



OPEN

# Force generation by protein–DNA co-condensation

Thomas Quail<sup>1,2,3,4</sup>, Stefan Golfier<sup>1,2,3,4</sup>, Maria Elsner<sup>1,2,3,4</sup>, Keisuke Ishihara<sup>1,2,3,4</sup>,  
Vasanthanarayan Murugesan<sup>1,2,3,4</sup>, Roman Renger<sup>1</sup>, Frank Jülicher<sup>1,2,3,4</sup> ✉ and Jan Brugués<sup>1,2,3,4</sup> ✉

**Interactions between liquids and surfaces generate forces<sup>1,2</sup> that are crucial for many processes in biology, physics and engineering, including the motion of insects on the surface of water<sup>3</sup>, modulation of the material properties of spider silk<sup>4</sup> and self-assembly of microstructures<sup>5</sup>. Recent studies have shown that cells assemble biomolecular condensates via phase separation<sup>6</sup>. In the nucleus, these condensates are thought to drive transcription<sup>7</sup>, heterochromatin formation<sup>8</sup>, nucleolus assembly<sup>9</sup> and DNA repair<sup>10</sup>. Here we show that the interaction between liquid-like condensates and DNA generates forces that might play a role in bringing distant regulatory elements of DNA together, a key step in transcriptional regulation. We combine quantitative microscopy, *in vitro* reconstitution, optical tweezers and theory to show that the transcription factor FoxA1 mediates the condensation of a protein–DNA phase via a mesoscopic first-order phase transition. After nucleation, co-condensation forces drive growth of this phase by pulling non-condensed DNA. Altering the tension on the DNA strand enlarges or dissolves the condensates, revealing their mechanosensitive nature. These findings show that DNA condensation mediated by transcription factors could bring distant regions of DNA into close proximity, suggesting that this physical mechanism is a possible general regulatory principle for chromatin organization that may be relevant *in vivo*.**

Compartmentalization is key to organizing cellular biochemistry. Biomolecular condensate formation underlies the compartmentalization of many cellular functions<sup>6</sup>. Considerable progress has been made towards understanding the biophysical properties of condensates in bulk. However, how these condensates interact with other cellular components such as polymers, membranes and chromatin remains unclear. Transcriptional hubs represent an example of compartments in the nucleus. These hubs involve the coalescence of transcription factors, biochemical regulators of transcription, and DNA<sup>11</sup>. The physical nature of these transcriptional hubs is under debate, though recent studies have proposed that transcriptional hubs can be understood as examples of biomolecular condensates<sup>12</sup>. In theory, the interactions between condensates composed of transcriptional machinery and the DNA polymer could deform DNA, potentially bridging distal regulatory elements, a critical step in gene regulation. However, we still lack a physical picture of how transcriptional regulators interact with each other and with the surface of the DNA polymer.

To investigate how transcription factors physically organize DNA, we attached linearized  $\lambda$ -phage DNA to a coverslip via biotin–streptavidin linkers (Fig. 1a). We used total internal reflection fluorescence microscopy to image the interactions between DNA and forkhead box protein A1 (FoxA1), a pioneer transcription

factor that regulates tissue differentiation across a range of organisms<sup>13</sup> (Fig. 1b). Upon addition of 10 nM FoxA1–mCherry (FoxA1) to the flow chamber in the presence of DNA, FoxA1 formed protein condensates that decorated the strand (Fig. 1c). In the absence of DNA, FoxA1 did not nucleate condensates in solution at concentrations ranging from 10 to 500 nM (Extended Data Fig. 1a). The requirement for DNA in condensate formation at low concentrations suggests that DNA mediates the condensation of a thin layer of FoxA1 on DNA.

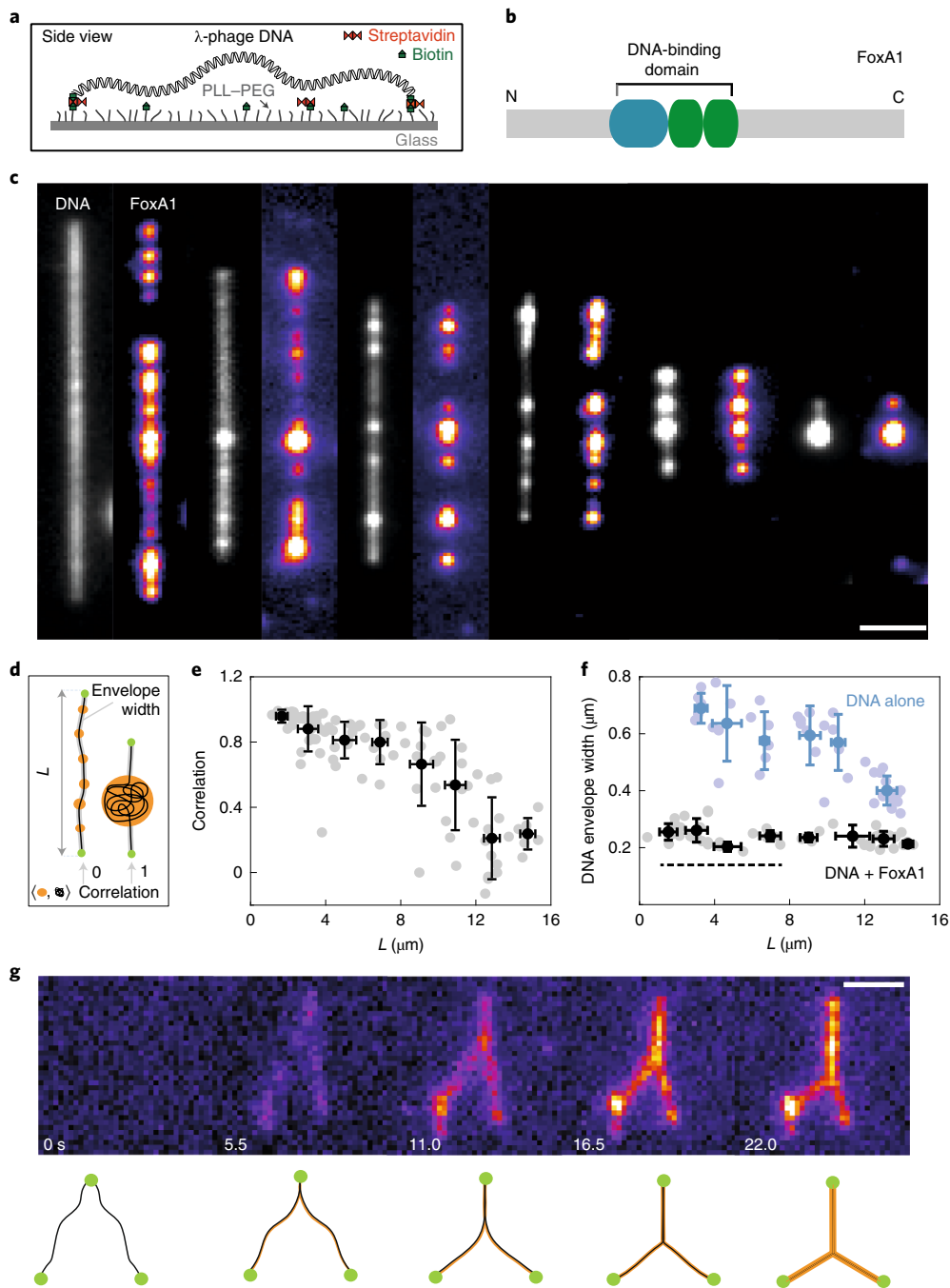
In our assay, DNA molecules displayed a broad distribution of end-to-end distances ( $L$ ), determined by the DNA–coverslip attachment points (Fig. 1c,d). This end-to-end distance tunes the tension of the DNA<sup>14</sup>. For DNA strands with end-to-end distances greater than approximately 10  $\mu\text{m}$ , FoxA1 generated protein condensates on DNA (Fig. 1c). However, FoxA1 condensation did not influence the DNA molecule (Fig. 1c, leftmost pair of images). Strikingly, for DNA molecules with end-to-end distances below 10  $\mu\text{m}$ , FoxA1 pulled the DNA into highly enriched condensates of FoxA1 and DNA (Fig. 1c and Extended Data Fig. 1b–e) with a density of roughly 750 molecules  $\mu\text{m}^{-3}$  (Methods and Extended Data Fig. 2a–d). To quantify FoxA1-mediated DNA condensation, we measured the cross-correlation of FoxA1–DNA intensities as a function of end-to-end distance (Methods, Fig. 1d,e and Extended Data Fig. 3a). Consistent with the ability of FoxA1 to form FoxA1–DNA condensates at low tensions, the cross-correlation decayed from one to zero with increasing end-to-end distance (Fig. 1e). Thus, FoxA1 mediates the formation of a protein–DNA-rich phase in a tension-dependent manner.

The observation that FoxA1 drives DNA condensation suggests that it can overcome the DNA molecule's entropic tension set by the end-to-end distance<sup>14</sup>. Incorporating DNA into the condensates increases the tension on the strand, thereby reducing the transverse DNA fluctuations of the non-condensed DNA. To quantify this, we measured the DNA envelope width of the non-condensed DNA fluctuations (Methods and Extended Data Fig. 3b). In buffer, the DNA envelope width decreased as a function of end-to-end distance, consistent with the corresponding increase of DNA strand tension for increasing end-to-end distance<sup>14</sup> (Fig. 1f). However, in the presence of FoxA1, the DNA envelope width remained constant for all end-to-end distances as FoxA1 pulled DNA into one or more condensates. The magnitude of the DNA envelope width was lower in the presence of FoxA1 than in buffer conditions for all end-to-end distances (Fig. 1f). Taken together, this suggests that FoxA1–DNA condensates generate forces that can overcome the entropic tension of the non-condensed DNA and buffer its tension.

The observation that FoxA1 can mediate DNA condensation suggests that it could bridge distant DNA strands. To investigate this possibility, we examined DNA molecules that were bound to

<sup>1</sup>Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany. <sup>2</sup>Max Planck Institute for the Physics of Complex Systems, Dresden, Germany. <sup>3</sup>Center for Systems Biology Dresden, Dresden, Germany. <sup>4</sup>Cluster of Excellence Physics of Life, TU Dresden, Dresden, Germany.

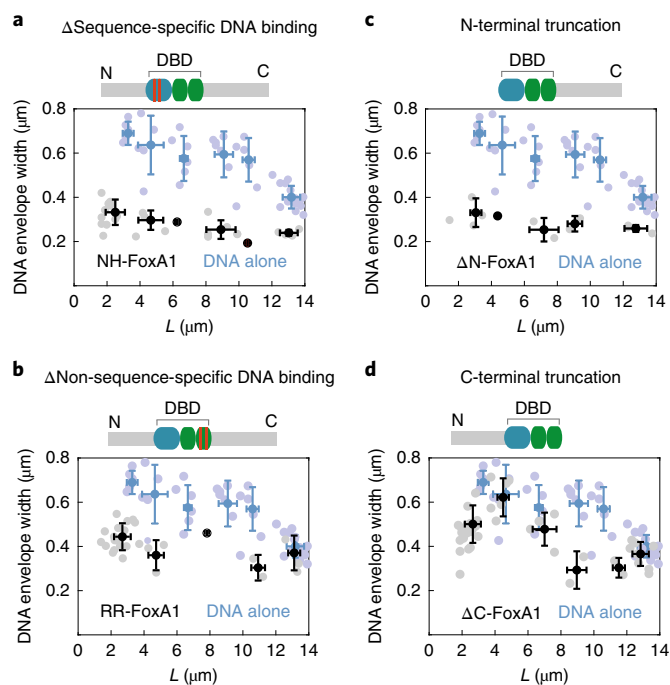
✉e-mail: [julicher@pks.mpg.de](mailto:julicher@pks.mpg.de); [brugues@mpi-cbg.de](mailto:brugues@mpi-cbg.de)



**Fig. 1 | FoxA1 forms DNA-FoxA1 condensates in a tension-dependent manner.** **a**, Schematic of single  $\lambda$ -phage DNA molecule assay. PLL, poly-L-lysine. **b**, Structure of FoxA1, consisting of a structured DNA-binding domain flanked by mostly disordered N and C termini. The DNA-binding domain has a sequence-specific binding region (blue) and two non-sequence-specific binding regions (green). **c**, Representative time-averaged projections of FoxA1 and DNA. The extent of FoxA1-mediated DNA condensation depends on the end-to-end distance of the strand. Note that the total amount of DNA is the same in each example. The DNA was imaged using 10 nM SYTOX Green. Scale bar, 2  $\mu\text{m}$ . **d**, Schematic displaying three main quantities used to characterize DNA-FoxA1 condensation:  $L$ , the DNA's end-to-end distance; cross-correlation of DNA and FoxA1 intensities; and DNA envelope width, a measure of transverse DNA fluctuations. **e**, Cross-correlation of FoxA1 and DNA signals shows that FoxA1 condenses DNA below a critical end-to-end distance. The grey dots represent individual strands,  $n=107$ . The data are binned every 2  $\mu\text{m}$  (black, mean  $\pm$  s.d. for both correlations and strand lengths). **f**, DNA envelope width measurements (Methods) reveal that FoxA1-DNA condensation buffers DNA tension (blue and black dots correspond to control and DNA + FoxA1 conditions,  $n=45$  and  $n=50$  respectively). The data are binned every 2  $\mu\text{m}$  (mean  $\pm$  s.d. for both the envelope widths and strand lengths). The dashed black line represents the theoretical diffraction limit. **g**, Representative images of FoxA1 zipping two independent DNA strands over time. Scale bar, 2  $\mu\text{m}$ .

the same streptavidin molecule at one end (Fig. 1g and Extended Data Fig. 3c). In the absence of FoxA1, these DNA molecules form a V-shaped morphology and fluctuate independently of one

another. Upon addition of FoxA1, however, we observed that the two strands zipped together, generating a Y-shaped morphology as the condensation of FoxA1 increased over time (Fig. 1g



**Fig. 2 | Mutant analysis reveals that the C terminus of FoxA1 drives DNA condensation.** DNA envelope width measurements for FoxA1 mutants. The data are binned every 2  $\mu\text{m}$  and the mean  $\pm$  s.d. (for both the envelope width and the strand length) are shown in black for each mutant and in blue for the control ( $n = 45$ ). **a**, Sequence-specific DNA-binding mutant NH-FoxA1 condenses DNA ( $n = 30$ ). DBD, DNA-binding domain. **b**, Non-sequence-specific DNA-binding mutant RR-FoxA1 condenses DNA ( $n = 28$ ). **c**, N-terminal truncation of FoxA1,  $\Delta\text{N}$ -FoxA1, condenses DNA ( $n = 13$ ). **d**, C-terminal truncation of FoxA1,  $\Delta\text{C}$ -FoxA1, inhibits DNA condensation ( $n = 44$ ). In all conditions, the protein concentration was 10 nM. See Supplementary Fig. 1 for representative protein–DNA images of the FoxA1 mutants.

and Extended Data Fig. 3c). Taken together, these data demonstrate that FoxA1 can physically bridge DNA strands in both *cis* and *trans* configurations.

