Empirical modeling of a single polymer collapsing under the entropic pressure of neutral crowders

Quentin Chaboche,[ac] Gerardo Campos-Villalobos,[e] Giuliana Giunta,[ef] Marjolein Dijkstra,[e] Marco Cosentino-Lagomarsino[cd‡] and Vittore F. Scolari[ab‡]

July 12, 2023

Abstract

Crowding by macromolecules in the surrounding environment can collapse a single long polymer to globular form through depletion forces of entropic nature. For instance, the genomic DNA of the bacterium *Escherichia coli* can be compacted by crowding, and the current consensus is that crowding is needed in order to fold it into the "nucleoid", occupying a well defined region of the cell. Motivated by the biological importance of this problem, several theoretical/computational studies have attempted to identify the main determinants of the polymer-crowders phases. However, we still lack a unique quantitative description of the collapse transition of a polymer driven by crowders. In particular, different studies based on simulations with explicit crowders proposed different candidate order parameters predicting this transition. Here, integrating results from published simulations from different sources, we show their agreement, and we derive a unifying phenomenological model with predictive power. By performing new simulations, we validate our predictions for polymers of different size and we address the role of jamming by crowders.

1 Introduction

The mechanics underlying the conformation of bacterial DNA and its functions involve the physics of polymer collapse.[1–3]. In *Escherichia coli*, a genome of millimeter scale occupies a region of less than a μm^3 called the nucleoid[4]. Entropic crowding, the depletion effect exerted by small neutral molecules such as polyethylene glycol on the DNA as a polymer, has been shown to be necessary to explain the formation of a compact nucleoid *in vitro*[5–8]. *In vivo*, osmotic shock and mechanical cell-size perturbations confirm this picture, and show that a 30% increase in the crowders (ribosomes and proteins) concentration can lead to a 3-fold decrease in nucleoid size[9], while association of DNA-binding proteins also plays a role in this process.

In order to explain these observations, several numerical simulation studies investigated the effects that neutral crowding agents can have on a polymer, described as a chain of beads. Since explicitly accounting for the crowder dynamics is computationally expensive, especially when crowders are small and many of them are needed, such models described this effect as an induced generic short-range attraction force[1,10]. In order to justify the effective attraction between the beads of the chain, and estimate its strength in function of the size and density of the crowders, these simulations relied on the Asakura-Oosawa theory, which is the classical framework in statistical physics to account for depletion forces[11,12].

Other numerical investigations instead explicitly accounted for the crowders as a gas of hard spheres[13–15], measuring their direct effects on the polymer folding. In all of these studies, the authors simulate the collapse of a polymer of monomer size D, mean distance between adjacent monomers $b \equiv D$, and total length of N monomers, in presence of crowders of diameter d and volume fraction of depletants ϕ. All the studies agree that an increase in crowding corresponds in a lowering of the solvent quality. In order to show this, they measure quantities related to the swelling ratio of the chain $\alpha = R(\phi)/R(\theta)$, defined as the ratio between the end-to-end distance of the polymer chain and its value in ideal conditions (when repulsive and attractive effects between monomers are balanced, and the polymer behaves as a ghost chain), as a function of ϕ. A sudden drop in α at increasing ϕ marks a polymer coil-to-globule transition, and this happens at different values for different d.

In chronological order, Shendruck and coworkers[13,16] conclude that the entropic crowding interaction is sufficient to trigger a second-order coil-to-globule transition, but that the Asakura-Oosawa theory is not sufficient to predict the form of the depletion pair potential. They propose, as a solution, the use of a multi-parameter morphometric thermodynamic approach for the pair potential. Kang et al.[14], instead, using simulations and scaling arguments,

[a] *Institut Curie, PSL Research University, Sorbonne Université, CNRS UMR168, Laboratoire Physico Chimie Curie, 75005 Paris, France.*
[b] *Institut Curie, PSL Research University, Sorbonne Université, CNRS UMR3664, Laboratoire Dynamique du Noyau, 75005 Paris, France.*
[c] *FIRC Institute of Molecular Oncology (IFOM), Milan, Italy.*
[d] *Physics Department, University of Milan, and INFN, Milan, Italy.*
[e] *Soft Condensed Matter, Debye Institute for Nanomaterials Science, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands.*
[f] *BASF SE, Carl-Bosch-Strasse 38, 67056 Ludwigshafen am Rhein, Germany*
‡ co-corresponding authors, E-mail: vittore.scolari@curie.fr and marco.cosentino-lagomarsino@ifom.eu

observe that the polymer collapse may or may not reach a globular phase. This leads to their scaling argument, suggesting that the swelling ratio of the chain is a unique function of a dimensionless control parameter defined as $x_{Kang} = R(0)\phi^{1/3}D/d$ (we will refer to this as the Kang parameter in the following). According to their prediction, all the curves of α as a function of ϕ that present the transition from coil to globule, collapse to one single curve if the abscissa is replaced by x_{Kang}. This analysis also implies that the effects of depletion are not compatible with the classic Asakura-Oosawa theory. Finally, Jeon et al.[15], produced a different set of simulation data which supports the thesis that the swelling of the chain is a unique function of a *different* dimensionless control parameter defined as $x_{Ha} = \phi D/d$ (which we will call the Ha parameter in the following). This order parameter is different with the one proposed by Kang and collaborators, and using approximations, they claim that this parameter is compatible with the Asakura-Oosawa theory; this contradicts the previous works simulating polymers with explicit crowders that had shown how Asakura-Oosawa alone cannot explain the full phenomenology of collapse. Additionally, they observe that the transition from coil to globule happens universally for values of x_{Ha} close to 1.

Interestingly, in these three different studies, the same kind of simulations giving the same quantitative results were evidently scrutinized using a different analytical paradigm by each author. This resulted in contradicting or incomparable interpretations. In this study, by gathering the simulation results from these studies and performing an integrated analysis, we understand what lead them to reach different conclusions. Further, we propose a simple effective mean-field theory that is able to describe all these simulation results and which can be generalized to different situations (confinement, ring polymers). Finally, we also perform molecular simulations of chains in explicit crowders to test our predictions in new regimes and to analyze the role of jamming of the crowders.

2 Methods

2.1 The Flory mean-field theory of real polymers

In polymer physics, the theory describing a chain of non-interacting monomers is called an ideal chain. An ideal chain is a random walk, and as such the mean end-to-end distance R is given by:

$$\langle R^2 \rangle = Nb^2. \tag{1}$$

A Flory-kind theory for the random walk, that is the one that would predict the mean end-to-end distance as the minimum of a free energy function, can be thus expressed by the following formula:

$$\beta F_{ideal} = \frac{3R^2}{2Nb^2} - 3\log\left(\frac{R}{Nb^2}\right), \tag{2}$$

Ref.	N	$\frac{D}{d}$	ϕ
Shendruck et al.[13]	15	3.00	< 0.45
Shendruck et al.[13]	15	4.00	< 0.45
Shendruck et al.[13]	15	5.00	< 0.45
Jeon et al.[15]	50	2.00	< 0.50
Jeon et al.[15]	50	2.50	< 0.50
Jeon et al.[15]	50	$3.\bar{3}3$	< 0.50
This paper	60	5.00	< 0.60
This paper	60	2.50	< 0.60
This paper	60	1.67	< 0.60
This paper	60	1.25	< 0.60
This paper	120	5.00	< 0.60
This paper	120	2.50	< 0.60
This paper	120	1.67	< 0.60
This paper	120	1.25	< 0.60

Table 1 List of simulation data used in this work. Each row indicates the study from which the simulation was taken, the number of monomers, the size ratio D/d between monomers and crowders and the range of crowder packing fraction values spanned in each work. In all the cases, the chains and crowders are modeled as described in Sec. 2.3.

where $\beta = (k_B T)^{-1}$ is the inverse temperature, with k_B the Boltzmann constant.

