Equilibrium mechanisms of self-limiting assembly

Michael F. Hagan

Martin Fisher School of Physics, Brandeis University, Waltham, MA 02454, USA

Gregory M. Grason

Department of Polymer Science and Engineering, University of Massachusetts, Amherst, MA 01003, USA

(Dated: March 16, 2021)

Self-assembly is a ubiquitous process in synthetic and biological systems, broadly defined as the spontaneous organization of multiple subunits (e.g. macromolecules, particles) into ordered multi-unit structures. The vast majority of equilibrium assembly processes give rise to two states: one consisting of dispersed disassociated subunits, and the other, a bulk-condensed state of unlimited size. This review focuses on the more specialized class of *self-limiting assembly*, which describes equilibrium assembly processes resulting in finite-size structures. These systems pose a generic and basic question, how do thermodynamic processes involving non-covalent interactions between identical subunits "measure" and select the size of assembled structures? In this review, we begin with an introduction to the basic statistical mechanical framework for assembly thermodynamics, and use this to highlight the key physical ingredients that ensure equilibrium assembly will terminate at finite dimensions. Then, we introduce examples of self-limiting assembly systems, and classify them within this framework based on two broad categories: *self-closing assemblies* and *open-boundary assemblies*. These include well-known cases in biology and synthetic soft matter — micellization of amphiphiles and shell/tubule formation of tapered subunits — as well as less widely known classes of assemblies, such as short-range attractive/long-range repulsive systems and geometrically-frustrated assemblies. For each of these self-limiting mechanisms, we describe the physical mechanisms that select equilibrium assembly size, as well as potential limitations of finite-size selection. Finally, we discuss alternative mechanisms for finite-size assemblies, and draw contrasts with the size-control that these can achieve relative to self-limitation in equilibrium, single-species assemblies.

Contents

I. INTRODUCTION

A. Overview

Self-assembly is a process in which multiple subunits, or "building blocks", spontaneously organize into collective and coherent structures. This process is ubiquitous in living systems, where it underpins a wide range of structures at the cellular and sub-cellular scale, from lipid membranes to multi-protein filaments and capsules (Alberts *et al.*, 2002). Inspired by biology's successful strate-

gies to build functional nanostructures, self-assembly is forming the basis of modern approaches to generate materials from the "bottom up" (Hamley, 2003). Chemical techniques enable synthesizing a bewildering array of small-molecule, macromolecular, or particulate subunits that are engineered to self-assemble into high-order architectures (Boles *et al.*, 2016; Klok and Lecommandoux, 2001; Stupp and Palmer, 2014). As in the biological context, the assemblies bridge between the scales of molecules and chemical function (nanometer and subnanometer) to size scales that are useful for controlling material properties (microns and beyond).

In different domains of science and engineering, the term "self-assembly" often connotes a range of distinct, if overlapping, physical processes. In its broadest usage, self-assembly implies the collective association of multiple elements into organized configurations, by dynamics that start from a relatively "disorganized" state and evolve with at least some degree of randomness. The great conceptual appeal of self-assembly in materials science is that the instructions for a desirable or useful structure may somehow be imprinted into the assembling subunits themselves, such that in a simple mixture, the desired target structures emerge from the random processes of Brownian motion and subunit association.

In this article, we focus on *self-limiting assembly* (SLA), defined as self-assembly processes that terminate at an equilibrium state in which superstructures have a well-defined and finite spatial extent in one or more dimensions. Many examples of SLA can be found in biological systems, where the assembly of identical subunits into larger, yet finite-sized, superstructures is common and functionally vital. As shown in Fig. 1, examples include i) the protein shells that enclose viruses (Caspar and Klug, 1962; Mateu, 2013; Perlmutter and Hagan, 2015) and microcompartments (Kerfeld *et al.*, 2010; Rae *et al.*, 2013; Tanaka *et al.*, 2008), ii) finite-size protein superstructures in photonic tissues (McPhedran and Parker, 2015; Prum *et al.*, 2009; Saranathan *et al.*, 2012), and iii) finite-diameter bundles and fibers of cytoskeletal or extracellular protein filaments (Fratzl, 2003; Neville, 1993; Popp and Robinson, 2012). Each of these examples shares the notable feature that the finite size of the assembled structure far exceeds the nanometer size scale of the protein subunits. Crucial to their biological roles, the functional properties of these protein superstructures are regulated through the control of their finite size: respectively, (i) selective encapsulation and transport; (ii) optical response; and (iii) stiffness and strength. In this way, Nature exploits self-assembly to deploy structures, built from the same or similar subunits, in diverse intracellular and extra-cellular environments, and adapts their performance and functions by controlling the size of the assembled structure.

In contrast to these examples of SLA, most typical mechanisms of self-assembly in synthetic systems result in *unlimited* organized states, such as crystalline or liquid crystalline mesophases. In these states, structure may be
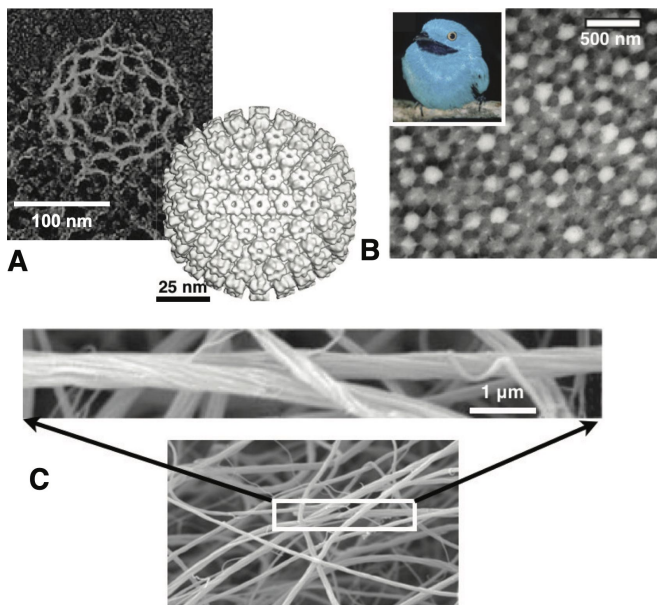

FIG. 1 Functional, finite-sized assemblies of proteins in biology: (A) protein shells of clathrin (left) and viral capsids of Herpes simplex (right); (B) photonic nanostructures form by keratin aggregates in feather barbs of Plum-throated Continga (inset); and (C) finite-diameter fibers in reconstituted fibrin clot. Figures are adapted from (Royle, 2012) (A, left) and (Baker *et al.*, 2000) (A, right); (Dufresne *et al.*, 2009) (B); and (Weisel, 2004) (C).

well-defined on some microscopic scale, such as the unit cell dimension, but its overall size is uncontrolled by assembly thermodynamics. This result, which may be described as bulk phase separation, is a generic consequence of the thermodynamic trade-off between entropic and energetic drives. In the most general case, once the net cohesive drive for a subunit to join an assembled structure exceeds the entropic penalty for giving up its higher configurational freedom as a disassociated unit, there is no thermodynamic reason to stop this process. Thus, subunits continually add to the aggregate until it reaches macroscopic proportions and the subunits are nearly depleted.

This review aims to describe the basic physical ingredients and common outcomes of assembly mechanisms that terminate at well-defined, finite sizes. We draw upon examples of SLA from biological systems, and consider the requirements to achieve such assemblies in synthetic systems. For clarity of presentation, we specifically focus on assemblies comprising a *single species of identical subunits*. Moreover, we restrict our definition of SLA to *equilibrium* assembly mechanisms, meaning that assembly terminates at a finite-sized *free energy* minimum structure.

Equilibrium assembly processes deserve special focus for both conceptual and practical reasons. A key advantage is that they are described by well-defined and generic statistical mechanical principles. This allows one, as we attempt to do in this article, to draw sharp distinctions between assemblies that either *are* or *are not*

self-limiting. Of course, reaching thermodynamic equilibrium requires subunits to associate and disassociate from aggregates sufficiently freely that a thermodynamically large collection of subunits behaves ergodically, sampling a sufficiently large ensemble of aggregation states in an experimentally relevant time. For systems at or near room temperature, such conditions are accessible when assembly is driven by non-covalent and reversible interactions, of the type that characterize physical association between macromolecules and colloidal particles in solutions (Israelachvili, 2011; Russel *et al.*, 1989), including van der Waals, electrostatic, hydrophobic, hydrogen-bonding, and depletion forces.

