Article:
Schaefer, Charley, De Bruijn, René A.J. and McLeish, Tom C.B. orcid.org/0000-0002-2025-0299 (2018) Ligand-regulated oligomerisation of allosterically interacting proteins. Soft Matter. pp. 6961-6968. ISSN 1744-683X

https://doi.org/10.1039/c8sm00943k

Journal Name

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

Ligand-Regulated Oligomerisation of Allosterically Interacting Proteins†

Charley Schaefer,*[a‡] René A. J. de Bruijn,[a§] and Tom C. B. McLeish [a‡]

Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

The binding of ligands to distinct sites at proteins or at protein clusters is often cooperative or anti-cooperative due to allosteric signalling between those sites. The allostery is usually attributed to a configurational change of the proteins from a relaxed to a configurationally different tense state. Alternatively, as originally proposed by Cooper and Dryden, a tense state may be achieved by merely restricting the thermal vibrations of the protein around its mean configuration. In this work, we provide theoretical tools to investigate fluctuation allostery using cooling and titration experiments in which ligands regulate dimerisation, or ring or chain formation. We discuss in detail how ligands may regulate the supramolecular (co)polymerisation of liganded and unliganded proteins.

1 Introduction

The allosteric binding of small molecules to proteins or to protein complexes vitally regulates a wide range of biological systems [1–5]. For instance, kinases and phosphatases may interact with the protein calmodulin (CaM), which is only in its active form after calcium has bound to it [6,7]. As shown very recently by Xu et al., in presence of the appropriate synthethic organic ligands CaM may reversibly self-assemble into helical microfilaments [7]. The concentration of the calcium ligand was found to not only control the self-assembly, but, owing to the conformational change of the protein, also tune the pitch of the helical assemblies. The crucial role of allostery in this relation between ligand binding, protein self-assembly and assembly structure begs the question to what extent the type of allostery may play a role. Indeed, Changeux-type enthalpic 'conformation allostery' and Cooper-Dryden-type entropic 'fluctuation allostery' may affect the concentration and temperature dependence of self-assembly differently [2,8,9]. In this work, we present a statistical physics model for ligand-induced copolymerisation (see Figure 1) that enable to investigate this.

Our work builds on a vast literature on supramolecular polymerisation, which distinguishes itself from regular polymerisation by the fact that the constituents of a polymer chain are bound reversibly rather than covalently. This implies that, provided that the time scales of polymerisation and depolymerisation are fast, a broad range of phenomena may be understood in terms of equilibrium thermodynamics [10–12]. Supramolecular *co*polymerisation may in general be summarised by the reaction equation [13]

$$mA + nB \overset{K_{mn}}{\rightleftharpoons} A_mB_n, \tag{1}$$

where m monomers A and n monomers B self-assemble into the copolymer A_mB_n. The equilibrium concentrations, indicated by $[.]$, are related via the equilibrium constant

$$K_{mn} = \frac{[A_mB_n]}{[A]^m\,[B]^n}, \tag{2}$$

which depends on the length and composition of the chain.

The length dependence of the equilibrium constant may in particular be of importance for polymers that grow cooperatively [10,14], which are characterised by a sharp transition from the monomeric to the polymeric state upon cooling or upon (monomer) titration [15]. However, in thermally and chemically

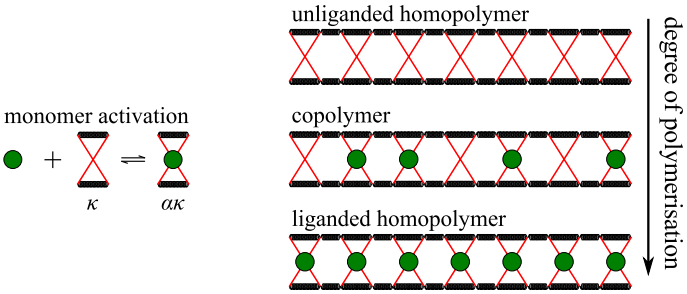

Fig. 1 (colour online) Left: The binding of a ligand to an 'allosteron' monomer with internal spring constant κ results in an activated monomer with internal spring constant $\alpha\kappa$. Right: Self-assembly of unliganded and liganded monomers into chains may result in homo- and copolymers.

[a] *Department of Physics, Durham University, South Road, Durham, DH1 3LE, UK. E-mail: charley.schaefer@york.ac.uk*

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/b000000x/

‡ Present address: Department of Physics, University of York, Heslington, York, YO10 5DD, UK

§ Present address: Theory of Polymers and Soft Matter, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

activated rather than in nucleated systems, cooperativity arises due to a conformational change of the monomer, while the free energy of any oligomerisation step is independent of the chain length [10,16–18]. The composition dependence of the equilibrium constant K_{mn} is responsible for a wide range of experimental observations, such as the sergeant-and-soldiers and the majority-rules effect [19,20]. These effects, where the properties of a copolymeric chain are dictated by either the minority or the majority component, may be tuned by the temperature [21], and is theoretically captured using nearest-neighbour interactions between adjacent monomers in the chain [17,22]. In the context of self-assembling proteins, the differences in interaction energies may be attributed to nearest neighbours with different conformations, i.e., the interaction energies between unliganded-unliganded, unliganded-liganded, and liganded-liganded proteins differ.

This conventional view is challenged by the now widely acknowledged phenomenon of 'fluctuation allostery' [23,24], where no such conformational changes are required. This concept was originally proposed theoretically by Cooper and Dryden [8], and has later been supported by compelling experimental evidence [25,26]. Fluctuation allostery originates from the coupling of the internal modes of the constituents in the chain, and carry over an 'elasticity screening length' beyond the size of a single monomer [9,27–29]. Moreover, this screening length implies that the interactions of dynamically allosteric constituents in a copolymer of unliganded and liganded proteins may not be described using nearest-neighbour interactions. Hence, the length and composition dependence of the equilibrium constant K_{mn} is crucially altered from the usual models with enthalpic nearest-neighbour interactions.

An additional level of complexity in our system, is that the transition of a protein from the unliganded to liganded state is regulated by effector binding through the chemical reaction

$$\mathrm{A} + \mathrm{E} \overset{K_\mathrm{e}}{\rightleftharpoons} \mathrm{B}, \tag{3}$$

where E is the ligand/effector, and A and B are the unliganded and liganded proteins, respectively. Finally,

$$K_\mathrm{e} \equiv \frac{[B]}{[E][A]} \tag{4}$$

is the chemical activation constant.

