fraction and increased mobility of EWS::FLI1 within hubs. Finally, we showed that increased O-GlcNAcylation suppressed EWS::FLI1-mediated transactivation at a synthetic GGAA microsatellite using the dual luciferase reporter assay.

Together, these findings support a model where O-GlcNAcylation regulates the dynamics and transactivation function of EWS::FLI1 without substantially altering its spatial distribution (Figure 5). In this model, excessive O-GlcNAcylation subtly decreases the multivalent interaction strength of EWS::FLI1 without substantially altering its spatial distribution, causing poorer EWS::FLI1 hub retention and increased hub fluidity, in turn resulting in suppressed transcription. This separation between spatial organization and interaction dynamics/function is particularly important in the context of current models of biomolecular condensation, where condensate formation is often treated as a binary transition between functional and nonfunctional states corresponding to the presence or absence of condensates. Here, consistent with previous findings, we observed a decoupling of nano- and mesoscale structure and interaction dynamics/function where despite no apparent change in hub morphology and spatial distribution, EWS::FLI1 underwent substantial changes in diffusion behavior, hub material properties, and transactivation function. This suggests that the functional state of EWS::FLI1 is not solely governed by their physical presence, but also by the interaction strength of their constituent molecules. That is, EWS::FLI1 hub formation alone is not sufficient for efficient transactivation; it also requires multivalent interaction strength sufficient to stabilize functional hubs that can retain EWS::FLI1 and maintain suitable viscosity. Excessive O-GlcNAcylation appears to shift this balance to a more weakly interacting state that is less able to drive productive transcription.

This work also motivates the importance of using single-molecule methods, as ensemble-averaged methods like FRAP may be unable to capture functionally relevant changes in dynamics. In our work, despite statistically significant increases in freely diffusing fraction and in-hub mobility, only SPT was able to resolve these clear shifts. This adds to a growing body of work highlighting the necessity of single-molecule methods in the study of nuclear

assemblies including transcription factor hubs, whose functional outputs may depend on subtle changes in dynamics or nanoscale structure[2,28,35].

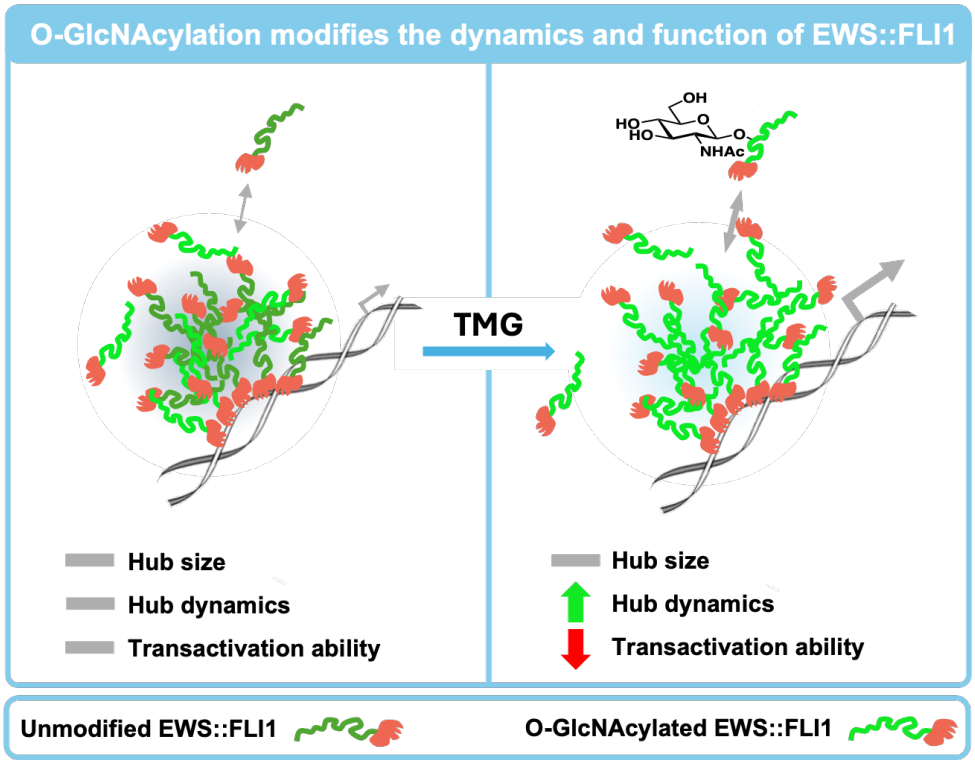

Figure 5: A summary of the putative role of O-GlcNAcylation in regulating EWS::FLI1 hub dynamics and transactivation function. 1) Under the increased O-GlcNAcylation achievable by using TMG in vivo, the spatial distribution of EWS::FLI1 does not appear to be altered. 2) However, upon increased O-GlcNAcylation, EWS::FLI1 becomes more mobile inside of hubs and is more poorly retained resulting in an increased freely diffusing fraction. 3) This leads to a decrease in EWS::FLI1 transactivation ability.

This study also has limitations that should be considered. First, TMG is an OGA inhibitor, thus increases O-GlcNAcylation globally rather than specifically on EWS::FLI1. Thus, while

our results are consistent with one another and with the literature, the observed effects may be confounded by the O-GlcNAcylation of other proteins. While currently, the glycosylation sites of EWS::FLI1 remain unknown, future studies should map them out and exploit site-specific mutagenesis, for example, by conversion of serine and threonine to alanine, to allow for the isolation of O-GlcNAcylation effects specific to EWS::FLI1. In addition, our SPT results imply a decrease in multivalent interaction strength, but do not demonstrate it directly. While we have used the LacO array assay to probe this explicitly, results have been inconsistent, as the effect may be subtle enough that it lives on the edge of what is detectable by such an ensemble method.

Overall, our findings establish O-GlcNAcylation as a regulator of EWS::FLI1 hub dynamics and transactivation function. Rather than modulating the spatial properties of hubs, O-GlcNAcylation modulates the interaction strength of EWS::FLI1, resulting in a decrease in hubs' abilities to retain molecules and their viscosity, which in turn suppresses their transcriptional output. More broadly, these results establish a model where the functional state of transcription factor hubs is not solely governed by their spatial organization, but also the interaction strengths and dynamics of their constituent molecules.

Methods

Cell Culture:

The knock-in A673 cell line expressing endogenous EWS::FLI1-Halo from Chong et al.[1] and Hek293T cells were cultured in 4.5 g/L high-glucose DMEM (Thermo Fisher, 10566016) with 10% FBS (Fisher Scientific, SH3039603) and 1% penicillin-streptomycin (Thermo Fisher, 15140122). The U2OS cell line containing ~50,000 LacO element repeats described in Janicki et al.[36] was cultured in 1 g/L low-glucose DMEM (ThermoFisher, 10567014) similarly with 10% FBS (Fisher Scientific, SH3039603) and 1% penicillin-streptomycin (Thermo Fisher, 15140122). During live-cell imaging, A673 and U2OS cells were instead cultured in identically supplemented phenol-red-free versions of high-glucose (Thermo Fisher, 31053028) and low-glucose (Thermo Fisher, 11054020) DMEM,

respectively. Both cell lines were cultured in a humidified incubator at 37°C and under 5% CO_2.

For confocal microscopy experiments, cells were cultured on 35 mm MatTek glass-bottom dishes (MatTek, P35G-1.5-20-C). Knock-in A673 cells were stained with 200 nM of either JFX549 or JFX646 by incubating in the appropriate ligand for 15 minutes, then being washed twice. For each wash, culture media was aspirated, cells were rinsed twice with warm PBS, fresh culture media was added back in, then cells were returned to the incubator for 15 minutes. For FRAP and puncta formation assays, wild type U2OS cells were transfected with EWS(LCD)-DsRed2 24 hrs before imaging using Lipofectamine 3000 (Fisher Scientific, L3000001) following the manufacturer's protocol. For LacO array assays, the U2OS cells containing ~50,000 LacO repeats were similarly transfected with both EWS(LCD)-EYFP-LacI and mCherry-EWS(LCD). An hour before imaging or fixation, cells were incubated in 50 μM OSMI-4, 200 μM Thiamet-G, or DMSO for 1 hr. For fixed-cell confocal imaging, cells were fixed with 4% paraformaldehyde (VWR, 100503-917) for 15 minutes, then washed thrice by rinsing twice with PBS then rocking for 5 min each wash.

For single-molecule imaging including PALM, cells were cultured on 25 mm #1.5H coverslips (Marienfeld Superior, 0117650). Coverslips were cleaned by immersing in 1 M KOH and sonicating for 1 hr, rinsing thrice with double-distilled water, immersing in pure ethanol and sonicating for 1 hr, then immersing in 20% ethanol diluted in double-distilled water for storage.

Confocal microscopy:

Confocal microscopy was performed on a Nikon Eclipse Ti2 AX R point-scanning confocal microscope setup with a x60/NA 1.42 (Nikon, CFI Plan Apochromat Lambda D 60X Oil, MRD71670). Sample excitation was performed using a Nikon LUA-S4 laser launch using 561 nm (JFX549 and DsRed2), and 640 nm (JFX646) beams. Z-stacks were acquired with 0.5 μm axial steps with a pinhole size of 1 Airy unit. For live-cell imaging, cells were kept humid at 37°C under 5% CO_2 using a stage incubator (Tokai Hit).

Fluorescence recovery after photobleaching (FRAP):

The FRAP dynamics of EWS(LCD)-DsRed2 were measured on the confocal microscope described above. The sample was imaged with the 561 nm beam line with a frame rate of 0.5 frames per second for three frames. Then, a 190 nm^2 circular region of interest (ROI) on the sample was bleached with the 561 nm laser at maximum power and the sample was imaged at 2.6 frames per second for another 774 frames (~5 minutes). The resulting ROI intensity trace was triple normalized following Chong et al.[1], and the resulting FRAP curve was fit against the below reaction-dominant, two-component model of recovery,

$$I_{FRAP}(t) = 1 - Ae^{-k_1 t} - Be^{-k_2 t}, k_1 \geq k_2$$

Here, k_1 and k_2 are fast and slow dissociation rate constants, respectively. As k_1 likely corresponds to transient or nonspecific interactions, it was not considered and k_2 was used in downstream analyses.

Single-molecule imaging:

Single-molecule imaging was performed on a Nikon Eclipse Ti2 N-STORM microscope setup with a x100/NA 1.49 oil-immersion objective (Nikon, CFI SR HP Apochromat TIRF 100XAC Oil, MRD01995). Samples were illuminated under a highly inclined and laminated optical sheet (HILO) with a Nikon LUN-F solid-state laser unit using 405 nm (PA-JF549 and PA-JF646 photoactivation), 561 nm (PA-JF549 and JFX549 excitation), and 640 nm (PA-JF646 and JFX646 excitation). Signal was split using a 635 nm single-edge dichroic beamsplitter (Semrock, Di02-R635-25x36), filtered using a 593/40 nm single-band bandpass filter (Semrock, FF01-593/40-25) for JFX549 and PA-JF549 or a 676/37 nm single-band bandpass filter (Semrock, FF01-676/37-25) for JFX646 and PA-JF646, then detected on iXon Ultra 897 EMCCD (Andor). Hardware was controlled using the NIS-Elements software (Nikon).