The ability of reversibly associating assemblies, if given sufficient time, to proceed toward one specific, thermodynamically-defined state, points to practical advantages of equilibrium assembly. As evidenced by the synthetic approaches to size-controlled structures referenced above, non-equilibrium control over finite-dimensions of assemblies requires extensive protocols to control the assembly environment, for example, precisely regulating the temporal sequence of temperatures and subunit concentrations. This makes it exceedingly difficult, if not impossible, to deploy these non-equilibrium size-control strategies in uncontrolled environments, such as the complex and dynamic milieu of living organisms. In such scenarios where assembly cannot be carefully "supervised", equilibrium mechanisms of assembly offer the distinct advantage that the final states may still be well defined. For example, viruses can exert only limited control over the inter-cellular media of their host organisms. Nevertheless, to be infectious, size-controlled capsid shells must assemble with high-fidelity from the capsomer subunits. While this assembly process is in general not purely equilibrium, biology often achieves such high fidelity by building upon equilibrium processes. For example, many viral capsids can spontaneously assemble from their purified components under (near) equilibrium conditions, with structures that are indistinguishable from capsids formed within a host cell (Fox *et al.*, 1998; Johnson and Speir, 1997; Wang *et al.*, 2015; Wingfield *et al.*, 1995), and in some cases are even infectious (e.g. (Fraenkel-Conrat and Williams, 1955)).

At the center of this specialized focus on equilibrium mechanisms for self-limiting single-species assembly is the puzzle: *how can equilibrium association processes "measure" the assembly to select a thermodynamic preferred state that is larger than a single subunit, yet less than infinite (i.e. bulk)?* Because thermodynamic equilibrium is independent of the history of system, this state cannot be defined by the temporal process in which subunits arrive to the aggregate. Nor do these identical subunits have specific "addresses" that prescribe where they are supposed to sit in a particular aggregation state. The answers, not surprisingly, lie in how the shape and interactions of subunits conspire to determine the dependence of assembly energetics on size. For example, in the canonical example of SLA, formation of spherical micelles from

amphiphillic molecules (Israelachvili *et al.*, 1976), the assembly motif favors individual subunits to span the assembly from the solvophobic core to the solvophillic surface. Hence, in this case it is intuitive that energetics favors aggregates that are limited to sizes that are comparable to the length of amphiphilies themselves. Far less intuitive is how single-species assemblies select finite equilibrium sizes that are much bigger than the subunit dimensions, or the range of their interactions. That is to say, what are "limits of self-limitation" — i.e., how large can a self-limited structure be, and how does this size limit depend on the physical characteristics (e.g. shape, interactions) of the subunits?

B. Outline

With these basic questions in mind, this review has two broad aims. We first overview the generic statistical and thermodynamic elements of SLA, and then present a broad classification for known mechanisms of SLA of identical subunits. The article is organized into two main sections based on these aims. In Sec. II, we present the thermodynamic principles of SLA based on the statistical mechanics of ideal aggregation of identical subunits. This begins with an introduction to ideal aggregation theory and illustration of the more generic case of *unlimited assembly*. Following this, we introduce a generic description of the ingredients for SLA. We review how the onset of aggregation – known as the *critical aggregation concentration* (CAC) – the self-limiting size of aggregates, and the statistics of aggregate size fluctuations depend on the functional form for the size-dependence of the intra-aggregate interaction energy. We then review the conditions for "polymorphic" SLA, in which assemblies exhibit multiple states of aggregation (some finite, some not). These systems are characterized by so-called secondary CACs, in which increasing concentration sufficiently far above the CAC leads to additional transitions between aggregation states.

Sec. III describes physical systems that exhibit SLA, and classifies the models that capture their behavior into two categories illustrated schematically in Fig. 2: *self-closing* and *open boundary* assembly. The former category describes assembly processes that terminate because they close upon themselves (Fig. 2A), and applies to shell and tubule formation, as well as the micellar assembly of surfactants, co-polymers, and other amphiphiles. The latter category applies to arguably lesser known classes of systems that have short-range attractions and long-range repulsions, or form geometrically-frustrated assemblies. These conditions enable assembly to terminate when the aggregate still has "open boundaries" characterized by a finite surface energy (Fig. 2B). For this case, we introduce a generic framework for understanding how the interplay between intra-aggregate stress accumulation and aggregate surface energy controls the finite-size of aggregates and the phase boundary between self-limiting and bulk

A self-closing assembly

e.g. tapered particles
with anisotropic binding

inter-subunit
"bond"

B open boundary assembly

e.g. "misfit" particles
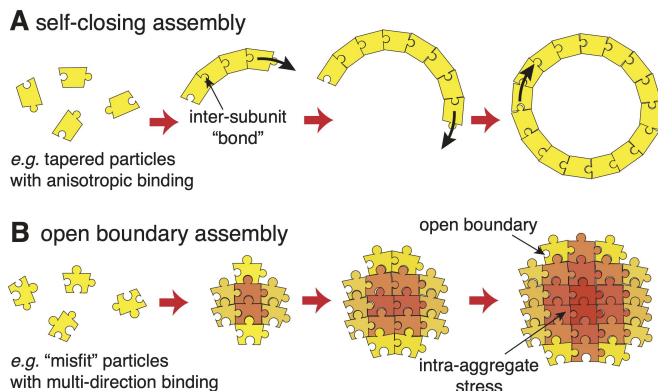with multi-direction binding

open boundary

intra-aggregate
stress

FIG. 2 Schematic illustrations of two classes of SLA described in Sec. III: (A) Self-closing assembly, in which inter-subunit rotations lead to cohesive assembly into closed, boundary-free aggregations; and (B) Open boundary (self-limiting) assembly, in which intra-aggregate stress accumulates with assembly and restrains the cohesive drive toward unlimited size.

aggregation.

While the core focus of this review is on the ingredients and outcomes of SLA from the point of view of thermodynamic equilibrium, the kinetic processes by which such systems reach equilibrium states (or in some cases, not) are essential to their study, particularly from the experimental point of view. A comprehensive review of kinetic limitations on assembly, which is relevant to both self-limited and unlimited assembly, is beyond the scope of this article. Nevertheless, we provide a basic introduction to some key considerations of assembly kinetics in Sec. IV. In that section, our purpose is to illustrate how those features of the assembly energetics that give rise to size-selection in equilibrium influence the principle kinetic pathways of their formation.

Before concluding, we provide a discussion in Sec. V of physical mechanisms leading to finite-size aggregates that fall outside of the major scope of the review, namely non-equilibrium and multi-species SLA, and questions these pose to the forgoing discussion for the more limited focus on single-species equilibrium SLA. We conclude with some remarks about open challenges in the application of the mechanisms and principles of SLA.

C. Scope of review

This review considers equilibrium assembly mechanisms that terminate at well-defined, finite sizes. As this focus suggests, we will leave out discussion of non-equilibrium processes in general, and more specifically, what might be called active-assembly processes, such as the steady-state length of treadmilling and severing cytoskeleltal filaments (Desai and Mitchison, 1997; Mohapatra et al., 2016; Pollard, 2016). Beyond that, we specifically consider assembly mechanisms of a *single species of identical subunits*. To be sure, this leaves out an emerging and fascinating area of research on so-called

"addressable assemblies (Jacobs and Frenkel, 2016; Zeravcic et al., 2017), where mixtures of multiple distinct subunit species may be "programmed" to assemble into a specifically defined 3D structure in equilibrium. In this article, we provide only a limited discussion about size-controlled multi-species assembly and possible trade-offs with single-species mechanisms, particularly how the number of required species increases with target size.

Although the fabrication and synthesis of finite, size-controlled structures is well-known in synthetic materials, for example, size-controlled nanoparticles of atoms (Cozzoli et al., 2006; Yin and Alivisatos, 2005) and macromolecules (Hiemenz and Lodge, 2007), these examples raise a key distinction between equilibrium and non-equilibrium assembly. The control over finite size in all of these foregoing examples relies on the non-equilibrium process by which they form. For example, the size distribution of metal nanoparticles (O'Brien et al., 2016; Yin and Alivisatos, 2005) is selected through spatio-temporal control of the physical-chemical factors that control nanocrystal growth (e.g. concentrations, temperature, ionic conditions). Indeed, as we discuss below, in generic conditions under which such assemblies form, allowing these assemblies to proceed to thermodynamic equilibrium would destroy the size control. Finite-sizes are only possible when these processes are driven, maintained, and arrested out of equilibrium. In this sense, we reserve the term "self-limiting" for those rarefied assembly processes that result in finite-size structures in thermodynamic equilibrium. The physical mechanisms of equilibrium assembly that achieve such size control are the central focus of this article.