For a collapsing chain, we can define the swelling parameter as the ratio between the end-to-end distance of the chain and the mean end-to-end distance of the corresponding ideal chain $\alpha = R/\left(bN^{1/2}\right)$, resulting in:

$$\beta F_{ideal} = 3\left(\frac{\alpha^2}{2} - \log(\alpha)\right). \tag{3}$$

The expression of the free energy for a real polymer is obtained by adding the interactions between monomers and between monomers and the solvent ($F_{tot} = F_{ideal} + F_{interactions}$). The $F_{interactions}$ term is taken analogously as in the theory of real gases, as a virial expansion[17]:

$$F_{interaction} = \frac{1}{2}N\left(v\rho + w\rho^2\right) + \mathcal{O}(\rho^3), \tag{4}$$

where $\rho = k\frac{N}{R^3}$ is the monomer number density (with geometrical prefactor $k = \frac{3}{4\pi}$) and v and w are respectively the 2-bodies and 3-bodies monomer interaction. Without any attraction between monomers, these terms account for the steric repulsion between beads, and reduce to the monomer volume: $v_0 = \frac{4}{3}\pi D^3$ and $w_0 = \frac{1}{6}v_0^2$. The 3-body interaction w is set as w_0. Then the total free energy can be written as:

$$\beta F = 3\left(\frac{\alpha^2}{2} - \log\alpha\right) + \frac{\sqrt{N}}{2\alpha^3}\left(\frac{D}{b}\right)^3\frac{v}{v_0} + \left(\frac{D}{b}\right)^6\frac{1}{6\alpha^6}, \tag{5}$$

and an implicit equation for the swelling parameter can be obtained by minimization $\frac{dF}{d\alpha} = 0$. This yields the equation of state of the Flory mean-field theory of the real polymer, first obtained by De Gennes in 1975[18,19]:

$$\alpha^5 - \alpha^3 - \frac{1}{6\alpha^3}\left(\frac{D}{b}\right)^6 = \frac{1}{2}\left(\frac{D}{b}\right)^3\sqrt{N}\frac{v}{v_0}. \tag{6}$$

2.2 Swelling collapse for existing simulation data

The Shendruck dataset[13] is described in Table 1. The observable reported in their plots is the radius of gyration R_g. In order to obtain the swelling ratio, we first normalize the radius of gyration with the initial value: $\alpha_0 = R_g(\phi)/R_g(\phi = 0)$. The difference between α_0 and the swelling ratio α is the denominator, thus the relationship between α and α_0 is only set by a prefactor, independent on ϕ and d/D, but dependent on the polymer length N and the ratio between the adjacent monomers mean distance and monomer size $b/D \equiv 1$. This prefactor is equal to the value of α in the absence of crowders (where $\phi = 0$, $\alpha_0 = 1$ and $v/v_0 = 1$). We obtain it by solving Eq. 6 with $v/v_0 = 1$, obtaining the following relation:

$$\left(\frac{\alpha}{\alpha_0}\right)^5 - \left(\frac{\alpha}{\alpha_0}\right)^3 - \frac{1}{6}\left(\frac{\alpha}{\alpha_0}\right)^{-3}\left(\frac{D}{b}\right)^6 = \frac{1}{2}\left(\frac{D}{b}\right)^3\sqrt{N}\,,$$

allowing to express the dataset in terms of swelling ratios α.

The Ha group dataset[15] is also described in Table 1. The observable they measure is the normalized radius of gyration α_0. We obtain the swelling ratios α using the same approach used for the Shendruck dataset.

2.3 Molecular simulations of chains with explicit crowders

Here, we briefly describe the model and molecular simulation method employed to generate more simulation data on the collapse of DNA chains in the presence of explicit crowders. For consistency, we employ the same models as in Refs.[13,15]. In particular, the DNA is modeled as a linear chain of N spherical beads of diameter D, while the crowders are modeled as spheres of diameter d. The interactions between all sites (monomer-monomer, crowder-crowder and monomer-crowder) are described via a truncated and shifted Lennard-Jones potential, also referred to as the Weeks-Chandler-Andersen (WCA) potential. This purely repulsive pair potential reads:

$$u(r) = \begin{cases} u_{\mathrm{LJ}}(r) - u_{\mathrm{LJ}}(r_c) & r < r_c \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

where $u_{\mathrm{LJ}}(r)$ is the Lennard-Jones potential

$$u_{\mathrm{LJ}}(r) = 4\varepsilon\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right], \tag{8}$$

with r the distance between centers of bead i and bead j, ε and σ_{ij} the energy- and length-scales of the interaction. Here, we define $\sigma_{\mathrm{mm}} \equiv D$, $\sigma_{\mathrm{cc}} \equiv d$ and $\sigma_{\mathrm{mc}} = (\sigma_{\mathrm{mm}} + \sigma_{\mathrm{cc}})/2$. The cut-off radius is set to $r_c = 2^{1/6}\sigma_{ij}$. With this choice of r_c, the potential is smoothly continuous and purely repulsive.

Adjacent monomers in a polymer chain are bonded together using the Finite Extensile Nonlinear Elastic (FENE) potential:

$$u_{\mathrm{FENE}}(r) = -\frac{1}{2}k_0 r_0^2 \ln\left[1 - \left(\frac{r}{r_0}\right)^2\right], \tag{9}$$

with k_0 the spring constant and r_0 the range of the potential. Here, we use the conventional values for polymers; $k_0 = 30\varepsilon/D^2$ and $r_0 = 1.5D$, [20,21]. This potential is purely attractive, and competes with the purely repulsive WCA potential. The equilibrium point between these two potentials is $r \approx 0.96D$. This justifies our assumption that $b/D = 1$.