Photoactivated localization microscopy (PALM):

Cells cultured on clean coverslips were stained with 200 nM PA-JF549 following the protocol outlined in Yoshida et al.[28], then were then rinsed twice with 1X PBS and fixed with a 4% paraformaldehyde (VWR, 100503-917) supplemented with 0.2% glutaraldehyde (Sigma Aldrich, 340855-25ML) for 15 minutes, then washed thrice with 1X PBS (Thermo Fisher, 18912014) for 5 min per wash. Fiducial markers were added to the coverslip by aspirating the PBS and adding 0.1 μm yellow-green Invitrogen Fluospheres (Thermo Fisher, F8803) diluted in DPBS supplemented with calcium and magnesium (Thermo Fisher, 14-040-141) and allowing the beads to adsorb to the coverslip for 10 minutes at room temperature. Coverslips were then transferred to Attofluor Cell Chambers (Thermo Fisher, A7816) and imaged.

The samples were imaged on the single-molecules imaging setup described in the preceding paragraphs under continuous and simultaneous photoactivation (405 nm) and excitation (561 nm). The 405 nm laser power was gradually increased to maintain relatively constant photoactivated molecule density. A 50 ms exposure time was used and the samples were imaged until depletion; that is, until single molecules were no longer detected. Fiducial marker images used for drift correction were acquired every 100 frames. Single-molecule localizations were then fitted using the NIS-Elements software using the following settings: minimum height: 1500, maximum height: 65535, CCD baseline: 100, minimum width: 200 nm, maximum width: 600 nm, initial fit width: 300 nm, maximum axial ratio: 1.2, maximum displacement: 0. Fiducial markers were fitted under nearly identical settings where only the minimum height was changed to 3000. Drift correction was performed in NIS-Elements using the fiducial marker localizations, and the drift-corrected localizations were corrected for photoblinking following the scheme outlined in Yoshida et al.[28]

Fast (200 Hz) single-particle tracking (SPT):

Cells cultured on a clean coverslip were stained with 20 nM nM PA-JF646 following a protocol outlined in Yoshida et al.[28] Then coverslips were transferred to Attofluor Cell chambers, DMEM was replaced with phenol-red-free DMEM, and cells were imaged on the single-molecule imaging setup described above using HILO illumination geometry and the

stroboscopic temporal illumination scheme described in Chong et al.[1] and Yoshida et al.[28] Briefly, single molecules were excited by the 640 nm laser in only the first 1 ms of the 5 ms exposure time largely preventing smearing of the point spread functions of highly mobile molecules. Single molecules were photoactivated during the 158.01 μs transition times between frames by a 405 nm laser whose power was gradually increased over the course of imaging to maintain a relatively constant population of photoactivated molecules. Each cell was imaged for 20000 frames.

Single molecules were then localized in each frame and assembled into trajectories using SLIMfast[37], a MATLAB based implementation of the multiple target tracking algorithm[38] using the following parameters: localization error: 10^{-5}, deflation loops: 2, maximum number of competitors: 3, maximum expected diffusion coefficient: $3.5\ \mu m^2/s$. The extracted trajectories were then analyzed using the state array SPT (saSPT) package developed in Heckert et al.[34] using a state array defined over one hundred logarithmically spaced diffusion coefficients from 10^{-2} to $10^{1}\ \mu m^2/s$, inclusive, and twenty linearly spaced localization errors from 0 to $0.07\ \mu m$, inclusive. The posterior occupations were then marginalized over localization error leaving a (marginalized) diffusion coefficient distribution.

25 Hz single-particle tracking (SPT):

20 Hz SPT was performed similarly only using a 40 ms exposure time without stroboscopic illumination. The 640 nm laser was used to excite the sample for the full exposure time, and the 405 nm laser was used to photoactivate single molecules only between frames as described above. Single molecules were localized in each frame of the single-molecule channel using SLIMfast using the following parameters: localization error: 10^{-6}, deflation loops: 2, maximum number of competitors: 3, maximum expected diffusion coefficient: $3.5\ \mu m^2/s$. Trajectories were analyzed using saSPT using a state array defined over one hundred logarithmically spaced diffusion coefficients from 10^{-4} to $10^{-1}\ \mu m^2/s$, inclusive, and twenty linearly spaced localization errors from 0 to $0.07\ \mu m$, inclusive. The posterior

occupations were then marginalized over localization error leaving a (marginalized) diffusion coefficient distribution.

Dual luciferase reporter assay:

The firefly luciferase reporter was generated as described in Chong et al.[1] and contains a 500-bp promoter region (-1.6 kb to -1.1 kb) from the endogenous human *NR0B1* gene upstream of a minimal SV40 promoter[39] itself containing a 23x GGAA microsatellite. Hek 293T cells were co-transfected with the firefly luciferase reporter, a pRL-TK *Renilla* reporter, and EWS::FLI1 (or an empty vector sharing the same backbone as a control) using Lipofectamine 3000 (Fisher Scientific, L3000001) following the manufacturer's protocol. The firefly and Renilla luciferase activity were then measured using a Dual-Luciferase® Reporter Assay System kit (Promega, E1960), again following the manufacturer's protocol, and the firefly luciferase signal was normalized against the *Renilla* luciferase signal to control for transfection efficiency.

Statistical analyses:

For normally distributed data with equal variance, Student's t-tests were used. For normally distributed data with unequal variance, Welch's t-tests were used. For skew data, Mann-Whitney U test was used. All statistical tests were performed in Python using the `scipy.stats` package. Significance is abbreviated in figures according to the following key: n.s. is not statistically significant, * is $p < 0.05$, ** is $p < 0.01$, and *** is $p < 0.001$.

References

1. Chong, S. *et al.* Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555 (2018).
2. Cai, D. *et al.* Phase separation of YAP reorganizes genome topology for long-term YAP target gene expression. *Nat. Cell Biol.* **21**, 1578–1589 (2019).
3. Lyons, H. *et al.* Functional partitioning of transcriptional regulators by patterned charge blocks. *Cell* **186**, 327-345.e28 (2023).

4. Lyons, H. *et al.* RNA polymerase II partitioning is a shared feature of diverse oncofusion condensates. *Cell* **188**, 3843-3862.e28 (2025).
5. Cho, W.-K. *et al.* Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).
6. Guo, Y. E. *et al.* Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* **572**, 543–548 (2019).
7. Boehning, M. *et al.* RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.* **25**, 833–840 (2018).
8. Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
9. Lin, Y.-H., Forman-Kay, J. D. & Chan, H. S. Theories for Sequence-Dependent Phase Behaviors of Biomolecular Condensates. *Biochemistry* **57**, 2499–2508 (2018).
10. Pappu, R. V., Wang, X., Vitalis, A. & Crick, S. L. A polymer physics perspective on driving forces and mechanisms for protein aggregation. *Arch. Biochem. Biophys.* **469**, 132–141 (2008).
11. Vocadlo, D. J. O-GlcNAc processing enzymes: catalytic mechanisms, substrate specificity, and enzyme regulation. *Curr. Opin. Chem. Biol.* **16**, 488–497 (2012).
12. Zachara, N. E. & Hart, G. W. The Emerging Significance of O-GlcNAc in Cellular Regulation. *Chem. Rev.* **102**, 431–438 (2002).
13. Chen, Q. & Yu, X. OGT restrains the expansion of DNA damage signaling. *Nucleic Acids Res.* **44**, 9266–9278 (2016).
14. Ohn, T., Kedersha, N., Hickman, T., Tisdale, S. & Anderson, P. A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly. *Nat. Cell Biol.* **10**, 1224–1231 (2008).
15. Patel, A. *et al.* A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066–1077 (2015).
16. Murray, D. T. *et al.* Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell* **171**, 615-627.e16 (2017).

17. Murray, D. T. *et al.* Structural characterization of the D290V mutation site in hnRNPA2 low-complexity-domain polymers. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E9782–E9791 (2018).
18. Franzmann, T. M. *et al.* Phase separation of a yeast prion protein promotes cellular fitness. *Science* **359**, eaao5654 (2018).
19. Feric, M. *et al.* Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **165**, 1686–1697 (2016).
20. Putnam, A., Cassani, M., Smith, J. & Seydoux, G. A gel phase promotes condensation of liquid P granules in Caenorhabditis elegans embryos. *Nat. Struct. Mol. Biol.* **26**, 220–226 (2019).
21. Kroschwald, S. *et al.* Different Material States of Pub1 Condensates Define Distinct Modes of Stress Adaptation and Recovery. *Cell Rep.* **23**, 3327–3339 (2018).
22. Nosella, M. L. *et al.* O-Linked-N-Acetylglucosaminylation of the RNA-Binding Protein EWS N-Terminal Low Complexity Region Reduces Phase Separation and Enhances Condensate Dynamics. *J. Am. Chem. Soc.* **143**, 11520–11534 (2021).
23. Bachmaier, R. *et al.* O-GlcNAcylation is involved in the transcriptional activity of EWS-FLI1 in Ewing's sarcoma. *Oncogene* **28**, 1280–1284 (2009).
24. Li, Y., Luo, H., Liu, T., Zacksenhaus, E. & Ben-David, Y. The ets transcription factor Fli-1 in development, cancer and disease. *Oncogene* **34**, 2022–2031 (2015).
25. Boulay, G. *et al.* Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. *Cell* **171**, 163-178.e19 (2017).
26. Johnson, K. M. *et al.* Role for the EWS domain of EWS/FLI in binding GGAA-microsatellites required for Ewing sarcoma anchorage independent growth. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 9870–9875 (2017).
27. Irgen-Gioro, S., Yoshida, S., Walling, V. & Chong, S. Fixation can change the appearance of phase separation in living cells. *eLife* **11**, e79903 (2022).
28. Yoshida, S., Zhong, Y., Kumagai, A., Dunphy, W. G. & Chong, S. Dynamic regulation of endogenous transcription factor hubs at single-molecule resolution. 2026.01.25.701606 Preprint at https://doi.org/10.64898/2026.01.25.701606 (2026).

29. Betzig, E. *et al.* Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science* **313**, 1642–1645 (2006).
30. Ripley, B. D. *Spatial Statistics*. (Wiley, 1981).
31. Thompson, J. W., Griffin, M. E. & Hsieh-Wilson, L. C. Methods for the Detection, Study, and Dynamic Profiling of O-GlcNAc Glycosylation. *Methods Enzymol.* **598**, 101–135 (2018).
32. Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R. & Darzacq, X. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* **6**, e25776 (2017).
33. Elf, J., Li, G.-W. & Xie, X. S. Probing transcription factor dynamics at the single-molecule level in a living cell. *Science* **316**, 1191–1194 (2007).
34. Heckert, A., Dahal, L., Tjian, R. & Darzacq, X. Recovering mixtures of fast-diffusing states from short single-particle trajectories. *eLife* **11**, e70169 (2022).
35. Chen, L. *et al.* Hormone-induced enhancer assembly requires an optimal level of hormone receptor multivalent interactions. *Mol. Cell* **83**, 3438-3456.e12 (2023).
36. Janicki, S. M. *et al.* From Silencing to Gene Expression: Real-Time Analysis in Single Cells. *Cell* **116**, 683–698 (2004).
37. Normanno, D. *et al.* Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. *Nat. Commun.* **6**, 7357 (2015).
38. Sergé, A., Bertaux, N., Rigneault, H. & Marguet, D. Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes. *Nat. Methods* **5**, 687–694 (2008).
39. Gangwal, K. *et al.* Microsatellites as EWS/FLI response elements in Ewing's sarcoma. *Proc. Natl. Acad. Sci.* **105**, 10149–10154 (2008).