II. THERMODYNAMIC ELEMENTS

We begin with a review of the elementary statistical mechanical framework to describe equilibrium aggregation. We then illustrate the statistical thermodynamics of aggregation in models of what we will call *canonical aggregation*, where assembly proceeds via cohesive (short-range and stress free) assembly of elemental units into 1D, 2D, and 3D aggregates. We illustrate how so-defined canonical assemblies do not exhibit self-limitation. We then describe the generic conditions for self-limiting (*finite*) equilibrium assembly, and give an overview of the concentration-dependent thermodynamics of self-limiting assembly. Finally, we discuss models of *competing* finite aggregates and *polymorphic* transitions between finite to unlimited assembly, both of which may by characterized multiple aggregation thresholds in the ideal theory.

A. Equilibrium principles

In this review we concern ourselves with equilibrium association of single subunits, or *monomers*, into states

with aggregation number n subunits, or *n-mers*. Our purpose is to describe the minimal ingredients of assembly dominated by structures with finite aggregation number n. To this end, we restrict our presentation to *ideal aggregation theory*, where interactions between distinct aggregates are neglected. This is not to say that interactions among subunits within the same aggregate are neglected. Quite the contrary, as we describe below, the intra-aggregate energetics, and its n-dependence, are critical for determining whether or not association leads to self-limited states, or instead, more canonical states of bulk aggregation.

1. Classical aggregation theory: fixed total concentration, non-interacting aggregates

Ideal aggregation theory is well established for certain classes of self-assembly systems, particularly in the context of amphiphiles and surfactants. As such, this theory is better described, and in greater depth, in references such as (Gelbart *et al.*, 1994; Israelachvili *et al.*, 1976; Safran, 1994; Tanford, 1974). Here, our purpose is to consider the application and implications of ideal aggregation theory to a broader class of self-limiting assembly systems. Hence, we only present a minimal introduction to the elements necessary to describe aggregation to self-limiting states.

We consider a solution of N total subunits in a fixed total volume V. In what follows, we refer to unassembled single subunits as *monomers* [1]. To describe the concentration of subunits, it is convenient to scale concentration by the reference state concentration v_0^{-1}; i.e., v_0 is the volume per subunit in the reference state [2]. In this way, we can define concentration in non-dimensional terms as the *total volume fraction* of subunits, $\Phi = N v_0 / V$ [3]. The N subunits are distributed among distinct n-mer aggregates, with the volume fraction of subunits in n-mers defined as ϕ_n. In the following, we refer to ϕ_n as the *subunit fraction distribution*. Throughout this review, it is also useful to introduce a separate variable for the *aggregate distribution* $\rho_n \equiv \phi_n / n$, which describes the relative

---

[1] In literature on amphiphile aggregation the term "unimer" is often used to describe the single subunit, to avoid overlap with the connotation of "monomer" as the chemical repeat of macromolecular chain, which is often a component of self-assembling molecular subunits.

[2] For concreteness, we may take v_0 to be the subunit volume in the disassociated state, which provides a convenient and non-dimensional measure of concentration which will be much less than unity for dilute conditions. However, the formalism is independent of choice of reference state; for example, $v_0 = 1/N_{\text{AV}}$ liter with the reference state concentration of 1 mol/liter commonly used in the life sciences. Changing the definition of the reference state only uniformly rescales the free energy of intra-aggregate interactions $\epsilon(n)$.

[3] Throughout, we also refer to a Φ more colloquially as the "concentration" of subunits

count of n-mers in the mixture. Defined in this way, ϕ_n and ρ_n are all less than unity for all n. Moreover, for the particular assumptions of ideal aggregation to hold (i.e. two-body contacts between aggregates are vanishingly rare), these quantities must all remain much less than unity. To simplify the nomenclature, throughout the review we refer to the non-dimensional concentration (volume-fraction) simply as the concentration.

We define $n\epsilon(n)$ as the free energy of intra-aggregate interactions; i.e., $\epsilon(n)$ is the *per subunit* aggregation free energy in an n-mer. The total free energy F for the ideal distribution of aggregates is given by

$$\frac{F}{(V/v_0)} = \sum_{n=1}^{\infty} \phi_n \Big(\epsilon(n) + \frac{k_{\text{B}}T}{n} \big[\ln(\phi_n/n) - 1 \big] \Big), \quad (1)$$

with the two terms in the parentheses respectively representing the intra-aggregate interaction free energy and translational entropy (in the ideal solution approximation) of n-mers, with the $1/n$ in the latter term reflecting the critical fact that all subunits of an n-mer share a common, single center-of-mass degree of freedom.

To obtain the equilibrium aggregate size distribution, we minimize F with respect to ϕ_n, subject to the constraint that the total subunit concentration is fixed:

$$\sum_{n=1}^{\infty} \phi_n = \Phi, \quad (2)$$

i.e.,

$$\frac{\partial}{\partial \phi_n} \Big[F + \mu \Big(\Phi - \sum_{n=1}^{\infty} \phi_n \Big) \Big] = 0 \quad (3)$$

with μ playing the role of a Lagrange multiplier. This yields

$$\mu = \epsilon(1) + k_{\text{B}}T \ln \phi_1 = \epsilon(2) + \frac{k_{\text{B}}T}{2} \ln(\phi_2/2) =$$

$$\ldots = \epsilon(n) + \frac{k_{\text{B}}T}{n} \ln(\phi_n/n) = \frac{\partial F}{\partial \phi_n} \quad (4)$$

showing that μ is the subunit chemical potential. This condition requires that subunits have the same chemical potential in all aggregates, and it derives from both the energetics of the assembly and the (ideal) translational entropy of the n-mer. Note that the prefactor of $1/n$ of the translational entropy, deriving from the sharing of a single center of mass in an n-mer, reflects a generically higher translational entropy of disaggregated states. The limit $n \to \infty$ gives $\mu = \epsilon(\infty)$, describing the equilibrium between monomers and a bulk phase-separated condensate which has no translational entropy. For simplicity of notation, throughout this article, we choose to define energies such that $\epsilon(1) = 0$, in which case $\epsilon(n)$ is defined as the *difference* of the per subunit energy between an n-mer and disassembled monomers.

It is convenient to use the first equality in eq. (4) to recast the chemical potential in terms of the (unknown)

monomer concentration, from which we can reformulate the generic chemical equilibrium conditions in terms of the *law of mass action*

$$\phi_n = n\Big(\phi_1 e^{-\beta\epsilon(n)}\Big)^n. \tag{5}$$

where $\beta^{-1} \equiv k_{\mathrm{B}}T$. Inserting the expression for $\phi_n(\phi_1)$ into the fixed number concentration eq. (2) and summing over n then results in an equation of state relating the total concentration Φ to the monomer concentration ϕ_1. This equation of state and the underlying distribution of assemblies, ϕ_n, derive from the specific n-dependence of aggregate interactions, with equilibrium states that are dominated by self-limited aggregates occurring only for certain forms of $\epsilon(n)$.

2. Unlimited assembly: short-range cohesive aggregation

Before describing models that give rise to SLA, we first consider the thermodynamics of the simplest models of physical association, described by short-range attraction between subunits. Crucially, while these models are relevant to a broad range of physical scenarios like colloidal crystallization (Manoharan, 2015; Morphew *et al.*, 2018), they do not exhibit SLA. Yet, they will serve as a useful reference point for illuminating the necessary conditions for SLA. Specifically, these models result in either a single dispersed state whose most populous aggregate state is $n = 1$ (the free monomer), or coexistence between the dispersed monomer-dominated state and an unlimited aggregate (macrophase separation). The absence of equilibrium finite-sized states will be traced to the generic size-dependence of the cost of open boundaries at the edges of cohesive aggregates. In subsequent sections we will show mechanisms that compete against the open boundary cost to enable SLA.