Two mechanisms can be postulated to explain FoxA1-mediated DNA condensation in our experiments: (1) direct crosslinking via the multiple DNA-binding activities of FoxA1<sup>15</sup> or (2) weak protein–protein interactions driven by disordered regions of FoxA1. FoxA1 consists of a winged helix–turn–helix DNA-binding domain and two amino- and carboxy-terminus domains that are mostly disordered<sup>15</sup>. The DNA-binding domain contains a sequence-specific binding region composed of three alpha helices and a non-sequence-specific binding region composed of two wings. Two point mutations known to affect sequence-specific DNA binding (NH-FoxA1<sup>15</sup>) had virtually no influence on DNA condensation activity (Fig. 2a). Although the presence of two point mutations known to affect non-sequence-specific DNA binding (RR-FoxA1<sup>15</sup>) partially inhibited FoxA1 localization to the strand (Fig. 2b), this mutant still condensed DNA. In this case, condensation occurred on a time scale of minutes rather than seconds (as in WT-FoxA1), which can be explained by the delay in condensing sufficient RR-FoxA1 to the strand. These data suggest that non-sequence-specific binding drives the localization of FoxA1 to DNA but does not mediate DNA condensation through crosslinking. Furthermore, the sequence-specific binding domain of FoxA1 is dispensable for its localization to DNA *in vitro*. To probe whether FoxA1 protein–protein interactions through disordered domains

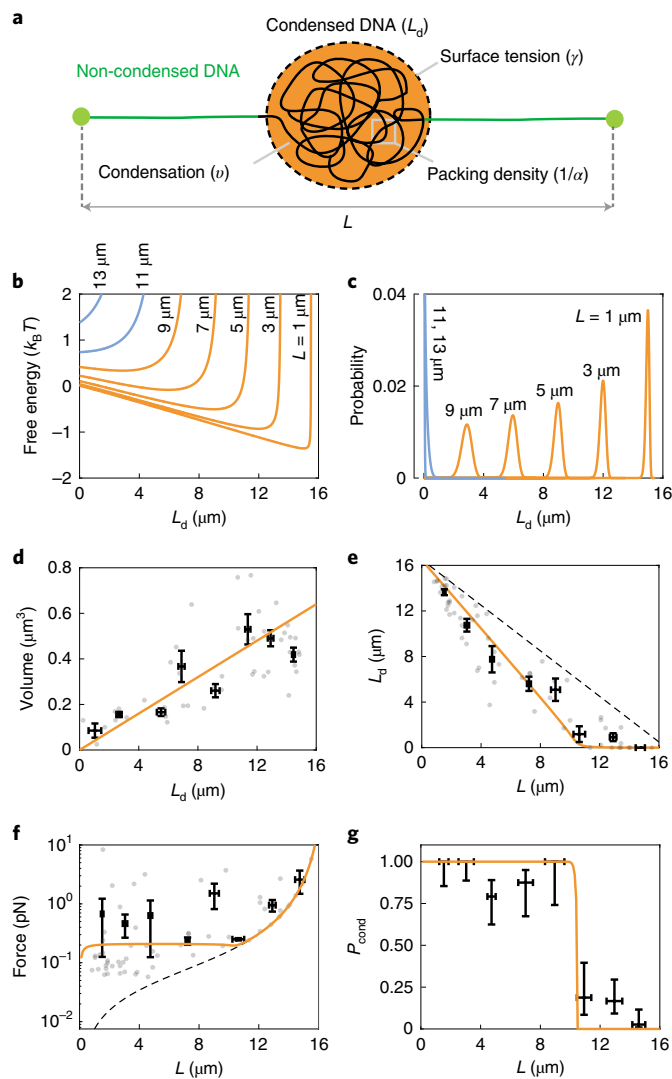
mediate DNA condensation, we truncated both the N and C termini of FoxA1. Although  $\Delta\text{N}$ -FoxA1 retained DNA condensation activity (Fig. 2c), truncating the disordered C terminus of FoxA1 largely inhibited DNA condensation activity (Fig. 2d). Additionally, we found that, at high FoxA1 concentrations in bulk (50  $\mu\text{M}$ ), 3% polyethylene glycol (PEG) (relative molecular mass 30,000–30K) nucleated highly enriched spherical FoxA1 condensates (Extended Data Fig. 4a), further suggesting the existence of weak FoxA1–FoxA1 interactions. Thus, non-sequence-specific binding drives FoxA1 localization to DNA, and the disordered C terminus of FoxA1 promotes DNA condensation.

Our results support the hypothesis that FoxA1 condenses onto DNA to generate a protein–DNA-rich condensate via weak protein–protein interactions that exerts a pulling force on the non-condensed strand (Thermodynamic description of DNA–protein condensation in Supplementary Information). To explore the thermodynamics of condensation, we developed a theoretical description based on a semiflexible polymer partially condensing into a liquid-like condensate. Here, the semiflexible polymer is DNA and the condensation is mediated by the transcription factor. The free energy of this process contains volume, ( $v\frac{4}{3}\pi R^3$ ), and surface contributions, ( $\gamma 4\pi R^2$ ), as well as a term representing the free energy of the non-condensed DNA (Fig. 3a), where  $v$  is the condensation free energy per volume,  $R$  is the condensate radius and  $\gamma$  is the surface tension of the condensate. We assume that the DNA is fully collapsed inside the condensate and thus its volume is proportional to the condensed DNA contour length,  $V = \alpha L_d$ , where  $1/\alpha$  describes the packing density given as DNA length per condensate volume. The free energy of the polymer,  $F_p(L, L_p) = \int_0^L f(L, L_p) dl$ , can be obtained from the force–extension curve of the polymer  $f(L, L_p)$ , where  $L_p$  is the contour length of the non-condensed polymer. Using  $L_p = L_c - L_d$ , where  $L_c$  is the contour length of  $\lambda$ -phage DNA (16.5  $\mu\text{m}$ ), the free energy is as follows:

$$F(L, L_d) = -v\alpha L_d + \gamma 4\pi \left(\frac{3\alpha}{4\pi}\right)^{\frac{2}{3}} L_d^{\frac{2}{3}} + \kappa \left( \frac{(L_c - L_d)^2}{4(L_c - L_d - L)} - \frac{L}{4} + \frac{L^2}{2(L_c - L_d)} - \frac{(L_c - L_d)}{4} \right)$$

where  $\kappa = \frac{k_B T}{P}$ ,  $k_B$  is the Boltzmann constant,  $T$  is the temperature and  $P$  is the persistence length of DNA (Thermodynamic description of DNA–protein condensation in Supplementary Information). For fixed  $L$ , the minimum of  $F(L, L_d)$  determines the preferred size of the condensate. This free energy predicts upon variation of  $L$  a stochastic first-order phase transition for the formation of protein–DNA condensates (Fig. 3b). The distribution of condensate sizes is then given by  $P(L_d) \approx e^{-\beta F(L, L_d)}$  for fixed  $L$  (Fig. 3c). This accounts for a sharp transition of DNA condensation controlled by the end-to-end distance and thus the tension of the DNA molecule. The first-order nature of this behaviour implies regimes of hysteresis and bistability. Our theory also predicts that the condensation forces exerted on the non-condensed DNA are kept roughly constant.

To test this theory, we first measured DNA condensate volumes and found that they increase linearly with  $L_d$ , with  $\alpha = 0.04 \pm 0.01 \mu\text{m}^2$  (Fig. 3d, Extended Data Fig. 4d and Methods). This confirms that DNA is in a collapsed conformation inside the condensates. Next, we simultaneously fitted the predictions to  $L_d$  and the probability of nucleating a DNA condensate ( $P_{\text{cond}}$ ) as a function of end-to-end distance (Methods). We calculated  $L_d$  (Fig. 3e and Extended Data Figs. 4e and 5) and  $P_{\text{cond}}$  (Fig. 3g and Extended Data Fig. 4f) using the Boltzmann probability distributions (Fig. 3c) from the free energy. Our fits agree quantitatively with the data and show that  $L_d$  decreases with  $L$  until a critical end-to-end distance beyond which DNA condensates do not form. Below this critical length, we observed that the force exerted by the condensate is buffered at 0.21 pN (0.18–0.30 pN



**Fig. 3 | Thermodynamic description of a liquid phase condensing onto a semiflexible polymer explains FoxA1-mediated DNA condensation.**

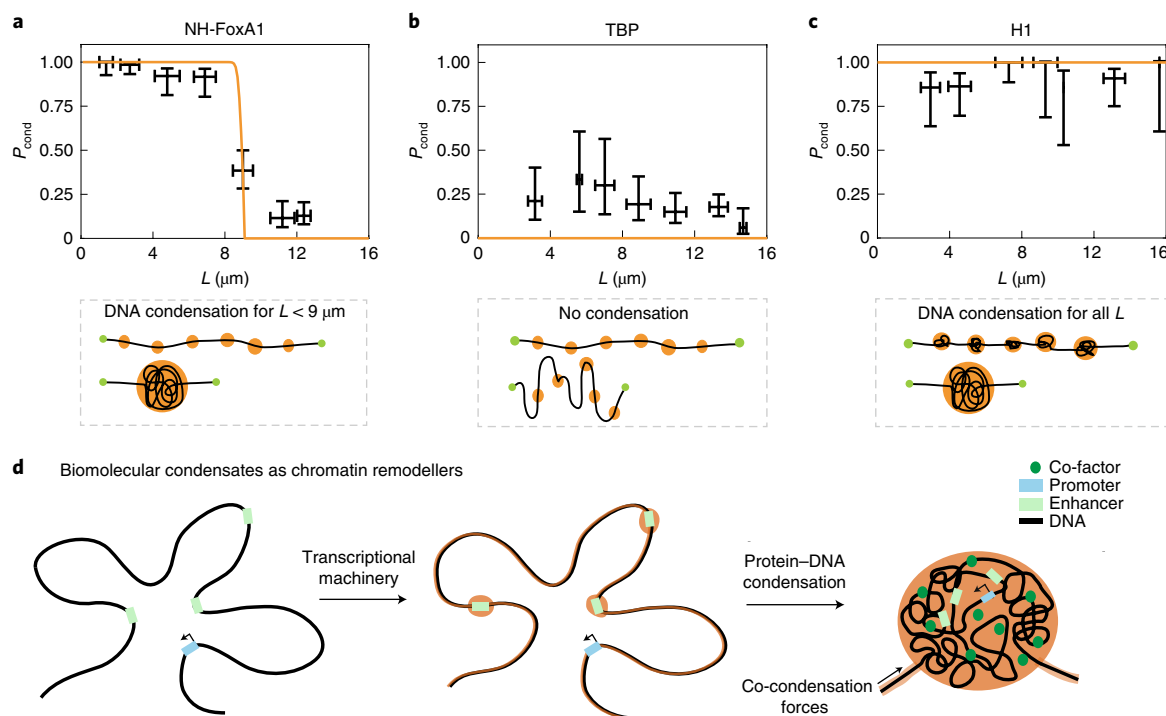
**a**, Schematic representing DNA-FoxA1 condensation (orange). DNA can be in a condensed state (black) or a non-condensed state (green). DNA condensation depends on  $\gamma$ ,  $\nu$  and  $\alpha$ . **b**, Free energy profiles as a function of  $L_d$  for different  $L$  reveal a first-order phase transition for protein-DNA condensation (orange and blue correspond to favourable and unfavourable condensation, respectively). **c**, Boltzmann distributions corresponding to the free energy profiles in **b**. **d**, Condensate volume increases linearly with  $L_d$ . The orange curve represents a linear fit to individual strands ( $n=47$ ). For **d-f**, individual strands are represented as grey dots and binned mean  $\pm$  s.e.m. is in black. **e**, The length of condensed DNA as a function of  $L$  ( $n=63$ ) reveals a sharp transition. The orange curve represents the optimal theoretical fit. The grey dashed line corresponds to the limit of maximum condensation where  $L_d$  is  $16.5 \mu\text{m}$  minus  $L$ . **f**, Condensation forces that DNA-protein condensates exert on non-condensed DNA are buffered ( $n=62$ ). The orange curve is the theoretical prediction. The grey dashed line represents the force when  $L_d=0$ . **g**,  $P_{\text{cond}}$  reveals a sharp transition at a critical end-to-end distance.  $P_{\text{cond}}$  is computed from binned local correlation data ( $n=181$  condensates). The end-to-end distance error bars are the s.d. and the  $P_{\text{cond}}$  error bars are the 95% confidence intervals from a beta distribution.

confidence interval), consistent with the theory (Fig. 3f). To complement our force measurements, we performed optical-tweezer measurements of FoxA1-mediated DNA condensation. Incubating

a single  $\lambda$ -phage DNA molecule at either  $L=6$  or  $8 \mu\text{m}$  in the presence of  $150 \text{ nM}$  FoxA1 generated forces of the order of  $0.4\text{--}0.6 \text{ pN}$ , consistent with the force measurements using fluorescence microscopy (Methods and Extended Data Figs. 6 and 7). Finally,  $P_{\text{cond}}$  exhibits a sharp transition at  $L=10.5 \mu\text{m}$  ( $9.4\text{--}10.9 \mu\text{m}$  confidence interval), in agreement with a stochastic first-order phase transition (Fig. 3g). We also observed a sudden force jump during the onset of condensate formation (as measured by the individual temporal force trajectories in the optical-tweezer experiments), consistent with a first-order phase transition (Extended Data Figs. 6c and 7). Close to the transition point FoxA1-mediated DNA condensation displayed bistability. This bistability was observed in strands that contained multiple FoxA1 condensates, but where only some of them condensed DNA (Extended Data Fig. 8a). Our fits allowed us to extract the physical parameters associated with condensate formation, namely  $\nu=2.6 \text{ pN } \mu\text{m}^{-2}$  ( $2.3\text{--}5.2 \text{ pN } \mu\text{m}^{-2}$  confidence interval) and  $\gamma=0.04 \text{ pN } \mu\text{m}^{-1}$  ( $0.04\text{--}0.28 \text{ pN } \mu\text{m}^{-1}$  confidence interval) (Methods). These parameters are consistent with previous measurements in vitro and in vivo condensates<sup>16,17</sup>.