Chapter 5

RNA POLYMERASE II PHOSPHORYLATION STATE DICTATES ITS ASSOCIATION WITH EWS::FLI1

Abstract

Eukaryotic transcription factor (TF) hubs formed through homotypic, multivalent low-complexity domain (LCD) interactions play a crucial role in driving the transcription of their target genes; despite this, how endogenous TF hubs associate with RNA Polymerase II (RNAP II) through the transcription cycle under physiological conditions remains poorly understood. Here, we use dual-color stimulated emission depletion (STED) microscopy in a patient-derived Ewing sarcoma cell line to characterize the spatial relationship between endogenous EWS::FLI1 and three functional phosphorylation states of RNAP II: hypophosphorylated RNAP II, serine-5 phosphorylated (Ser5P) RNAP II associated with initiation, and serine-2 phosphorylated (Ser2P) RNAP II associated with elongation. We found that EWS::FLI1 hubs are strongly enriched with Ser5P-RNAP II while remaining largely depleted of hypophosphorylated or Ser2P-RNAP II. We also found that this Ser5P-RNAP II association was largely independent of EWS::FLI1 hub size. Together, these results suggest that endogenous EWS::FLI1 hubs primarily function at the initiation stage of transcription and that nascent hub nucleation events are sufficient to engage with initiating RNAP II, contributing to a mechanistic understanding of how oncogenic TF hubs drive gene expression.

Introduction

Eukaryotic transcription factors (TFs) are typically composed of well-structured DNA-binding domains and intrinsically disordered, transactivating low-complexity domains (LCDs) that often drive the formation of local, high-concentration hubs at target gene loci. Recent studies have shown that these hubs serve to activate their target genes by concentrating other critical transcriptional machinery components including TFs[1,2], RNA polymerase II (RNAP II)[3–7], and other coactivators[3,4,8].

A well-studied model TF is EWS::FLI1, an oncofusion protein that causes Ewing sarcoma and is caused by a t(11;22)(q24;q12) chromosomal translocation that fuses the intrinsically disordered LCD of the RNA-binding protein EWSR1 with the DNA-binding domain of the TF FLI1[9]. EWS::FLI1 forms sub-diffraction-limit hubs at GGAA microsatellites driven by homotypic LCD-LCD interactions; these hubs are necessary for transactivation of target genes[1,10,11]. EWS::FLI1 and other transcriptional players have also been shown to undergo "Goldilocks" behavior where an optimal hub size is needed for efficient transactivation[12].

Despite this progress, it is still unknown how endogenous EWS::FLI1 hubs interface with RNAP II through the various stages of transcription. RNAP II undergoes sequential carboxy-terminal domain (CTD) phosphorylation as it transitions between different functional states; hypophosphorylated RNAP II is present in the nucleoplasm, and at promoters prior to initiation, serine-5 phosphorylated (Ser5P) RNAP II is associated with initiation and early elongation, and serine-2 phosphorylated (Ser2P) RNAP II is associated with productive elongation along the gene body[13]. Prior work using optogenetically induced TAF15 condensates showed that they are enriched with hypophosphorylated RNAP II while Ser5P- and Ser2P-RNAP II accumulated on the droplet periphery, suggesting that these synthetic TF droplets interact preferentially with pre-initiation or early-initiation RNAP II[14]. However, it is unknown whether this relationship holds for endogenous TF hubs largely owing to their sub-diffraction-limit size.

To address this gap, we used dual-color stimulated emission depletion (STED) microscopy[15] to characterize the nanoscale spatial relationship between endogenous EWS::FLI1 and three distinct phosphorylation states of RNAP II in a patient-derived Ewing sarcoma cell line. This allowed us to examine the association of different functional species of RNAP II with EWS::FLI1 hubs at the individual hub level and extract any hub-size-dependence of these associations. Our findings here provide new insight into how endogenous TFs interact with RNAP II at different stages of the transcription cycle.

Results

EWS::FLI1 hubs are specifically enriched with initiation RNA Polymerase II

In order to compare the associations of endogenous EWS::FLI1 and phosphorylation states of RNAP II, we performed dual-color STED microscopy on our previously established knock-in A673 Ewing sarcoma cell line expressing endogenous EWS::FLI1-Halo. We simultaneously labeled EWS::FLI1-Halo and immunostained RNAP II with antibodies targeting either hypophosphorylated RNAP II (hereafter RNAP II CTD for brevity), Ser5P-RNAP II associated with initiation, or Ser2P-RNAP II associated with elongation and found that EWS::FLI1 hubs most strongly colocalized with Ser5P-RNAPII (Figure 1a). EWS::FLI1 colocalization with hypophosphorylated RNAP II and Ser2P was limited. Representative intensity line profiles highlight this preferential association, with EWS::FLI1 and Ser5P-RNAP II intensities appearing very correlated (Figure 1b). In order to quantify this association across the whole nucleus, we computed a pixelwise Pearson's correlation coefficient (PCC)[16], or r, between the EWS::FLI1 and RNAP II intensities across the three RNAP II phosphorylation states (Figure 1c). Confirming what we observed in the line profiles, we found that the PCC between EWS::FLI1 and Ser5P-RNAP II was $r = 0.86$, indicating very strong linear correlation compared to $r = 0.39$ and $r = 0.51$ for EWS::FLI1 vs hypophosphorylated RNAP II and Ser2P-RNAP II, respectively.

In order to more precisely examine the association between EWS::FLI1 hubs and RNAP II puncta rather than overall correlation, we next computed the fractions of EWS::FLI1 hubs that contained an RNAP II punctum for each phosphorylation state as described further in the methods. Briefly, we identified individual EWS::FLI1 hubs by thresholding the EWS::FLI1 channel and identifying clusters of pixels corresponding to hubs, identified RNAP II puncta by finding local maxima, and computed the fraction of EWS::FLI1 hubs that had a corresponding RNAP II punctum within them. We found that, consistent with the correlation analysis, the phosphorylation state of RNAP II most enriched in EWS::FLI1 hubs was Ser5P, with 84.2% of hubs containing RNAP II puncta, significantly higher than both RNAP II CTD (53.1%) and Ser2P (56.8%). Together, these data show that EWS::FLI1 hubs most strongly associate with initiation-associated rather than preinitiation- or elongation-associated RNAP II species.

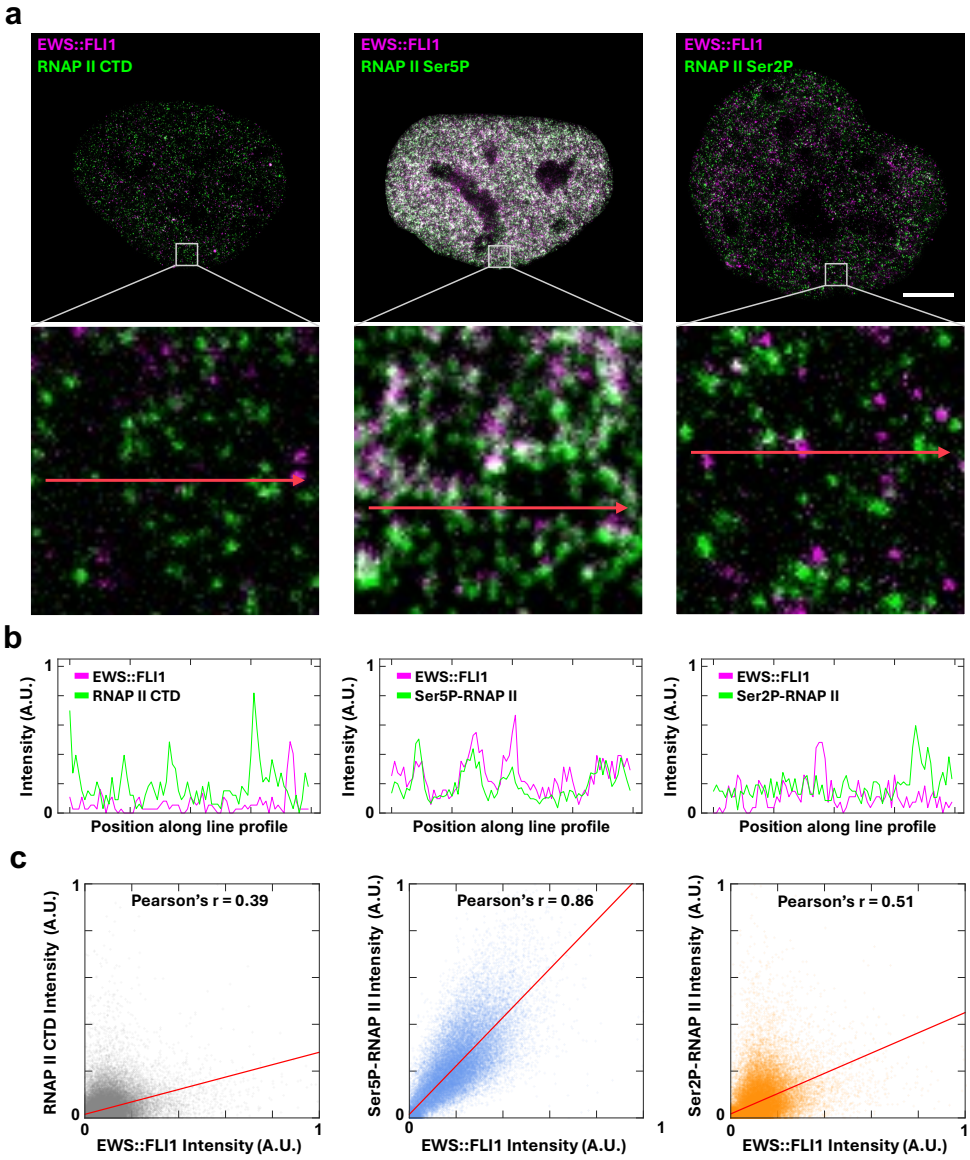

Figure 1: EWS::FLI1 preferentially associates with serine 5 phosphorylated RNAP II. (a) Dual-color STED images of endogenous EWS::FLI1-Halo (magenta) and RNAP II CTD, Ser5P, and Ser2P (green) show that EWS::FLI1 most strongly colocalizes with Ser5P-RNAP II. Scale bar is 5 µm. (b) Representative intensity line profiles highlights this preferential association between EWS::FLI1 and Ser5P-RNAP II. (c) Plots of EWS::FLI1 vs RNAP II intensity per pixel and Pearson's correlation coefficient, r, show

that EWS::FLI1 and Ser5P-RNAP II intensity are very strongly correlated. Gaussian jitter ($\sigma = 0.02$) was added to aid in point-density visualization of discrete data.

Ser5P-RNAP II recruitment to EWS::FLI1 hubs is size-independent

We next asked whether the association between EWS::FLI1 hubs and Ser5P-RNAP II was dependent on hub size, which could imply a concentration-dependent interaction or recruitment mechanism. To investigate this, we used the aforedescribed EWS::FLI1 hub identification method and this time quantified the Ser5P-RNAP II signal intensity within each EWS::FLI1 hub. We then plotted Ser5P-RNAP II intensity as a function of EWS::FLI1 hub diameter and found no apparent size-dependence, with the mean RNAP II intensity remaining roughly constant across the detected range of hub diameters (Figure 2). This suggests that size-dependent RNAP II recruitment to hubs is likely not the mechanism underlying the "Goldilocks" principle observed in Chong et al.[12] and shows that even the smallest detectable EWS::FLI1 hubs can associate with initiating RNAP II.