Here, we consider models where inter-subunit association promotes uniform d-dimensional aggregates, e.g. 1-dimensional chain-like or 2-dimension sheet-like aggregates. Every internal subunit forms, on average, z attractive bonds of strength $-u_0$, and subunits at the free boundary have δz fewer contacts (e.g. Fig. 3A). For example, in $d = 1$ chain-like assembly $z = 2$ and $\delta z = 1$. For the generic dimensionality, the interaction free energy takes the form

$$\epsilon(n) = -\epsilon_0 + \frac{\Delta_0}{n^{1/d}}, \tag{6}$$

where $\epsilon_0 = u_0 z/2$ and represents the per subunit cohesive free energy in the bulk (i.e. $n \to \infty$) structure. The second term derives from the growth of number of particles at the boundary ($\sim n^{(d-1)/d}$) and their deficit of cohesive bonds (δz), so that Δ_0 is equal to $u_0(\delta z)$ times a geometric factor accounting for the mean bond geometry at the boundary. Notably, the bonding-geometry in these assemblies permits the structure to grow uniformly without disrupting this local contact structure at any size

scale, a condition that we revisit when describing examples of self-limiting assembly in Sec. III.

We define the concentration $\phi_{\mathrm{s}} \equiv e^{-\beta\epsilon_0}$, so that the law of mass action, eq. (5), takes the form

$$\phi_n = n\big(\phi_1/\phi_{\mathrm{s}}\big)^n e^{-\beta\Delta_0 n^\alpha}, \tag{7}$$

where $\alpha = 1 - 1/d$ is an exponent that characterizes the geometric growth of the exposed boundary with n. As shown in Fig. 3B, $\phi_1 \leq \phi_{\mathrm{s}}$ and the distribution $\rho_n = \phi_n/n$ decreases exponentially with aggregate size for large n, and for any d. Hence, under these conditions Φ, the sum over the subunit fraction distribution in eq. (2), is finite, implying the existence of conditions where the concentration of subunits achieves equilibrium in the suspension. However, no such equilibrium exists for $\phi_1 > \phi_{\mathrm{s}}$, implying that $\phi_1 \to \phi_{\mathrm{s}}$ is an upper limit to concentrations that may be in equilibrium in a dispersed state. In other words, when Φ is sufficiently large that $\phi_1 = \phi_{\mathrm{s}}$ the solution is *saturated*, and additional subunits (further increasing Φ) must phase separate to the macroscopic state (i.e. $n \to \infty$).

First consider the linear case ($d = 1$), where the equation of state can be ready computed from eq. (7) with $\alpha = 0$ and the geometric series,

$$\Phi = e^{-\beta\Delta_0} \frac{\phi_1\phi_{\mathrm{s}}}{(\phi_{\mathrm{s}} - \phi_1)^2} \quad \text{for } d = 1 \tag{8}$$

which notably diverges as $\phi_1 \to \phi_{\mathrm{s}}$. As plotted in Fig. Fig. 3C, this divergence indicates that the monomer concentration increases with total concentration, but never reaches the point of saturation (i.e. $\phi_1 < \phi_{\mathrm{s}}$ for any finite Φ). Hence, for all subunit concentrations the system maintains $\phi_1/\phi_{\mathrm{s}} < 1$, implying that the distribution of linear aggregates is always exponential, $\phi_n/n \propto e^{-n/\langle n \rangle}$, where the number-average length is $\langle n \rangle = 1/\ln(\phi_{\mathrm{s}}/\phi_1)$. [4] Noting that $\epsilon_0 = \Delta_0$ for a 1D chain assembly, the growth of mean length with end energy in the limit of high concentration $\langle n \rangle \simeq e^{\beta\Delta_0/2}\sqrt{\Phi}$, which is well-known for equilibrium polymers (Hiemenz and Lodge, 2007) and cylindrical micelles (Gelbart *et al.*, 1994; Safran, 1994), highlights the mechanism that prevents "bulk" assembly for 1D aggregation. In this dimension, the probability to introduce a free end remains finite $\sim e^{\beta\Delta_0}$ in the $n \to \infty$ limit, analogous to the statistics of domain walls in the 1D Ising model at finite temperature (Fisher, 1984). However, while the mean-size is finite, this case is distinct from what we will describe as *self-limiting assembly*, in both the strong dependence of $\langle n \rangle$ on total concentration, and perhaps more significantly, the fact that fluctuations in aggregate size are always comparable to the mean; that is, $\langle (n - \langle n \rangle)^2 \rangle^{1/2} \propto \langle n \rangle$.

---

[4] The *number-average* aggregate size is $\langle n \rangle = (\sum_n n\rho_n)/(\sum_n \rho_n) = (\sum_n \phi_n)/(\sum_n \phi_n/n)$, where ϕ_n/n is the distribution of n-mers, while the *mass-average* aggregate size is $\langle n \rangle_{\mathrm{M}} = (\sum_n n\phi_n)/(\sum_n \phi_n)$.
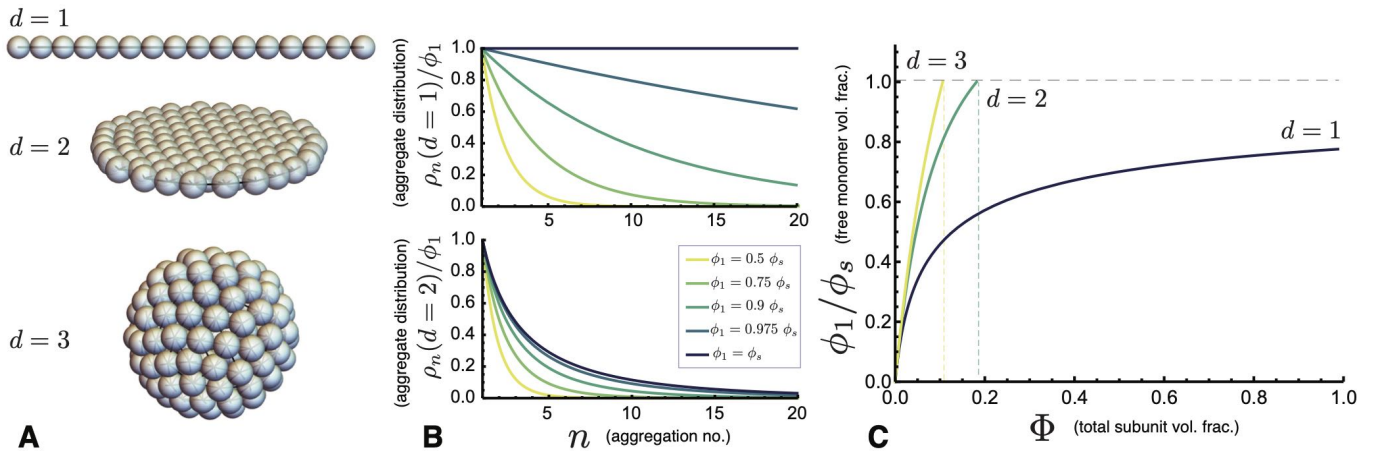
FIG. 3 (A) Examples of d-dimensional (linear, planar and spherical) short-range cohesive aggregation. (B) Plots of the aggregation distributions (relative counts of n-mers) for $d = 1$ (top) and $d = 2$ (bottom) for monomer concentrations increasing to saturation (i.e. $\phi_1 = \phi_s$) for $\Delta_0 = 1\ k_BT$. While the dispersity (and mean size) of linear aggregates diverges as $\phi_1 \to \phi_S$, it remains finite for $d \geq 2$ at saturation. (C) The equation of state (ideal aggregation theory) for the free monomer population ϕ_1 as a function of the total concentration Φ for linear, planar and spherical aggregates for $\Delta_0 = 2.75\ k_BT$. For $d = 2$ and $d = 3$ the free monomer concentration saturates at a finite Φ where $\phi_1 = \phi_S$. For linear aggregation, saturation is not reached in the ideal theory.