In this work, we will investigate how the ligands may induce the copolymerisation of dynamic-allosteric building blocks. In order to do so, in the following section we develop the statistical mechanics of self-assembled polymers and copolymers in solution. Subsequently, we use that model to calculate titration curves numerically, which indicate how ligand addition affects the composition of the copolymers as well as their degree of polymerisation.

2 Theory

2.1 Copolymers of dynamic-allosteric constituents

Dynamic allostery may be captured within the parsimonious 'allosteron' model [9,29–32]. Within this model, the subunits of a protein are coarse-grained to the maximal level by endowing them with a single internal dynamic structural mode, characterised by harmonic springs with a spring constant κ. The interactions between the different subunits is captured by connecting the harmonic springs through a direct-coupling spring stiffness, κ_d. As a final ingredient to the model, ligand binding to a subunit is presumed to stiffen the internal spring through $\kappa \mapsto \alpha\kappa$, which implies the vibrations around the average are reduced. This stiffening is carried over to the neighbouring monomers, because their internal modes are connected through the coupling spring. Consequently, the binding of a first ligand pre-stiffens the system prior to the binding of a second neighbouring ligand. The binding of the second ligand therefore requires a smaller entropic penalty to overcome than the first, implying cooperative binding. In this section, we formalise this description of ligand binding to a chain-like configuration of monomers within a statistical physics description.

We consider a chain consisting of N monomers of which $N-N_\mathrm{e}$ are unliganded and N_e are liganded, i.e., N_e ligands are bound to the chain. How the ligands are distributed along the backbone of the chain is given by the vector $\vec{\eta}$, where $\eta_i = 0$ if monomer i is unliganded and $\eta_i = 1$ if it is liganded. We write the Hamiltonian of this chain as

$$\mathscr{H}(\mathbf{x};\vec{\eta},\alpha,K_\mathrm{d},\varepsilon,\varepsilon_\mathrm{e}) = \mathbf{x}^\mathrm{T}\kappa\hat{H}_N\mathbf{x} + \varepsilon(N-1) + \varepsilon_\mathrm{e}N_\mathrm{e}, \tag{5}$$

with $\hat{H}_N(\vec{\eta},\alpha,K_\mathrm{d})$ the tridiagonal Hessian matrix with $\alpha_1 + K_\mathrm{d},\ldots,\alpha_i+2K_\mathrm{d},\ldots,\alpha_N+K_\mathrm{d}$ as diagonal elements, $-K_\mathrm{d}$ as the first lower and upper off-diagonal elements, and 0 otherwise. The symbol $\alpha_i \equiv (1+(\alpha-1)\eta_i)$ equals 1 if $\eta_i = 0$ and equals α if $\eta_i = 1$. The internal modes of the allosterons are coupled by the direct-coupling spring constant K_d, which, as we show below, is associated with an entropic penalty of polymerisation that non-linearly depends on the chain length and on the number of ligands bound to the chain. Interestingly, an indirect coupling (i.e., via resolved fluctuations in their separation distance) of the internal modes to fluctuations of the shape of monomers and bond-length fluctuations between the monomers does not lead to such a dependence (see ESI). To compensate for the entropic penalties generated by restriction to the structural fluctuations the monomers are bound through a binding energy ε, which we take independent of the state of the proteins. The more general case may be implemented using an Ising-type model as in Ref. [17]. The binding of ligands to the monomers is driven by the binding energy ε_e.

Using the Hamiltonian in Eq. (5), we calculate the partition function of a chain of length N as

$$Z_N(N_\mathrm{e}) = \sum_{\eta_1=0}^{1}\cdots\sum_{\eta_i=0}^{1}\cdots\sum_{\eta_N=0}^{1}\delta\left(\sum_{\eta_i=1}^{N}\eta_i - N_\mathrm{e}\right)\int \mathrm{d}\mathbf{x}\exp\left(-\beta\mathscr{H}\right), \tag{6}$$

with $\mathscr{H} = \mathscr{H}(\mathbf{x};\vec{\eta},\alpha,K_\mathrm{d},\varepsilon,\varepsilon_\mathrm{e}))$, where we sum over all possible configurations and use the dirac delta δ to sift the configurations with N_e ligands bound to the chain. In the SI we show that this partition function may be cast in the usual form [9,16]

$$Z_N(N_\mathrm{e}) = Z_1^N(0)W_N(N_\mathrm{e})\mathrm{e}^{-\beta\varepsilon(N-1)}W_\mathrm{e}^{N_\mathrm{e}}\mathrm{e}^{-\beta\varepsilon_\mathrm{e}N_\mathrm{e}} \tag{7}$$

with $Z_1(0) = \sqrt{\pi/\beta\kappa}$ the partition function of an inactive

monomer, and $W_e \equiv \alpha^{-1/2}$ measures the entropic penalty of monomer activation. The entropic penalty of polymerisation ($k_B \ln W_N(N_e)$) may be approximated using

$$W_N(N_e) \approx \alpha^{N_e/2} \binom{N}{N_e} \frac{1}{\sqrt{\sum_{k=1}^{N} \langle d_k^N(N_e,\alpha)\rangle K_d^{k-1}}}, \qquad (8)$$

where $\langle d_k^N(N_e,\alpha)\rangle$ are configuration-averaged polynomial coefficients that can be calculated analytically (see ESI). The quality of this approximation is determined by our approximation of the determinant $\det \hat{H}_N$. The left panel of Figure 2 shows that our approximation describes this quantity accurately as a function of the chain length, chain composition and allosteron parameters K_d and α.