We perform Molecular Dynamics (MD) simulations in the canonical ensemble with a fixed number of beads (monomers and crowders) $N_{\mathrm{t}} = N + N_{\mathrm{c}}$, system volume V and temperature T. The mass of the monomers and crowders is considered as the mass unit m, while the units of length, energy and time are D, ε and $\tau_0 = D\sqrt{m/\varepsilon}$, respectively. The equations of motion are integrated using the velocity Verlet method with a timestep $\delta t = 0.005\tau_0$. We use a Langevin thermostat with a damping constant of $0.1\tau_0^{-1}$ to keep the temperature at $T = 1.0\varepsilon/k_B$. In all the simulations, we consider cubic simulation boxes of side length $L = 3R_g(\phi = 0)$, where $R_g(\phi = 0)$ is the radius of gyration of the chain in the absence of crowders. This box size, which is dependent on the polymer length, is selected so as to prevent artificial finite-size effects during the MD simulations, which are performed under three-dimensional periodic boundary conditions. We perform simulations of single chains of length $N = 60$ and 120 immersed in solutions where the volume fraction of the crowders is fixed by the χ_{Ha} parameter. In particular, we study systems where $0.15 \leq \chi_{\mathrm{Ha}} \leq 1.50$, for size ratios between crowders and monomers of $d/D = [0.2, 0.4, 0.6, 0.8]$. Unless stated otherwise, the total number of MD steps of the simulations is 5×10^7, with statistics on the radius of gyration and end-to-end distance of the chains collected every 1000 steps.

3 Results

3.1 Previously proposed dimensionless control parameters do not fully capture the collapse curves

The size of a polymer under the effect of molecular crowding depends a priori on a large set of variables jointly determining its macroscopic state (Fig. 1A). If there exist an unique dimensionless parameter, function of all these variables, unambiguously defining the size of the polymer, then the problem of predicting the compaction of the polymer would be greatly simplified, since all the swelling ratios would fall on a single master curve when plotted as a function of this parameter. Following this line of reasoning, we collected all the swelling ratio from the simulation results of Shendruck et al.[13] and Jeon et al.[15] (Table 1) and plotted them in function of the two order parameters x_{Kang} or x_{Ha} identified by previous literature (see Introduction). Fig. 1 shows that both the Kang parameter (Fig. 1B) and the Ha parameter (Fig. 1C) cannot fully define the swelling ratio of both datasets, since in each plot the swelling ratios of different simulations fall on different curves when plotted in function of the respective order parameter.

On the other hand, we observe that the swelling ratios

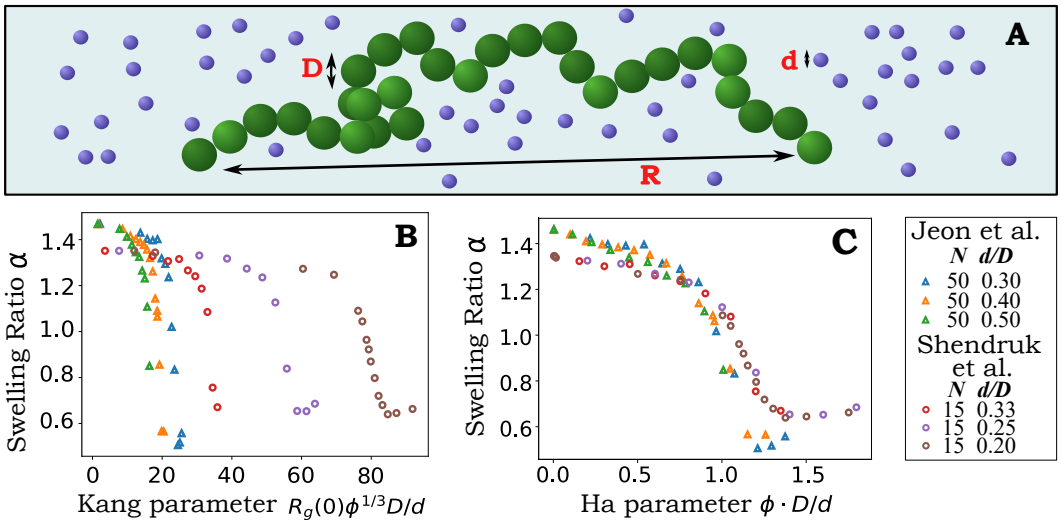

Figure 1 Swelling ratio in simulated crowder-induced single polymer collapse. **(A)** Sketch of the bead-spring polymer in a crowded medium of hard spheres, illustrating the main simulation parameters common to all the studies considered here. Monomers have size D and crowders have size d and volume fraction ϕ. **(B)** Swelling ratio α (simulations from refs.[13,15]), plotted as a function of the parameter x_{Kang} proposed by Kang and coworkers[14]. **(C)** Same data, plotted as a function of the the control parameter x_{Ha} proposed in refs.[15,22].

of polymers that share the same chain length N do collapse to a single curve as a function of x_{Ha}, but when comparing polymers of different lengths each of them fall on a different curve. Hence, we conclude that by simulating only the $N = 50$ polymer length, Jeon and coworkers did not account for the role of the chain length N in the swelling ratio of the crowder-induced polymer collapse. In fact, both at low concentration of crowders and in the conditions where crowding induces polymer collapse, the radius of gyration depends on the chain length N. This dependency is predicted by standard polymer theory, that at the mean-field level sets the radius of gyration scaling respectively as $\propto N^{3/5}$ for the polymer in good solvent, $\propto N^{1/2}$ at the transition, and $\propto N^{1/3}$ for the globular state[23–25]. In terms of swelling ratio, this corresponds to $\alpha(\phi = 0) \propto N^{1/10}$ and $\alpha(\phi \to 1) \propto N^{-1/6}$. We can thus rule out the existence of a single parameter such as x_{Ha} and x_{Kang} setting the value of the swelling ratios across the collapse transition: no single rescaling of the crowders volume fraction can account for the *different* scaling in polymer length N in the swollen compared to the globular phase.

3.2 The second-order virial coefficient is approximately set by $x_{Ha} = \phi D/d$ as the relevant dimensionless parameter

In face of the results of the previous section, we wondered if there could be an intensive (*i.e.* independent on the size N) quantity that could be captured, even approximately, by a single dimensionless parameter. It is possible to extract the quality of the solvent, evaluated by the second-order virial coefficient v, from the swelling ratios through the Flory mean-field theory[18,19]. This theory establishes the relationship between the polymer size N, α and v through

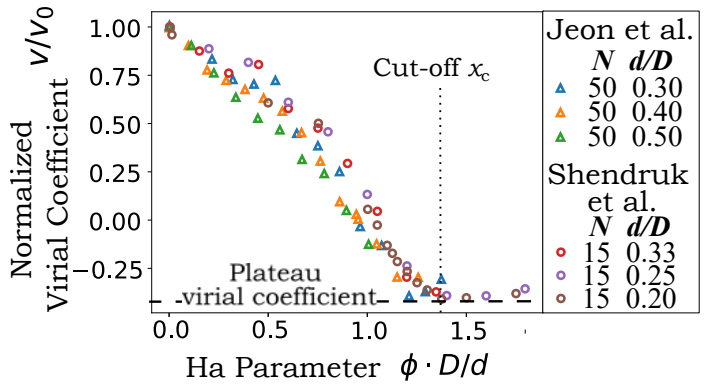

Figure 2 Second-order virial coefficient extracted from published simulation data (refs.[13,15]) plotted as a function of $x_{Ha} = \phi D/d$, identified by Ha and coworkers[15,22?]. This shows a unique behavior across the coil-globule transition.