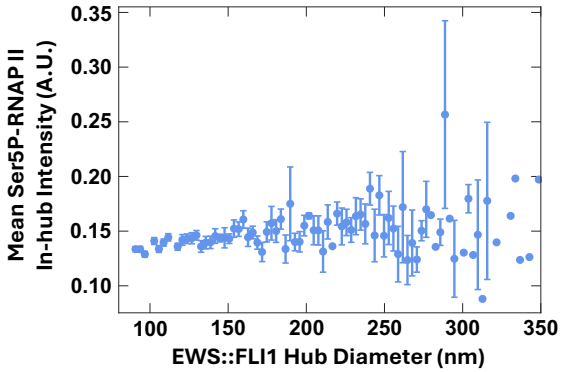

Figure 2: RNAP II Ser5P recruitment to EWS::FLI1 hubs is not hub-size dependent. (a) Plot comparing the Ser5P-RNAP II intensity corresponding to EWS::FLI1 hubs of varying size shows no obvious maximum. Data taken from n = 8 cells and error bars are SEM.

Discussion

Our results provide a high-resolution view into the functional relationship between endogenous EWS::FLI1 and the various phosphorylation states of RNAP II involved in the transcription cycle. Using dual-color STED, we found that EWS::FLI1 strongly associates with Ser5P-RNAP II rather than unphosphorylated RNAP II or Ser2P-RNAP II. This is consistent with a model where EWS::FLI1 hubs activate target genes by concentrating initiating RNAP II at the gene locus.

This behavior of EWS::FLI1 hubs contrasts with that of TAF15 condensates, which were shown to preferentially associate with hypophosphorylated RNAP II, while Ser5P- and Ser2P-RNAP II formed concentric rings at the droplet periphery[14]. The difference in preferred interacting RNAP II species despite both EWS::FLI1 hubs and TAF15 droplets being driven by FET-family LCDs is likely due to differences in the flavor of condensates formed by the two proteins owing to their unique sequence features and also may have been impacted by differences in expression level with EWS::FLI1 studied at endogenous levels here and TAF15 studied under overexpression in Wei et al.[14] Despite these differences, both results are consistent with a more general model where TFs recruit RNAP II before productive elongation whereby they are released from hubs.

The finding that Ser5P-RNAP II recruitment to EWS::FLI1 hubs is not size-dependent also carries significant mechanistic implications. Were recruitment dependent on exceeding a critical hub size or generally concentration dependent, we would expect to see RNAP II association increase with larger hubs. The size-independent distribution we observe instead suggests that even the smallest detected EWS::FLI1 hubs were sufficient to recruit RNAP II, supporting a model where nascent hub nucleation events, rather than large, stable droplets, are sufficient to drive transcription. This is consistent with our previous findings that the EWS::FLI1 hub population is dynamically maintained through constant nucleation and dissolution of hubs rather than stable growth. It also suggests that RNAP II recruitment is not the mechanism underlying the Goldilocks principle observed in Chong et al.[12], who

showed that EWS::FLI1 transactivation requires a finely tuned optimum of LCD-LCD interaction levels.

Together, our findings provide a clearer picture of how EWS::FLI1 hubs interact with RNAP II at endogenous expression levels. The preferential association of EWS::FLI1 with Ser5P-RNAP II, combined with the size-independent nature of this interaction, supports a model where EWS::FLI1 hubs serve to concentrate pre-elongation transcriptional machinery at target genes, consistent with their established role in driving the aberrant transcriptional program characteristic of Ewing sarcoma. Further investigations elucidating the nature of the interaction driving RNAP II recruitment to EWS::FLI1 hubs may help guide targeted therapeutic strategies aimed at disrupting this association.

Methods

Cell Culture:

The knock-in A673 Ewing sarcoma cell line expressing endogenous EWS::FLI1-Halo originally from Chong et al.[1] was cultured in 4.5 g/L high-glucose DMEM (Thermo Fisher, 10566016) supplemented with 10% FBS (Fisher Scientific, SH3039603) and 1% penicillin-streptomycin (Thermo Fisher, 15140122) in a 37°C, humidity-controlled incubator under 5% CO_2.

Immunostaining:

Cells were cultured on 18 mm #1.5 coverslips (Electron Microscopy Sciences, 7222201) coverslips cleaned by immersion in 1 M KOH with sonication for 1 hr, three rinses with double-distilled water, immersion in 200 proof ethanol with sonication for 1 hr, then immersion in 20% ethanol diluted with double-distilled water for storage. EWS::FLI1 was stained with JFX646 HaloTag ligand for 1 hr, followed by three, 10-minute washes (each wash entails aspiration of media, two rinses with 1X PBS, and fresh media). Cells were then fixed with 4% paraformaldehyde (VWR, 100503-917), rinsed thrice with 1X PBS, permeabilized with 0.5 ml 2.5% Triton X-100 (Sigma-Aldrich, X100-1L) in PBS for 3

minutes on a rocker, washed twice more with 1X PBS for 10 minutes per wash, and blocked with 5% BSA (Merck Millipore, 12659-100GM) in PBS for 20 minutes on a rocker. Cells were then incubated with primary antibody targeting the RNAP II CTD (Abcam, ab26721), RNAP II Ser5P (Abcam, ab5408), or RNAP II Ser2P (Abcam, ab193468) with 0.1% Tween 20 (VWR, M147-1L), 3% BSA in PBS at room temperature for 1 hr in the dark. While the RNAP II CTD antibody is designed to target all CTD regardless of phosphorylation state, it has been shown to strongly preferentially target unphosphorylated CTD[17]. Cells were then washed twice with 1X PBS, 5 minutes per wash in the dark then incubated in a 1:200 dilution of secondary antibody in 0.1% Tween 20, 3% BSA in PBS at room temperature for 1 hr in the dark. For secondary antibodies, Goat Anti-rabbit Abberior STAR ORANGE (Abberior, STORANGE-1002-500UG) was used for RNAP II CTD and RNAP II Ser2P and Goat Anti-mouse Abberior STAR ORANGE (Abberior, STORANGE-1001-500UG) was used for RNAP II Ser5P. Samples were then washed thrice with 1X PBS with 5 minutes per rinse. Fiducial markers for channel alignment were added to the coverslip by adding 0.1 μm yellow-green Invitrogen Fluospheres (Thermo Fisher, F8803) diluted in DPBS supplemented with calcium and magnesium (Thermo Fisher, 14-040-141) and allowing the beads to adsorb to the coverslip for 10 minutes at room temperature. Samples were then mounted to frosted glass slides (Sail Brand, 7105) with Abberior mount solid antifade (Abberior, MM-2013-2X15ML) and allowed to cure for 24 hrs before imaging.

STED:

STED was performed on an Abberior STEDYCON unit installed on an inverted Zeiss microscope with a Zeiss x100/NA 1.46 oil-immersion objective. A 561 nm laser was used to excite Abberior STAR ORANGE, and a 640 nm laser was used to excite JFX646, both using a pulsed 775 nm depletion beam for STED. Images were acquired with a 25 nm pixel size.

STED Image Analysis:

EWS::FLI1 hubs were first segmented by thresholding the at 95% of non-zero pixel intensities and only retaining connected pixels with areas greater than 10 pixels. Regions over 300 pixels likely corresponding to multiple hubs merged due to resolution limits were

binary eroded thrice with a 1 pixel radius and treated as independent hubs if multiple features emerged. To determine if there was a corresponding RNAP II punctum inside the hub, colocalization was tested between EWS::FLI1 hub masks with 1 pixel binary dilation and RNAP II punctum local maxima, which were detected by local maxima search with a minimum peak separation of 2 pixel and an intensity threshold of greater than 3 standard deviations above the mean nuclear intensity. For each EWS::FLI1 hub, area, diameter, integrated RNAP II intensity, and mean RNAP II intensity were recorded. The plot of EWS::FLI1 size vs mean RNAP II intensity was generated using binned EWS::FLI1 hub diameters with 3 nm bins. Pearson's r was used as a metric for colocalization and was computed on robust min-max normalized images clipped at the 99.99 percentile to account for bright pixel outliers. Scramble controls were performed by dividing the EWS::FLI1 channel into 10 by 10 pixel blocks (excluding fully non-nuclear blocks), randomly permuting them, and recomputing their correlation coefficients.

References

1. Chong, S. *et al.* Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555 (2018).
2. Cai, D. *et al.* Phase separation of YAP reorganizes genome topology for long-term YAP target gene expression. *Nat. Cell Biol.* **21**, 1578–1589 (2019).
3. Lyons, H. *et al.* Functional partitioning of transcriptional regulators by patterned charge blocks. *Cell* **186**, 327-345.e28 (2023).
4. Cho, W.-K. *et al.* Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).
5. Guo, Y. E. *et al.* Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* **572**, 543–548 (2019).
6. Boehning, M. *et al.* RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.* **25**, 833–840 (2018).
7. Lyons, H. *et al.* RNA polymerase II partitioning is a shared feature of diverse oncofusion condensates. *Cell* **188**, 3843-3862.e28 (2025).

8. Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
9. Li, Y., Luo, H., Liu, T., Zacksenhaus, E. & Ben-David, Y. The ets transcription factor Fli-1 in development, cancer and disease. *Oncogene* **34**, 2022–2031 (2015).
10. Johnson, K. M. *et al.* Role for the EWS domain of EWS/FLI in binding GGAA-microsatellites required for Ewing sarcoma anchorage independent growth. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 9870–9875 (2017).
11. Boulay, G. *et al.* Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain. *Cell* **171**, 163-178.e19 (2017).
12. Chong, S. *et al.* Tuning levels of low-complexity domain interactions to modulate endogenous oncogenic transcription. *Mol. Cell* **82**, 2084-2097.e5 (2022).
13. Komarnitsky, P., Cho, E.-J. & Buratowski, S. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* **14**, 2452–2460 (2000).
14. Wei, M.-T. *et al.* Nucleated transcriptional condensates amplify gene expression. *Nat. Cell Biol.* **22**, 1187–1196 (2020).
15. Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19**, 780–782 (1994).
16. Dunn, K. W., Kamocka, M. M. & McDonald, J. H. A practical guide to evaluating colocalization in biological microscopy. *Am. J. Physiol. - Cell Physiol.* **300**, C723–C742 (2011).
17. Anastasakis, D. G. *et al.* Nuclear PKM2 binds pre-mRNA at folded G-quadruplexes and reveals their gene regulatory role. *Mol. Cell* **84**, 3775-3789.e6 (2024).

Chapter 6

PROGRESS AND CHALLENGES WITH SINGLE-MOLECULE COUNTING USING PHOTOACTIVATED LOCALIZATION MICROSCOPY

Abstract

The accurate quantification of the absolute number or stoichiometry of biomolecules in their native cellular contexts remains a challenge in biological microscopy. While fluorescence intensity- and stepwise photobleaching-based approaches are increasingly commonly employed, these methods are limited in resolution due to the diffraction limit and in accuracy due to photophysical and optical effects that lead to nonlinear scaling between copy number and fluorescence intensity. Photoactivated localization microscopy (PALM)-based molecular counting strategies are a potential solution that temporally separate individual fluorophores' emissions, thus enabling single-molecule localization and, in principle, absolute molecular counting. However, such methods nearly universally make the tacit assumption that the emitter photophysics are invariant to local emitter concentration. Here, we sought to establish PALM-based absolute molecular counting with photoactivatable Janelia Fluor dye, PA-JF549, and benchmark the method by labeling well-established standard candle oligomers, FTH1, a 24-mer, and I3-01, a 60-mer. However, in the process, we found very strong coupling of the photophysical behaviors of proximal PA-JF549 emitters tethered to the same oligomer, where rather than exhibiting temporally dispersed emissions throughout the duration of the acquisition, fluorophores belonging to the same oligomer appeared to synchronize, nearly all activating and emitting within a narrow time window. The generality of the finding that densely packed PA-JF549 can exhibit highly correlated photophysical behavior should be investigated and introduces a significant caveat in the application of PALM-based molecular counting to densely labeled biological assemblies. More broadly, our findings suggest that PALM-based molecular counting approaches are not inherently immune to the proximity-based photophysical effects that complicate conventional confocal- or widefield-based approaches.