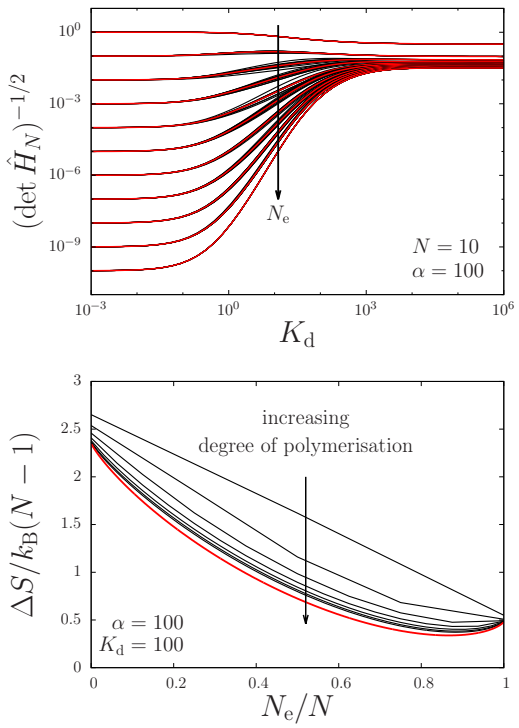

Fig. 2 (colour online) Top: Value of the determinant of the Hessian matrix, $\hat{H}_N$, as a function of the coupling parameter K_d for a chain of length $N = 10$ with $N_e = 0,1,2,\ldots,N$ ligands bound to it. The black lines represent various chain configurations, and the red lines represent the approximate configuration average, see Eqs. (S.3) and (8). Bottom: Entropic penalty ΔS per monomer-monomer bond as a function of the fraction N_e/N of ligand-activated monomers for various chain lengths N using Eq. (8) (black lines). The red line represents the long-chain approximation in Eq. (9). See main text.

In turn, this implies that we can accurately calculate the entropy of chain elongation as a function of these parameters as $\Delta S = -k_B \ln W_N(N_e)/(N-1)$, shown in the right panel of Figure 2. As expected[9], there is weak cooperative growth, because the entropic penalty decreases to a constant value if the chain length becomes much larger than $\sqrt{K_d}$. As shown in the ESI, this long-chain limit can reasonably well be approximated by

$$\frac{\Delta S}{k_B(N-1)} = -\phi \ln \frac{\phi}{\sqrt{w(K_d/\alpha)}} - (1-\phi)\ln \frac{1-\phi}{\sqrt{w(K_d)}}, \qquad (9)$$

with $w(x) = 1/2 + x + (1/2)\sqrt{(1+4x)}$. Here, the first two terms in the right-hand side of the equation represent the usual mixing entropy of the monomers within a copolymer and the last two terms represent the entropic allosteric penalty of monomer coupling. While the mixing entropy favours a 1 : 1 ratio of active and inactive monomers, the allosteric penalty results in a bias of the minimum in entropic penalty towards a large fraction of active monomers.

2.2 Equilibrium statistics of fibres containing bound ligands

The partition function for copolymeric chains that we calculated above describes the interactions between the species in solution. The second contribution that affects the ligand-induced polymerisation of proteins is the chemical potential of all species, which originates from the entropies of translation that may experimentally be controlled by the concentrations. In this section, we derive an equilibrium-statistics model that combines these contributions.

As a starting point, we rely on the statistical-physics recipe presented in Ref.[16], and consider a system of volume V in which an overall number density of monomers ρ^0 and ligands ρ_e^0 are dissolved. The monomers are distributed over free monomers and chains of various lengths with number density $\rho(N)$. For this solution the grand potential is given by

$$\frac{\beta\Omega'}{V} = \frac{\beta\Omega_e}{V} + \frac{\beta\Omega}{V}, \qquad (10)$$

where the first term in the right-hand side of the equation is the contribution by the unbound ligands and the second term is the contribution by the monomers. The latter contribution includes free and bound monomers, as well as bound ligands. The two contributions to the grand potential are described as follows.

The grand potential of the free ligands in solution is given by

$$\frac{\beta\Omega_e}{V} = \rho_e \{\ln(\rho_e \upsilon) - 1 - \beta\mu_e\}, \qquad (11)$$

with υ the interaction volume, which is about the size of a solvent molecule, and μ_e the chemical potential of a free ligand. ρ_e is the number density of the ligands/effectors free in solution. The contribution of the ligands that are bound to monomers is implicit in the grand potential of all monomers.

The grand potential of the monomers is given by[33]

$$\frac{\beta\Omega}{V} = \sum_{N=1}^{\infty}\sum_{N_e=0}^{N} \rho_N(N_e) \times$$
$$\{\ln[\rho_N(N_e)\upsilon] - 1 - \ln Z_N(N_e) - \beta\mu N - \beta\mu_e N_e\}, \qquad (12)$$

with the partition function $Z_N(N_e)$ given by Eq. (7), with μ the chemical potential of a monomer, and with μ_e the chemical potential of a ligand.

We obtain the length distribution of our chains in thermal equilibrium by setting $\delta\Omega'/\delta\rho_e$ and $\delta\Omega'/\delta\rho_N(N_e)$ to zero for all combinations of N and N_e. This gives

$$\rho_e = \frac{1}{\upsilon} e^{\beta\mu_e} \qquad (13)$$

for the number density of free ligands and

$$\rho_N(N_\mathrm{e}) = \frac{1}{\upsilon} Z_N(N_\mathrm{e}) \mathrm{e}^{\beta\mu N + \beta\mu_\mathrm{e} N_\mathrm{e}} \tag{14}$$

for the number density of polymers of length N with N_e ligands bound to them. These expressions indicate that all number densities are determined by the chemical potentials of the ligands and monomers.

Indeed, the number densities of the chains are related to the number densities of the free ligands and free monomers via the (here dimensionless) equilibrium constants as earlier introduced in Eq. (4) and Eq. (2) by

$$K_\mathrm{e} = \upsilon^{-1} \frac{\rho_1(1)}{\rho_\mathrm{e}\rho_1(0)} = \mathrm{e}^{-\beta\varepsilon_\mathrm{e} + \ln W_\mathrm{e}}, \tag{15}$$

and by

$$K_N(N_\mathrm{e}) \equiv \upsilon^{-(N-1)} \frac{\rho_N(N_\mathrm{e})}{\rho_1^{N-N_\mathrm{e}}(0)\rho_1^{N_\mathrm{e}}(1)} = \mathrm{e}^{-\beta\varepsilon(N-1) + \ln[W_N(N_\mathrm{e})]}. \tag{16}$$

Using these equilibrium constants, the number density of each type of oligomer, $\rho_N(N_\mathrm{e})$, can be expressed as a function of the number densities of free inactive monomers, $\rho_1(0)$, and ligands, ρ_e, as $\rho_N(N_\mathrm{e}) = \upsilon^{N+N_\mathrm{e}-1} K_N(N_\mathrm{e}) K_\mathrm{e}^{N_\mathrm{e}} \rho_\mathrm{e}^{N_\mathrm{e}} \rho_1^N(0)$.