the following equation of state:

$$\alpha^5 - \alpha^3 - \frac{1}{6\alpha^3}\left(\frac{D}{b}\right)^6 = \frac{1}{2}\left(\frac{D}{b}\right)^3 \sqrt{N}\frac{v}{v_0}, \qquad (10)$$

see Methods 2.1 for its derivation. The two first terms arise from the free energy of the ideal chain and the third term as well as the right hand side are a virial expansion accounting for the interactions between monomers. This formula considers solely second- and third-order terms in the virial expansion of an ideal chain. Both terms are independent of N, and can be taken as unspecified functions of D, d and ϕ when the solvent is solely composed by neutral crowders. Furthermore, since three-body collisions are infrequent in a polymer in a good solvent, the only relevant part of the third-order term is the contribution of the repulsive terms,

so that this term is relevant only when the polymer transitions to a globular form in a very bad solvent. Equally, higher-order interactions between chain monomers are neglected as they should be less relevant. While the derivation of Eq. 10 is not rigorous (since it has been shown that long-distance correlations in a polymer cannot strictly be ignored[24,26]), it is often a good approximation, for reasons that are more generic than its derivation[27]. Eq. 10 interpolates in one compact phenomenological formulation the three following asymptotic behaviors (each corresponding to a universality class): $R \propto N^{1/2}$ for the ideal chain phenomenology found close to the collapse transition ($v \approx 0$), $R \to N^{3/5}$ for the self avoiding chain found in good solvent conditions ($v \gg 0$) and $R \to N^{1/3}$ for the globule in a bad solvent ($v \ll 0$). Hence, we used Eq. 10 to derive v/v_0 from the simulated end-to-end distances.

Fig. 2 shows that all the v/v_0 extracted from the swelling-ratios in simulations using Eq. 10 are well defined by the Ha parameter (x_{Ha}), falling all approximately on a single curve if plotted against it. Most notably, the discrepancies in the asymptotic scaling for different values of N that were present in the plot of α *vs* x_{Ha} (Fig. 1C) are resolved when plotting of v/v_0 *vs* x_{Ha}. The value of v/v_0 is set unity for $x_{Ha} \to 0$ by definition (see Methods), and plateaus to a fixed value that we identify from the simulated data as $(v/v_0)_c \simeq -0.35$ for values of x_{Ha} greater than $x_c \simeq 1.25$, independently from any simulation parameters considered. In conclusion, the collapse of all datapoints on a single curve appears to be very satisfactory, although not perfect. However, we expect the numerical values for $(v/v_0)_c$ and x_c to be non-universal, and to depend on details in the potentials defining the simulated polymers and crowders, although no discrepancies were observed in any of the available simulations.

3.3 Derivation of an effective potential for the implicit solvent

The "inverse" analysis presented in Fig. 2 can be used to generate an effective "direct" mean-field theory. As we already mentioned in the Introduction, the theory of depletion attraction on spherical macromolecules in diluted conditions due to crowding developed by Asakura-Oosawa[11,12] is not applicable directly to the collpase of a polymer in presence of crowders. In that theory, entropic crowding is effectively included in an attractive pairwise short-range potential energy. Assuming that an effective, different, short-ranged potential can describe our system, we derived a simple formula for the effective potential energy between each pair of monomers. Specifically, we assumed a square-well potential $U(r)$, of range $D+d$ with hard core repulsion of size D as represented in Fig. 3A. While this discontinuous potential might seem like a crude approximation, it has the advantage of depending only on one parameter, βE_0, corresponding to its depth. From a fundamental perspective, the details of the shape of the potential are expected to be irrelevant as long as it is short ranged, and we made the reasonable assumption that the range of interactions corresponds to the size of the crowding agent d. The functional form for $E_0 = E_0(\phi, d/D)$ is then derived from the Mayer formula for the second-order virial coefficient[28]:

$$v/v_0 = \frac{1}{v_0}\int_{\mathbb{R}^3} \mathrm{d}\vec{r}\; 1 - e^{-\beta U(\vec{r})} = 1 + \frac{(d+D)^3 - D^3}{D^3}\left(1 - e^{\beta E_0}\right) \; ; \tag{11}$$

noting that if we assume that the virial coefficient collapses with x_{Ha}, then the interaction energy E_0 cannot be written as a function only one parameter x_{Ha}, but it needs to be a function of of d/D and ϕ separately.

From the plot of E_0, calculated from v/v_0 through Eq. 11, as a function of ϕ (Fig. 3B) we conclude that E_0 can be modeled as a linear function of ϕ that plateaus at the coil-globule transition, when $x_{Ha} = x_c$. From this observation, valid for the range of parameters considered in the simulations, and by the definition of the second order coefficient at the plateau transition $v/v_0 = (v/v_0)_c$, we can then calculate the first-order approximation in d/D of the value of this plateau,

$$\beta E_0^{\text{plateau}} = \log\left[1 - \frac{(v/v_0)_c}{(1+d/D)^3 - 1}\right] \approx C\frac{D}{d} \; . \tag{12}$$

By fitting equation 12 to the simulated data we find the constant $C = 0.223 \pm 0.003$. The numerical value of the constant C measured from the available simulations, combined with the fact that in absence of crowders there is no attractive interaction ($E_0 = 0$), suggests as a model the following empirical linear formulation:

$$\beta E_0 \approx \begin{cases} \beta E_0^{\text{plateau}} \cdot \frac{1}{x_c}\frac{D}{d}\phi & \text{for } \phi < x_c \frac{d}{D} \\ \beta E_0^{\text{plateau}} & \text{otherwise,} \end{cases} \tag{13}$$

and its approximation obtained through the first-order approximation on the right hand side of Eq. 12 is:

$$\beta E_0 \approx \begin{cases} \frac{C}{x_c}\left(\frac{D}{d}\right)^2 \cdot \phi & \text{for } \phi < x_c \frac{d}{D} \\ C\frac{D}{d} & \text{otherwise.} \end{cases} \tag{14}$$

Comparing this approximation to available simulations, we see that it holds in the range of parameters observed ($0.20 < d/D < 0.50$), since it well represent the data for both E_0 as a function of ϕ for different d/D (Fig. 3B, dashed lines) as well as E_0 as a function of d/D for different ϕ (Fig. 3C, dashed lines). Interestingly, this result corresponds to a scaling of the interaction energy $E_0 \propto \phi(\frac{D}{d})^2$ that does not correspond to the classic Asakura-Oosawa energy scaling (see Appendix B for a rationalization of why this is the case).