For higher assembly dimensionality, $(d > 1)$, the geometric growth of the free boundary cost restrains the $n \to \infty$ divergence in the distribution eq. (7) as the solution approaches saturation. For $\phi_1 = \phi_s$ the distribution takes the form $\phi_n(\phi_1 \to \phi_s) = ne^{-\beta\Delta_0 n^\alpha}$, the sum over which converges for $\alpha > 0$ when $d > 1$. For example, we can approximate the sum for planar aggregates $(d = 2)$ by replacing the sum over aggregation number n with dimensionless aggregate radius $n = \pi r^2$, i.e. $\sum_n \phi_n \to 2\pi \int dr\, r\, \phi(r)$. At saturation, aggregates sizes are exponentially distributed, $\phi(r) \simeq \pi r^2 e^{-\sqrt{\pi}\beta\Delta_0 r}$, yielding a saturation concentration

$$\Phi_s(d = 2) \simeq 12\Big(\frac{k_BT}{\Delta_0}\Big)^4 \tag{9}$$

at which point the ideal solution of aggregates reaches equilibrium with the bulk condensate. Thus, with the exception of the special case of $d = 1$ assembly, short-range cohesive aggregation is characterized by a finite saturation concentration, $\Phi_s = \Phi(\phi_1 = \phi_s)$, above which subunits phase separate into an unlimited bulk structure.

The thermodynamics of these examples are plotted in Fig. 3C in terms of the equation of state $\phi_1(\Phi)$ for each dimensionality. Note that while the distributions of short-range interacting systems have finite mean sizes in the absence of macrophase separation, their distributions are dominated by monomers; i.e., the concentration of n-mers ρ_n is always maximal for $n = 1$, a property that sharply contrasts with the SLA behavior described next in Sec. II.B.

B. Self-limiting assembly: Elements and outcomes

In this section, we describe the generic ingredients and thermodynamic outcomes of assembly models that ex-

hibit *self-limitation*. That is, unlike the short-range cohesive models described above, these systems undergo ideal assembly into self-limiting states dominated by aggregates with a finite size n_* that is larger than 1, yet smaller than bulk (unlimited) states.

The physical mechanisms that give rise to this behavior will be discussed in detail in Sec. III. Here, we give an overview of the essential thermodynamic ingredients and behavior based on a generic description of the energetics of a self-limiting system. We consider the assembly behavior in terms of a generic function for the interaction free energy per subunit, $\epsilon(n)$, which, as sketched in Fig. 4A, favors aggregation at a particular finite size, or possibly several distinct finite sizes. To highlight its distinct role from the translational entropy of aggregation, we use the term *aggregation energetics* to refer to $\epsilon(n)$, but we note that this describes a free energy per subunit, as interactions in general have both energetic and entropic contributions.

Given a known form of the $\epsilon(n)$, the discussion of this section addresses several key questions. First, what determines the onset of aggregation from the dispersed state? Second, what selects the (dominant) size of finite aggregates? And third, what are the conditions for driving transitions between different states of self-limiting aggregation, or between self-limited and unlimited aggregation states?

1. Aggregation threshold

We first describe the simplest picture of the concentration dependence of ideal aggregation. We begin with the assumption of an energy per subunit $\epsilon(n)$ of the form shown in Fig. 4A, which has a single energy minimum at a finite aggregation number n_T, which we call the *target*
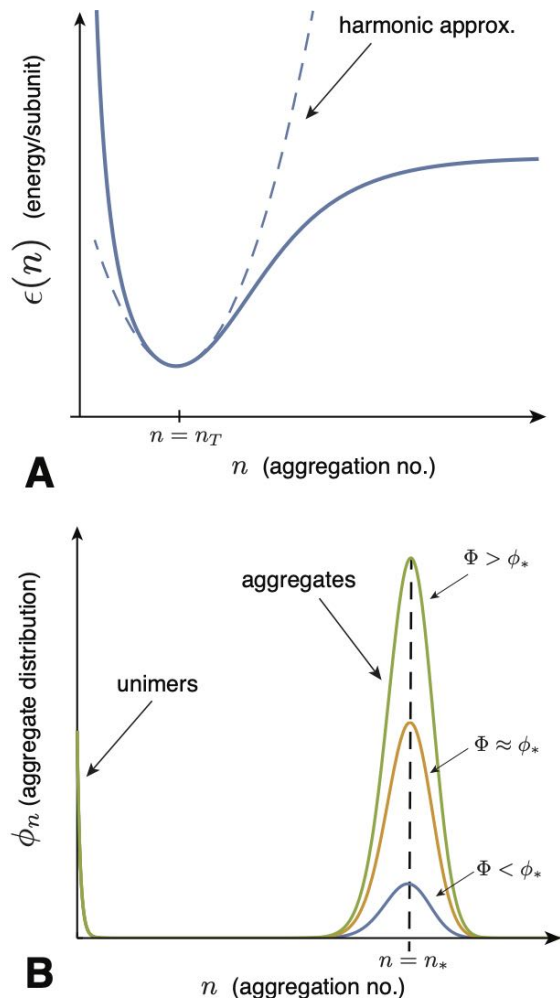
FIG. 4 (A) A schematic plot of the aggregation (free) energy per subunit is shown as a continuous function of aggregation number n. The dashed line shows a harmonic expansion around a local minimum at the *target size* $n = n_T$. (B) Schematic plots of the aggregate distribution for a model of the type sketched in (A).

size. The basic dependence of aggregation on concentration for such a model is sketched in Fig. 4B. There are two dominant populations of aggregates, *monomers* and *n-mers*, with the n-mers narrowly distributed around the most populous state $n_* \approx n_T$.

For large enough n_*, the thermodynamics of aggregation can be captured, to a first approximation, by a *two-state*, or bimodal, distribution, in which fluctuations around free monomers and the n-mer aggregate peak are neglected. We will self-consistently test the validity of this approximation below (see eq. 18). When subunits are distributed strictly between the $n = 1$ and $n = n_*$ states, the conservation of subunit mass is simply $\Phi = \phi_1 + \phi_{n_*}$. Chemical equilibrium then gives the concentration in preferred aggregates $\phi_{n_*} = n_* (\phi_1 e^{-\beta \epsilon_*})^{n_*}$, where $\epsilon_* \equiv \epsilon(n_*) < 0$ is the per subunit energy gain upon aggregation into the optimal size. Defining the concen-

tration scale

$$\phi_* \equiv \left[n_* e^{-n_* \beta \epsilon_*} \right]^{-1/(n_*-1)} \approx e^{\beta \epsilon_*} / n_*^{1/n_*}, \qquad (10)$$

yields the following equation of state, relating total concentration to monomer concentration

$$\frac{\Phi}{\phi_*} = \frac{\phi_1}{\phi_*} + \left(\frac{\phi_1}{\phi_*} \right)^{n_*}, \qquad (11)$$

a result derived originally by Debye to explain scattering in soap solutions (Debye, 1949). The dependence of the populations of monomers and n_*-mers on total concentration is plotted in Fig. 5A and can be summarized as follows. For low concentration, $\Phi \ll \phi_*$, and additional subunits added to the system go predominantly to monomers, since $\phi_1 \approx \Phi$ while $\phi_{n_*} \approx \phi_* (\Phi/\phi_*)^{n_*} \ll \Phi$. Notably, the population of aggregates in this regime ϕ_{n_*} is simply proportional to the random probability of n_* free subunits to spatially coincide, Φ^{n_*}, times the enhanced Boltzmann factor for aggregation, $e^{-n_* \beta \epsilon_*} \approx \phi_*^{-n_*}$, and hence is diminishingly small. In the large concentration regime $\Phi \gg \phi_*$, the dominant populations are reversed: the n_*-mer population increases in proportion to total concentration, $\phi_{n_*} \approx \Phi$, while monomers increase much more slowly, $\phi_1 \approx \phi_* (\Phi/\phi_*)^{1/n_*} \ll \Phi$. These two regimes are characterized by a crossover near $\Phi \approx \phi_*$, which is known as the *critical aggregation concentration* (CAC), although it is not strictly a phase transition for finite n_*.[5] As illustrated in Fig. 5A, the aggregation crossover becomes increasingly sharp as the aggregation number n_* increases. Fig. Fig. 5B shows an example of CAC behavior observed in experiments of hepatitis B virus capsid protein assembly (Ruan *et al.*, 2018).