We now obtain the final result that describes the equilibrium statistics of the chains in the form of a mass balance for the monomers and a mass balance for the ligands. The overall number density of the monomers is the sum of all monomers free in solution and the ones bound in a chain, $\rho^0 = \sum_{N=1}^{\infty}\sum_{N_\mathrm{e}=0}^{N} N\rho_N(N_\mathrm{e})$. Similarly, the overall number density of ligands is $\rho_\mathrm{e}^0 = \rho_\mathrm{e} + \sum_{N=1}^{\infty}\sum_{N_\mathrm{e}=1}^{N} N_\mathrm{e}\rho_N(N_\mathrm{e})$. If we insert the expressions for the number densities $\rho_N(N_\mathrm{e})$ derived above, we finally arrive at

$$\frac{\rho^0}{\rho^\ddagger} = \sum_{N=1}^{\infty} N \left(\sum_{N_\mathrm{e}=0}^{N} K_N(N_\mathrm{e}) \left[\frac{\rho_\mathrm{e}}{\rho_\mathrm{e}^\ddagger}\right]^{N_\mathrm{e}} \right) \left[\frac{\rho_1(0)}{\rho^\ddagger}\right]^N, \tag{17}$$

$$\frac{\rho_\mathrm{e}^0}{\rho_\mathrm{e}^\ddagger} = \frac{\rho^\ddagger}{\rho_\mathrm{e}^\ddagger} \sum_{N=1}^{\infty} \left(\sum_{N_\mathrm{e}=1}^{N} N_\mathrm{e} K_N(N_\mathrm{e}) \left[\frac{\rho_\mathrm{e}}{\rho_\mathrm{e}^\ddagger}\right]^{N_\mathrm{e}} \right) \left[\frac{\rho_1(0)}{\rho^\ddagger}\right]^N + \frac{\rho_\mathrm{e}}{\rho_\mathrm{e}^\ddagger}, \tag{18}$$

which are polynomials in $\rho_1(0)$ and ρ_e. We have rendered the concentrations dimensionless using the 'transition' concentrations $\rho^\ddagger$ and $\rho_\mathrm{e}^\ddagger$. We have defined $\rho^\ddagger$ as the number density of monomers at which half of the monomers is aggregated in absence of ligands[9,16]. Further, $\rho_\mathrm{e}^\ddagger$ is the ligand concentration at which half of the free monomers are activated at small monomer concentrations, hence $K_\mathrm{e}\upsilon\rho_\mathrm{e}^\ddagger = \rho_1(1)/\rho_1(0) = 1$. This choice of scaling allows us to calculate a universal phase diagram that indicates the conditions under which (ligand-activated) polymerisation takes place.

3 Results

Some universal features of supramolecular polymerisation in presence of effectors may be predicted from the mass balances in Eq. (18) without solving that equation numerically ¶. These features are the transitions between four regions in the phase diagram of Figure 3. This Figure shows that depending on the overall concentrations of monomers and ligands, the monomers may be in an active or inactive state, and may be either free in solution or bound in chains. In the following, we quantify and discuss the transition lines in the phase diagram. Following that discussion, we numerically calculate the concentrations of self-assembled species along straight horizonal and vertical lines through the phase diagram, which experimentally represent monomer- and ligand-titration curves.

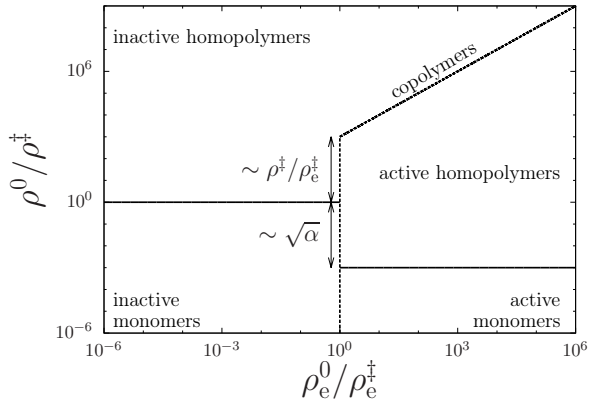

Fig. 3 Phase diagram for ligand-induced self-assembly of monomers in a dilute solution. The axes represent the scaled monomer and ligand concentration $\rho^0/\rho^\ddagger$ and $\rho_\mathrm{e}^0/\rho_\mathrm{e}^\ddagger$, respectively. The solid lines indicate the phase transition from (in)active monomers to homopolymers, which are displaced through the ligand parameter α. The dashed line (given by $K_\mathrm{e}\upsilon\rho_\mathrm{e}^0 = 1 + \upsilon K_\mathrm{e}\rho^0$ with K_e the activation constant) represents the phase transition from inactive to ligand-activated material. Near this transition copolymers may be formed.

The transition lines can partially be understood in terms of the self-assembly of a single type of either inactive or active monomers into homopolymers[9]. For inactive species, i.e., at low ligand concentrations, the polymerisation concentration is proportional to the square root of the allosteron coupling parameter K_d. This factor originates from the coupling of monomers with an internal mode that is described by the allosteron spring constant κ, which is in actived monomers increased to $\alpha\kappa$. As a consequence the entropic penalty of coupling monomers decreases and, while the interaction range between the monomers in the chain remains unaltered, the polymerisation concentration decreases by a factor $\sqrt{\alpha}$. These different polymerisation concentrations are represented by the horizontal solid lines in the phase diagram of Figure 3.

The dashed line in the phase diagram indicates the transition from inactive to active species. As we later show when discussing the titration curves, this transition may be understood in terms

¶A numerical solution can be obtained using a common multivariate bisection method. This method makes use of the fact that $\rho_1(0)$ monotonically increases and ρ_e decreases upon an increase of the overall monomer concentration, ρ^0, while $\rho_1(0)$ monotonically decreases and ρ_e increases upon an increasing overall ligand concentration, ρ_e^0