Our theory and experiments show that two key parameters govern protein-DNA co-condensation, namely  $\nu$  and  $\gamma$ . We reasoned that different DNA-binding proteins may exhibit a range of behaviours depending on these parameters. First, we investigated the sequence-specific DNA-binding region mutant (NH-FoxA1), which also condensed DNA but to a lesser extent (Fig. 2a). Quantitatively, we found that the surface tension of condensates formed with this mutant was roughly unchanged compared with WT-FoxA1,  $\gamma=0.065 \text{ pN } \mu\text{m}^{-1}$  ( $0.05\text{--}0.07 \text{ pN } \mu\text{m}^{-1}$  confidence interval), but the free energy per volume of condensation was reduced, consistent with reduced DNA binding,  $\nu=1.05 \text{ pN } \mu\text{m}^{-2}$  ( $0.9\text{--}1.1 \text{ pN } \mu\text{m}^{-2}$  confidence interval) (Extended Data Fig. 9 and Fig. 4a). This was also reflected in a decrease in the extent of DNA packing, with  $\alpha=0.09 \pm 0.02 \mu\text{m}^2$  (Extended Data Fig. 9a). We also observed that NH-FoxA1-mediated condensates generated a force of  $0.17 \text{ pN}$  ( $0.16\text{--}0.19 \text{ pN}$  confidence interval), lower than that for WT-FoxA1. In addition, NH-FoxA1 displayed bistable protein-DNA condensation activity in the neighbourhood of the transition point (Extended Data Fig. 8b). Next, we examined the interactions of a different transcription factor, TATA-binding protein (TBP), with DNA. We found that TBP also formed small condensates on DNA, but did not condense DNA even at the lowest imposed DNA tensions (Fig. 4b). Instead, TBP performed a diffusive motion along the DNA strand (Extended Data Fig. 10c), suggesting that protein-DNA condensation is not thermodynamically favoured. Another transcription factor, Gal4-VP16, formed condensates on DNA and condensed DNA in a tension-dependent manner consistent with FoxA1 (Extended Data Fig. 10e). Finally, we analysed somatic linker histone H1, a protein that is structurally similar to FoxA1. However, in contrast to FoxA1, one of the known functions of H1 is to compact chromatin<sup>18</sup>, so we expected H1 to strongly condense DNA. Consistent with this, we found that H1 displayed a stronger DNA condensation activity compared with FoxA1, condensing DNA for all measured end-to-end distances (Fig. 4c). Interestingly, the *Xenopus* embryonic linker histone B4 condensed DNA in a tension-dependent manner but not to the same extent as H1 (Extended Data Fig. 10f). Thus, we propose that the competition between condensation free energy per volume of the protein-DNA phase and surface tension regulates a spectrum of DNA condensation activities, which may be tuned by the structure of transcription factors.

Here, we show that FoxA1 can condense DNA under tension to form a protein-DNA-rich phase that nucleates through a force-dependent first-order transition for forces below a critical value. This critical force, which is of the order of  $0.2\text{--}0.6 \text{ pN}$  for FoxA1, is set by co-condensation forces that the protein-DNA phase exerts on the non-condensed DNA. These forces are similar in magnitude to those recently measured for DNA loop extrusion, of



**Fig. 4 | Universality of protein-DNA co-condensation. a–c**, Condensation probability quantification for sequence-specific DNA-binding mutant NH-FoxA1 (**a**), TBP (**b**) and somatic linker histone H1 (**c**).  $P_{\text{cond}}$  is computed from local correlation data with  $n = 361$  condensates for NH-FoxA1 (**a**),  $n = 247$  condensates for TBP (**b**) and  $n = 101$  for H1 (**c**). The error bars for the end-to-end distance are s.d. and the  $P_{\text{cond}}$  error bars are the 95% confidence intervals from a beta distribution. We found that NH-FoxA1 condensed DNA less strongly than WT-FoxA1, TBP could not condense DNA for any end-to-end distance and H1 condensed DNA for all measured end-to-end distances. **d**, Biomolecular condensates generate condensation forces that could serve to recruit transcriptional regulators, and potentially remodel chromatin at physiologically relevant force scales to properly regulate transcription. See Supplementary Fig. 2 for representative protein-DNA images of NH-FoxA1, TBP and H1.

the order of 0.2–1 pN (refs. <sup>19,20</sup>), and those estimated in intact nuclei from nuclear condensate fusion<sup>21</sup>. Thus, we speculate that these weak forces we find in vitro may be of relevance to the mechanics of chromatin organization, though future studies are necessary to show this. Taken together, our work suggests that co-condensation forces may act as an additional mechanism to remodel chromatin in addition to molecular motors that extrude loops and complexes that remove or displace nucleosomes (Fig. 4d).

Transcription-factor-mediated protein-DNA condensation represents a possible mechanism by which transcription factors coordinate enhancer-promoter contacts in transcriptional hubs<sup>12</sup>. In this context, protein-DNA condensates could act as scaffolds, pulling co-factors into the droplet (Fig. 4d). Our theoretical description reveals that these protein-DNA condensates are formed via a first-order phase transition, suggesting that they can be assembled and disassembled rapidly by changing mechanical conditions. Near the transition point, assembly and disassembly of these in vitro protein-DNA condensates becomes highly stochastic, reminiscent of the rapid dynamics associated with the initiation and cessation of transcriptional bursts observed in vivo<sup>22</sup>.

We have demonstrated that protein-DNA co-condensation is associated with a difference in chemical potential between the condensed and non-condensed DNA. This difference in chemical potential is transduced by the condensate to perform mechanical work on the non-condensed DNA strand. Capillary forces represent another example of forces that involve liquid-surface interactions<sup>1,2,23</sup>. With both co-condensation and capillary forces, attractive interactions give rise to the transduction of free energy into work. Such forces may also be relevant beyond chromatin in other biological contexts, including membranes and the cytoskeleton.

Protein-DNA co-condensation not only provides mechanisms to facilitate enhancer-promoter contacts, but could also play a more general role in DNA compaction and maintenance of bulk chromatin rigidity in processes such as mitotic chromatid compaction<sup>24</sup> and the formation of chromatin compartments<sup>8,25,26</sup>. Owing to the tension-dependent nature of protein-DNA co-condensation, our work suggests that these forces could play a key, and, as yet, underappreciated role in genome organization and transcriptional initiation. It is appealing to imagine that transcriptional outputs respond not only to concentrations of transcription factors in the nucleus, but also to mechanical cues from chromatin.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41567-021-01285-1>.

Received: 2 November 2020; Accepted: 2 June 2021;  
Published online: 08 July 2021

### References

- De Gennes, P.-G., Brochard-Wyart, F. & Quere, D. *Capillarity and Wetting Phenomena* (Springer, 2013).
- Bico, J., Reyssat, É. & Roman, B. Elastocapillarity: when surface tension deforms elastic solids. *Annu. Rev. Fluid Mech.* **50**, 629–659 (2017).
- Hu, D. L. & Bush, J. W. M. The hydrodynamics of water-walking arthropods. *J. Fluid Mech.* **664**, 5–33 (2010).
- Eletto, H., Neukirch, S., Vollrath, F. & Antkowiak, A. In-drop capillary spooling of spider capture thread inspires hybrid fibers with

- mixed solid–liquid mechanical properties. *Proc. Natl Acad. Sci. USA* **113**, 6143–6147 (2016).
5. Mastrangeli, M. et al. Self-assembly from milli- to nanoscales: methods and applications. *J. Micromech. Microeng.* **19**, 083001–083038 (2009).
  6. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid–liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **30**, 39–58 (2014).
  7. Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
  8. Larson, A. G. et al. Liquid droplet formation by HP1 $\alpha$  suggests a role for phase separation in heterochromatin. *Nature* **547**, 236–240 (2017).
  9. Feric, M. et al. Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165**, 1686–1697 (2016).
  10. Patel, A. et al. A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* **162**, 1066–1077 (2015).
  11. Furlong, E. E. M. & Levine, M. Developmental enhancers and chromosome topology. *Science* **361**, 1341–1345 (2018).
  12. Cramer, P. Organization and regulation of gene transcription. *Nature* **573**, 45–54 (2019).
  13. Zaret, K. S. & Mango, S. E. Pioneer transcription factors, chromatin dynamics, and cell fate control. *Curr. Opin. Genet. Dev.* **37**, 76–81 (2016).
  14. Bustamante, C., Marko, J. F., Siggia, E. D. & Smith, S. Entropic elasticity of lambda-phage DNA. *Science* **265**, 1599–1600 (1994).
  15. Sekiya, T., Muthurajan, U. M., Luger, K., Tulin, A. V. & Zaret, K. S. Nucleosome-binding affinity as a primary determinant of the nuclear mobility of the pioneer transcription factor FoxA. *Genes Dev.* **23**, 804–809 (2009).
  16. Jawerth, L. M. et al. Salt-dependent rheology and surface tension of protein condensates using optical traps. *Phys. Rev. Lett.* **121**, 258101 (2018).
  17. Brangwynne, C. P. et al. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729–1732 (2009).
  18. Botao, X. et al. Histone H1 compacts DNA under force and during chromatin assembly. *Mol. Biol. Cell* **23**, 4864–4871 (2012).
  19. Golfer, S., Quail, T., Kimura, H. & Brugués, J. Cohesin and condensin extrude DNA loops in a cell-cycle dependent manner. *eLife* **9**, e53885 (2020).
  20. Banigan, E. J. & Mirny, L. A. Loop extrusion: theory meets single-molecule experiments. *Curr. Opin. Cell Biol.* **64**, 124–138 (2020).
  21. Shin, Y. et al. Liquid nuclear condensates mechanically sense and restructure the genome. *Cell* **175**, 1481–1491 (2018).
  22. Fukaya, T., Lim, B. & Levine, M. Enhancer control of transcriptional bursting. *Cell* **166**, 358–368 (2016).
  23. Jang, J., Schatz, G. C. & Ratner, M. A. Capillary force in atomic force microscopy. *J. Chem. Phys.* **120**, 1157–1160 (2004).
  24. Yoshimura, S. H. & Hirano, T. HEAT repeats—versatile arrays of amphiphilic helices working in crowded environments? *J. Cell. Sci.* **129**, 3963–3970 (2016).
  25. Strom, A. R. et al. HP1 $\alpha$  is a chromatin crosslinker that controls nuclear and mitotic chromosome mechanics. *eLife* **10**, e63972 (2021).
  26. Williams, J. F. et al. Phase separation enables heterochromatin domains to do mechanical work. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.07.02.184127> (2020).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

## Methods

**Cloning and protein purification.** FoxA1–mCherry was introduced into a bacterial expression vector with an N-terminal His<sub>6</sub> tag using Gateway cloning. Unlabelled FoxA1 was cloned and purified the same way. This vector was transformed into T7 Express cells (enhanced BL21 derivative, NEB C2566I), grown to optical density (OD)  $\approx$  0.4–0.8, whereupon we added 1 mM isopropyl- $\beta$ -D-thiogalactoside and expressed His<sub>6</sub>–FoxA1–mCherry for 3–4 h at 37 °C. We thawed frozen pellets in binding buffer that contained 20 mM Tris-HCl (pH = 7.9), 500 mM NaCl, 20 mM imidazole and 1 mM MgCl<sub>2</sub>, supplemented with protease inhibitors and Benzonase. The redissolved pellets were lysed and clarified via centrifugation. Discarding the supernatant, we resuspended the pellets in binding buffer + 6 M urea, spun, collected the supernatant and poured it over an immobilized metal ion affinity chromatography column, eluting the protein with binding buffer + 6 M urea + 250 mM imidazole. We dialysed overnight into storage buffer, 20 mM HEPES (pH = 6.5), 100 mM KCl, 1 mM MgCl<sub>2</sub>, 3 mM dithiothreitol (DTT) and 5 M urea. Multiple dialysis rounds reduced the concentration of urea. Finally, the protein was dialysed into storage buffer + 2 M urea, spin-concentrated to 4–5 mg ml<sup>-1</sup> ( $\approx$  50  $\mu$ M), and then snap-frozen in nitrogen and stored at –80 °C. NH-FoxA1–mCherry and RR-FoxA1–mCherry were obtained following ref. <sup>15</sup> using a Q5 site-directed mutagenesis kit. The truncation constructs were generated using restriction digestion–ligation approaches coupled with PCR. We used Alexa-488-labelled somatic linker histone H1 purified from calf thymus (H-13188, Thermo Fisher). To purify mCherry–B4, the gene (GenScript) was cloned into a bacterial expression vector with N-terminal His<sub>6</sub> and mCherry tags, transformed into T7 Express cells, grown to OD  $\approx$  0.7, supplemented with 0.8 mM isopropyl- $\beta$ -D-thiogalactoside and expressed at 37 °C for 4 h. Resuspending the pellets in lysis buffer, 1 $\times$ PBS with 500 mM NaCl, 1 mM DTT plus protease inhibitors and Benzonase, we then lysed the cells, collected the supernatant, ran the supernatant over an immobilized metal ion affinity chromatography column and eluted the protein with lysis buffer + 250 mM imidazole. The protein was dialysed into 1 $\times$ PBS + 500 mM NaCl overnight, spin-concentrated, snap-frozen and stored at –80 °C. We purified labelled versions of TBP and Gal4–VP16 using similar purification strategies. Both vectors—His<sub>6</sub>–MBP–eGFP–zTBP and His<sub>6</sub>–Gal4–GFP–VP16—were transformed into T7 Express cells and grown to OD  $\approx$  0.6, whereupon we added 0.2 mM isopropyl- $\beta$ -D-thiogalactoside, and expressed overnight at 18 °C. We lysed the cells into buffer containing 50 mM Tris-HCl (pH = 8.0), 1 M NaCl, 10% glycerol, 1 mM DTT and 1 mM MgCl<sub>2</sub> supplemented with protease inhibitors. For subsequent steps, 10  $\mu$ M ZnSO<sub>4</sub> was added to buffers for the Gal4–VP16 purification. After lysis, we added NP40 to 0.1% and clarified via centrifugation. We performed a polyethyleneimine precipitation to precipitate DNA and then an ammonium sulfate precipitation to recover the protein, resuspending the precipitated proteins in buffer containing 50 mM Tris-HCl (pH = 8.0), 1 M NaCl, 10% glycerol, 1 mM DTT, 0.1% NP40 and 20 mM imidazole and clarified the soluble fraction via centrifugation. We poured the lysate over an immobilized metal ion affinity chromatography column and eluted the protein using 2 $\times$ PBS, 250 mM imidazole, 10% glycerol and 1 mM DTT. We pooled protein fractions and dialysed TBP overnight into 20 mM HEPES (pH = 7.7), 150 mM KCl, 10% glycerol and 1 mM DTT and Gal4–VP16 into 20 mM HEPES (pH = 7.7), 100 mM KCl, 50 mM sucrose, 0.1 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 1 mM DTT and 10  $\mu$ M ZnSO<sub>4</sub>. We then spin-concentrated the proteins, snap-froze them using liquid nitrogen and stored at –80 °C.