Underlying the transition is a thermodynamic trade off between translational entropy and the interaction free energy that drives aggregation. Maximizing the ideal translational entropy of aggregates favors maximizing the number of independent translational degrees of freedom, i.e. the number of independent centers of mass in the mixture. To form an aggregate, n_* monomers must give up their n_* centers of mass, for the single center of mass of the aggregate. Only when the aggregation free energy is sufficient to "pay" this entropic price (i.e. when the concentration of "excess" monomers is sufficiently large), does aggregation become thermodynamically favorable. Hence, the CAC depends not only on aggregation energetics, but also the aggregation number n_*. According to eq. (11), ϕ_* exhibits a modest increase with n_* (as $\sim n_*^{-1/n_*}$) due to the increased translational entropy loss

---

[5] The CAC is commonly referred to as the *critical micelle concentration* (CMC) in the amphiphile literature. In the virus assembly literature it is often called the *pseudo-critical concentration* to emphasize that it does not correspond to a true phase transition for finite n_*, and because the CAC observed in finite-time experiments typically exceeds the equilibrium CAC due to nucleation barriers.
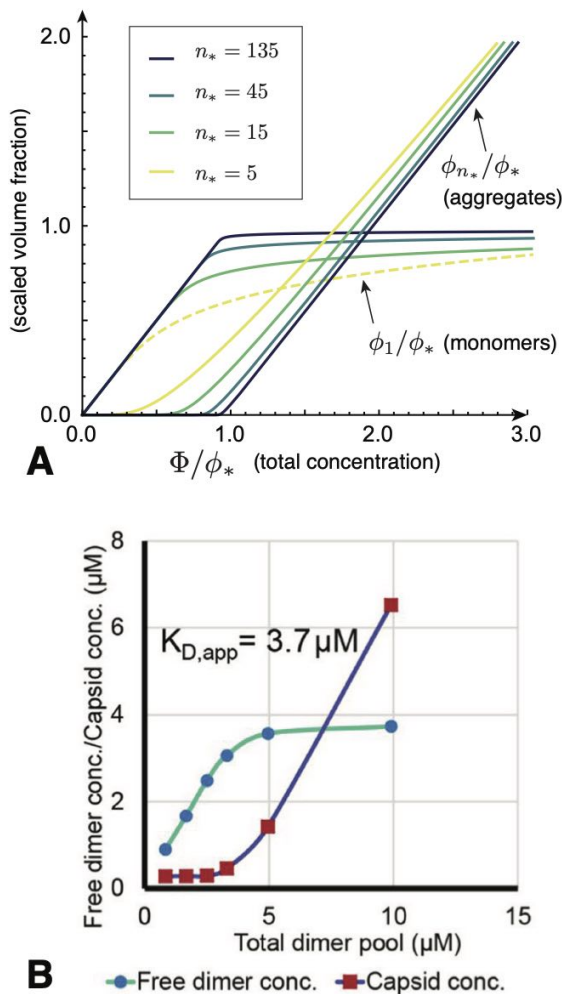
FIG. 5 (A) Plots of a "two-state" model composed of only monomers ($n = 1$) and aggregates of a single peak size ($n = n_* \approx n_T$) as functions of the total concentration Φ and for different finite aggregate sizes. The *critical aggregation concentration* (CAC), here ϕ_*, characterizes the concentration range beyond which aggregates dominate the subunit population. (B) Assembly behavior of hepatitis B virus (HBV) capsid protein in vitro as a function of concentration. The plot shows the concentration-dependence of the fraction of subunits (protein dimer) in two states: free subunits and assembled capsids composed of 120 subunits. Notice that as the total concentration crosses the CAC (labeled $K_{D,app}$), the concentration of free subunits is nearly constant, with almost all additional subunits assembling into capsids. Figure in (B) reprinted from (Ruan *et al.*, 2018).

for when joining larger aggregates. We return to the implications of the n_*-dependence of aggregation thresholds in the discussion of competing aggregate states below. Notice that although the model ignores physical interactions between distinct aggregates, the change in translational entropy couples n_* units, making aggregation a cooperative process. For this reason, the CAC becomes progressively sharper and tends toward a thermodynamic transition as $n_* \to \infty$ (Fig. 5A).

2. Finite aggregates: Mean size and size dispersity

Here we review the conditions for the most probable, or *optimal aggregate size* n_* given a known form of $\epsilon(n)$, which we assume for the moment to have a single minimum at target size n_T. The optimal size n_* corresponds to the maximum in the aggregate distribution $\rho_n = \left(e^{\beta(\mu - \epsilon(n))}\right)^n$, or equivalently, the minimum in the free energy $n[\epsilon(n) - \mu]$, which includes the total interaction free energy and entropy cost of forming an n-mer from free monomers. However, except under conditions where monomers are buffered to a fixed concentration, the chemical potential $\mu = \ln \phi_1$ varies as the equilibrium monomer population changes with total concentration. Naively, this might suggest that the optimal aggregate size should strongly vary with total concentration. Here, we illustrate why, notwithstanding the variation of μ with concentration, n_* is nearly independent of Φ and almost entirely determined by the form of aggregation energetics, $\epsilon(n)$. Following this, we summarize the effects of dispersity (i.e. finite width of the aggregation peaks in Fig. 4B) on aggregation thermodynamics, which is necessary to account for the (weak) concentration dependence of the optimal aggregate size.

Two-state aggregation: As a first approximation, consider the two-state aggregation model deep into the aggregation regime, i.e. well above the CAC ($\Phi \gg \phi^*$). The optimal aggregate size derives from the condition $\frac{d\rho_n}{dn}\big|_{n_*} = 0$, or

$$\epsilon(n_*) + n_* \epsilon'(n_*) - \mu = 0, \qquad (12)$$

where $\epsilon' = \frac{d\epsilon}{dn}$. Using the fact that $\phi_1(\Phi \gg \phi_*) \simeq \phi_*(\Phi/\phi_*)^{1/n_*} = e^{\epsilon_*}(\Phi/n_*)^{1/n_*}$ from eq. (11) in the previous section, this transforms to the condition for the optimal (peak) aggregate size

$$\epsilon'(n_*) = \frac{k_B T}{n_*^2} \ln\left(\Phi/n_*\right) \qquad (\text{two} - \text{state}). \qquad (13)$$

From eq. (13) we may draw two key conclusions. First, in the limit of large target size $n_* \gg 1$, the optimal size corresponds to a minimum of $\epsilon(n)$. That is, since $\epsilon'(n_*) \to 0$, $n_* \to n_T$ and the aggregate peak is selected by minimizing per sub unit aggregation energy, independent (to a first approximation) of concentration. Second, the right-hand side of eq. (13), which is proportional to the translational free energy of a dilute concentration of n_*-mers, is negative, and hence $\epsilon'(n_*) < 0$. Combining this with eq. (12), we find the inequality,

$$\mu < \epsilon(n_*). \qquad (14)$$

This last condition shows that the equilibrium chemical potential approaches from below, but never quite reaches, the interaction energy of the optimal aggregate $\epsilon(n_*)$ (excepting the unphysical limit $\Phi/n_* \to 1$).

Finally, the fact that n_* corresponds to a *maximum* in the aggregate size distribution, suggests the following

condition from $\frac{d^2\rho_n}{dn^2} < 0$

$$\epsilon''(n_*) > -\frac{2\epsilon'(n_*)}{n_*} > 0. \qquad (15)$$

As the righthand side goes to zero as $\sim n_*^{-3}$, the aggregation energetics must be convex in the vicinity of the optimal size. Strictly speaking however, a stronger condition than convexity alone is needed to justify the neglect of aggregation number fluctuations in the 2-state approximation, as discussed next.

Gaussian approximation: We now consider the effect of *convexity* of the aggregation energetics, characterized by the second-derivative of $\epsilon(n)$ at the peak aggregate size. As above, we restrict our analysis to the case of a single, well-defined minimum in $\epsilon(n)$ occurring at a finite target size $n_{\mathrm{T}} > 1$. Close to the minimal-energy size, the energetics have the form

$$\epsilon(n) \simeq \epsilon_{\mathrm{T}} + \frac{\epsilon''_{\mathrm{T}}}{2}(n - n_{\mathrm{T}})^2, \qquad (16)$$

where $\epsilon_{\mathrm{T}} < 0$ and $\epsilon''_{\mathrm{T}} > 0$ respectively characterize the minimum energy and convexity of the target aggregate, as illustrated in the harmonic approximation in Fig. 4A. Physically, ϵ''_{T}, which we call the *convexity*, quantifies (twice) the energetic cost (in $k_{\mathrm{B}}T$) to alter the aggregate number from its target by ± 1. In the following section, we will describe the physical effects that control convexity in different models of self-limiting assembly. Here, we see that the concentration-dependence of the mean (or peak) self-limiting size, as well as the size-dispersity, are controlled by a single combination of ϵ''_{T} **and** n_{T}.