Introduction

The quantification of the absolute number or stoichiometry of biomolecules in their native cellular environments is critical for the development of quantitative models of biological processes; despite this, it remains a significant challenge in biological microscopy. Perhaps the most widely applied standard involves using multiple standard candles, oligomers with a known number of fluorescently tagged subunits, to generate a calibration curve that can be used to infer the number of biomolecules from fluorescence intensity measurements. While generally suitable and relatively easily implemented, it still faces a number of limitations exacerbated in the study in biomolecular condensates and other relatively dense cellular assemblies, namely that fluorescence intensity often scales nonlinearly with molecular copy number due to a number of phenomena including self-quenching, out-of-focus fluorescence, concentration-dependent changes in emitter photophysics, and more. In combination, the imaging itself if diffraction limited, thus, it generally cannot readily distinguish between increased fluorescence due to true stoichiometry or sub-diffraction-limit clustering resulting in modified photophysics. Other more robust methods exist that rely on inferring copy number from stepwise photobleaching traces[1], but these methods are typically extremely low throughput, requiring the measurement of assemblies one-by-one. More generally, these methods are all diffraction-limited, limiting their applicability in the study of many biologically relevant nanoscale assemblies.

Photoactivated localization microscopy (PALM) lends itself as a natural solution to this application by temporally separating fluorescence emissions from different fluorophores, allowing them to be identified as discrete molecules[2]. However, extracting absolute copy number from PALM localizations requires the correction of two broad classes of phenomena: undercounting and overcounting. Undercounting has a number of causes depending on the fluorophore used but is generally due to incorrect or incomplete folding of fluorescence proteins, or limited ligand binding efficiency for self-labeling systems like HaloTag and SNAP-tag[3,4]. This has been well studied in prior work[5] and in the study of biomolecular condensates, and can be accounted for when measuring copy number in sufficiently large condensates simply by measuring an effective detection efficiency and applying its inverse

as a correction. Overcounting is the appearance of multiple localizations all corresponding to a single emitter and is largely due to phenomena like frame discretization (the splitting of a single emission event across multiple frames), variable photobleaching times, and photoblinking (single emitters transitioning randomly between emissive and temporary dark states before permanently photobleaching)[6–12]. There are a number of methods extensively discussed elsewhere that can be used to correct for these effects, but they almost all universally rely on the assumption that the photophysics for individual molecules are independent such that the stochastic blinking-filtering models, including their fluorophore-specific parameterizations extracted from sparse emitters, can be readily applied to densely labeled assemblies. In other words, these models are almost always based on the photophysics of an individual emitter and are generally verified on experimental systems where the emitter density is sparse enough such that individual emitters are both spatially and temporally distinct. They are then applied to biological systems where emitters may be more clustered, with the tacit assumption that the photophysics of the emitters remains independent even when they are in close proximity with one another.

Recent advances in highly photostable fluorophores, including photoactivatable Janelia Fluors like PA-JF549 and PA-JF646[13,] have enabled increasingly quantitative analyses of nanoscale biomolecular condensates and allowed for increasingly precise modeling of their physical underpinnings[14]. PA-JF549 in particular features very limited blinking that also takes place on a very short timescale, making it very ideal for PALM-based molecular counting measurements[5,14]. Despite this, we are not aware of it being used for absolute molecular counting, and its photophysical behaviors in dense molecular assemblies have not been characterized. Because many biological assemblies of interest contain proteins at high local concentrations, deviations from independent blinking behavior could fundamentally limit its application to molecular counting.

Here, we initially aimed to establish PALM-based absolute molecular counting using the well-behaved PA-JF549 and benchmark its performance against standard candles but instead identified evidence for coupled fluorophore photophysics in dense assemblies and demonstrated an important caveat for PALM-based molecular counting.

Results

In order to benchmark our PALM-based molecular counting, we generated two standard candle oligomers with known stoichiometries: HaloTag-SNAPtag-FTH1 based on a ferritin heavy-chain 24-mer[15], and HaloTag-SNAPtag-I301 based on a hyperstable 60-mer nanocage designed by the Baker Lab[16]. The SNAP-tag was added for intended future benchmarking of stoichiometry; for the remainder of this study, we focused on using HaloTag.

Before performing PALM, we first verified successful oligomerization of both constructs by labeling the constructs with JFX549 HaloTag ligand and using stimulated emission depletion (STED) microscopy[17]. Imaging revealed successful and high-efficiency assembly of both the FTH1-based 24-mer (Figure 1a) and I301-based 60-mer (Figure 1b).

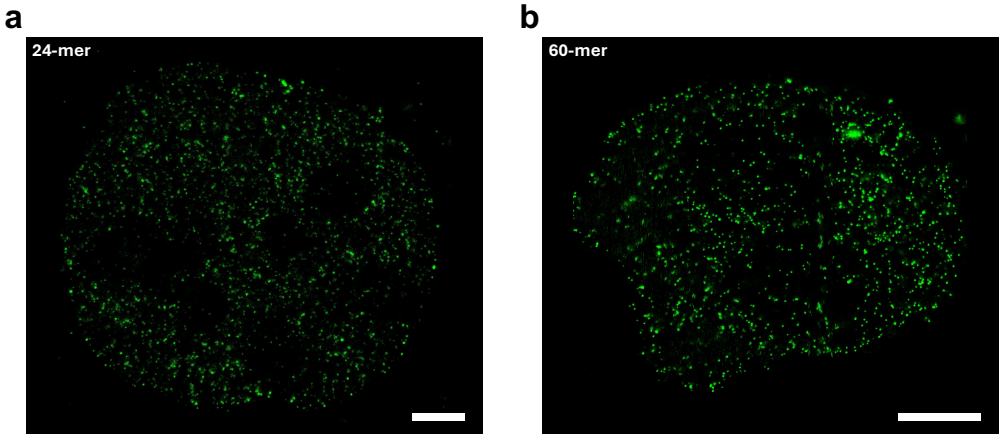

Figure 1: Fluorescently tagged 24-mer and 60-mer assemblies both successfully assemble in U2OS. (a) STED microscopy shows that JFX549 HaloTag-ligand-stained (a) HaloTag-SNAPtag-FTH1 and (b) HaloTag-SNAPtag-I3-01 assemble correctly. Scale bar is 5 μm.

We next labeled both constructs with PA-JF549 HaloTag ligand and performed PALM. However, inspection of localization time-series, plots of frames at which emitters were localized, revealed that PA-JF549 behavior was highly coupled in the oligomers. In previous

work studying endogenous EWS::FLI1 hubs, we found that the assumption critical to blinking correction, that emitter photophysics remained independent, was largely preserved even in hubs. This is evident upon examination of a representative localization time-series for an individual hub; localization events take place throughout the duration of the acquisition, and the waiting-time distribution between localization events was approximately exponential, consistent with independent stochastic activation behavior, enabling accurate inference (Figure 2a).

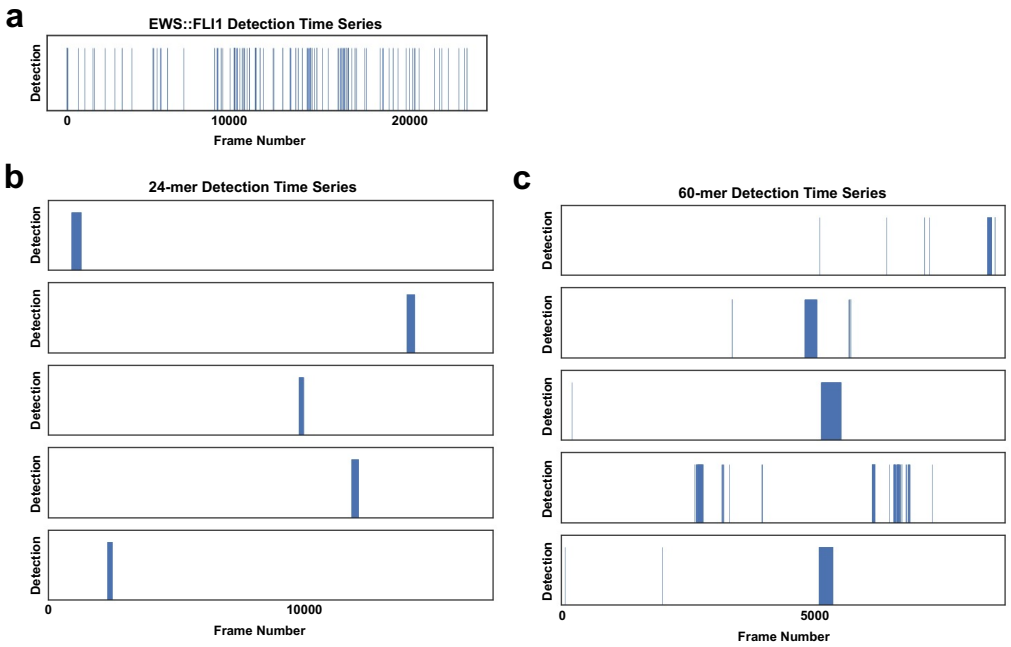

Figure 2: Dense oligomeric labeled exhibit synchronized PALM emission behavior. (a) A representative time series of localization events from a single PA-JF549-labeled EWS::FLI1 hub with shows apparently independent individual localization events distributed throughout the acquisition. (b) Multiple representative time series of localization events from a single PA-JF549-labeled (b) 24-mer FTH1 assembly or (c) 60-mer I3-01 assembly. Localization events occur in one or a few long, dense temporal bursts rather than occurring randomly throughout the acquisition showing strongly coupled photophysical behavior between proximal molecules.

In contrast, the oligomers both displayed significantly different temporal localization distributions, with fluorophores labeling the 24-mers all appearing to photoactivate within a narrow time window, resulting in single dense bursts of localizations rather than being approximately evenly distributed throughout the acquisition (Figure 2b). The 60-mers exhibited similar behavior, albeit with a few fluorophores appearing to have remained independently photoactivated (Figure 2c). These observations suggest that the PA-JF549 fluorophores can experience strongly coupled photophysics when confined to extremely compact assemblies like the 24-mer and 60-mer, which are both substantially more densely labeled than EWS::FLI1 hubs. Under these conditions, the assumption of independent emitter photophysics is violated, and standard PALM-based molecular counting methods cannot be used.

Discussion

Accurate PALM-based absolute molecular counting requires that emitter photophysics remain independent even in the denser assemblies present in biological systems. Here, using compact oligomeric standard candles, we found that densely packed PA-JF549 molecules can exhibit highly correlated photophysical behavior incompatible with modern PALM-based copy number inference strategies. These synchronized photoactivation/emission behaviors were not observed in less dense EWS::FLI1 hubs, where localization events remained broadly distributed throughout the acquisition, suggesting that it is the close proximity of multiple PA-JF549 molecules that triggers this photophysical coupling. Although the physical/chemical mechanism underlying this coupling remains unclear, it seems likely due to some local energy transfer effects, collective environmental effects, or some other proximity-dependent interaction.