**DNA functionalization, coverslip PEGylation and DNA microchannel preparation.** To biotinylate DNA purified from  $\lambda$ -phage, we followed the protocol given in ref. <sup>19</sup>. Each end of the biotinylated  $\lambda$ -phage DNA had two biotin molecules. To PEGylate the cover slips and prepare the DNA microchannels we followed the protocol given in ref. <sup>19</sup>.

**DNA and protein imaging.** We fluorescently stained immobilized DNA strands with 10 nM SYTOX Green in Cirillo buffer (20 mM HEPES, pH = 7.8, 50 mM KCl, 2 or 3 mM DTT, 5% glycerol, 100  $\mu$ g ml<sup>-1</sup> BSA). For experiments with H1 and TBP, we imaged DNA using 25 nM SYTOX Orange. We used protein concentrations of 10 nM. We used a Nikon Eclipse microscope with a Nikon  $\times$ 100/numerical aperture (NA) 1.49 oil SR Apo total internal reflection fluorescence microscope and an Andor iXon3 EMCCD camera using a frame-rate of 100–300 ms. A highly inclined and laminated optical sheet was established using a Nikon Ti-TIRF-E unit mounted on the microscope stand.

**Optical-tweezer measurements.** We performed optical-tweezer experiments using a C-Trap G2 system (LUMICKS) in a microfluidics flowcell (LUMICKS), providing separate laminar flow channels. For each experiment, we trapped two streptavidin-coated polystyrene beads (Spherotech SVP-40-5). Once trapped, we moved these beads to a channel containing biotinylated  $\lambda$ -phage DNA (LUMICKS) at a concentration of 0.5  $\mu$ g ml<sup>-1</sup>, whereupon we used an automated ‘tether-finder’ routine to capture a single molecule between the two beads. Once a single  $\lambda$ -phage DNA molecule was attached to the two beads, we moved the trapped beads to a buffer-only channel (containing Cirillo buffer with 3 mM DTT). In the buffer-only channel, we fixed  $L$  at either 6 or 8  $\mu$ m. We then moved the tethered DNA to a channel containing 150 nM FoxA1 in Cirillo buffer or another buffer-only

channel (as a control) and tracked the force and imaged the FoxA1–mCherry fluorescence for 100 s.

**Bulk phase-separation assays.** We performed bulk phase-separation assays with FoxA1–mCherry, NH-FoxA1–mCherry and somatic linker histone H1. The storage buffer for FoxA1 and NH-FoxA1 was 20 mM HEPES (pH = 6.5), 100 mM KCl, 1 mM MgCl<sub>2</sub>, 3 mM DTT and 2 M urea. The storage buffer for H1 was 1 $\times$ PBS. For FoxA1, we combined 6  $\mu$ l of FoxA1 (at 50  $\mu$ M) and 1  $\mu$ l of 20% 30K PEG. For NH-FoxA1, we combined 9  $\mu$ l and 1  $\mu$ l of 20% 30K PEG. For H1, we combined 9  $\mu$ l H1 and 1  $\mu$ l 100  $\mu$ M 32-base-pair single-stranded DNA. We prepared flow channels with double-sided tape on the cover slide and attached a PEGylated coverslip to the tape. We imaged the condensates using spinning disc microscopy and a  $\times$ 60 objective.

**FoxA1 molecule number estimation.** To estimate the number of FoxA1–mCherry molecules per condensate, we quantified the intensity of single FoxA1–mCherry molecules bound non-specifically to the slide. Around each segmented spot of DNA-independent FoxA1 intensity, we cropped an area of 10 pixels  $\times$  10 pixels, performed a background subtraction and summed the remaining intensity in the cropped area. To determine the contribution of the background, the same method was applied to 10 pixel  $\times$  10 pixel areas void of FoxA1 signal intensity. The resulting distribution of these integrated signal intensities reveals consecutive peaks that are evenly spaced by an average intensity of about 400 a.u., allowing us to calculate the number of molecules. This approach should be interpreted as a lower-bound estimate of the number of FoxA1–mCherry molecules per condensate, as it neglects effects such as fluorescent quenching<sup>27</sup>.

**Hydrodynamic stretching of DNA.** DNA molecules bound at only one end to the slide were hydrodynamically stretched using a constant flow rate of 100  $\mu$ l min<sup>-1</sup> of 0.5 nM FoxA1–mCherry in Cirillo buffer with 10 nM SYTOX Orange. The flow rate was sustained for tens of seconds using a programmable syringe pump (Pro Sense NE-501).

**Strand length calculation.** To calculate the end-to-end distance, we generated time-averaged projections of FoxA1 and DNA and integrated these projections along the strand’s orthogonal axis. To find the profile’s ‘left’ edge, we computed the gradient of the signal and determined the position where the gradient went through a threshold (defined as 0.2). We then took all the points from the start of the signal to this position, performed a background subtraction, and fitted an exponential to these points. To ensure that we included the entire DNA signal, we defined the fitted threshold for both the left and the right edges as three-quarters of the value of the fitted exponential value at the point when the gradient had gone through the intensity threshold. Using this fitted threshold, we computed the position values for the left and the right sides, and computed the end-to-end distance as the difference between these two positions.

**Global cross-correlation analysis.** We generated time-averaged projections from videos of both FoxA1 and DNA, and then summed the intensities in the orthogonal axis to the strand, generating line profiles. We then calculated the strand length and cropped both the FoxA1 and DNA line profiles from the edges of the strand. We then subtracted the mean value from these cropped line profiles, normalized the amplitudes of the signals by their Euclidean distances and computed the zero-lag cross-correlation coefficient of the normalized signals, which we defined as ‘correlation’:  $R(\tau = 0) = \sum_{n=1}^N \bar{x}_n \bar{y}_n$ , where  $\tau$  is the number of lags,  $N$  is the number of points in the normalized FoxA1 and DNA signals,  $\bar{x}_n$  is the  $n$ th entry of the normalized FoxA1 signal and  $\bar{y}_n$  is the  $n$ th entry of the normalized DNA signal. In general, correlation values range from –1 to 1, but in our experimental data the values range from roughly 0 to 1, where 1 represents the formation of DNA–FoxA1 condensates and 0 represents the formation of only FoxA1 condensates (no DNA condensation).

**DNA envelope width calculation.** To compute the DNA envelope width, we first generated time-averaged projections from videos of FoxA1 and DNA. We then selected segments of the strand that did not contain FoxA1—regions of non-condensed DNA. Using these segments, we extracted a line profile of the DNA signal orthogonal to the strand that gave the maximum width. We then subtracted the background of the DNA profile, normalized the signal’s amplitude using the Euclidean distance and fitted a Gaussian. We defined the DNA envelope width as  $\sqrt{2}\sigma$ , which represents the square root of twice the s.d. of the fitted Gaussian. The theoretical diffraction limit is calculated using the Rayleigh criterion, a measure of the minimal resolvable distance between two point sources in close proximity for a given set of imaging conditions:  $d = \frac{0.61\lambda}{NA}$ , where  $\lambda$  represents the imaging wavelength. For our imaging set-up,  $d = 0.2 \mu$ m, which is approximately  $2\sigma$  of the fluorescent source from the DNA. As the DNA envelope width is defined as  $\sqrt{2}\sigma$ , our ‘diffraction limit’ as given by the dashed line in Fig. 1f is 0.14  $\mu$ m.

**Condensate volume analysis.** To calculate condensate volumes, we generated time-averaged DNA–FoxA1 projections and then localized the peaks of the DNA condensates. Using the peak locations, we extracted background-subtracted

one-dimensional profiles of the DNA condensates in the orthogonal axis to the strand—these profiles went through the peak location. We fitted Gaussians to these profiles without normalizing the amplitude. To define the radii of the condensates, we computed the gradients of the fitted Gaussians and defined the condensate ‘edges’ as when the absolute value of the gradient of the Gaussian function gradient went through a threshold value (defined as unity, determined by comparing with fluorescence). Assuming that condensates are spherical, we computed the condensate volume as  $V = \frac{4}{3}\pi R^3$ . To compute a condensate volume for strands with multiple condensates, we simply added up the volumes for each condensate.

**Condensed DNA length analysis.** To compute  $L_d$ , we generated time-averaged projections of DNA and FoxA1 signals, integrating the DNA signal in the orthogonal direction to the strand. We then defined condensed versus non-condensed DNA by  $\text{Threshold}_{\text{drop}}$ : the median value of the profile plus a tolerance. Intensity values below  $\text{Threshold}_{\text{drop}}$  were defined as pixels of non-condensed DNA, and intensity values above  $\text{Threshold}_{\text{drop}}$  were defined as pixels of condensed DNA. This assumption was also consistent with the measured FoxA1 signal, where FoxA1 signals clearly localized to regions of condensed DNA, as defined by  $\text{Threshold}_{\text{drop}}$ . The tolerance value was used to suppress artefactual fluctuations of the non-condensed DNA signal in the neighbourhood of the median. To optimize the tolerance value, we assume that  $L_d$  as a function of  $L$  is linear for lower values of  $L$  ( $<5\mu\text{m}$ ) with a  $y$  intercept equal to the contour length of the DNA molecule ( $16.5\mu\text{m}$ ), as this is consistent with our theoretical description. We plotted the  $y$  intercepts of the linear fits as a function of tolerance and found that tolerance = 500 gives a  $y$  intercept equal to 16.5 and generates DNA–FoxA1 condensates up to  $10\mu\text{m}$ , consistent with our data and analysis (Extended Data Fig. 5). To calculate the DNA length contained within the droplet, we integrated the intensities from pixels above  $\text{Threshold}_{\text{drop}}$  divided this value by the sum of the total intensity of the profile and then multiplied this ratio by the contour length of  $\lambda$ -phage DNA,  $16.5\mu\text{m}$ . The non-condensed DNA length was calculated as simply the contour length minus  $L_d$ . We used the same tolerance = 500 for the NH-FoxA1 mutant analysis.

**Force analysis.** To calculate the force that the condensate exerts on the non-condensed DNA, we used the worm-like-chain model, which relates  $\lambda$ -phage DNA's extension and force. Upon addition of FoxA1, the amount of non-condensed DNA reduces, and the extension changes as follows:  $E = \frac{L}{L_c - L_d}$ . We then directly compute the force using the worm-like-chain model,

$$F = \kappa \left( \frac{1}{4} (1 - E)^{-2} - \frac{1}{4} + E \right).$$

**Condensation probability analysis.** To calculate the probability of the formation of a protein–DNA condensate as a function of end-to-end distance, we localized the peaks of the FoxA1 condensates from time-averaged projections of FoxA1 and DNA. We then extracted  $0.9\mu\text{m} \times 0.5\mu\text{m}$  windows centred around the localized FoxA1 peaks of both the FoxA1 and DNA signals—with the window's long axis going with the strand and the short axis orthogonal to the strand. We then computed the zero-lag normalized cross-correlation coefficient as follows:

$$C_{\text{loc}} = \frac{\sum (f(x, y) - \mu_f) (g(x, y) - \mu_g)}{\sqrt{\sum f(x, y)^2 - \mu_f^2} \sqrt{\sum g(x, y)^2 - \mu_g^2}}$$

where  $f(x, y)$  is the DNA,  $g(x, y)$  is FoxA1,  $\mu_f$  is the mean of the DNA image and  $\mu_g$  is the mean of the FoxA1 image. This generates values from  $-1$  to  $1$ . For FoxA1-mediated DNA condensation, the values for particular condensates are close to  $1$ . When FoxA1 fails to condense DNA, owing to the morphology of the underlying DNA strand and the small number of pixels, we obtain values that range from  $-1$  to roughly  $0.5$ . To obtain a value for  $P_{\text{cond}}$  as a function of end-to-end distance, we selected a threshold of  $0.75$ — $C_{\text{loc}}$  values above the threshold are considered ‘condensed’ and values below ‘non-condensed’. We binned the  $C_{\text{loc}}$  data in  $2\mu\text{m}$  increments as a function of end-to-end distance, and calculated  $P_{\text{cond}}$  by taking the number of condensed condensates and dividing it by the total number of condensates within the bin. The confidence intervals for  $P_{\text{cond}}$  in each respective bin are computed by computing the 95% confidence interval of a beta distribution, which represents the probability distribution for a Bernoulli process that takes into account the total number of successes with respect to the total number of attempts.