Finally, considering that E_0 and v/v_0 are linked by Eq. 11 and the fact that the interaction energy E_0 depends on d/D and ϕ separately, the model predicts that the empirical collapse of v/v_0 by a single dimensionless parameter x_{Ha} reproduced in the previous section is, in fact, not universal. We should instead consider this collapse as the simplest empirical approximation that is valid for the range of parameters

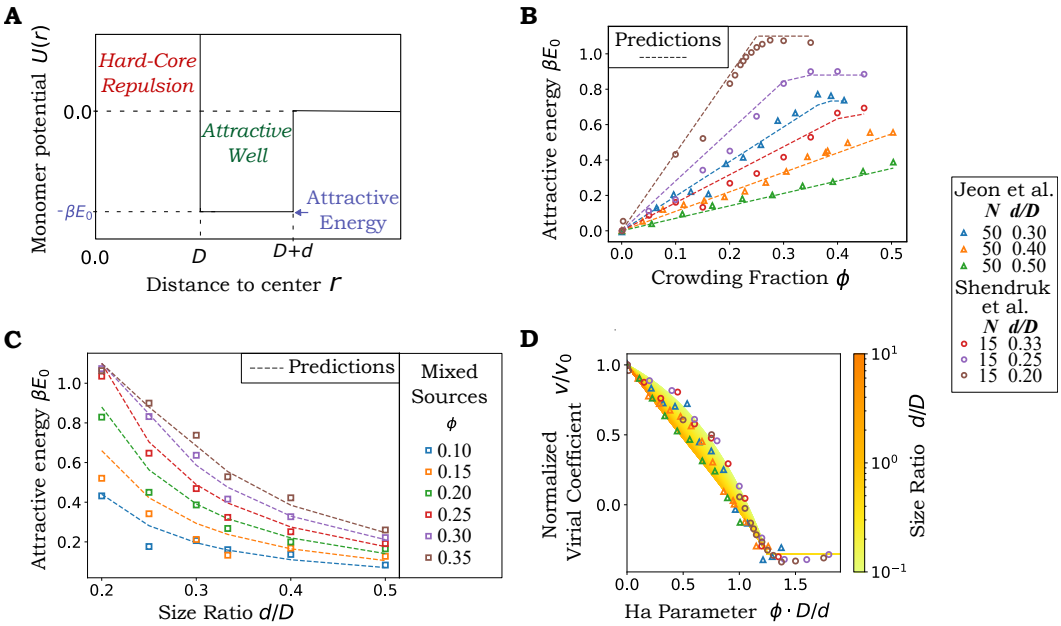

Figure 3 A Flory-like mean-field model captures the crowder-induced coil-globule transition. **(A)** The model approximates the interaction between monomers as the sum of a hard-core repulsion and a monomer-monomer square-well attraction with a range equal to the crowder size d. **(B)** Depth of the attraction energy well βE_0, plotted as a function of crowder density ϕ for model (dashed lines) and simulation data (symbols). **(C)** Depth of the attraction energy well, plotted as a function of relative crowder size d/D. In panels B triangles represent simulation data for different values of respectively relative crowder size d/D and crowder density ϕ, dashed lines are fit the empirical potential derived in the main text (Eq. 14), in panel C simulation data (squares) is interpolated. **(D)** Excluded volume as a function of the Ha parameter, datapoints are the same as Fig. 2B, theory (shaded colors) highlight the non-linear effects in function of d/D (combining eqs. and 11 and 14).

considered in the available simulations. In fact, by numerically combining Eq. 11 with Eq. 13, the model gives a different approximation, which gives a richer and more precise description of the observed data for v/v_0 as a function of x_{Ha} (Fig. 3D). Note that our prediction is not a single universal function of x_{Ha}, but the mean-field model predicts a "gradient of shapes" for different crowder sizes d/D, as, shown in Fig. 3D. The same figure shows that the predicted gradient reproduces a trend that is visible in the simulated data sets.

3.4 The identified effective potential generates correct predictions for simulations of confined polymers and ring polymers.

So far, we have focused on the case of a linear unconfined polymer. However, as stated in the Introduction, motivated by the organization of bacterial genomes there are several simulations in the literature that consider the case of a ring polymer and/or a polymer in confinement. Therefore, we asked whether our effective model could be generalized in a straightforward way to these situations. These modifications influence the overall free energy landscape of the polymer, but should only marginally change the monomer-monomer interaction (which is represented in our model as a local interaction). The main correction to the current model would arise from surface adsorption on the confining edges, which we neglected for simplicity. Under this assumption, the effective pairwise interaction energy that characterizes our model should also be valid for ring-shaped and/or confined polymers, and only the Flory formula from De Gennes (Eq. 10) should be modified. Since ring polymers tend to adopt a linear two-strand configuration[30], the ring polymer free energy can be approximated as a chain polymer with half the number of monomers. Hence $F_{pol}^{ring}(N) = F_{pol}^{chain}(N/2)$, and we can write

$$F_{tot}^{ring}(N) = F_{ideal}(N/2) + F_{interactions}(N) \ . \qquad (15)$$

We adopted this simple approximation and compared our prediction with numerical simulations. Chauhan et al.[29] simulated a ring polymer in crowded environment. Fig. 4A compares their simulation data (triangles) to the predictions of our modified model. The comparison (which involves no free parameters) shows an excellent qualitative agreement and a satisfactory quantitative agreement.

To describe a confined polymer, we extended our model with the classic blob model from Pincus[31], for a cylindrical confinement of diameter L. Specifically, we used the non-confined free energy to compute the number of monomers N_L that a free chain would need to have a size $R(N_L) = L$. Then if $N < N_L$, we used the free energy of free chain, otherwise we described the polymer as a chain of $N_c = N/N_L$ confinement blobs. In this case, the total end-to-end distance is the number of confinement blobs multiplied by the blob diameter, hence the confinement diameter, $R(N) = N_c L$.

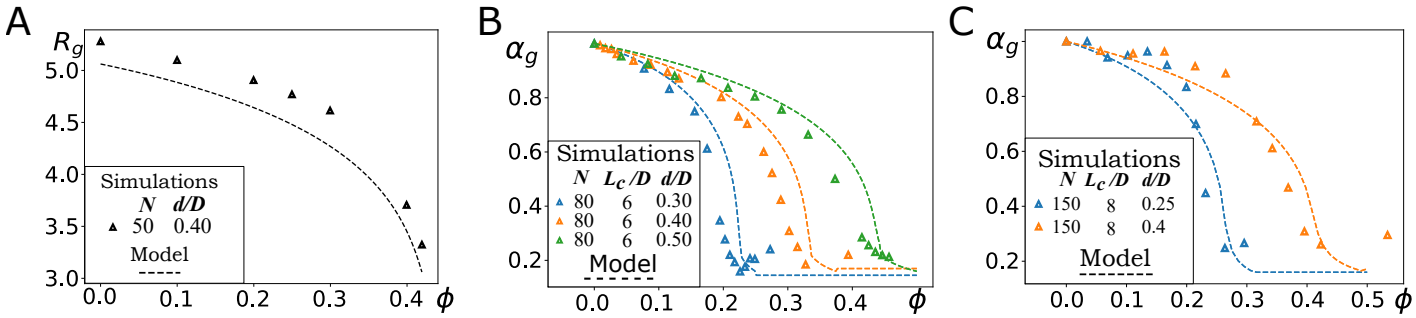

Figure 4 Simple modifications of our model make effective predictions for a confined linear polymer, a free ring polymer and a confined ring polymer. **(A)** Prediction of the collapse curve for a confined linear polymer compared with simulation data from ref.[22]. **(B)** Prediction of the collapse curve for a free ring polymer compared with simulation data from ref.[29]. **(C)** Prediction of the collapse curve for a confined ring polymer, compared with simulation data from ref[9]. The above comparisons involve no free or adjusted parameters, as all parameters in the model are fixed from simulation data.