The effect of finite convexity is to allow fluctuations in n around the peak size n_*. When ϵ''_{T} and n_{T} are sufficiently large, the aggregate distribution follows a Gaussian,

$$\rho_n(n \gg 1) \simeq e^{n_*\beta(\mu - \epsilon_*)} e^{-\frac{(n - n_*)^2}{2\langle\Delta n^2\rangle}}, \qquad (17)$$

where $\langle\Delta n^2\rangle$ characterizes the variance of aggregate sizes relative to n_*. Assuming that the Gaussian distribution of aggregates is well separated from the monomer peak, the size fluctuations around n_* may be summed in $\phi_n = n\rho_n$, yielding the same mass-action formula as eq. (11), but with a redefined CAC,

$$\phi_* \approx \frac{e^{\beta\epsilon_*}}{\left(n_*\sqrt{2\pi\langle\Delta n^2\rangle}\right)^{1/n_*}}, \qquad \text{(Gaussian).} \qquad (18)$$

Compared to the two-state approximation, ϕ_*, is depressed by a factor proportional to $\langle\Delta n^2\rangle^{1/2n_*}$ due to the comparative increase in the number of aggregates states and associated entropy. Likewise, well above the CAC, the monomer population is depressed (relative to the 2-state approximation) by the same factor. Combining this effect into the chemical potential with the peak aggregate condition in eq. (12) gives the following prediction for peak (mean) aggregate

$$n_* \simeq n_{\mathrm{T}}\left(1 + \frac{k_{\mathrm{B}}T}{n_{\mathrm{T}}^3\epsilon''_{\mathrm{T}}}\ln\left[\frac{\Phi}{n_{\mathrm{T}}\sqrt{2\pi\langle\Delta n^2\rangle}}\right]\right), \quad \text{(Gaussian)} \qquad (19)$$

where we are considering only the leading correction to $n_* - n_{\mathrm{T}}$. In this same limit, aggregate dispersity becomes,

$$\frac{\langle\Delta n^2\rangle^{1/2}}{n_{\mathrm{T}}} \simeq \frac{1}{\left(n_{\mathrm{T}}^3\beta\epsilon''_{\mathrm{T}}\right)^{1/2}}, \qquad \text{(Gaussian)}, \qquad (20)$$

where again averages are taken with respect to aggregate distribution ρ_n (i.e. number average).

The results of the Gaussian approximation in eqs. (19) and (20) highlight two physical effects of convexity. First, in eq. (19), the mean aggregate size always falls *slightly below* the minimal-energy target size (i.e. because $\Phi < 1$, and hence, the logarithmic factor is always negative). This "sub-optimal" aggregate size derives from the (translational) entropic preference for smaller-n aggregates, and hence, this weak depression of n_* decreases with increasing supersaturation as the translation free energy of aggregates $(k_{\mathrm{B}}T/n_{\mathrm{T}})\ln(\Phi/n_{\mathrm{T}})$ tends toward zero. Second, the relative shift of mean aggregation number $(n_* - n_{\mathrm{T}})/n_{\mathrm{T}}$ *and* the relative size variance $\langle\Delta n^2\rangle^{1/2}/n_{\mathrm{T}}$ decrease with the reciprocal of $\epsilon''_{\mathrm{T}}n_{\mathrm{T}}^3/k_{\mathrm{B}}T$. Hence, corrections from size variations become small, in relative terms, either for sharp minima, when $\epsilon''_{\mathrm{T}} \gg k_{\mathrm{B}}T$, or for larger aggregate number, when $n_{\mathrm{T}} \gg (k_{\mathrm{B}}T/\epsilon''_{\mathrm{T}})^{1/3}$. While at first glance, this might suggest a generic tendencies toward monodisperse aggregation in the large n_{T} limit, SLA models described in Sec. III show convexity to be a *decreasing* function of n_{T}. Hence, as it turns out, the decrease of relative size fluctuations with target size becomes non-trivially dependent on the geometric sensitivity of aggregation energy.

Self-limitation without minima: Before considering landscapes with more complex equilibria, we briefly note that it is possible to construct functional forms of $\epsilon(n)$ which exhibit self-limitation without local minima. As an example, consider a variation of the general type of energetics described in Sec II.A.2,

$$\epsilon(n) = -\epsilon_0 + \Delta(n) \qquad (21)$$

where $\Delta(n)$ is a monotonically decreasing, and convex, function of n so that the minimal energy per particle occurs for infinite aggregates, i.e. $\epsilon(n \to \infty) = -\epsilon_0$. The condition for a maximum in ρ_n in eq. (12) gives

$$\mu + \epsilon_0 \equiv -(\Delta\mu)_\infty = \Delta(n_*) + n_*\Delta'(n_*) \qquad (22)$$

Because the chemical potential is bounded from above by $-\epsilon_0$ (from eq. (14)) and hence $(\Delta\mu)_\infty > 0$, the conditions for finite optimal aggregate size can be satisfied (at some accessible μ) provided that $\Delta(n)$ decreases *faster* than $1/n$, or more specifically, that

$$\frac{\Delta'(n)}{\Delta(n)} < -\frac{1}{n}, \qquad (23)$$

for some range of finite n. For example, for any model that approaches the bulk energy as a power-law $\Delta(n) = \Delta_0/n^\gamma$ where $\gamma > 1$, it is straightforward to show the peak size obeys $n_* = \left[(\gamma - 1)\Delta_0/(\Delta\mu)_\infty\right]^{1/\beta}$, which increases continuously with concentration as $\Delta\mu_\infty \to 0$ from above.

While the previous argument shows that it is possible to construct mathematical examples of minima-free energy densities that result in finite-n peaks in ρ_n, physical cases of SLA fall outside of this category. For example, generic physical grounds suggest that aggregates possess a boundary, or surface, at which the assembly energy is different (generally higher) than in the interior, as described for cases of short-ranged cohesive interactions in Sec. II.A.2. In the limit of large n, the asymptotic contribution to the energy density from this boundary, $n^{-1/d}$ where $d \geq 1$, will dominate over other possible terms falling off faster than n^{-1} (such as the $\Delta(n)$ term in eqs. (21) and (23)), leading to unlimited assembly. Hence, we exclude such anomalous cases, and focus the discussion on situations where self-limitation is directly associated with well-defined minima of $\epsilon(n)$.

3. Competing states of aggregation

Secs. II.B.1 and II.B.2 above describe the simplest case of self-limiting assembly: a concentration-controlled crossover, or "pseudo-transition", from a monomer-dominated state to a state dominated by aggregates of one finite size. The finite aggregate size corresponds to a single minimum in $\epsilon(n)$, and the transition occurs at a single CAC. In this section, we overview the thermodynamics of cases in which assembly is characterized by multiple local minima, or instead, by transitions between self-limiting and unlimited aggregation states. In these cases, the aggregation thermodynamics can exhibit a more complex dependence on concentration, corresponding to *secondary CACs* between different aggregation states. While concentration-dependent transitions between different aggregates are commonly attributed to interactions between aggregates (Israelachvili, 2011), it is less widely appreciated that they can also occur in ideal aggregation models, which strictly neglect inter-aggregate interactions. As we review in Sec. III.A.2, the possibility of an "ideal" secondary CAC was first considered in the context of transitions between spherical and cylindrical surfactant micelles (May and Ben-Shaul, 2001; Porte *et al.*, 1984). In this section, we describe this behavior as a generic consequence of the translational entropy preferences for smaller aggregate sizes, and as such, ideal secondary CACs can occur in a much broader class of SLA models.