**Parameter fitting of the thermodynamic description and confidence intervals.**

To fit  $\alpha$ , we used a linear fit of the condensate volumes for individual strands as a function of  $L_d$ . The confidence intervals are the 95% confidence intervals generated from directly fitting the points. To fit  $\gamma$  and  $\nu$ , we minimized the error of the average  $\overline{L_d}(L)$  and  $P_{\text{cond}}(L)$  with respect to the data to optimize the parameter values. We used the normalized Boltzmann distribution  $P(L_d) = \frac{e^{-\beta F(L_d)}}{\int_0^{L_c-L} e^{-\beta F(l)} dl}$  to calculate  $\overline{L_d} = \int_0^{L_c-L} l P(l) dl$ . To compute  $P_{\text{cond}}(L)$ , we localized the position of the local maximum in the free energy,  $L_d^{\text{max}}$ , for a given  $L$  and then computed

the probability to ‘not’ nucleate a droplet from the Boltzmann distribution  $\int_0^{L_d^{\text{max}}} P(l) dl$ , which gives  $P_{\text{cond}} = 1 - \int_0^{L_d^{\text{max}}} P(l) dl$ . To minimize the error, we binned the data in  $2\mu\text{m}$ -width bins. For each ‘binned’ mean for both condensed DNA and condensation probability, we computed the squared residual of the mean value with respect to the theoretical expression. For residuals calculated from  $\overline{L_d}(L)$ , we normalized each residual by the squared standard error of the mean, and then summed the normalized residuals to obtain the error. For residuals calculated from  $P_{\text{cond}}(L)$ , we normalized each residual by the variance of the beta distribution,  $P_{\text{nuc cond}}^{\sigma^2} = \frac{(1+k)(1-k+N)}{(2N^2(3+N))}$  and then summed the normalized residuals to obtain the error. For the global error, we simply added the error from both deviations in  $\overline{L_d}(L)$  and  $P_{\text{cond}}(L)$ . We then iterated through a range of values for  $(\gamma, \nu)$  and computed the total error associated with each set of parameter values, exponentiated the negative values of the total error matrix and computed the largest combined value to select the parameter values. To calculate the parameters' confidence intervals, we obtained one-dimensional profiles of the integrated exponentiated total error for  $\nu$  as a function of  $\gamma$  and  $\gamma$  as a function of  $\nu$ . The peaks of these profiles represented the values that we selected for our best-fit parameters. We assumed that these profiles represented probability distributions for parameter selection, and then calculated the left and right bounds, where the area under the curve between these bounds represented 95% of the area. These left and right bounds represent the lower and upper values of our confidence intervals. To compute the 95% confidence interval for the force for each end-to-end distance value, we scanned through  $(\gamma, \nu)$  parameter space and computed the value of  $L_d$  for each set of parameters. We then plotted these values against the probability that these parameter values were the ‘true’ values—simply the probability from the exponentiated error matrix. Integrating the points under the probability versus  $L_d$  curve and dividing this by the total area under this curve, we generated a probability distribution function from which we could compute the 95% confidence intervals for  $L_d$ . Because the force was constant, to compute the confidence intervals for the force we calculated the force using the worm-like-chain model with corresponding  $L_d$  values for an end-to-end distance that retained FoxA1-mediated DNA condensation. To compute the confidence intervals for  $L_{\text{crit}}$ , we scanned through  $(\gamma, \nu)$  parameter space and computed  $L_{\text{crit}}$  for each set of parameters. We then plotted  $L_{\text{crit}}$  values with the corresponding values from the probability that these parameter values were true (again, the exponentiated error matrix). Integrating the points under the probability versus  $L_{\text{crit}}$  curve and dividing this by the total area under this curve, we generated a probability distribution function from which we could compute the 95% confidence intervals for  $L_{\text{crit}}$ .

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Data generated and analysed supporting the findings of this manuscript will be made available upon reasonable request. Source data are provided with this paper.

## Code availability

Code generated supporting the findings of this manuscript will be made available upon reasonable request.

## References

- McCall, P. M. et al. Quantitative phase microscopy enables precise and efficient determination of biomolecular condensate composition. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.10.25.352823> (2020).
- Shakya, A., Park, S., Rana, N. & King, J. T. Liquid–liquid phase separation of histone proteins in cells: role in chromatin organization. *Biophys. J.* **118**, 753–764 (2020).
- Turner, A. L. et al. Highly disordered histone H1–DNA model complexes and their condensates. *Proc. Natl Acad. Sci. USA* **115**, 11964–11969 (2018).

## Acknowledgements

We thank P. Tomancak, A. Hyman, S. Grill, I. Patten, M. Loose, D. Oriola, B. Dalton, P. McCall and C. Meyer for helpful feedback and stimulating discussions. We thank A. Bogdanova for discussions and help with cloning as well as both the Protein Expression and Purification Facility and the Light Microscopy Facility at the Max Planck Institute of Molecular Cell Biology and Genetics. We acknowledge and thank N. Vastenhouw for discussions and the construct containing the TBP template, and C. Zechner for help with statistical analyses. Finally, we thank K. Zaret for sending constructs, information on mutant FoxA1 proteins and advice on FoxA1 protein purification. This work was supported by an EMBO long-term fellowship (ALTF-1456-2015) (T.Q.), DFG project BR 5411/1-1 (J.B., V.M.) and a Volkswagen ‘Life’ grant number 96827 (J.B., T.Q.).

## Author contributions

T.Q. and J.B. conceived the project. T.Q. and S.G. performed imaging experiments. S.G. established the single-strand DNA assay. T.Q. purified proteins, made constructs



and performed data analysis. T.Q., J.B. and F.J. performed theoretical calculations. M.E. made the TBP and Gal4–VP16 constructs and purified the proteins. V.M. purified B4. T.Q. and R.R. performed optical-tweezer measurements. S.G. and R.R. performed data analysis and contributed to writing Methods. K.I. made the initial FoxA1 construct and provided key biochemical support. J.B. and F.J. supervised the work. T.Q., J.B. and F.J. wrote the manuscript, and all authors contributed ideas and reviewed the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

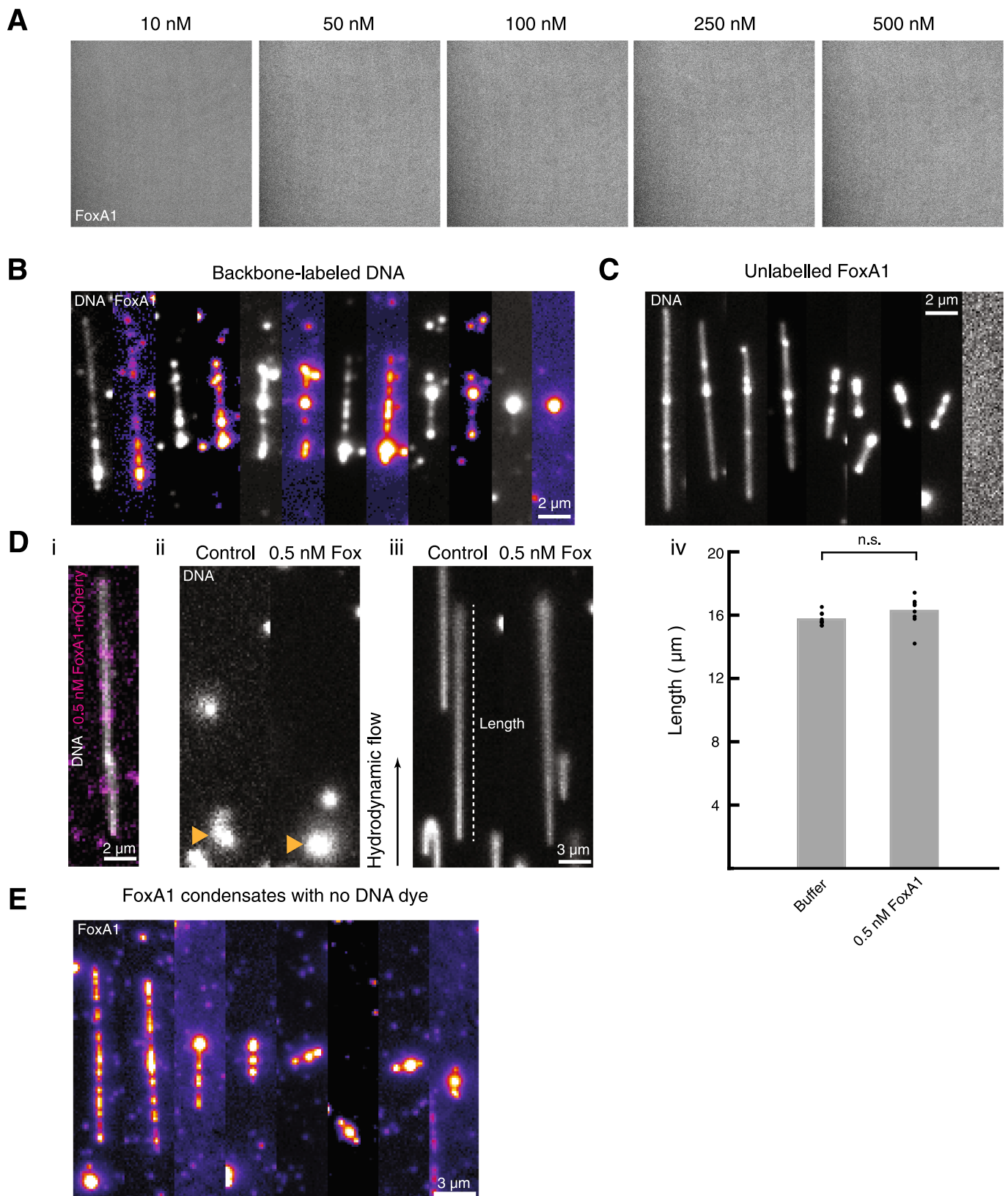
**Extended data** is available for this paper at <https://doi.org/10.1038/s41567-021-01285-1>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41567-021-01285-1>.

**Correspondence and requests for materials** should be addressed to F.J. or J.B.

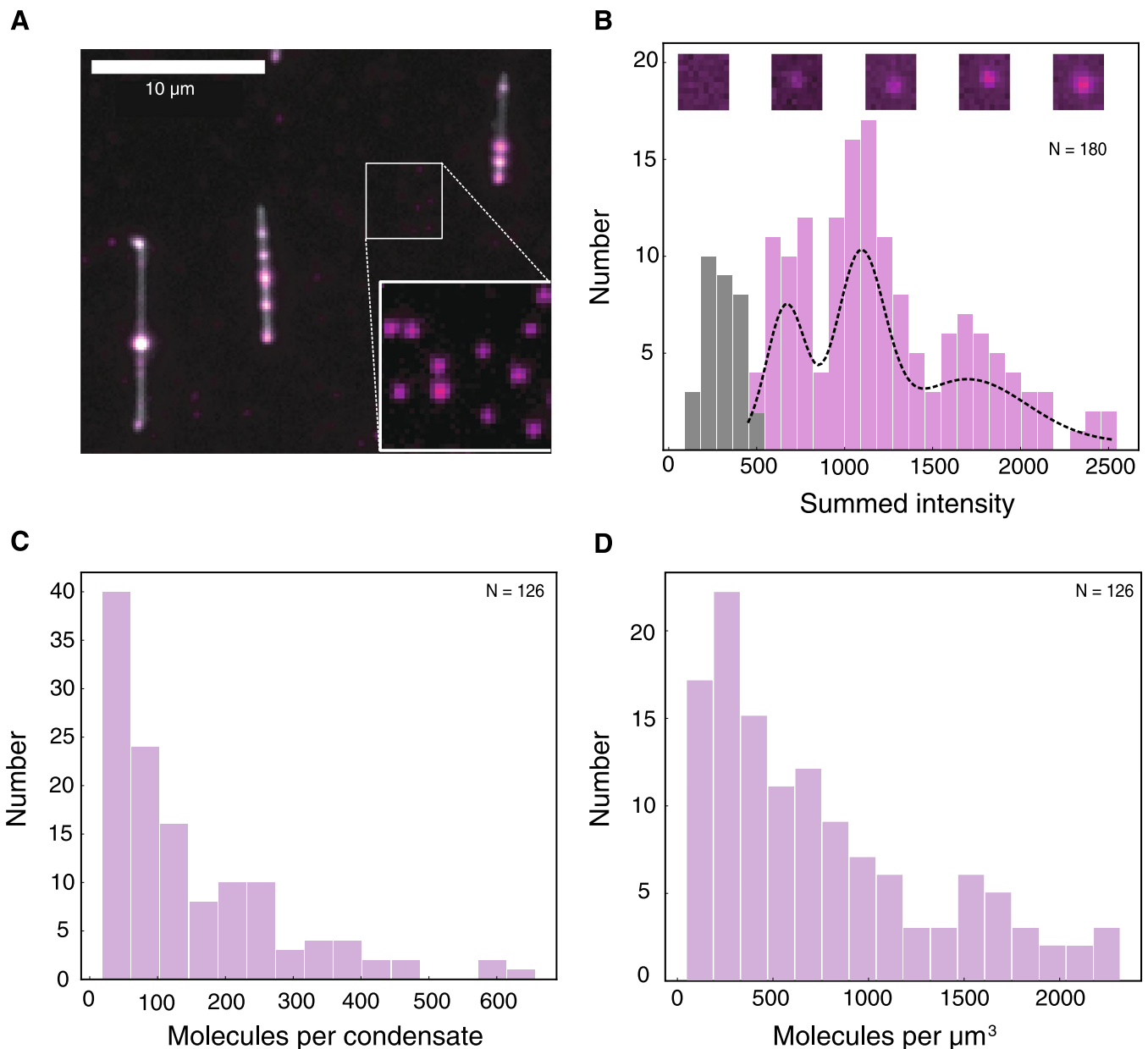
**Peer review information** *Nature Physics* thanks the anonymous reviewers for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