This straightforward modification of our model allows us to formulate another prediction, which can be compared with simulations. Kim et al.[22] simulated polymers in cylindrical confinement with explicit crowding. Fig. 4B compares the simulation data (triangles) to the predictions of our model, once again finding good agreement. Finally, Yang et al.[9] simulated ring polymers in cylindrical confinement subjected to crowding. In order to describe this situation, we combined the two above modifications, finding again a good match between model predictions and simulated data, with no adjustable parameters (Fig. 4C). These results confirm the validity of the phenomenological form of the dependency of the effective interaction energy from the crowders volume fraction, as well as its flexibility for the use in different systems.

3.5 Polymer collapse does not take place for crowders above a critical size, as predicted by the theory

As described in section 3.3, in all the simulations considered in this study the polymer reaches the collapsed state when the densities of crowders are above the critical value defined by $\phi > x_c \frac{d}{D}$. It is evident that since ϕ is bound by definition ($\phi \leq 1$), there is a physical limit to the size d of crowders that can collapse a polymer chain. To investigate this limit, we produced a new set of simulations (table 1) that could explore values of d/D larger then what was available in the literature. By plotting the swelling parameter α at increasing d/D, we can show the polymer collapse transition shifts to increasing ϕ. For large size ratios $d/D \geq 0.6$, collapse was not observed in the range of ϕ investigated in our simulations (Fig. 5A). At large values of ϕ, collapse was prevented by a slowing down in the dynamics of crowders and their eventual fluid-to-solid phase transition. From these observations we conclude that it is possible to obtain collapse transitions in the kind of simulations considered only when $\phi \leq \phi_c$, with $\phi_c \simeq 0.6$, *i. e.* lower than the theoretical limit of 1 and closer to the random close-packing density[32]. This sets a limit on the size of crowders that can collapse a polymer ($d/D < \phi_c/x_c$), while for larger d/D one would obtain a polymer in a jammed solvent instead.

The system under investigation has some analogies with the phase transition observed in 3D glass-forming supercooled† binary mixtures[33–37], where jamming transitions occur at values of ϕ_c above 0.49. In binary mixtures, the larger spheres go into glassy states at lower density than the smaller spheres, as such, as long as the d/D ratio is sufficiently small, the polymer made by these larger spheres collapses before jamming happens. Conversely, for higher d/D, jamming will happen before the collapse of the polymer. In light of this literature, and analogously to what was proposed by Kang and coworkers[14], we propose a phase diagram for the polymer-crowders mixture as a function of d/D and ϕ, (Fig. 5B), which reveals the existence of a triple point $(\phi, d/D) = (\phi_c, \phi_c/x_c)$. This triple point separates the region where the polymer is swollen from two other regions: the one where it is collapsed (and possibly jammed), and the one where it cannot collapse because the solvent is jammed or has experienced a phase transition to a solid phase.

4 Discussion

In this work we proposed a novel mean-field approach using a two-body effective interaction between monomers for a polymer in a bath of crowders, assuming a square-well potential with range equals to the size of the crowders. The Mayer function connects the depth of this potential to the second-order virial coefficient, that we extract from simulations with explicit crowders to construct our model. The resulting picture (Fig. 3BC) suggests a direct model for the effective interaction that interpolates an "inelastic" regime (the globule) where the short-range attraction energy is a constant with the crowders density, with an "elastic" regime (the coil) where the energy increases linearly with the crowders density ϕ starting at $E = 0$ for $\phi = 0$. The density at which the polymer collapses is then defined by the intersec-

† *we mean here that in a simulation the system is prepared as liquid and then undergoes relaxation and phase separation*

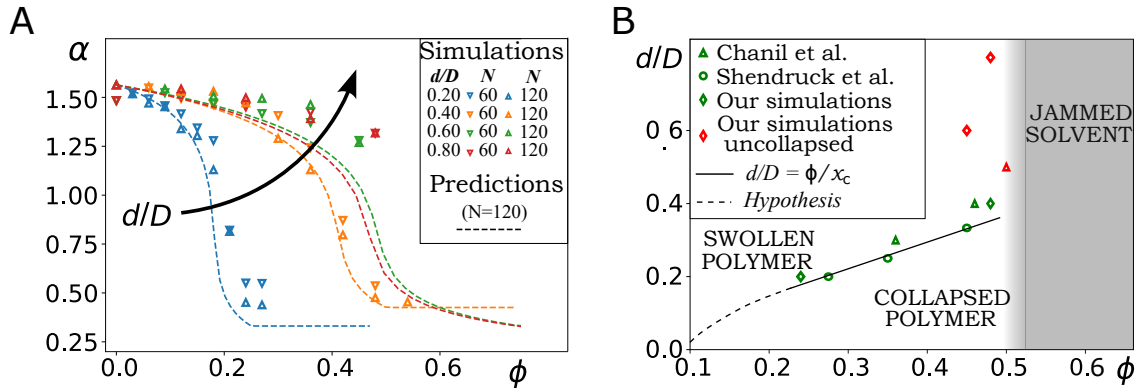

Figure 5 Polymer collapse can be prevented by jamming of the crowders for large volume fractions and d/D. **(A)** Simulations for $N = 60, 120$ at different d/D with predictions of eqs. 10, 11 and 14 for $N = 120$. The arrow indicates that swelling ratios marginally increase with increasing d/D. For high d/D and high ϕ, the collapse point is not reached and the model prediction no longer matches the simulation. **(B)** Phase diagram showing the conditions for which the chain is collapsed (green) and high-crowding simulations where equilibration was not reached (red). The solid line reports the prediction for the phase separation between the collapsed and swollen phases (Eq. 14), the shaded area indicates the region where the phase diagram of a hard-sphere solid exits from the fluid phase.

tion between these two curves. Since the second curve can be linearly approximated according to data, the mean-field model is then completely determined by the two regimes. Clearly, depletion is the reason driving polymer collapse, but, in contrast to a bottom-up approach, our theory does not explicitly account for that. Instead, it interpolates two well-defined polymer phases and their specific mechanical properties in the pure spirit of the Flory mean-field theory.

Our approach has the advantage of being generic and simple, and succeeds where the classic AO theory fails. Indeed, as discussed earlier[13,14], the AO energy potentials are derived (and applicable) only in the regime of very low solvent concentration $\phi < 0.2$, and this condition is not satisfied by any of the simulations available in the literature, because it is very far from the "interesting" regime where the polymer can collapse. Instead, many simulations are often close to the jamming transition, for the solvent as well as the globular phases (Fig. 5B). For this reason, we speculate that a theory valid even for jammed fluids would perhaps be appropriate to extend our mean-field observations. Other approaches available in the literature for the regime considered here are typically more complex and parameter-rich than our model. These include so-called morphometric thermodynamics[38,39], 3rd order virial expansions[17], and the "scaled particle theory" (SPT)[40]. All these approaches result in theories with a higher number of parameters than the AO theory, and, as such, of harder interpretation and with limits of applicability that are more difficult to estimate.