Our findings carry significant implications for PALM-based molecular counting methods; namely, the assumption that the kinetics of fluorophore activation, emission, and blinking measured in sparse systems remain invariant under crowding should be directly verified for the relevant emitter using our oligomers or others before performing copy number inference. More broadly, these results emphasize that PALM imaging and photoactivatable dye-based

molecular counting approaches are not inherently immune to the proximity-dependent photophysical effects that complicate traditional confocal/widefield-based molecular counting measurements.

Methods

Cell Culture:

U2OS cells were cultured in 1 g/L low-glucose DMEM (Thermo Fisher, 10567014) supplemented with 10% FBS (Fisher Scientific, SH3039603) and 1% penicillin-streptomycin (Thermo Fisher, 15140122) at 37°C in a humidity-controlled incubator with 5% CO_2. Cells were cultured on 25 mm #1.5H coverslips (Marienfeld Superior, 0117650) that were cleaned as follows: sonication in 1 M KOH for 1 hr, three rinses with double-distilled water, sonication in 200 proof ethanol for 1 hr, and immersion in 20% ethanol diluted in double-distilled water for storage. Cells were transfected with either the HaloTag-SNAPtag-FTH1 or the HaloTag-SNAPtag-I301 construct 24 hrs before imaging using Lipofectamine 3000 (Fisher Scientific, L3000001) following the manufacturer's protocol.

For STED, cells were stained for 30 min with 500 nM JFX549 HaloTag ligand and 10 μM JFX646 SNAP-tag ligand and rinsed twice. Each rinse entailed aspiration of media, three washes with ligand-free media, and incubation with 1 g/L low-glucose DMEM supplemented with 10% FBS and 1% penicillin-streptomycin followed by incubation for 15 minutes. Cells were then rinsed twice with 1X PBS and fixed using 4% paraformaldehyde (VWR, 100503-917) for 10 minutes, then rinsed thrice with 1X PBS for 5 min (Thermo Fisher, 18912014). Coverslips were then transferred into Attofluor Cell Chambers (Thermo Fisher, A7816) and 0.1 μm yellow-green Invitrogen Fluospheres (Thermo Fisher, F8803) to be used as channel-alignment fiducial markers were diluted in DPBS with calcium and magnesium (Thermo Fisher, 14-040-141), and added to the coverslips for 10 minutes at room temperature to allow the beads to adsorb to the coverslip.

For PALM, cells were then stained for 1 hr with 500 nM PA-JF549 HaloTag ligand and/or 10 μM PA-JF646 SNAP-tag ligand and rinsed four times. Each rinse entailed aspiration of

media, three washes with ligand-free media, and incubation with 1 g/L low-glucose DMEM supplemented with 20% FBS rather than 10% and 1% penicillin-streptomycin followed by incubation for 15 minutes. Cells were then rinsed four times with 1X PBS and fixed using 4% paraformaldehyde (VWR, 100503-917) supplemented with 0.2% glutaraldehyde (Sigma Aldrich, 340855-25ML) for 15 minutes, then rinsed thrice with 1X PBS for 5 min (Thermo Fisher, 18912014). Coverslips were then transferred into Attofluor Cell Chambers (Thermo Fisher, A7816) and 0.1 μm yellow-green Invitrogen Fluospheres (Thermo Fisher, F8803) to be used as drift-correcting fiducial markers were diluted in DPBS with calcium and magnesium (Thermo Fisher, 14-040-141), and added to the coverslips for 10 minutes at room temperature to allow the beads to adsorb to the coverslip.

STED:

STED was performed on an Abberior STEDYCON unit with an inverted Zeiss microscope body using a Zeiss x100/NA 1.46 oil-immersion objective. 561 nm and 640 nm beams were used to excite JFX549 and JFX646, respectively, with both using a pulsed 775 nm depletion beam for STED. Images were acquired with a 25 nm pixel size.

PALM:

PALM was performed on a Nikon Eclipse Ti2 N-STORM system with a x100/NA 1.49 oil-immersion objective (Nikon, CFI SR HP Apochromat TIRF 100XAC Oil, MRD01995). Samples were illuminated under a highly inclined and laminated optical sheet (HILO)[18] using a 405 nm beam for photoactivation and 561 nm and 640 nm beams to excite PA-JF549 and PA-JF646, respectively. Emission fluorescence was split using a 635 nm single-edge dichroic beamsplitter (Semrock, Di02-R635-25x36) and filtered through 593/40 nm single-band bandpass filter (Semrock, FF01-593/40-25) for PA-JF549 or 676/37 nm single-band bandpass filter (Semrock, FF01-676/37-25) for PA-JF646 before being detected on a pair of iXon Ultra 897 EMCCD (Andor). NIS Elements software was used to control the microscope. Images were analyzed as discussed in the methods of Yoshida et al.[19]

Acknowledgements

We thank Dr. David Glover for generously allowing us to use his lab's STED microscope.

References

1. Tsekouras, K., Custer, T. C., Jashnsaz, H., Walter, N. G. & Pressé, S. A novel method to accurately locate and count large numbers of steps by photobleaching. *Mol. Biol. Cell* **27**, 3601–3615 (2016).
2. Betzig, E. *et al.* Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science* **313**, 1642–1645 (2006).
3. Annibale, P., Scarselli, M., Greco, M. & Radenovic, A. Identification of the factors affecting co-localization precision for quantitative multicolor localization microscopy. *Opt. Nanoscopy* **1**, 9 (2012).
4. Durisic, N., Laparra-Cuervo, L., Sandoval-Álvarez, Á., Borbely, J. S. & Lakadamyali, M. Single-molecule evaluation of fluorescent protein photoactivation efficiency using an in vivo nanotemplate. *Nat. Methods* **11**, 156–162 (2014).
5. Thevathasan, J. V. *et al.* Nuclear pores as versatile reference standards for quantitative superresolution microscopy. *Nat. Methods* **16**, 1045–1053 (2019).
6. Platzer, R. *et al.* Unscrambling fluorophore blinking for comprehensive cluster detection via photoactivated localization microscopy. *Nat. Commun.* **11**, 4993 (2020).
7. Annibale, P., Vanni, S., Scarselli, M., Rothlisberger, U. & Radenovic, A. Identification of clustering artifacts in photoactivated localization microscopy. *Nat. Methods* **8**, 527–528 (2011).
8. Annibale, P., Vanni, S., Scarselli, M., Rothlisberger, U. & Radenovic, A. Quantitative photo activated localization microscopy: unraveling the effects of photoblinking. *PloS One* **6**, e22678 (2011).
9. Lee, S.-H., Shin, J. Y., Lee, A. & Bustamante, C. Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM). *Proc. Natl. Acad. Sci. U. S. A.* **109**, 17436–17441 (2012).
10. Veatch, S. L. *et al.* Correlation functions quantify super-resolution images and estimate apparent clustering due to over-counting. *PloS One* **7**, e31457 (2012).

11. Sengupta, P. *et al.* Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. *Nat. Methods* **8**, 969–975 (2011).
12. Annibale, P., Scarselli, M., Kodiyan, A. & Radenovic, A. Photoactivatable Fluorescent Protein mEos2 Displays Repeated Photoactivation after a Long-Lived Dark State in the Red Photoconverted Form. *J. Phys. Chem. Lett.* **1**, 1506–1510 (2010).
13. Grimm, J. B. *et al.* Bright photoactivatable fluorophores for single-molecule imaging. *Nat. Methods* **13**, 985–988 (2016).
14. Yoshida, S., Zhong, Y., Kumagai, A., Dunphy, W. G. & Chong, S. Dynamic regulation of endogenous transcription factor hubs at single-molecule resolution. 2026.01.25.701606 Preprint at https://doi.org/10.64898/2026.01.25.701606 (2026).
15. Hoelzgen, F. *et al.* Structural basis for the intracellular regulation of ferritin degradation. *Nat. Commun.* **15**, 3802 (2024).
16. Hsia, Y. *et al.* Design of a hyperstable 60-subunit protein dodecahedron. [corrected]. *Nature* **535**, 136–139 (2016).
17. Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19**, 780–782 (1994).
18. Tokunaga, M., Imamoto, N. & Sakata-Sogawa, K. Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nat. Methods* **5**, 159–161 (2008).
19. Yoshida, S., Zhong, Y., Kumagai, A., Dunphy, W. G. & Chong, S. Dynamic regulation of endogenous transcription factor hubs at single-molecule resolution. 2026.01.25.701606 Preprint at https://doi.org/10.64898/2026.01.25.701606 (2026).

Chapter 7

EMERGENT PRINCIPLES OF TRANSCRIPTION FACTOR HUB BIOLOGY FROM SINGLE-MOLECULE IMAGING

Abstract

Eukaryotic gene expression depends on the ability of transcription factors (TFs) to locate and bind their targets, and locally concentrate the molecular machinery needed for transcription. Many TFs form local, high-concentration hubs via multivalent low-complexity domain (LCD) interactions that are essential for the transactivation of target genes. TF hubs are of particular interest in modern transcriptional biology because they offer a physical explanation for how dilute TFs rapidly bind their target sites and achieve sufficient local concentration to activate transcription. Despite this, the mechanisms underlying their formation and regulation remained largely unknown. Here, we synthesize the novel findings from quantitative single-molecule imaging studies of the oncofusion TF EWS::FLI1 at endogenous levels. EWS::FLI1 forms sub-diffraction-limit hubs distinct from bona fide liquid-liquid phase-separated (LLPS) droplets with energetics that imply an active dissolution mechanism. That hub formation ability is a neomorphic, gain-of-function property of the fusion that is not directly conferred by EWSR1 or FLI1, its parental proteins. In addition, physical properties of the hubs, including their spatial distribution and dynamics, can be tuned by numerous regulatory axes, including during mitosis, by RNA, and by post-translational modification by O-GlcNAc, with functional impacts on EWS::FLI1-mediated transcription. Interestingly, EWS::FLI1 appears to preferentially associate with initiating, Serine 5-phosphorylated RNA polymerase II (RNAP II) in a hub size-independent manner. Finally, endogenous EWS::FLI1 hubs in live Ewing sarcoma cells were shown to be dissolved and mislocalized by small molecules, resulting in suppressed transcription, establishing the validity of perturbing pathological TF hubs as a therapeutic strategy. Together, these findings paint a picture of EWS::FLI1 hubs as dynamic assemblies poised in a regime where they are highly and uniquely sensitive to regulation during mitosis, by RNA, and by O-GlcNAcylation, and that while able to associate with initiating RNAP II across a range of sizes, remain sensitive to pharmacological perturbation. Further, our studies

introduce a CRISPR and single-molecule imaging-based platform that can be applied more broadly to the study of other oncofusion proteins to probe the generality of our findings.

Introduction

In eukaryotes, DNA-binding transcription factors (TFs) can be thought of as being composed of well-structured DNA-binding domains (DBDs) and intrinsically disordered low-complexity domains (LCDs) that are responsible for transactivation. Many of these TFs undergo dynamic, multivalent LCD-LCD interactions that drive the formation of local, high-concentration hubs at target gene loci[1,2]. These hubs can act as *de novo* enhancers that recruit transcriptional coactivators, and optimal levels of LCD-LCD interactions have been shown to maximally drive the transcription of downstream genes[2,3]. Further, mutations in the LCDs of TFs have been heavily implicated in a number of diseases, particularly cancers including Ewing sarcoma[4] and rhabdomyosarcoma[5]. A commonly studied TF in this context is the chimeric oncoprotein EWS::FLI1, which causes Ewing sarcoma[4]. It is caused by a t(11;22)(q24;q12) chromosomal translocation that fuses the well-structured DNA-binding domains of FLI1 with the intrinsically disordered LCD of EWSR1, resulting in a neomorphic fusion TF that preferentially binds to GGAA microsatellites where it forms hubs that can drive downstream transcription[6].