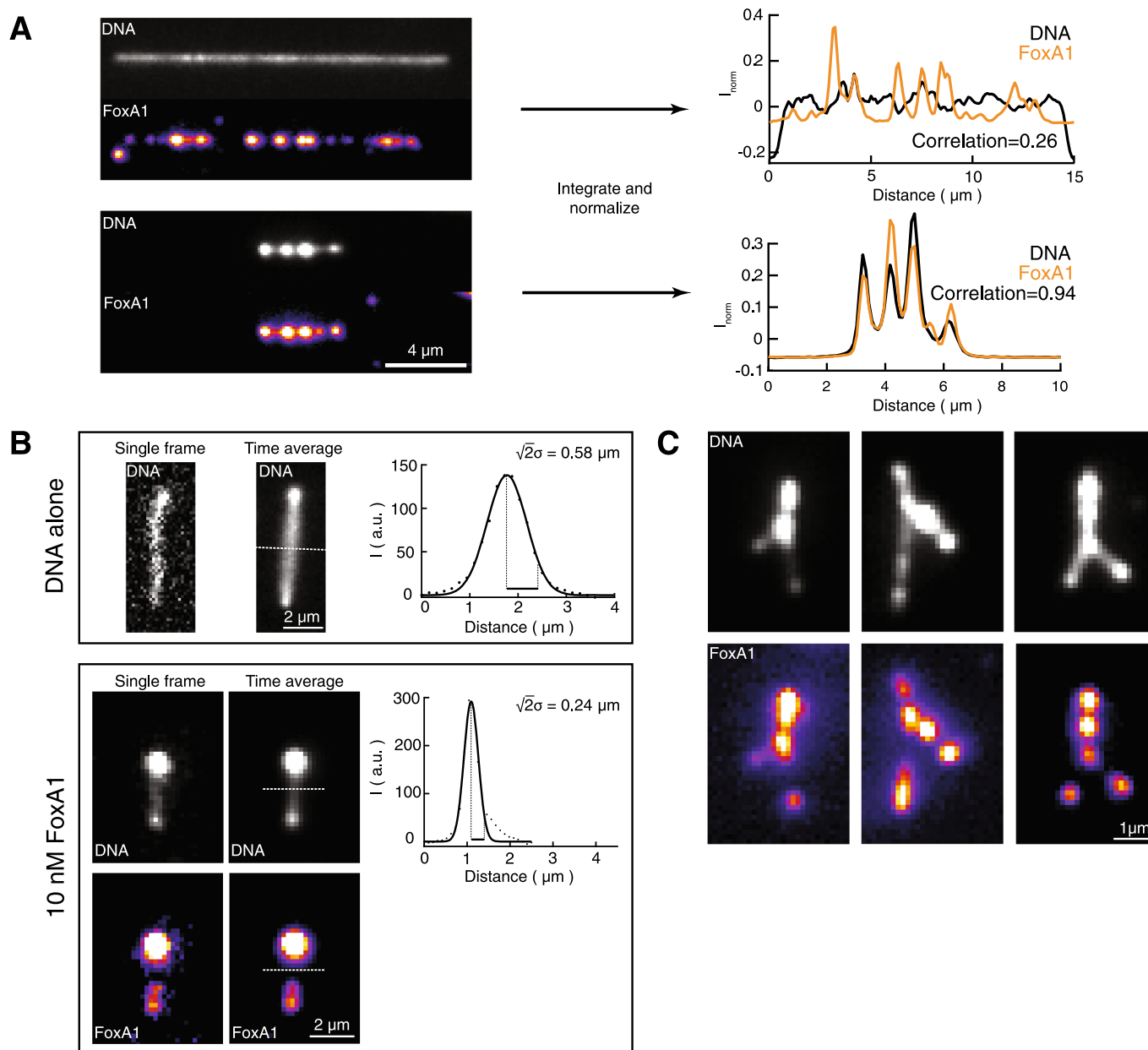


Extended Data Fig. 1 | See next page for caption.

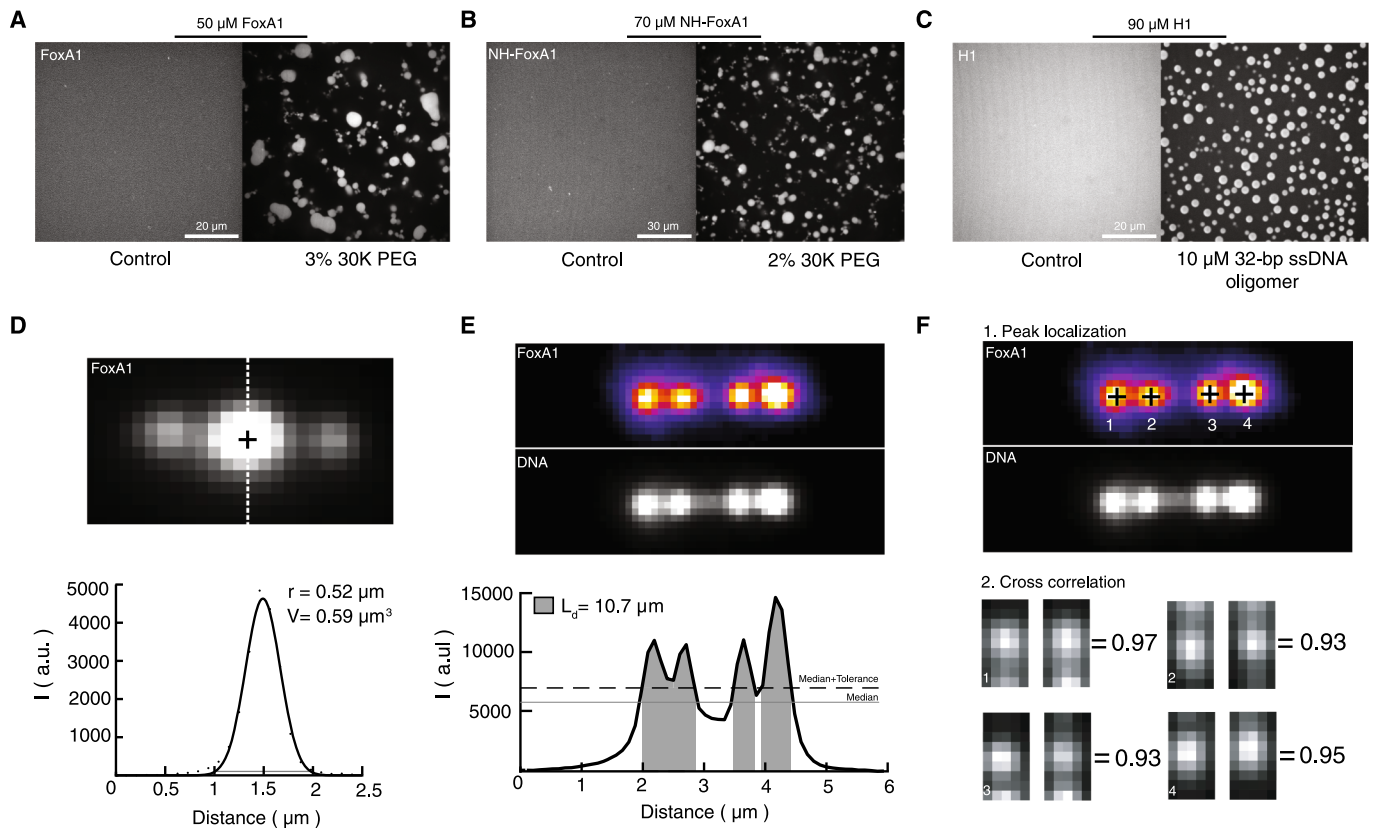
**Extended Data Fig. 1 | Experimental controls for FoxA1-mediated DNA condensation.** **a**, Representative fluorescent images of FoxA1-mCherry in buffer (20 mM HEPES, pH=7.8, 50 mM KCl, 2 mM DTT, 5% glycerol, 100  $\mu\text{g}/\text{ml}$  BSA) at different concentrations, 10–500 nM, in the absence of DNA reveals that FoxA1 does not form condensates in bulk at these concentrations. Using spinning disk microscopy and a 60x objective, we acquired images  $70\ \mu\text{m} \times 70\ \mu\text{m}$  in size with an exposure time of 250 msec and a time stamp of 500 msec to generate movies 30 seconds in duration. For all measured concentrations we generated  $n = 3$  movies and did not observe any FoxA1 condensation. **b**, FoxA1-mCherry condenses  $\lambda$ -phage DNA molecules with Cy5 dye covalently attached to the phosphate backbone of DNA (Label-IT Nucleic Acid Labeling Kit, Cy5, Mirus). **c**, Unlabeled FoxA1 condenses DNA (visualized with 10 nM Sytox Green). The rightmost panel is a representative image of the mCherry 561 nm imaging channel, revealing that the FoxA1 molecule does not have a mCherry fluorophore. **d**, Sparse labeling of FoxA1 (0.5 nM) does not influence the persistence length and contour length of  $\lambda$ -phage DNA, as determined by hydrodynamic stretching (see Methods). (i) FoxA1 (purple) is sparsely bound to DNA (in grey), visualized with 10 nM Sytox Green. (ii) Snapshots of unstretched DNA molecules bound at only one end to the coverslip before hydrodynamic stretching in both control and 0.5 nM FoxA1 conditions. The yellow arrows point to the DNA molecules. (iii) Snapshots of stretched DNA molecules bound at one end to the coverslip during hydrodynamic stretching in both control and 0.5 nM FoxA1 conditions. (iv) Quantification of stretched DNA lengths in both control ( $n = 10$ ) and 0.5 nM FoxA1 ( $n = 9$ ) conditions reveals that there is no significant difference in the length under hydrodynamic stretching (unpaired t-test,  $p = 0.11$ ). **e**, FoxA1 condensates imaged in the absence of DNA dye are consistent in size with that of FoxA1 condensates formed in the presence of DNA dye.



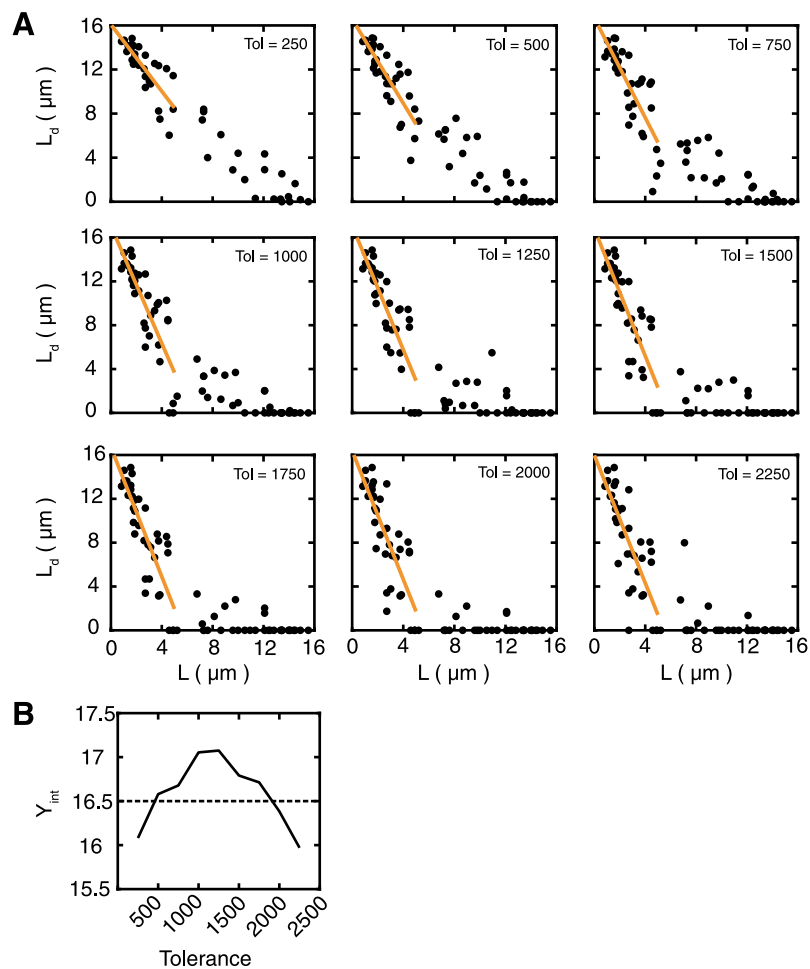
**Extended Data Fig. 2 | Counting FoxA1 molecules in condensates.** **a**, Representative image of three DNA strands with FoxA1 condensates. The inset shows an area of the PEGylated glass slide void of DNA. Increased contrast reveals the presence of individual spots of FoxA1 non-specifically bound to the coverslip. **b**, Histogram of integrated intensities of these DNA-independent FoxA1 to calibrate the amount of fluorescence per molecule. The grey bars represent the integrated background intensity of areas where no FoxA1 signal could be detected (maximum at 289 a.u.). Pink bars represent the integrated intensity of individual spots of DNA-independent FoxA1 signal. Black dotted line is a multi-Gaussian fit to the pink histogram, indicating consecutive peaks in the histogram at intensities of 683, 1096 and 1706 (a.u.), suggesting an integrated intensity of 400 a.u. per FoxA1 molecule. Representative images (10×10 pixels) of background (left) and individual DNA-independent FoxA1 spots used in this analysis are placed above the histogram according to their integrated signal intensity. **c**, Histogram of the number of FoxA1 molecules in FoxA1 condensates on DNA, calculated on the basis of an integrated intensity of 400 a.u. per FoxA1 molecule, determined in **(b)**. The mean number of molecules is 150 per condensate. **d**, Histogram of the density of FoxA1 molecules in the FoxA1-DNA condensates analyzed in **(c)**. The mean value is 750 molecules per  $\mu\text{m}^3$ . These estimates represent lower bounds as previous studies have demonstrated that fluorescent-based methods for estimating the number of molecules neglect effects such as quenching and can underestimate the number of molecules by as much as 50 fold<sup>27</sup>.