Another quality of our model is that it allows us to qualitatively reinterpret the phase diagram proposed by Kang and coworkers[14] for the crowders-induced collapse of polymers, as we explained in section 3.5. Kang and colleagues address the question of why molecular crowders do not significantly affect the conformation of disordered proteins but they affect the collapse of DNA molecules. Their conclusion is that "size matters", or in other words since proteins are generally smaller than DNA molecules, they are less affected by crowding. Instead, our model (confirmed by simulations) suggests that it is not the size of the polymer, but the size of its monomers compared to that of crowders that matters the most. Consequently, since DNA is made of large units (in the 10nm-100nm size range), it can be collapsed by protein-sized crowders due to depletion interactions. Since proteins are composed by amino acids having a contour length in the sub-nanometer range[41], this results in high d/D, and our phase diagram predicts that in these conditions jamming of the solvent prevents collapse. On the other end, our model does not exclude that crowding could affect protein folding in certain cases: for instance, a large protein made of an alternation of disordered and poorly interacting globular domains might be affected by crowding according to our model, since the large globular domains could effectively behave as large monomers linked by the disordered domains. In this case, our theory predict that globular domains themselves would assemble into a larger folded state or disassemble into individual domains depending on the effect of neutral crowding. In the theory by Kang and coworkers, instead, the presence of multiple globular domains would not change the behavior at fixed length. Hence, we speculate that the effects of neutral crowding on proteins with multiple globular domains could be used as an experimental testable discriminant between the two theories.

In general, our theory is currently limited by the fact that the monomer aspect ratio should be relevant, particularly for DNA, and we expect that this adds an extra dimensionless variable to the problem. This will alter the numerical values of all the observed dimensionless constants extracted from the simulations, and modeling this will require an extension to the theory proposed here. This addresses the following question: what is the size of a monomer? Indeed DNA is known to be a slender polymer, and one's first intu-

ition would be to match the monomer with a rod, long as the persistence length and wide as 2 nucleotides. This does not correspond to what we modeled in this paper. Here we aim to take advantage of the coarse-grain statistics of the polymer, and to consider an effective monomer inspired by *pincus approach* for polymers under tension. The idea would be to define a characteristic length (hence the monomer size) for which a spheroid of this size would contain a region of the polymer being sterically repulsive to other spheroid regions. These regions could be treated as independent monomers. Similar ideas have been developed in Ref.[42]. To fully understand this point, simulations of polymers with rod-shaped monomers or worm-like chains in explicit crowders will be required.

While our theory is a step forward towards descriptions of crowder-induced bacterial chromosome compaction[2,6,9,15], there are still many aspects to address. First, it is based on simulations of polymers and embedding media that lack any biological activity. Chromosomes are found in a constant state of crowding where out-of-equilibrium effects that are relevant for polymers larger then the ones simulated so far could slow down the polymer collapse and lock it in a glassy state. Such states can result from topological barriers (*i. e.* knots)[43] and pearling[44,45]. Glassy states may survive for time scales comparable to the cell cycle rendering the theory of globular polymers inadequate to describe bacterial chromosomes. These out-of-equilibrium effects might be mitigated by opposed active processes fueled by ATP, which can fluidize the dynamics of the cytoplasm[46]. Finally, our model focuses only on the effect of neutral crowders, while *in vivo* chromosomes are surrounded by both neutral and charged crowders, as well as relying on further interactions, e.g. from bridging proteins and active loop extruders[2,47,48]. A full theory or simulation encompassing the effects of crowding as well as a quantitative characterization of the effects of each of these elements on the compaction of bacterial nucleoids is still missing.

Author Contributions

Conceptualization: MCL, VFS, QC. Data curation and analysis - model development: QC. Software - simulations: MD, GG, GCV. Supervision: MCL, VFS and MD. Writing - original draft: QC and VFS. Writing - review and editing: VFS and MCL.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We are grateful to Ignacio Pagonabarraga, Jean-François Joanny, and Daan Frenkel for useful discussions. MCL was supported by AIRC - IG grant no. 23258.

Appendix

A Data analysis

Data were analyzed by custom-made Python jupyter notebooks with numpy, scipy, available at the following link: https://github.com/QuentinChab/PolymerCollapseWithCrowding

B The Asakura Oosawa theory doesn't fit the simulations

Asakura and Oosawa[11,12] predicted the existence of an attractive short-range depletion force between any two plates submerged in a bath of colloids. This force can be generalized for other geometries. The Asakura-Oosawa (AO) potential is in all cases equal to the osmotic pressure of the crowder's gas multiplied by the gain in volume $\delta V(r)$ accessible to this gas due to the proximity of the attracted bodies,

$$\Delta F = U(r) = \begin{cases} \infty, & \text{for } r < D \\ -\Pi \cdot \delta V(r), & \text{for } r > D \text{ and } r < D+d \\ 0. & \text{for } r > D+d \end{cases} \quad (16)$$

For spherical bodies, this volume consists of the overlap between two spherical shells[49], $\delta V = \frac{\pi}{6}(D+d-r)^2(2D+2d+r)$. Eq. 16 includes the infinite hard-core potential that models the hard spheres of radius D, and is zero for distances greater than $r > D+d$. Hence, the AO theory predicts short ranged effective interactions. Because of this, Eq. 16 can be approximated by a square well potential[17] of depth $-E_0$ delimited by the hard-core potential and the interaction range $D+d$, the same shape introduced by us in section 3.3 and Fig. 3A. Then, the value of E_0 is commonly taken equal to $\delta V(r) \cdot \Pi$ at $r = D$,[17,49]. As a first approximation, the osmotic pressure is taken linear with the crowders volume fraction $\Pi = \phi \frac{k_B T}{4/3\pi d^3}$, since in diluted conditions ($\phi < 30\%$) the interactions between crowder particles can be neglected[49]. This results in the following effective interaction energy,

$$E_0^{AO} = k_B T \left(1 + \frac{3D}{2d}\right) \phi. \quad (17)$$

In the limit of small crowders (when $D/d \ll 1$), this formula depends on x_{Ha} only, and this fact has been used in support for x_{Ha} being the key dimensionless parameter[15]. However, section 3.3 shows that in order to reproduce the simulated data, and thus the approximate scaling with x_{Ha}, E_0 should scale as $E_0 \approx \frac{C}{x_c}\left(\frac{D}{d}\right)^2 \cdot \phi$. Hence, we conclude that the scaling relation predicted by this approximation of the AO theory (Eq. 17) is fundamentally not compatible with the relation between D/d and E_0 obtained in simulations.

References

[1] V. F. Scolari and M. Cosentino Lagomarsino, *Soft Matter*, 2015, **11**, 1677–1687.

[2] V. G. Benza, B. Bassetti, K. D. Dorfman, V. F. Scolari, K. Bromek, P. Cicuta and M. C. Lagomarsino, *Reports on Progress in Physics. Physical Society (Great Britain)*, 2012, **75**, 076602.