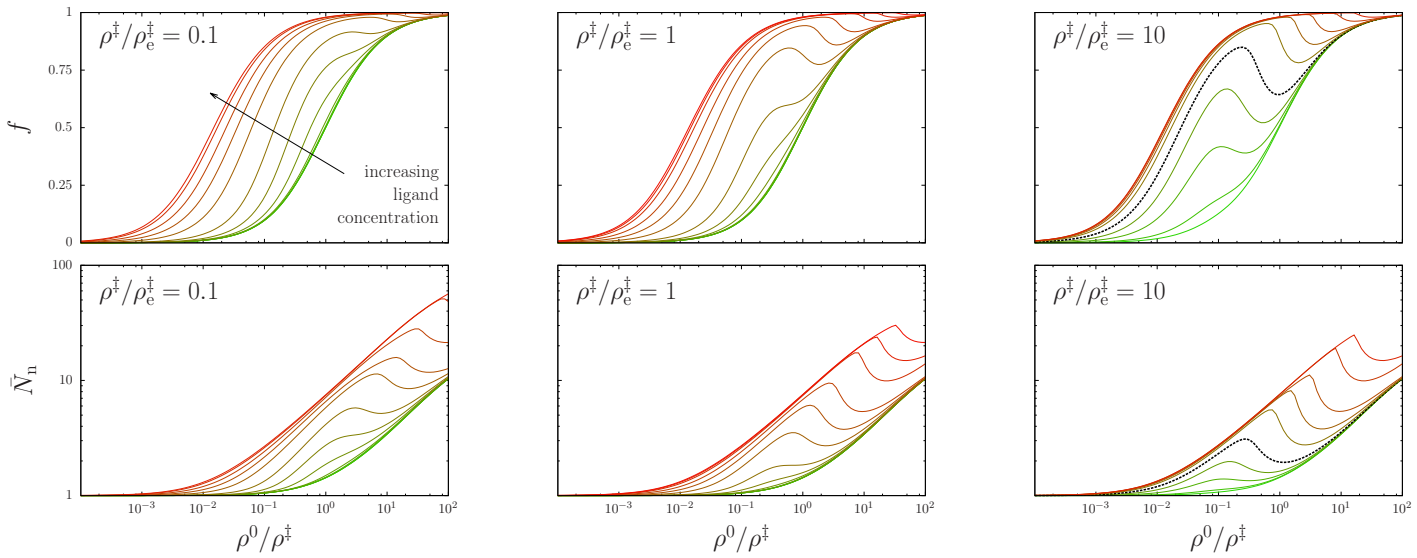

Fig. 4 (colour online) Fraction of aggregated material, f (top panels), and number-averaged degree of polymerisation, $\bar{N}_n$ (bottom panels), as a function of the scaled monomer concentration $\rho^0/\rho^\ddagger$ for a varying ligand concentration $\rho_e/\rho_e^\ddagger$. From left to right, the ratio between polymerisation and activation concentration is varied from $\rho^\ddagger/\rho_e^\ddagger = 0.1$ to 10. The shift of the polymerisation concentration upon ligand addition is governed by the allosteron parameters $K_d = \alpha = 10^4$. We have subjected the dashed line for further analysis in Figure 5.

of monomer activation in absence of dimers or larger oligomers. Within that approximation the mass balances in Eq. (18) reduce to $\rho^0 = \rho_1(0) + \rho_1(1)$ and $\rho_e^0 = \rho_e + \rho_1(1)$, with the constraint $K_e \upsilon = f_e/\rho_e(1-f_e)$ and $f_e = \rho_1(1)/\rho^0$ the fraction of activated material. It follows that the fraction of activated material may be expressed as the quadratic equation $f_e = (1/2a)(\sqrt{b^2+4a}-b)$ with $a = \rho^0/\rho_e^\ddagger$ and $b = 1+(\rho_e^0-\rho^0)/\rho_e^\ddagger$. In the asymptotic limit of low monomer concentrations, the fraction of activated material is

$$f_e = 1 - \frac{1}{1+\rho_e^0/\rho_e^\ddagger}, \tag{19}$$

independent of the monomer concentration. At high monomer concentrations and sufficiently low ligand concentration ($\rho_e^0 \leq \rho^0$), the fraction of activated material increases linearly with an increasing ligand concentration as

$$f_e = \left(1+\frac{\rho^0}{\rho_e^\ddagger}\right)^{-1} \frac{\rho_e^0}{\rho_e^\ddagger} \tag{20}$$

up to to sharp transition at $\rho_e^0 \leq \rho^0$, where it reaches unity. The condition $\rho_e^0/\rho_e^\ddagger = 1+\rho^0/\rho_e^\ddagger$ not only defines the transition from inactive to active species at high monomer concentrations, but is also consistent with the condition $\rho_e^0\rho_e^\ddagger = 1$ at low monomer concentrations at which half of the material is activated. Therefore, we consider this condition, represented by the dashed line in Figure 3, as a reasonable definition of the transition line from inactive to active species.

Now that we have discussed the transition lines in the phase diagram in general terms, we investigate the consequences of these transition lines for experimentally measurable titration curves. We first focus on the monomer-titration experiments in which the concentration of monomers is varied at a fixed ligand concentratio (Figure 4). In that Figure, we have numerically calculated how the fraction of aggregated material, $f \equiv 1-[\rho_1(0)+\rho_1(1)]/\rho^0$, and the number-averaged degree of polymerisation, defined by

$$\bar{N}_n = \frac{\sum_{N,N_e} N\rho_N(N_e)}{\sum_{N,N_e} \rho_N(N_e)}, \tag{21}$$

depend on the monomer concentration for three values $\rho^\ddagger/\rho_e^\ddagger \propto K_e$. In each panel, the ligand concentration is varied from low (green) to high (red) ligand values around the transition concentraiton (increments by factors of $0, 10.2, 0.5, 1, 2, 5, \ldots$).

The two main phenomena that occur due to an increasing monomer concentration are (i) a transition from liganded to unliganded self-assembly, and (ii) the transition through a maximum. At different fixed ligand concentrations, the self-assembly curve varies as follows. For low ligand concentrations (green lines), the unliganded monomers self-assemble into chains for monomer concentrations well above the polymerisation concentration $\rho^\ddagger \propto \sqrt{K_d}$[9]. This is associated with the increasing fraction of aggregated material, as well as the degree of polymerisation that increases with the square root of the monomer concentration. At high ligand concentrations, the shape of the self-assembly curve is almost identical, but shifted to a lower polymerisation concentration by a factor $\sqrt{\alpha}$, provided that $\alpha \leq K_d$. At intermediate ligand concentrations, this results in a maximum in the polymerisation curve (see e.g. the black dashed lines).

We inspect one of those non-monotonic curves, i.e., the black dashed line in Figure 4, in more detail in Figure 5. In the left panel, we present the chain-length distribution, $\rho_N \equiv \sum_{N_e} \rho_N(N_e)$, from low (green) to high (red) monomer concentration. This Figure shows that sufficiently long chains obey the usual exponential distribution and that its width, related to the degree of polymerisation, increases non-monotonically. This is due to the fact that the nucleus size depends on the monomer concentration, shown in the right panel.

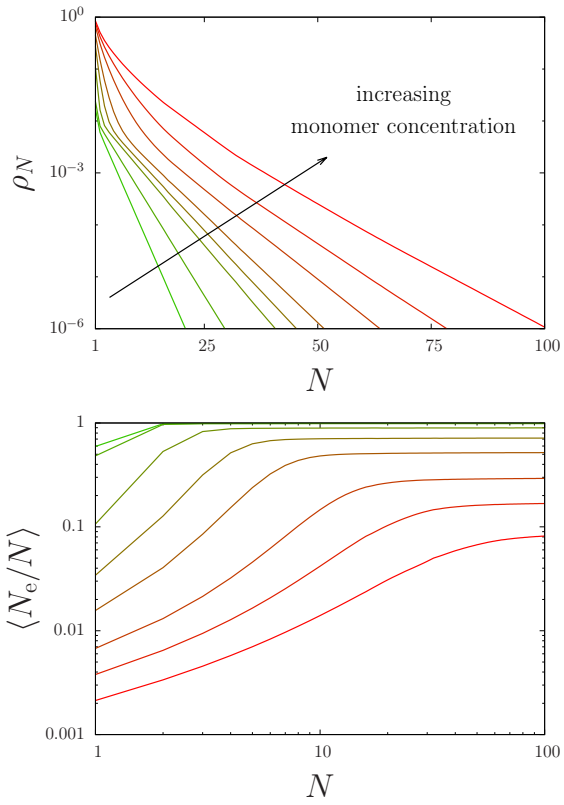

Fig. 5 (colour online) Fraction of ligands (top) and number density (in arbitrary units) (bottom) as a function of the chain length for a monomer concentration ranging from $\rho/\rho^{\ddagger} = 10^{-1}$ (green) to 10 (red) at a fixed ligand concentration of $\rho_e/\rho_e^{\ddagger} = 2$, corresponding to the black dashed line in Figure 4. The fixed parameters are $\rho^{\ddagger}/\rho_e^{\ddagger} = 10$, and allosteron parameters $K_d = \alpha = 10^4$.

The right panel of Figure 4 shows fraction of activated monomomers in an assembly, $\langle N_e/N \rangle$, for the same concentrations as in the left panel. As expected, for an increasing monomer concentration the fraction of ligands decreases. In all cases, this fraction is low for small oligomers below the nucleus size (or elasticity screening length), but reaches a plateau for long chains. We find that the nucleus size increases with a decreasing ligand fraction, that can be reasonably-well approximated by $\sqrt{K_d/\alpha^{\phi}}$ with ϕ the fraction of ligands in the long-chain limit.

It is this change in nucleus size that we hold responsible for the peak in the black dashed titration curve of Figure 4: For small monomer concentrations, the ligand fraction is large and most monomers are activated. Due to the small nucleus size, an increasing monomer concentration may lead to larger oligomers. However, the nucleus size increases, and at a further increased monomer fraction the oligomers become unstable and lead to a decreased fraction of aggregated material and a decreased degree of polymerisation. Upon further increase of the monomer concentration, fraction of ligands in the chain tends to zero, and the nucleus size becomes constant and usual non-activated polymerisation may take place.

We conclude that ligand-activated copolymerisation of monomers leads to distinct phenonoma that are discernible as non-monotonic monomer-titration curves, which may be measured experimentally. As we discuss below, ligand binding also leads to notable features in ligand-titration curves at fixed monomer concentrations, see Figure 6.

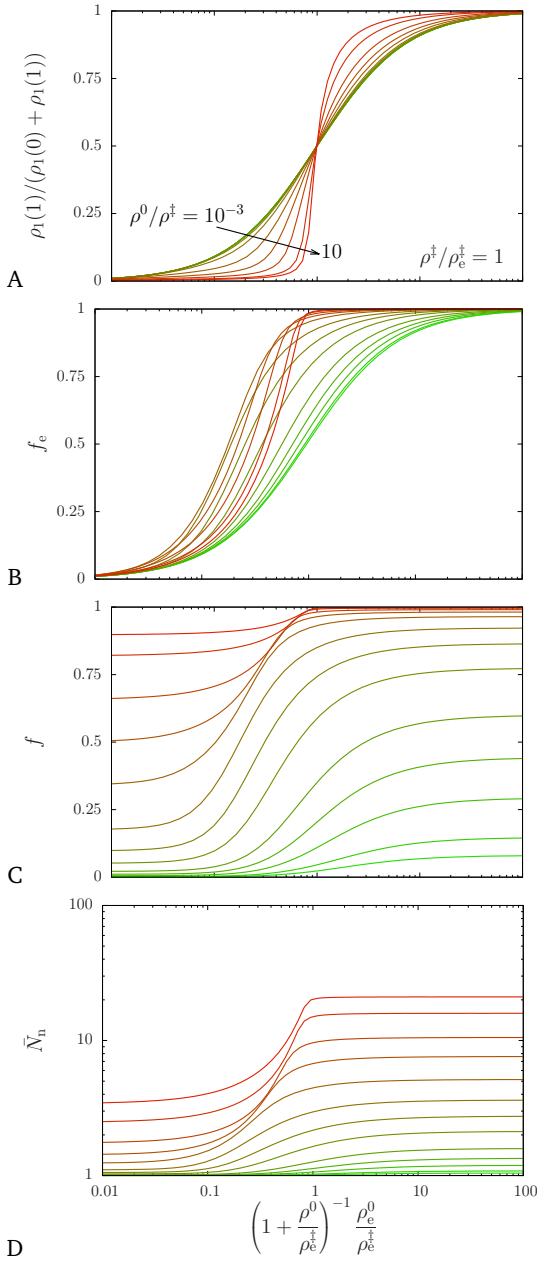

Fig. 6 (colour online) Titration of ligands from a scaled ligand concentration of $\rho_e^0/\rho_e^{\ddagger} = 10^{-2}$ to 10^2 for scaled monomer concentrations varying from $\rho^0/\rho^{\ddagger} = 10^{-3}$ (green) to 10 (red), for $K_d = \alpha = 10^4$ and $\rho^{\ddagger}/\rho_e^{\ddagger} = 1$. The plots show how ligand titration affects A) the fraction of free monomers that are active, B) the total fraction of activated material, C) the fraction of aggregated monomers, and D) the number-averaged degree of polymerisation.

Figure 6 shows the fraction of active free monomers, $\rho_1(1)/(\rho_1(0)+\rho_1(1))$, the fraction of active material, f_e, the fraction of aggregated material, f, and the number-averaged degree of polymerisation $\bar{N}_n$ as a function of the scaled ligand concentration. This concentration is scaled using the ligand concentration that defines the dashed transition line between regions dom-

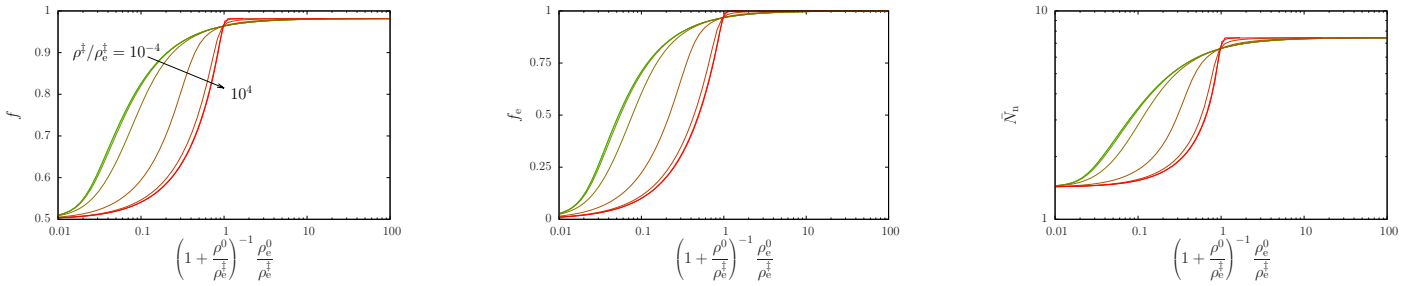

Fig. 7 (colour online) Ligand-titration curves for a activation factor $\rho^\ddagger/\rho_e^\ddagger$ varying from 10^{-4} (green) to 10^4 (red) at a monomer concentration fixed at the polymerisation concentration, $\rho^0 = \rho^\ddagger$. From left to right: Fraction of aggregated material, f, fraction of activated material, f_e, degree of polymerisation, $\bar{N}_n$.

inated by inactive and active material in the phase diagram of Figure 3. This concentration is given by the condition where f_e in Eq. (20) equals unity.

As expected, for small monomer concentrations (green), the fraction of activated monomers in plot A and the overall fraction of activated material in plot B are identical, and obey Eq. (19) independently of the monomer concentration. For a large monomer concentration (red lines), we find that the fraction of activated material, f_e, tends towards Eq. (20), where f_e linearly increases with an increasing ligand concentration and sharply crosses over to unity at the transition concentration. However, the fraction of free active monomers remains low up to the transition. This implies that the active species predominantly reside in oligomers.

Indeed, the fraction of aggregated material (Figure 6C) and degree of polymerisation (Figure 6D) increase with an increasing fraction of monomers below the transition concentration. For larger ligand concentrations, both quantities increase as well, indicating ligands may actually be employed to switch between two different degrees of polymerisation at a fixed monomer concentration. While the transition becomes sharper with an increasing monomer concentration, it may also be controlled by the activation constant through $\rho^\ddagger/\rho_e^\ddagger \propto K_e$, see Figure 7.

In Figure 7 we have fixed the overall monomer concentration, ρ^0, at the polymerisation concentration, $\rho^\ddagger$, while varying the ligand concentration by four orders of magnitude around the activation transition (the activation factor $\rho^\ddagger/\rho_e^\ddagger \propto K_e$ varies from 10^{-4} to 10^4). This increase results in a crossover from a smooth to a sharp transition between a small and large fraction of aggregated material and between inactive and active species, as shown in the left and middle panels of Figure 7, respectively. In turn, the transition from small to large oligomers becomes sharper, see right panel of Figure 7. Hence, we have shown that while monomer titration can result in a non-monotonic increase of the fraction of aggregated material and of the degree of polymerisation, ligand titration gives rise to a monotonic increase that strongly depends on both the monomer concentration and the activation constant.

4 Conclusions

We have theoretically investigated the supramolecular polymerisation of allosterically interacting proteins to which ligands may bind. We view this is a representative system for an entire family of ligand-regulated self-assemblies of proteins, which also includes dimers and rings.

In practice, fluctuation allostery is associated with a spectrum of short- and long-wavelength fluctuations on the protein conformation, which may be assessed, e.g., using a Gaussian network model[32,34]. While the short-wavelength fluctuations presumably only give rise to short-ranged interactions between proteins, the long-wavelength fluctuations may affect proteins in a chain at long distances. In the present work, we viewed proteins at the maximum level of coarse graining and investigated the influence of long-wavelength fluctuations exclusively.

In this entropic 'allosteron' model, the internal modes of proteins are coupled upon polymerisation, which leads to an elasticity screening length $\sqrt{K_d}$. The polymerisation concentration is proportional to this screening length, but may however be decreased by ligand binding if they stiffen the internal modes of the monomers. Indeed, if their internal mode becomes stiffer by a factor α, the elasticity screening length decreases to $\sqrt{K_d/\alpha^\phi}$. Hence, the polymerisation concentration can be reduced by a factor $\sqrt{\alpha} \leq \sqrt{K_d}$ if sufficient ligands are added to the solution.

The width of the transition may be tuned using the monomer concentration and the activation constant. The transition itself may be monotonic or non-monotonic, which is also controlled by the activation constant. Indeed, for large activation constants and small monomer concentrations, small ligand-activated oligomers may be formed. Upon monomer addition, however, the oligomers will harbour a smaller fraction of ligands due to which the elasticity screening length surpasses the oligomer length, which in turn destabilises the oligomers. Consequently, with an increasing monomer concentration the number of oligomers may decrease.

While we have focused exclusively on the calculation of titration curves, our method may be directly applied to calculate cooling curves. Those curves are a superposition of our monomer- and ligand-titration curves, in a way that explicitly depends on the monomer-monomer and ligand-monomer binding energies ε and ε_e, respectively. Following the usual approach[14,35], the model may be fitted to experimental titration or cooling curves. This should provide direct information about fluctuation allostery in the experimental system.