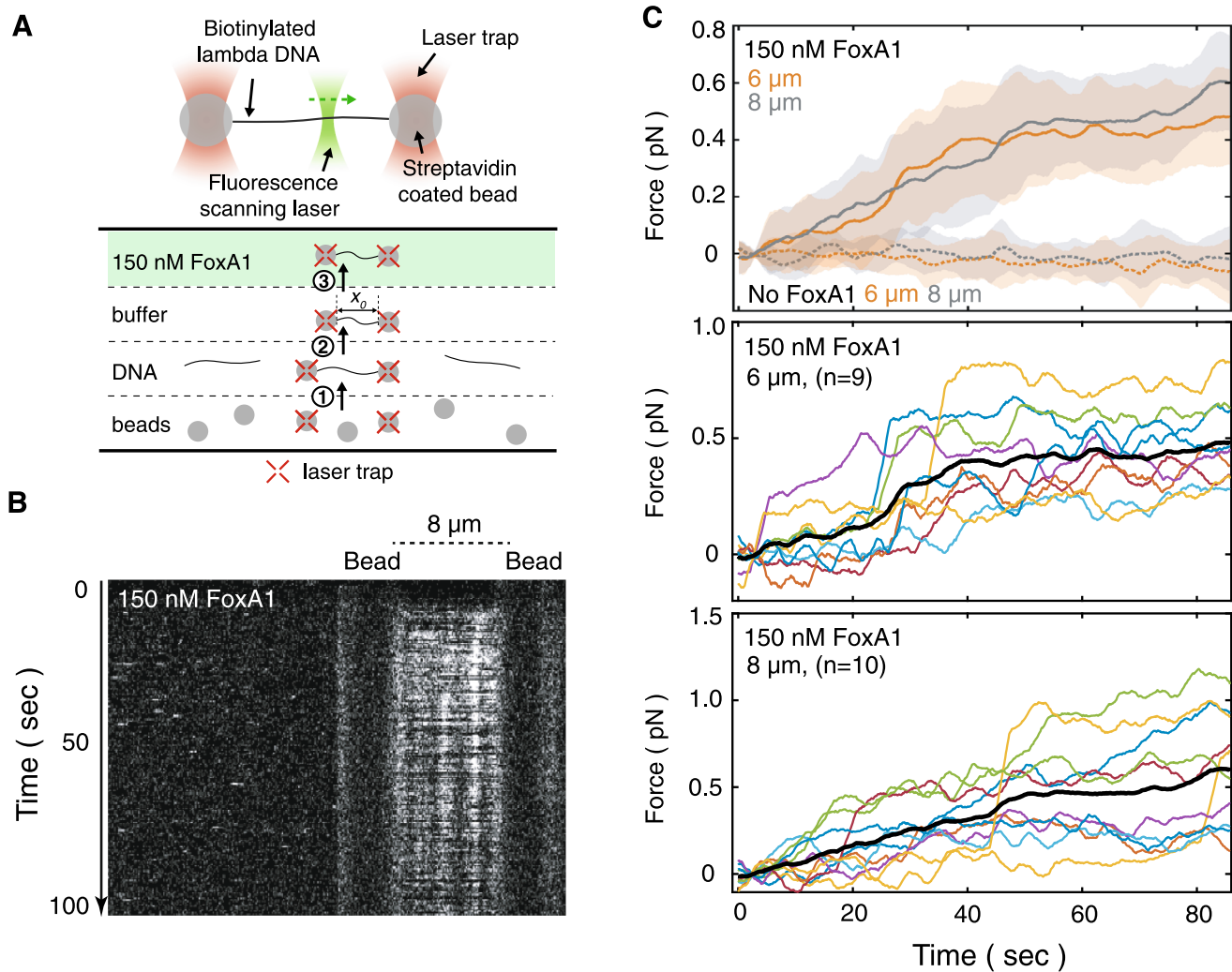
**Extended Data Fig. 3 | Quantification of FoxA1-mediated DNA condensation.** **a**, Global cross-correlation between FoxA1 and DNA reveals FoxA1-mediated DNA condensation. Left, representative fluorescent time-averaged projections of DNA and FoxA1 at two different end-to-end distances. Integrating both the DNA and FoxA1 signals along the axis orthogonal to the long axis of the strand gave rise to line profiles, which we normalized and then plotted as a function of distance (DNA in black and FoxA1 in orange). We then computed the zero-lag cross-correlation coefficient defined as ‘Correlation’ (see Methods). **b**, DNA envelope width measure measures the transverse fluctuation of non-condensed DNA. Top box: DNA alone condition. Bottom box: DNA + FoxA1 condition. For both conditions, we display representative fluorescent images of single frames and time-averaged projections of the DNA and FoxA1 signals. The white dashed line represents the maximum width of the DNA signal along the orthogonal axis of the non-condensed DNA. The black dots in the profile represent the background-subtracted points from the white dashed line, and the black line represents a Gaussian fit. The DNA envelope width was defined as  $\sqrt{2}\sigma$ , where  $\sigma$  is the standard deviation of the Gaussian fit. **c**, Three representative examples of FoxA1-mediated zipping. These images are time-averaged projections of both FoxA1 and DNA.



**Extended Data Fig. 4 | Bulk biomolecular condensate formation and quantification of condensate volume, condensed DNA length, and condensation probability.** **a**, Three per cent 30 K PEG triggers FoxA1 condensate formation in bulk at 50  $\mu\text{M}$  in storage buffer: 20 mM HEPES (pH = 6.5), 100 mM KCl, 1 mM  $\text{MgCl}_2$ , 3 mM DTT, and 2 M Urea. **b**, Two per cent 30 K PEG triggers NH-FoxA1 condensate formation in bulk at 70  $\mu\text{M}$  in storage buffer. **c**, The addition of 10  $\mu\text{M}$  32-BP ssDNA oligomers nucleated droplets of H1 in bulk at 90  $\mu\text{M}$  that exhibited features of liquid-like droplets consistent with literature<sup>28,29</sup>. These data demonstrate that H1-DNA form liquid-like condensates, which could be driven via transient crosslinking of H1 and DNA or H1-H1 interactions. Both mechanisms are accounted for in our free energy description. **d**, Condensate volume quantification of a representative time-averaged projection of a FoxA1-DNA condensate, where the black cross is the condensate peak location and the white dashed line is the intersecting profile to measure the volume. Lower panel: the black dots are the profile's background-subtracted values and the solid black line is a Gaussian fit. The gray line represents the threshold value computed from the gradient of the Gaussian function that defines the edges of the condensate (see Methods). **e**, Condensed DNA length quantification of a representative time-averaged projection of FoxA1 and DNA. Below: the integrated one-dimensional DNA profile is defined into condensed versus non-condensed regions using the median of the profile's median (gray) plus a tolerance (black dashed). **f**, Local correlation quantification of a representative time-averaged projection of FoxA1 and DNA. The condensates were localized (black crosses) and then  $0.9 \mu\text{m} \times 0.5 \mu\text{m}$  boxes centered around these peaks were cropped. The correlations between the cropped regions of FoxA1 (left) and DNA (right) were then computed.

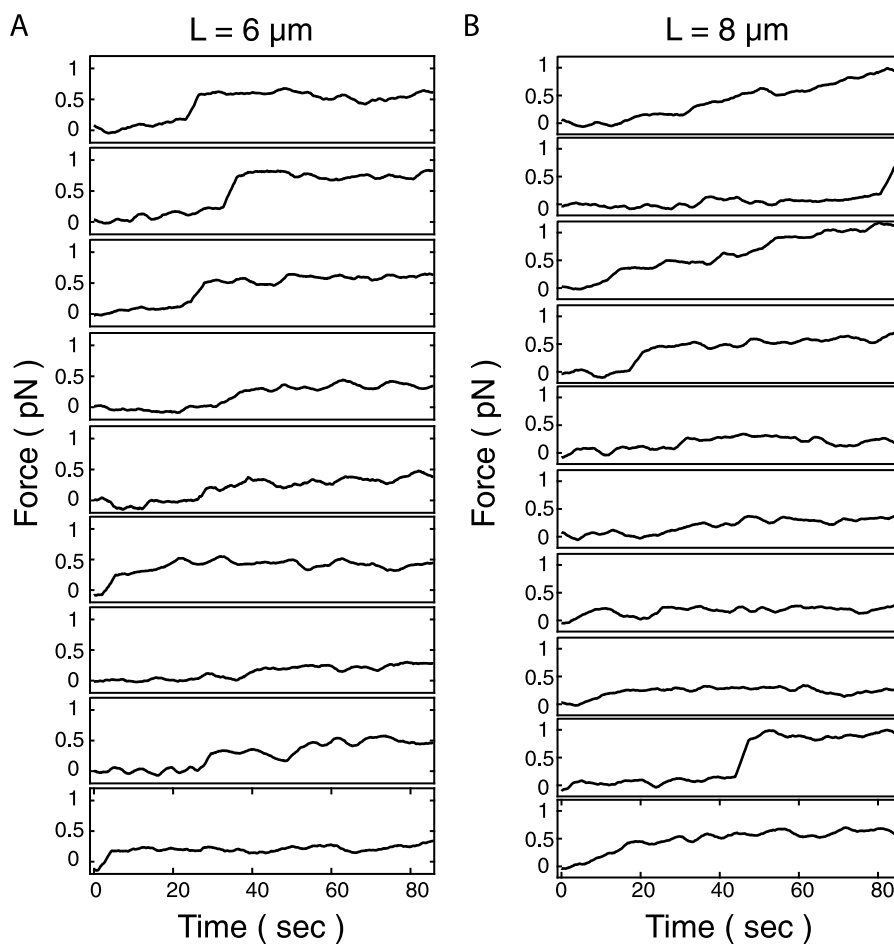


**Extended Data Fig. 5 | Tolerance value calculation.** Quantification of the condensed DNA length as a function of end-to-end distance for a range of tolerance values. Condensed DNA length is computed by defining regions of condensed versus non-condensed DNA using a threshold composed of the signal's median value plus a tolerance. **a**, Condensed DNA length is plotted as a function of end-to-end distance  $L$  for tolerance values from 250 to 2250 where the black dots represent the condensed DNA length for individual strands and the orange curve represent linear fits to these points for end-to-end distances below  $5 \mu\text{m}$ . **b**, Y intercept of the fitted linear curves. A tolerance=500 was selected as the y intercept was equal to the contour length of  $\lambda$ -phage DNA ( $16.5 \mu\text{m}$ ) and gave FoxA1-DNA condensate formation up to approximately  $10 \mu\text{m}$ , consistent with experimental observations (see Methods).

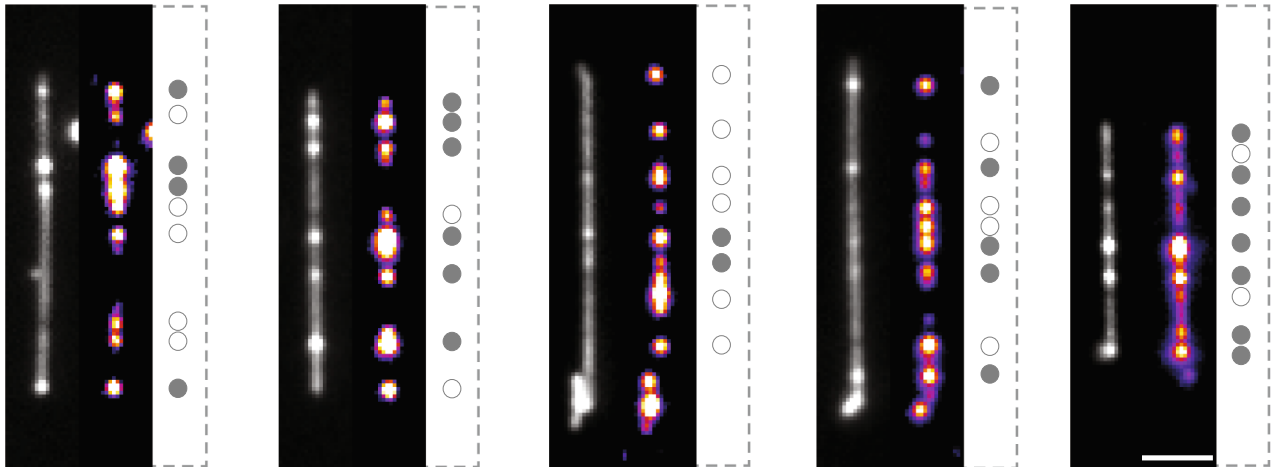
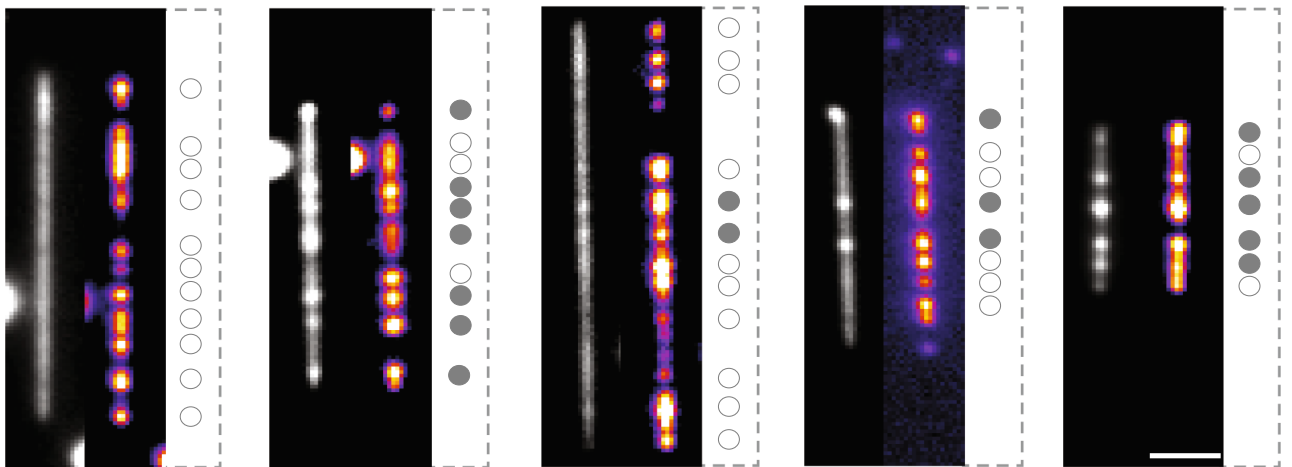