Despite the biological and clinical significance of TF hubs, their biophysical properties and the mechanisms underlying their formation and regulation, particularly under endogenous conditions, remain relatively unexplored. This is due to a number of limitations, including the use of overexpression of fluorescently tagged TFs to represent physiological conditions, when in fact hub formation is a highly concentration-dependent phenomenon[7]; the use of immunofluorescence and other fixation-based methods without controls to ensure the absence of fixation artifacts[8,9]; and resolution limits imposed by the diffraction-limited nature of widely available fluorescence imaging modalities like widefield and confocal imaging. Here, we synthesize findings from our recent studies that apply live-cell, single-molecule, and super-resolution imaging methods to characterize hubs formed by endogenously labeled EWS::FLI1.

Discussion

The physical properties of EWS::FLI1 hubs

To begin dissecting the physical properties of TF hubs, their dimensions must be accurately measured, as their size and morphology have direct physical implications on how they can plausibly affect transcription; for example, hub sizes place explicit constraints on the number of molecules they can concentrate at a given genomic locus as well as the degree to which the local material properties or chemical environment can be altered. Characterization of hub dimensions is traditionally limited by a reliance on overexpression, which can drastically alter the phase behavior of a protein, and the diffraction limit, which makes the accurate measurement of hubs with diameters lower than a few hundred nanometers difficult. Our recent super-resolution and single-molecule imaging studies of endogenously labeled EWS::FLI1 hubs simultaneously overcame both limitations and revealed that on average, hubs have diameters around 50 – 100 nm[10]. This clearly differentiates them from larger, micron-scale droplets of various coactivators observed at super-enhancers[11].

In addition to establishing the length scale for the system, characterization of the distribution of hub sizes has the useful side effect of pointing to a specific physical mechanism underlying hub formation. While the extensive debates over nomenclature in the biomolecular condensate field may seem largely semantic, the rationale for the discourse is anything but; each name for a local, high-concentration region, be it biomolecular condensate, droplet, hub, punctum, assembly, etc, without clarification, may naturally carry with it an implied physical mechanism. Further, distinguishing between the flavors of biomolecular condensates (here, using it as a mechanism-agnostic term), including complex liquid-liquid phase-separated droplets, droplets formed by phase-separation coupled to percolation, subsaturated clusters, etc., is crucial, as each term implies a set of physical properties that may otherwise be assumed without sufficient evidence[12].

Notably, many of the physical processes underlying the formation of biomolecular condensates have an associated (and often distinct) size distribution that can be used to empirically distinguish them. Many of the condensate size distributions also feature

characteristic time evolution, further enabling identification. In the case of EWS::FLI1 hubs, their size distributions were well fit within the framework of classical nucleation theory (CNT)[10], which describes the initial steps of the formation of a new phase – not necessarily liquid – to great success in biology[13,14]. In this framework, monomers stochastically associate and dissociate from subcritical clusters. Once the subcritical clusters exceed a critical size, they begin to grow effectively irreversibly, taking us into the domain of a newly formed phase[15]. EWS::FLI1 hubs have been shown to have a positive chemical potential difference, showing that EWS::FLI1 is supersaturated under endogenous conditions, and feature a modest free energy barrier to nucleation, which is consistent with frequent nucleation events[10]. Further, the nature and energetics of nucleation indicated the presence of a dissolution mechanism for supercritical hubs that can act on the order of ~34 hubs per second, pointing to an active rather than passive mechanism, though the underlying process remains elusive[10]. This framework for identifying physical properties by their size distribution can be used to complement existing methods for other condensate-forming proteins.

Hub formation is an emergent, neomorphic property of the EWS::FLI1 fusion and a therapeutic vulnerability

A separate but related question is related questions is from where the hub formation ability of EWS::FLI1 is derived. Our recent super-resolution study showed that of EWS::FLI1, EWSR1, and FLI1, EWS::FLI1 formed significantly greater and more prominent hubs, compared to EWSR1 and FLI1[10]. This shows that the hub formation ability of EWS::FLI1 is a neomorphic gain-of-function property and implies that EWS::FLI1 combines two activities that are each present in only one parental protein; i.e., by combining the LCD-LCD interaction-mediated self-association ability of EWSR1 with the sequence-specific DNA-binding ability of FLI1, EWS::FLI1 is able to accumulate at GGAA microsatellites throughout the genome, where further LCD-LCD interactions drive the formation of bona fide hubs.

This finding itself carries two implications. First, disruption of EWS::FLI1 hubs has been shown to prevent oncogenic transformation[1], and because hub formation is a property of

EWS::FLI1 and not its parental proteins, it should only exist in cancerous cells harboring the translocation; thus, therapeutically targeting hub formation, rather than a fragment of the protein sequence, enables an avenue to selectively target cancer cells. This, in addition to the notoriously undruggable nature of TFs, highlights the potential for TF hubs as therapeutic targets. Second, it more generally points to a hypothesis about fusion oncoproteins in general, which are numerous and readily produced by chromosomal translocations across a number of different cancers[16]. If the combination of DNA-binding domain and multivalent, homotypic interaction-capable LCD is a more general feature of many fusion oncoproteins, then neomorphic hub formation might be a selected consequence of the translocation itself rather than an incidental property of the EWS::FLI1 fusion in particular.

EWS::FLI1 hubs are regulated by various mechanisms

A recurring theme across our recent studies is that disparate cellular processes ranging from progression through the cell cycle and the transcription of nascent RNA to post-translational modification by O-GlcNAc all serve to regulate EWS::FLI1 hub properties and downstream transcription through a unique underlying biochemical mechanism. During mitosis, as chromatin condenses, EWS::FLI1 hubs dissolve, and the freely diffusing fraction of EWS::FLI1 increases substantially[10]. Despite this, EWS::FLI1 still retains its ability to bind chromatin, albeit more transiently[10], a property consistent with the phenomenon of mitotic bookmark, where certain TFs retain their ability to bind to chromatin during mitosis and mark specific gene loci to enable the rapid reactivation of the intended transcriptional program in the daughter cells upon reentry into interphase[17–20]. Nascent RNA produced at EWS::FLI1 target genes regulates EWS::FLI1 hubs through a different axis entirely, appearing to destabilize EWS::FLI1-hub binding, resulting in increased binding-unbinding kinetics without significantly affecting the spatial distribution of EWS::FLI1[10]. While it shares some similarities with the RNA-feedback observed in Mediator condensates, they are distinct mechanisms[21]. Finally, O-GlcNAcylation of EWS::FLI1 was shown to subtly increase the bound fraction of EWS::FLI1 and make hubs less viscous, resulting in suppressed EWS::FLI1-mediated transcription; however, similarly to nascent RNA

feedback, O-GlcNAcylation of EWS::FLI1 did not significantly affect its spatial distribution as measured by PALM.

Taken together, these findings have implications beyond the obvious; they suggest that the spatial organization and binding dynamics of EWS::FLI1 hubs are separable and can be independently regulated; i.e., a hub can look structurally unchanged under super-resolution imaging while its function can be altered or impaired by changes in residence time and material properties. While further experiments like live-cell imaging of EWS::FLI1 hubs and active transcription at target genomic loci are needed to develop a physical and mechanistic model of this decoupling, the observation alone meaningfully argues against a binary picture of EWS::FLI1 hub-regulated transcription whereby the presence or absence of a hub is the sole dictator of transcription; rather, it suggests that transcriptional control by hubs can be tuned independently along orthogonal spatiotemporal axes.

The timing of EWS::FLI1 hubs in the transcription cycle

The presence of a hub at a target gene is not sufficient to identify in which step of transcription a hub is functioning. RNA polymerase II (RNAP II) is heavily phosphorylated along its intrinsically disordered C-terminal domain in a transcriptional timing-dependent manner. Roughly, during pre-initiation it is hypophosphorylated, and through initiation and early elongation it is phosphorylated on serine 5, and during productive elongation, it is phosphorylated on serine 2[22]. Our dual-color super-resolution imaging study used stimulated emission depletion (STED) microscopy to show that EWS::FLI1 hubs are strongly and selectively enriched for initiation-associated, serine-5 phosphorylated RNAP II. This suggests that EWS::FLI1 hubs primarily function at the initiation stage of transcription, consistent with its literature-reported roles in recruiting transcriptional coactivators to GGAA microsatellite *de novo* enhancers. Notably, the association of serine-5 phosphorylated RNAP II to EWS::FLI1 hubs did not depend on hub size, ruling it out as a candidate to explain the hub size-dependent transcription levels observed in the literature[2].

Broader context and outlook

It has become increasingly clear that in the study of transcriptional control, a process mediated by the exquisite spatiotemporal coordination of hundreds of proteins all in the correct functional state, at tightly controlled local concentrations, in exact nanoscale subcellular localizations at precisely the right time, these systems should be studied under conditions that are as close to physiology as possible. This is especially true in the case of investigations centered on transcription factor hubs, biomolecular assemblies whose very formation is a concentration-dependent phenomenon. To that end, here, we have applied an array of live-cell, single-molecule, and super-resolution imaging strategies in order to dissect how transcription factor hubs form, their physical and material properties, and how they are regulated.

Several important questions arose and are natural extensions of this work. Firstly, the identity of the hub-dissolution mechanism, which serves to maintain the non-equilibrium steady state distribution of hubs, remains unknown. Extracting the expected hub dissolution rate of ~34 hubs per second can only serve as a guidepost that sets the timescale and suggests that the process is rapid and active. Significant additional exploratory work is needed to even identify candidate dissolution mechanisms for further study.

The mitotic bookmarking result also raises further questions. During mitosis, EWS::FLI1 retains its ability to bind mitotic chromatin, suggesting a potential mechanism for propagating the oncogenic transcriptional program across cell divisions even in the absence of hubs; however, are the points in the genome where EWS::FLI1 remains bound primarily at GGAA microsatellites upstream of target genes, and is this bookmarking even functionally necessary for re-establishing the oncogenic transcriptional program? A combination of ChIP-seq and live-cell imaging with perturbations that can abrogate EWS::FLI1 mitotic chromatin binding could address this.

The O-GlcNAcylation result, in addition to raising the obvious questions of the roles of other post-translational modifications in transcription factor hub regulation, together with the nascent RNA result, raises a broader question: to what extent can the functional output of transcription factor hubs be governed by the interaction strength and dynamics of their

constituent biomolecules independently of hub structure? The decoupling of hub morphology from transactivation function has significant implications; namely, it suggests that imaging the hub distribution alone is not sufficient to predict transcriptional output.

The finding that LY2835219 and trabectedin dissolve and mislocalize EWS::FLI1 hubs, respectively, also warrants significant further attention as it provides direct mechanistic evidence supporting the therapeutic angle of targeting oncogenic transcription factor hubs composed of otherwise untargetable proteins. The finding that partial hub disruption even in the absence of complete hub dissolution is sufficient to suppress transcription of target genes is also an informative finding for downstream screening for small molecules that can similarly disrupt hubs.