[3] V. B. Teif and K. Bohinc, *Progress in Biophysics and Molecular Biology*, 2011, **105**, 208–222.

[4] N. Kleckner, J. K. Fisher, M. Stouf, M. A. White, D. Bates and G. Witz, *Current Opinion in Microbiology*, 2014, **22**, 127–137.

[5] T. Odijk, *Biophysical Chemistry*, 1998, **73**, 23–29.

[6] J. Pelletier, K. Halvorsen, B.-Y. Ha, R. Paparcone, S. J. Sandler, C. L. Woldringh, W. P. Wong and S. Jun, *Proceedings of the National Academy of Sciences*, 2012, **109**, E2649–E2656.

[7] S. K. Ramisetty and R. S. Dias, *Journal of Molecular Liquids*, 2015, **210**, 64–73.

[8] A. F. Jorge, S. C. Nunes, T. F. Cova and A. A. Pais, *Current Opinion in Colloid & Interface Science*, 2016, **26**, 66–74.

[9] D. Yang, J. Männik, S. T. Retterer and J. Männik, *Molecular Microbiology*, 2020, **113**, 1022–1037.

[10] M. Negri, M. Gherardi, G. Tiana and M. C. Lagomarsino, *Soft Matter*, 2018, **14**, 6128–6136.

[11] S. Asakura and F. Oosawa, *The Journal of Chemical Physics*, 1954, **22**, 1255–1256.

[12] S. Asakura and F. Oosawa, *Journal of Polymer Science*, 1958, **33**, 183–192.

[13] T. N. Shendruk, M. Bertrand, H. W. de Haan, J. L. Harden and G. W. Slater, *Biophysical Journal*, 2015, **108**, 810–820.

[14] H. Kang, P. A. Pincus, C. Hyeon and D. Thirumalai, *Physical Review Letters*, 2015, **114**, 068303.

[15] C. Jeon, Y. Jung and B.-Y. Ha, *Soft Matter*, 2016, **12**, 9436–9450.

[16] T. N. Shendruk, M. Bertrand, J. L. Harden, G. W. Slater and H. W. de Haan, *The Journal of Chemical Physics*, 2014, **141**, 244910.

[17] L. Reichl, *A modern course in statistical physics*, Wiley, New, 1998.

[18] P. De Gennes, *Journal de Physique Lettres*, 1975, **36**, 55–57.

[19] P. de Gennes, *Journal de Physique Lettres*, 1978, **39**, 299–301.

[20] G. S. Grest and K. Kremer, *Physical Review A*, 1986, **33**, 3628–3631.

[21] G. Giunta, C. Svaneborg, H. A. Karimi-Varzaneh and P. Carbone, *ACS Applied Polymer Materials*, 2019, **2**, 317–325.

[22] J. Kim, C. Jeon, H. Jeong, Y. Jung and B.-Y. Ha, *Soft Matter*, 2015, **11**, 1877–1888.

[23] P. J. Flory, *Principles of Polymer Chemistry*, Cornell University Press, 1953.

[24] P. G. d. Gennes, *Scaling Concepts in Polymer Physics*, Cornell University Press, 1979.

[25] A. Y. Grosberg, A. R. Khokhlov and H. E. Stanley, *Statistical Physics of Macromolecules*, AIP-Press, 1994.

[26] J. S. Kim and I. Szleifer, *The Journal of Physical Chemistry C*, 2010, **114**, 20864–20869.

[27] S. M. Bhattacharjee, A. Giacometti and A. Maritan, *Journal of Physics: Condensed Matter*, 2013, **25**, 503101.

[28] D. A. McQuarrie, *Statistical Mechanics*, Harper & Row, 1975.

[29] G. Chauhan, M. L. Simpson and S. M. Abel, *Soft Matter*, 2021, **17**, 16–23.

[30] Y. Jung, C. Jeon, J. Kim, H. Jeong, S. Jun and B.-Y. Ha, *Soft Matter*, 2012, **8**, 2095–2102.

[31] P. Pincus, *Macromolecules*, 1976, **9**, 386–388.

[32] G. Parisi and F. Zamponi, *Reviews of Modern Physics*, 2010, **82**, 789–845.

[33] M. Dijkstra, R. van Roij and R. Evans, *Physical Review E*, 1999, **59**, 5744–5771.

[34] L. Berthier, *Physical Review E*, 2004, **69**, 020201.

[35] G. Biroli and J.-P. Bouchaud, *Journal of Physics: Condensed Matter*, 2007, **19**, 205101.

[36] A. S. Keys, A. R. Abate, S. C. Glotzer and D. J. Durian, *Nature Physics*, 2007, **3**, 260–264.

[37] S. Pillitteri, G. Lumay, E. Opsomer and N. Vandewalle, *Scientific Reports*, 2019, **9**, 7281.

[38] E. López-Sánchez, C. D. Estrada-Álvarez, G. Pérez-Ángel, J. M. Méndez-Alcaraz, P. González-Mozuelos and R. Castañeda-Priego, *The Journal of Chemical Physics*, 2013, **139**, 104908.

[39] C. D. Estrada-Alvarez, E. López-Sánchez, G. Pérez-Ángel, P. González-Mozuelos, J. M. Méndez-Alcaraz and R. Castañeda-Priego, *The Journal of Chemical Physics*, 2014, **140**, 026101.

[40] E. Thiele, *The Journal of Chemical Physics*, 1963, **39**, 474–479.

[41] S. R. K. Ainavarapu, J. Brujic, H. H. Huang, A. P. Wiita, H. Lu, L. Li, K. A. Walther, M. Carrion-Vazquez, H. Li and J. M. Fernandez, *Biophysical Journal*, 2007, **92**, 225–233.

[42] S. Kadam, K. Kumari, V. M, S. Dutta, M. K. Mitra and R. Padinhateeri, *Bioarxiv*, 2022.

[43] A. Rosa and R. Everaers, *PLoS computational biology*, 2008, **4**, e1000153.

[44] A. Buguin, F. Brochard-Wyart and P. De Gennnes, *Comptes rendus de l'Académie des sciences. Série II, Mécanique, physique, chimie, astronomie*, 1996, **322**, 741–746.

[45] V. F. Scolari, G. Mercy, R. Koszul, A. Lesne and J. Mozziconacci, *Physical Review Letters*, 2018, **121**, 057801.

[46] B. R. Parry, I. V. Surovtsev, M. T. Cabeen, C. S. O'Hern, E. R. Dufresne and C. Jacobs-Wagner, *Cell*, 2014, **156**, 183–194.

[47] M. S. Luijsterburg, M. C. Noom, G. J. L. Wuite and R. T. Dame, *Journal of Structural Biology*, 2006, **156**, 262–272.

[48] R. L. Ohniwa, Y. Ushijima, S. Saito and K. Morikawa, *PloS One*, 2011, **6**, e19172.

[49] D. Marenduzzo, C. Micheletti and P. R. Cook, *Biophysical Journal*, 2006, **90**, 3712–3721.