**Extended Data Fig. 6 | Optical tweezer measurements reveal that FoxA1 generates forces on the order of 0.4–0.6 pN.** **a**, Schematic outlining optical tweezer experimental design (see Methods). **b**, Representative kymograph reveals that FoxA1 condensates co-localize with a single molecule of  $\lambda$ -phage DNA trapped between two beads at an end-to-end distance of 8  $\mu\text{m}$ . **c**, Force trajectories for single DNA molecules reveal forces on the order of 0.4–0.6 pN when in FoxA1-containing buffer. (Top panel) This panel displays the mean  $\pm$  STD of force trajectories for each condition ( $n=9$  for +FoxA1 with  $L=6\ \mu\text{m}$ ,  $n=10$  for +FoxA1 with  $L=8\ \mu\text{m}$ ,  $n=10$  for control with  $L=6\ \mu\text{m}$ , and  $n=13$  for control with  $L=8\ \mu\text{m}$ ). This average force is slightly higher than what we measured in Fig. 3F using fluorescence, though a comparison of the relative errors reveals that both measurements give rise to comparable forces close to their respective detection limits and within the error bars. Additionally, the optical tweezer measurements were performed at a higher FoxA1 concentration—this was due to the large amount of tubing from the entry port to the flowcell in the custom-built Lumicks system, representing a considerable amount of surface for the protein to non-specifically bind to. We found that 150 nM FoxA1 was necessary to elicit a force response and to observe FoxA1 condensate formation on DNA. We conducted these measurements in the presence of 150 nM FoxA1 in Cirillo buffer 20 mM HEPES, pH=7.8, 50 mM KCl, 3 mM DTT, 5% glycerol, 100  $\mu\text{g}/\text{ml}$  BSA (solid lines) and in the presence of Cirillo buffer only (hatched lines) at end-to-end distances of  $L=6$  (orange) or 8  $\mu\text{m}$  (grey). Individual force trajectories for  $\lambda$ -phage DNA in the presence of buffer containing 150 nM FoxA1 with an initial end-to-end distance of 6  $\mu\text{m}$  (middle panel) and 8  $\mu\text{m}$  (bottom panel) reveal jumps in force, consistent with a first-order phase transition. These trajectories are re-plotted for clarity in Extended Data Fig. 7.

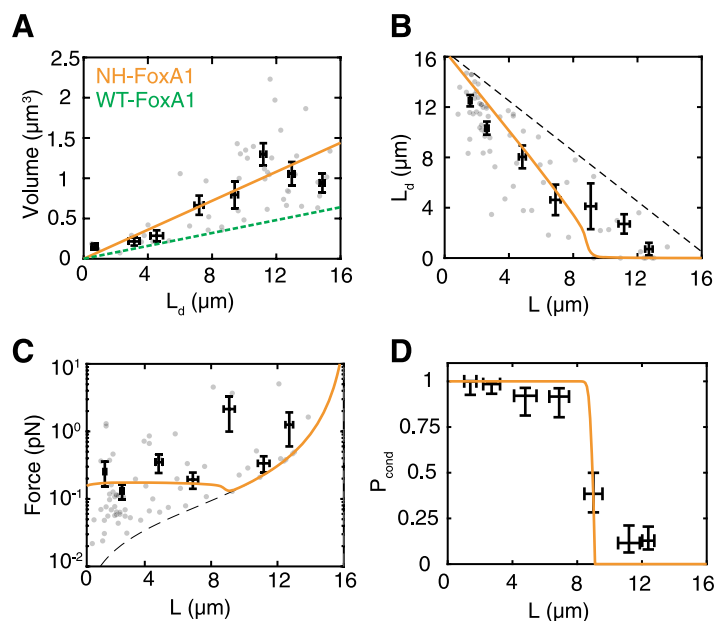




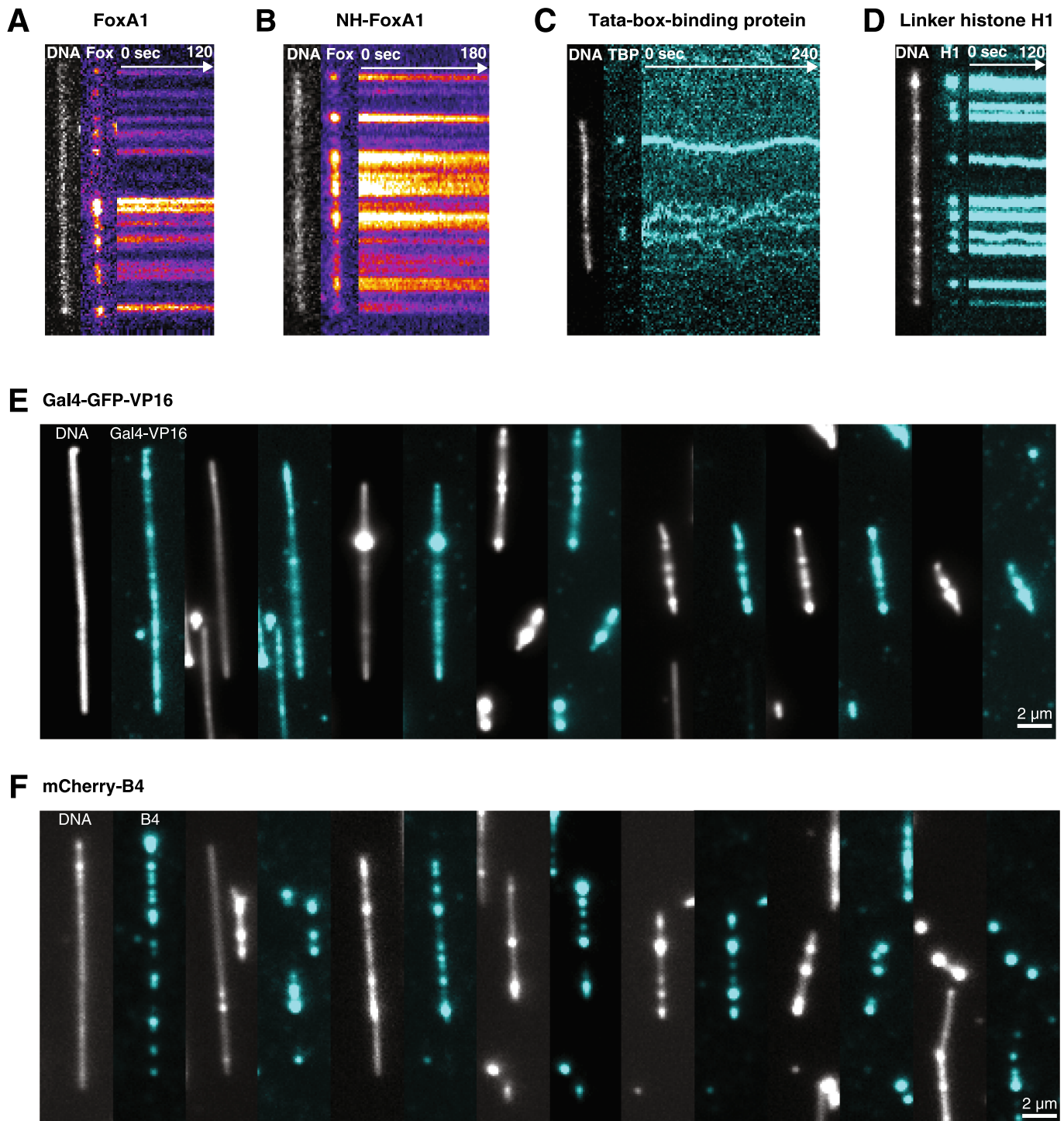
**Extended Data Fig. 7 | Individual temporal optical tweezer force measurements.** Temporal force measurements from optical tweezers with an initial end-to-end distance of  $6\ \mu\text{m}$  ( $n=9$  strands) (**a**) and  $8\ \mu\text{m}$  ( $n=10$  strands) (**b**) in the presence of  $150\ \text{nM}$  FoxA1. These data are the same as in Extended Data Fig. 6c, and are re-plotted individually for clarity.

**A** FoxA1**B** NH-FoxA1

**Extended Data Fig. 8 | Bistability of FoxA1-mediated DNA condensation.** **a**, Representative time-averaged projections of DNA and FoxA1 signals show that FoxA1 condenses DNA in an all-or-nothing manner. On the right side of each pair of images, we localized the FoxA1 condensates and showed whether FoxA1 condenses DNA (filled-in gray circle) or not (open circle). Interestingly, there is a mixed population, revealing the bistable nature of the condensation process. **b**, Representative images of condensation bistability for the sequence-specific DNA-binding mutant, NH-FoxA1. Scale bars = 2  $\mu\text{m}$ .



**Extended Data Fig. 9 | Quantification of NH-FoxA1-mediated DNA condensation.** **a**, Condensate volume as a function of condensed DNA length ( $L_d$ ). The grey dots represent individual strands ( $n=47$ ) and the data is binned every  $2\ \mu\text{m}$  (mean  $\pm$  SEM). The individual data are points are fit with a linear curve with a slope of  $0.09\ \mu\text{m}^2$  given in orange. The green dashed line is the WT-FoxA1 fit (slope =  $0.04\ \mu\text{m}^2$ ). **b**, Condensed DNA length as a function of end-to-end distance. The black dots represent individual strands ( $n=70$ ) and the data is binned every  $2\ \mu\text{m}$  (mean  $\pm$  SEM). The orange curve is the expression computed from the theoretical description with parameter values determined through error minimization (see Methods). The black hatched line represents the DNA's contour length ( $16.5\ \mu\text{m}$ ) minus the end-to-end distance. **c**, The force that the condensate exerts on the non-condensed DNA as a function of end-to-end distance. The grey dots represent individual strands ( $n=68$ ) and the data is binned every  $2\ \mu\text{m}$  (mean  $\pm$  SEM). The orange curve is the expression computed from the theoretical expression of  $L_d$  versus  $L$  from panel B for the force. NH-FoxA1 generates forces at roughly  $0.17\ \text{pN}$ . The dashed black line represents the force exerted on the non-condensed strand when  $L_d=0$ . **d**, Probability for NH-FoxA1 to form a DNA-FoxA1 condensate reveals a sharp transition at a critical end-to-end distance. Local correlations of individual FoxA1 condensates with DNA (Extended Data Fig. 4c) are calculated, binned into  $2\text{-}\mu\text{m}$ -width bins, and  $P_{\text{cond}}$  is calculated (see Methods). There are a total number of  $n=361$  condensates used for this analysis. The dashed lines represent the  $P_{\text{cond}}$  value as computed within the bin with  $\pm$  SD for the strand's end-to-end distance. The confidence intervals for  $P_{\text{cond}}$  are computed by computing the 95% confidence interval of a beta-distribution (see Methods). The orange curve represents  $P_{\text{cond}}$  computed from the theoretical description with parameter values determined through error minimization.



**Extended Data Fig. 10 | Dynamics of DNA-binding proteins.** **a**, Representative images of FoxA1 condensates on DNA. The kymograph reveals FoxA1 condensates do not move on DNA. **b**, NH-FoxA1 condensates remain stable on DNA and do not move. **c**, TBP condensates exhibit diffusive-like behavior on DNA. **d**, Similar to FoxA1 condensation, H1 condensates do not exhibit diffusive-like behavior on DNA. **e**, Representative images of Gal4-GFP-VP16-mediated DNA condensation. DNA was imaged with 10 nM Sytox Orange. **f**, Representative images of mCherry-B4-mediated DNA condensation. DNA was imaged with 10 nM Sytox Green.

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

**Data collection** Using a Nikon eclipse Ti-E Microscope, we extracted time lapses of DNA and fluorescently labeled FoxA1 (as well as mutant FoxA1 proteins, TBP, and H1) using NIS-AR Software. Optical tweezer data was collected using a C-Trap G2 system (Lumicks).

**Data analysis** Images were cropped using the image analysis tool FIJI. FIJI was also used sometimes for generating intensity profiles. Custom-written code was generated using MATLAB to analyze these images (and Python to analyze Extended Data Fig. 2). Theoretical fitting was also performed using scripts written in MATLAB.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Our manuscript contains a data availability statement. In addition, source data files are available for Figures 1-4 and Extended Data Figs. 1,2,5,6,7,9.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample size was determined by ensuring that we had enough single strand examples to sample the end-to-end distances from approximately 1 micron to 16 um per condition. This allowed us to bin the data as a function of the end-to-end distance and calculate the mean and the standard error (or standard deviation, depending on the panel) for this binned data. The optical tweezer data was collected in order to visualize the difference between FoxA1 treatment and control defined by non-overlapping standard deviations after 80 seconds incubation.
Data exclusions	Single strands of DNA were excluded if they were stuck to the coverslip or if they were broken, which we determined via clear morphological features.
Replication	Reproducibility was ensured by repeating these experiments with FoxA1 on several different days and using FoxA1 purified the same way but from different cultures. Furthermore, mutant FoxA1 experiments were performed as well as experiments with structurally similar proteins to understand our main finding. In total, we analyzed nine separate proteins. Comparing these experiments with the wild-type protein allowed us to ensure that the main finding was not simply an artefact that arose from problems with the protein. All attempts to replicate our data were successful.
Randomization	This was not relevant for our study.
Blinding	This was not relevant for our study.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging