
Supplementary information

Force generation by protein–DNA co-condensation

In the format provided by the
authors and unedited

Supplementary Information

1 Thermodynamic description of DNA-protein condensation

We consider the free energy associated with nucleating a condensate that contains DNA and FoxA1. The free energy of this process contains volume and surface contributions of the DNA-protein condensate as well as the free energy of the DNA polymer outside the condensate,

$$F(L, L_d) = F_d(L_d) + F_p(L, L_d), \quad (1)$$

where L is the end-to-end distance of the DNA, L_d is the length of condensed DNA, F_d is the free energy of the condensate, and F_p is the free energy of the DNA polymer outside the condensate. Assuming that the DNA co-condenses with the protein to form a dense condensed phase with defined volume fraction of DNA, the droplet volume and the length of condensed DNA are linearly related, $V = \alpha L_d$, or $R^3 = \frac{3\alpha}{4\pi} L_d$, where $1/\alpha$ describes the DNA packing density given as DNA length per condensate volume. We can then obtain the condensate free energy of nucleating a condensate as a function of L_d and end-to-end distance L as

$$F_d(L_d) = -\nu\alpha L_d + \gamma 4\pi \left(\frac{3\alpha}{4\pi}\right)^{2/3} L_d^{2/3}, \quad (2)$$

where ν is the condensation free energy per volume, and γ is the surface tension of the condensate. The free energy of the polymer $F_p(L, L_d)$ is related to the external force applied to pin the free DNA polymer and its associated chemical potential by

$$f_{\text{ext}} = \left. \frac{\partial F_p}{\partial L} \right|_{L_d}, \quad \mu_p = -a \left. \frac{\partial F_p}{\partial L_d} \right|_L, \quad (3)$$

where a is the length of a base pair. The force-extension relation for λ -phage DNA has been extensively studied previously, and here we use the phenomenological force-extension curve of the worm-like-chain model for λ -phage DNA (14) with contour length, L_c (for λ -phage DNA $L_c = 16.5 \mu\text{m}$). If a length L_d of the DNA is condensed, the extension of the non-condensed strand is $x = \frac{L}{L_c - L_d}$. The force on the strand then can be expressed as

$$f_{\text{ext}} = \frac{k_B T}{P} \left(\frac{1}{4} \frac{1}{(1-x)^2} - \frac{1}{4} + x \right), \quad (4)$$

where k_B is the Boltzmann constant, T is the temperature, and P is the persistence length of DNA. For what follows we define $\kappa = \frac{k_B T}{P}$. From this expression of the force and its relation to the free energy of the DNA polymer (equation 3), we can obtain the free energy of the DNA polymer outside the condensate as $F_p(L, L_d) = \int_0^L dL' f_{\text{ext}}(L', L_d)$, leading to

$$F_p(L, L_d) = \kappa \left(\frac{1}{4} \frac{(L_c - L_d)^2}{L_c - L_d - L} - \frac{1}{4} L + \frac{1}{2} \frac{L^2}{L_c - L_d} - \frac{1}{4} (L_c - L_d) \right) \quad (5)$$

The total free energy associated with nucleating a FoxA1-DNA condensate on a DNA strand reads:

$$F(L, L_d) = -\nu \alpha L_d + \gamma 4\pi \left(\frac{3\alpha}{4\pi} \right)^{2/3} L_d^{2/3} + \kappa \left(\frac{1}{4} \frac{(L_c - L_d)^2}{L_c - L_d - L} - \frac{1}{4} L + \frac{1}{2} \frac{L^2}{L_c - L_d} - \frac{1}{4} (L_c - L_d) \right) \quad (6)$$

The equilibrium between condensate and polymer is given by $\frac{\partial F}{\partial L_d} = 0$, which is equivalent to equilibrating the chemical potentials of the condensate and free polymer,

$$a \frac{\partial F}{\partial L_d} = \mu_d(L_d) - \mu_p(L, L_d) = 0, \quad (7)$$

with

$$\mu_p(L, L_d) = -a\kappa \left(\frac{1}{4} + \frac{L^2}{2(L_c - L_d)^2} + \frac{(L_c - L_d)^2}{4(L_c - L - L_d)^2} - \frac{L_c - L_d}{2(L_c - L - L_d)} \right) \quad (8)$$

Using the expression for the total free energy, we can vary the length L_d of condensed polymer and obtain profiles for the free energy as a function of L_d , which depend on the end-to-end distance L (see Fig. 3b). For L values close to 0—where the strand is not under tension—we observe that there is a minimum of F for L_d close to L_c . This means that, at this end-to-end distance, FoxA1 has mediated the generation of a FoxA1-DNA condensate using almost all of the DNA in the strand. As L increases, however, the local minimum shifts to lower values of L_d and ultimately F at the minimum becomes higher than the free energy without condensate $F(L_d = 0)$, giving rise to a branch of metastable states. For even higher L values, the metastable state disappears and the global minimum is at $L_d = 0$ (Fig. 3b). This sharp transition corresponds to a first-order phase transition. Simple scaling arguments are useful to generate intuition for the conditions necessary for condensate formation, and for the condensate to pull DNA. Briefly, there are three energy scales associated with this problem: the energy associated to create a droplet, which is $\nu\alpha L$; the surface energy of scale $4\gamma\pi \left(\frac{3\alpha}{4\pi}\right)^{2/3} L^{2/3}$; and lastly the energy scale associated to the non-condensed polymer $L \frac{k_B T}{P}$. First, to create a droplet, $\frac{\nu(4\pi)^{2/3}(\alpha L)^{1/3}}{4\pi 3^{2/3}\gamma} > 1$. Once condensation is favorable, in order for the droplet to pull DNA, $\frac{\nu\alpha P}{k_B T} > 1$. Notably, fitting the parameter values (see Methods) demonstrated that, at low L , the free energy gained by the system is on order of 1-2 $k_B T$, implying that stochasticity

is relevant for the condensation process. To account for the inherent stochastic nature of the condensation, we compute the probability of nucleating a DNA-protein condensate of size L_d using Boltzmann distributions from the corresponding energy profiles,

$$P(L_d) = \frac{e^{-\beta F(L, L_d)}}{\int_0^{L_c-L} dL'_d e^{-\beta F(L, L'_d)}} \quad (9)$$

where $\beta = \frac{1}{k_B T}$. To determine the relationship between L_d and L , we compute the mean L_d value of these Boltzmann distributions: $\bar{L}_d = \int_0^{L_c-L} L_d P(L_d) dL_d$ which then allows us also to calculate the magnitude of the condensation forces using the worm-like chain model given in Eq. (4).