An extremely promising protein of which the self-assembly may potentially be affected, and regulated, by fluctuation allostery is the calmodulin (CaM) protein[7]. There is strong evidence that, apart from conformational changes, calcium-regulated fluctua-

tion allostery plays a key role to the binding of CaM to target molecules[36,37]. How this regulation takes place in detail, and what the role of fluctuation allostery is to CaM-CaM dimerisation remains unknown. However, in recent works a range of different types of titration curves were obtained from coarse-grained molecular dynamics[38], as well as from experiment[7].

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was funded by the Engineering and Physical Sciences Research Council [grant number EP/N031431/1].

Notes and references

1 S. Jusuf, P. J. Loll and P. H. Axelsen, *J. Am. Chem. Soc.*, 2003, **125**, 3988–3994.

2 J.-P. Changeux and S. J. Edelstein, *Science*, 2005, **308**, 1424–28.

3 A. Aggeli, I. A. Nyrkova, M. Bell, R. Harding, L. Carrick, T. C. B. McLeish, A. N. Semenov and N. Boden, *PNAS*, 2001, **98**, 11857–11862.

4 R. P. Davies, A. Aggeli, N. Boden, T. C. McLeish, I. A. Nyrkova and A. N. Semenov, *Mechanisms and principles of 1D self-assembly of peptides into β-sheet tapes*, Elsevier, Amsterdam, 2009, vol. 35.

5 K. Hashimoto, H. Nishi, S. Bryant and A. R. Panchenko, *Phys. Biol.*, 2011, **8**, 035007.

6 D. Chin and A. R. Means, *Trends cell. Biol.*, 2000, **10**, 322–328.

7 M. Xu, L. Liu and Q. Yan, *Angew. Chem.*, 2018, **130**, 5123–5126.

8 A. Cooper and D. T. F. Dryden, *Eur. Biophys. J.*, 1984, **11**, 103–109.

9 T. C. B. McLeish, C. Schaefer and A. C. von der Heydt, *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 2018, **373**, xx.

10 T. F. A. D. Greef, M. M. J. Smulders, M. Wolffs, A. P. H. J. Schenning, R. P. Sijbesma and E. W. Meijer, *Chem. Rev.*, 2009, **109**, 5687–5754.

11 C. Rest, R. Kandanelli and G. Fernández, *Chem. Soc. Rev.*, 2015, **44**, 2543–2572.

12 P. Besenius, *J. Polymer Sci. A: Polymer Chem.*, 2017, **55**, 34–78.

13 A. Das, G. Vantomme, A. J. Markvoort, H. M. M. ten Eikelder, M. Garcia-Iglesias, A. R. A. Palmans and E. W. Meijer, *J. Am. Chem. Soc.*, 2017, **139**, 7036–7044.

14 C. Kulkarni, E. W. Meijer and A. R. A. Palmans, *Acc. Chem. Res.*, 2017, **50**, 1928–1936.

15 R. F. Goldstein and L. Stryer, *Biophys. J.*, 1986, **50**, 583–599.

16 P. van der Schoot, *Theory of supramolecular polymerization*, CRC Press, Boca Raton, 2nd edn, 2005.

17 S. Jabbari-Farouji and P. van der Schoot, *J. Chem. Phys.*, 2012, **137**, 064906.

18 P. A. Korevaar, C. Grenier, A. J. Markvoort, A. P. H. J. Schenning, T. F. A. de Greef and E. W. Meijer, *PNAS*, 2013, **110**, 17205–17210.

19 M. M. Green, M. P. Reidy, R. J. Johnson, G. Darling, D. J. O'Leary and G. Wilson, *J. Am. Chem. Soc.*, 1989, **111**, 6452–6454.

20 M. M. Green, B. A. Garetz, B. Munoz and H. P. Chan, *J. Am. Chem. Soc.*, 1995, **117**, 4181–4182.

21 M. M. J. Smulders, I. A. W. Filot, J. M. A. Leenders, P. van der Schoot, A. R. A. Palmans, A. P. H. J. Schenning and E. W. Meijer, *J. Am. Chem. Soc.*, 2010, **132**, 611–619.

22 A. J. Markvoort, H. M. M. ten Eikelder, P. A. Hilbers, T. F. A. de Greef and E. W. Meijer, *Nature Comm.*, 2011, **2**, 509.

23 R. G. Smock and L. M. Gierasch, *Science*, 2009, **324**, 198–203.

24 H. N. Motlagh, J. O. Wrabl, J. Li and V. J. Hilser, *Nature*, 2014, **508**, 331–339.

25 N. Popovych, S. Sun, R. H. Ebright, Charalampos and Kalodimos, *Nat. Struct. Mol. Biol.*, 2006, **13**, 831–838.

26 C. M. Petit, J. Zhang, P. J. Sapienza, E. J. Fuentes and A. Lee, *Nature*, 2009, **106**, 18249–18254.

27 V. J. Hilser, J. O. Wrabl and H. N. Motlagh, *Annu. Rev. Biophys.*, 2012, **41**, 585–609.

28 M. Goulian, R. Bruinsma and P. Pincus, *Europhys. Lett*, 1993, **22**, 145–150.

29 T. C. B. McLeish, T. L. Rodgers and M. R. Wilson, *Phys. Biol.*, 2013, **10**, 056004 (9pp).

30 R. J. Hawkins and T. C. B. McLeish, *Phys. Rev. Lett.*, 2004, **93**, 098104.

31 R. J. Hawkins and T. C. B. McLeish, *Biophys. J.*, 2006, **91**, 2055–2062.

32 H. Toncrova and T. C. B. McLeish, *Biophys. J.*, 2010, **98**, 2317–2326.

33 M. E. Cates and S. J. Candau, *J. Phys.: Condens. Matter*, 1990, **2**, 6869–6892.

34 T. L. Rodgers, P. D. Townsend, D. Burnell, M. L. Jones, S. A. Richards, T. C. B. McLeish, E. Pohl, M. R. Wilson and M. J. Cann, *PLoS Biol*, 2013, **11**, e1001651.

35 C. Schaefer, I. Voets, A. Palmans, E. Meijer, P. van der Schoot and P. Besenius, *ACS Macro Lett.*, 2012, **1**, 830–833.

36 K. K. Frederick, M. S. Marlow, K. G. Valentine and A. J. Wand, *Nature*, 2007, **448**, 325–330.

37 M. S. Marlow, J. Dogan, K. K. Frederick, K. G. Valentine and A. J. Wand, *Nat. Chem. Bio.*, 2010, **6**, 352–358.

38 P. Nandigrami and J. J. Portman, *J. Chem. Phys.*, 2016, **144**, 105101.