More broadly, a key remaining question is how generalizable the hub properties and regulation strategies of EWS::FLI1 are to other transcriptional regulators, both healthy and pathological. Of particular interest are other oncogenic fusion transcription factors caused by chromosomal translocations, including PAX3::FOXO1, which causes fusion-positive alveolar rhabdomyosarcoma, and SS18:SSX, which causes synovial sarcoma. The combined framework of CRISPR-based endogenous fluorescent labeling and quantitative live-cell and single-molecule imaging introduced here provides a template for further studies targeting any number of transcriptional regulators implicated in disease.

References

1. Chong, S. *et al.* Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555 (2018).
2. Chong, S. *et al.* Tuning levels of low-complexity domain interactions to modulate endogenous oncogenic transcription. *Mol. Cell* **82**, 2084-2097.e5 (2022).
3. Guillon, N. *et al.* The oncogenic EWS-FLI1 protein binds in vivo GGAA microsatellite sequences with potential transcriptional activation function. *PloS One* **4**, e4932 (2009).
4. Kovar, H. Dr. Jekyll and Mr. Hyde: The Two Faces of the FUS/EWS/TAF15 Protein Family. *Sarcoma* **2011**, 837474 (2011).

5. Linardic, C. M. PAX3-FOXO1 fusion gene in rhabdomyosarcoma. *Cancer Lett.* **270**, 10–18 (2008).
6. Li, Y., Luo, H., Liu, T., Zacksenhaus, E. & Ben-David, Y. The ets transcription factor Fli-1 in development, cancer and disease. *Oncogene* **34**, 2022–2031 (2015).
7. McSwiggen, D. T., Mir, M., Darzacq, X. & Tjian, R. Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev.* **33**, 1619–1634 (2019).
8. Irgen-Gioro, S., Yoshida, S., Walling, V. & Chong, S. Fixation can change the appearance of phase separation in living cells. *eLife* **11**, e79903 (2022).
9. Yoshida, S. R., Maity, B. K. & Chong, S. Visualizing Protein Localizations in Fixed Cells: Caveats and the Underlying Mechanisms. *J. Phys. Chem. B* **127**, 4165–4173 (2023).
10. Yoshida, S., Zhong, Y., Kumagai, A., Dunphy, W. G. & Chong, S. Dynamic regulation of endogenous transcription factor hubs at single-molecule resolution. 2026.01.25.701606 Preprint at https://doi.org/10.64898/2026.01.25.701606 (2026).
11. Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
12. Choi, J.-M., Holehouse, A. S. & Pappu, R. V. Physical principles underlying the complex biology of intracellular phase transitions. *Annu. Rev. Biophys.* **49**, 107–133 (2020).
13. Zandi, R., van der Schoot, P., Reguera, D., Kegel, W. & Reiss, H. Classical nucleation theory of virus capsids. *Biophys. J.* **90**, 1939–1948 (2006).
14. Narayanan, A. *et al.* A first order phase transition mechanism underlies protein aggregation in mammalian cells. *eLife* **8**, e39695 (2019).
15. Clouet, E. Modeling of Nucleation Processes. Preprint at https://doi.org/10.48550/arXiv.1001.4131 (2010).
16. Tripathi, S. *et al.* Defining the condensate landscape of fusion oncoproteins. *Nat. Commun.* **14**, 6008 (2023).
17. Caravaca, J. M. *et al.* Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor to mitotic chromosomes. *Genes Dev.* **27**, 251–260 (2013).

18. Lerner, J. *et al.* Human mutations affect the epigenetic/bookmarking function of HNF1B. *Nucleic Acids Res.* **44**, 8097–8111 (2016).
19. Lodhi, N., Ji, Y. & Tulin, A. Mitotic bookmarking: maintaining post-mitotic reprogramming of transcription reactivation. *Curr. Mol. Biol. Rep.* **2**, 10–16 (2016).
20. Teves, S. S. *et al.* A dynamic mode of mitotic bookmarking by transcription factors. *eLife* **5**, e22280 (2016).
21. Henninger, J. E. *et al.* RNA-Mediated Feedback Control of Transcriptional Condensates. *Cell* **184**, 207-225.e24 (2021).
22. Komarnitsky, P., Cho, E.-J. & Buratowski, S. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* **14**, 2452–2460 (2000).

Appendix A

SUPPORTING INFORMATION: DYNAMIC REGULATION OF ENDOGENOUS TRANSCRIPTION FACTOR HUBS AT SINGLE-MOLECULE RESOLUTION

Adapted from: **Yoshida, S.**, Zhong, Y., Kumagai, A., Dunphy, W. G. & Chong, S. Dynamic regulation of endogenous transcription factor hubs at single-molecule resolution. 2026.01.25.701606

Preprint at https://doi.org/10.64898/2026.01.25.701606 (2026).

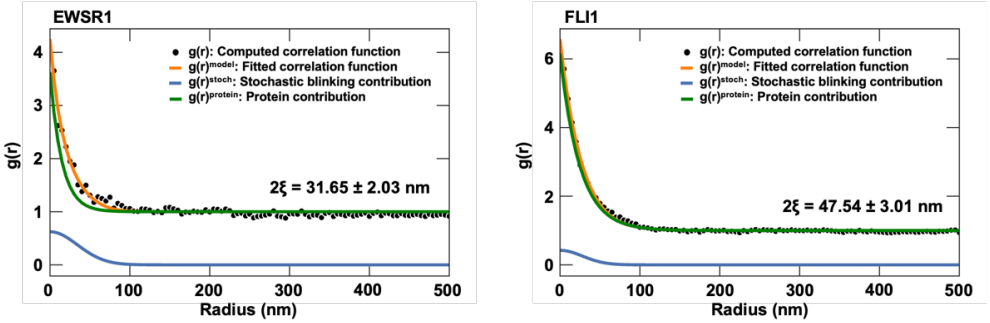

Figure S1: Pair-correlation PALM characterizes the approximate size of EWSR1 and FLI1 puncta. The computed pair-wise autocorrelation function of the raw PALM localizations (black) for (a) EWSR1 and (b) FLI1 were fit to a model that accounts for the contributions of both true protein clustering and the multiple appearances of the same fluorophores due to stochastic blinking (orange). The protein correlation term and stochastic blinking correlation term are shown in green and blue, respectively. The correlation length of protein clusters, ξ, is proportional to the cluster size and 2ξ can be used as an estimate of average cluster diameter. EWSR1 and FLI1 puncta are significantly smaller than EWS::FLI1 hubs with approximate diameters 2ξ = 31.65 ± 2.03 nm and 2ξ = 47.54 ± 3.01 nm, respectively. Data is from n = 10 cells.

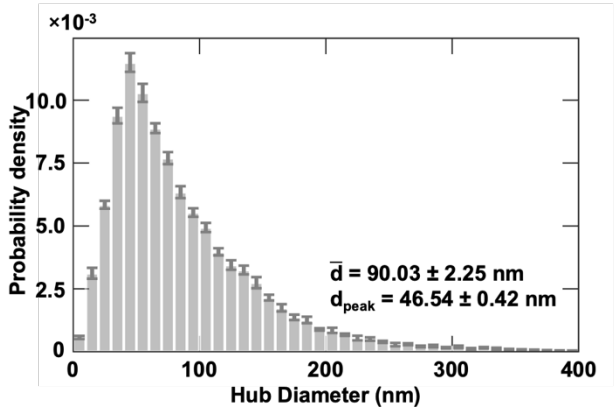

Figure S2. PALM microscopy characterizes the dimensions of EWS::FLI1 hubs after treatment with triptolide. The distribution of mean hub diameters generated by clustering photoblinking-corrected localizations with DBSCAN.

Supplementary Note A:

In the mean-field treatment of homogeneous nucleation, the free energy corresponding to the nucleation of a spherical hub of radius, R, is written as

$$\Delta G = 4\pi R^2\gamma - \frac{4}{3}\pi R^3 \nu^{-1} k_B T \log(S) \tag{S1}$$

where γ is the surface energy at the interface, ν is the volume occupied by a single molecule, and $S = c/c^*$ is the supersaturation, or the ratio of the actual concentration to the equilibrium or saturation concentration. Then, for n molecules, the total volume occupied by the hub is $n\nu = \frac{4}{3}\pi R^3$ and rearranging to obtain expression for R and substituting it into Equation S1 yields

$$\Delta G = (36\pi)^{1/3} v^{2/3} \gamma n^{2/3} - n k_B T \log(S) \quad \text{(S2)}$$

Defining $\varepsilon \coloneqq (36\pi)^{1/3} v^{2/3} \gamma$ and recognizing that $\Delta\mu = k_B T \log(S)$, we obtain Equation 1, the compact form of Equation S1 used in the main text, Equation 1.

$$\Delta G = \varepsilon n^{2/3} - \Delta\mu n \quad \text{(S3)}$$

The critical hub size n_c, the corresponding critical radius R_c, and free energy barrier to nucleation ΔG_c, are set by the maximum of the free energy profile; thus, we set $\frac{d}{dn}\Delta G = 0$ and find $n_c = \left(\frac{2\varepsilon}{3\Delta\mu}\right)^3$, corresponding to

$$R_c = \frac{2\varepsilon}{3\Delta\mu} \quad \text{(S4)}$$

and

$$\Delta G_c = \frac{4\varepsilon^3}{27\Delta\mu^2} \quad \text{(S5)}$$

Supplementary Note B:

Assuming Becker and Döring kinetics where individual molecules (rather than multimers) join and leave hubs, the steady-state nucleation rate per unit volume follows the Zeldovich form given by[1]

$$J = \rho\beta Z exp\left(\frac{-\Delta G_c}{k_B T}\right) \quad \text{(S6)}$$

Where ρ is the number density of molecules in m^{-3}, β is the impingement rate of molecules in s^{-1}, and Z is the Zeldovich factor, which accounts for the curvature of the free energy profile around the critical radius.

We can estimate that $\rho = 3.20 \times 10^{19}\ m^{-3}$ using 200 nM the concentration of EWS::FLI1 from Chong et al. 2018[2] with Avogadro's number. The impingement rate for diffusion-

limited (rather than reaction-limited) incorporation can be estimated as $\beta = 4\pi R_c D\rho = 77.98\ s^{-1}$, where $D = 1.97 \times 10^{-12}\ m^2 s^{-1}$ is the diffusion coefficient of EWS::FLI1 from the main text. The Zeldovich factor is a function of the curvature of ΔG at the critical size and is given as

$$Z = \sqrt{\frac{-1}{2\pi k_B T}\left(\left.\frac{\partial^2 \Delta G}{\partial n^2}\right|_{n=n_c}\right)} \tag{S7}$$

We can rewrite this using Equation 1 yielding

$$Z = \frac{1}{3k_B T}\pi^{-1/2}\varepsilon^{1/2} n_c^{-2/3} \tag{S8}$$

Notably, Z (hence J) is not insensitive to constant multiplicative factors of n_c. Thus, rather than using the fit value, we take the average number of localizations in hubs of size R_c and divide by 0.21, the effective localization efficiency for PA-JF549 to estimate that $n_c = 153.21$. Using this value of n_c together with the fitted value $\varepsilon = 0.0015 k_B T$ yields $Z = 2.54 \times 10^{-4}$. We can also use the fitted value of $\frac{\Delta G_c}{k_B T} = 4.88$. Using our measurements with Equation S6, we estimate that $\mathrm{J} = 6.83 \times 10^{16}\ m^{-3} s^{-1}$.

References:

1. Becker, R. & Döring, W. Kinetische Behandlung der Keimbildung in übersättigten Dämpfen. Ann. Phys. 416, 719–752 (1935).

2. Chong, S. et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science 361, eaar2555 (2018)