2 Supplementary Information Figures

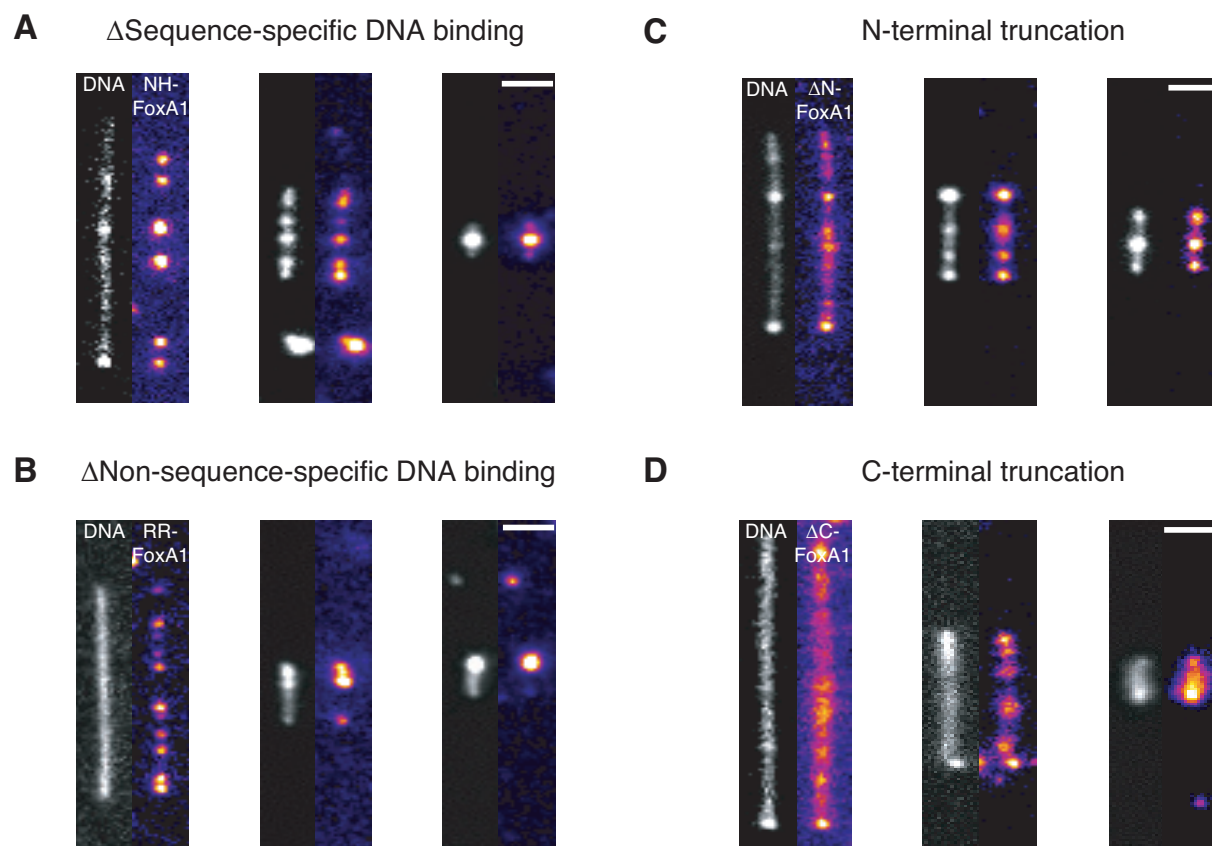


Figure 1: Representative images of sequence-specific-binding NH-FoxA1 mutant (A) non-sequence-specific-binding RR-FoxA1 mutant (B) N-terminal FoxA1 truncation (C) and C-terminal FoxA1 truncation (D). The scale bars are 2 μ m. DNA is imaged with 10 nM Sytox Green. Note that the contour length of each DNA molecules is constant (16.5 μ m) but the end-to-end distance is different.

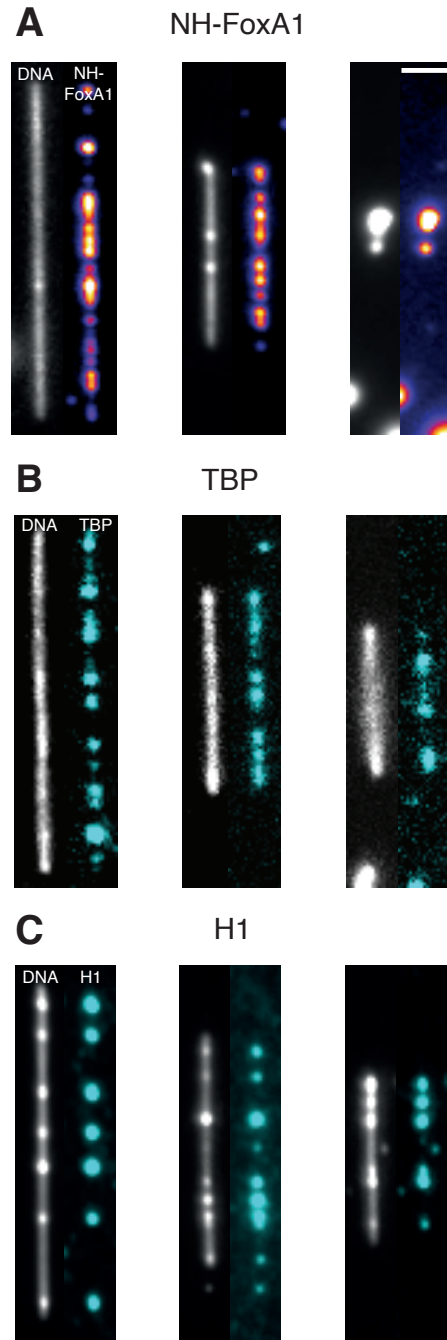


Figure 2: Representative images for NH-FoxA1 (A) Tata-box-binding protein (B) and somatic linker histone H1 (C). The images are time-averaged projections of movies for NH-FoxA1 and H1 but single images for TBP owing to TBP's diffusivity. The scale bar is 2 μm .