Two finite aggregate states: We first describe a simple model with only 2 states of finite aggregates in equilibrium with a population of free monomers, for simplicity ignoring number fluctuations around these three states. We consider two states of *small* and *large* aggre-

gates, corresponding to two well-defined local minima of $\epsilon(n)$, at n_S and $n_L > n_S$ subunits, respectively, as shown schematically in Fig. 6A. In this case, aggregation is controlled by not only the difference in the respective energy minima ϵ_S and ϵ_L, but also the difference in the aggregation number. To understand aggregation in the presence of multiple minima, it is convenient to define the nominal CACs corresponding to either aggregate state, from eq. (10),

$$\phi_\nu^* \equiv \left[n_\nu e^{-\beta n_\nu \epsilon_\nu}\right]^{-1/(n_\nu - 1)} \approx \frac{e^{\beta\epsilon_\nu}}{n_\nu^{1/n_\nu}} \quad \text{for } \nu = S, L \quad (24)$$

These are concentrations at which aggregates of either type would overtake free monomers, were it not for the additional equilibrium between the populations of small and large aggregates. In terms of these concentration scales, the law of mass action takes the form

$$\Phi = \phi_1 + \phi_S^*\left(\frac{\phi_1}{\phi_S^*}\right)^{n_S} + \phi_L^*\left(\frac{\phi_1}{\phi_L^*}\right)^{n_L}, \quad (25)$$

where the last two terms represent the respective populations of subunits in n_S-mers and n_L-mers , which we denote as Φ_S and Φ_L. As above for the case of a single minimum in $\epsilon(n)$, in the limit of high concentration, aggregation always proceeds towards the state with the lowest energy. However, in this case, there are two possible thermodynamic scenarios for the concentration dependence, depending on the relative energy difference between small and large aggregates:

i) $\epsilon_S < \epsilon_L$: Because $n_S < n_L$, in this case it is always true that $\phi_S^* < \phi_L^*$, which means that as concentration increases, $\phi_1 \to \phi_S^*$ before reaching ϕ_L^*. Above the threshold where $\phi_1 \approx \Phi_S$, monomers remain effectively "buffered" at $\phi_1 \approx \phi_S^*$, and it is straightforward to show that $\Phi_L \ll \Phi_S$ [6]. Thus, when the smaller aggregate has a lower per subunit aggregation energy, aggregation proceeds as if there was only a single target state with a CAC at ϕ_S^* and never yields a significant number of large aggregates.

ii) $\epsilon_S > \epsilon_L$: When the large aggregates are energetically favored, there are two possibilities. Consider first the case of large energy differences, such that $e^{\beta(\epsilon_L - \epsilon_S)} < \frac{n_L^{1/n_L}}{n_S^{1/n_S}}$. For this first regime $\phi_L^* < \phi_S^*$, and there is only a single CAC at the (lower) critical concentration for large aggregates. For the second regime, when the energy difference between large and small aggregates is smaller, in the range $1 > e^{\beta(\epsilon_L - \epsilon_S)} > \frac{n_L^{1/n_L}}{n_S^{1/n_S}}$, the order of the CACs reverses: $\phi_S^* < \phi_L^*$, and leads to two CACs. As shown in Fig. 6B, for this case upon increasing concentration from the dilute limit, the concentration reaches a

---

[6] The assumptions of ideal aggregation require that Φ must remain below unity, a condition that requires $\phi_1 \ll \phi_L^*$ for case (i).
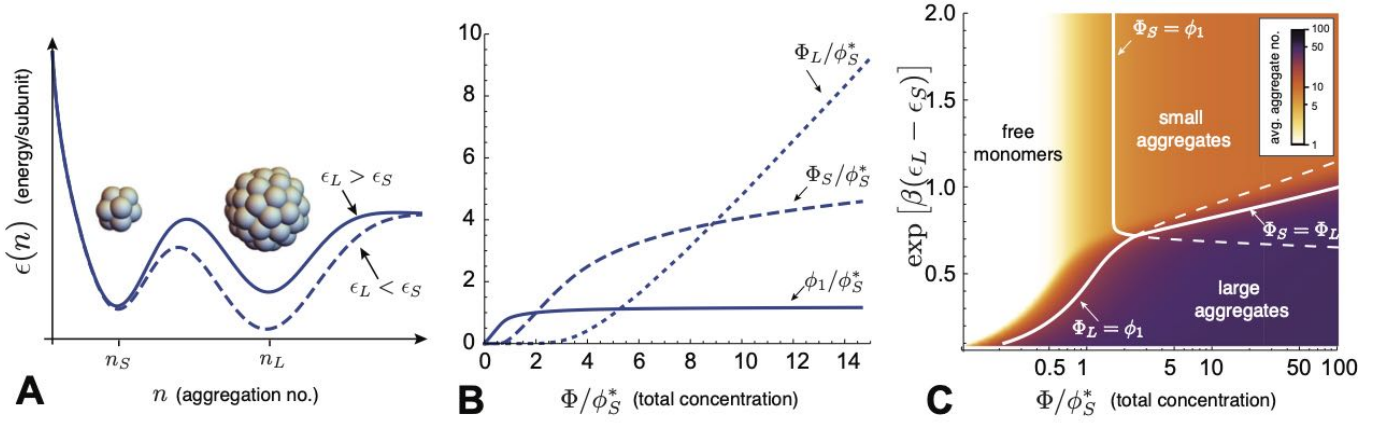
FIG. 6 (A) A schematic plot of the aggregation energetics for a case with two local minima corresponding to *small* and *large* finite aggregates, containing n_S and n_L respective subunits in their target sizes. The solid curve shows a case where small aggregates are the global minimum of $\epsilon(n)$, while the dashed curve shows a case where the larger aggregate is the global minimum. The latter case can lead to secondary CAC transitions between small and large aggregates, as shown in (B), which plots the monomer, small aggregate and large aggregate populations as functions of total concentration for large aggregates that are slightly energetically more favorable than small ones ($e^{\beta(\epsilon_L - \epsilon_S)} = 0.8$). In this case, small aggregates dominate at intermediate concentrations, but ultimately are overtaken by a second population of large aggregates above a second CAC. (C) An assembly state map for the two finite aggregate model, as a function of the (per subunit) energy difference between small and large aggrates and total concentration. The color scale shows the mean aggregate size, while the solid lines indicate boundaries between states dominated by monomers, n_S-mers, and n_L-mers. Dashed lines indicate where subdominant aggregates reach the monomer concentration. Plots in (B) and (C) show results of the two finite aggregate model for $n_S = 10$ and $n_L = 50$.

first CAC at $\Phi \approx \phi_S^*$, with a transition to a state dominated by small aggregates (i.e. $\Phi_S \gg \phi_1, \Phi_L$). This state persists until reaching a *second CAC* at $\Phi \approx \phi^{**}$, defined by a crossover in dominant aggregation state to $\Phi_L > \Phi_S$. The concentration threshold condition can be estimated by solving for the monomer concentration at which $\Phi_S(\phi_1) = \Phi_L(\phi_1) \gg \phi_1$ from eq. (25)

$$\phi^{**} \approx \phi_L^* \left(\frac{\phi_L^*}{\phi_S^*} \right)^{n_S/(n_L - n_S)}, \qquad (26)$$

which is larger than the "bare" value ϕ_L^* owing to the depletion of free monomers by small aggregates, and corresponds to a total concentration $\Phi^{**} \approx 2\phi^{**}$. The high-concentration regime above the second CAC is dominated by minimal-energy large aggregates, but maintains a sizeable amount of *small* aggregates (i.e. $\Phi_L \gg \Phi_S \gg \phi_1$), approximately buffered at the second CAC concentration.

To summarize this 2-aggregate model, it is possible to have *two* pseudo-critical transitions (as in Fig. 6B), first from disassembled monomers to small aggregates and then from small to large aggregates, provided the energy difference between aggregates is sufficiently small, in the window

$$0 < \epsilon_S - \epsilon_L < \frac{k_B T}{n_S} \ln n_S - \frac{k_B T}{n_L} \ln n_L. \qquad (27)$$

This is consistent with the secondary CAC behavior shown in the assembly state diagram in Fig. 6C, calculated for the case of $n_S = 10$ and $n_L = 50$.

The physical origin of this double CAC behavior can be traced to a competition between the higher cohesive energy of large aggregates pitted against the higher

(per subunit) translational entropy of smaller aggregates. This can be cast in terms of the chemical equilibrium between large and small aggregates, which requires

$$\mu = \epsilon_S + \frac{k_B T}{n_S} \ln \rho_S = \epsilon_L + \frac{k_B T}{n_L} \ln \rho_L. \qquad (28)$$

Energetically favorable large aggregates, $\epsilon_L < \epsilon_S$, require aggregate concentrations to adjust to maintain a suitably higher translational entropy of *small* aggregates, namely $\frac{1}{n_S} \ln \rho_S < \frac{1}{n_L} \ln \rho_L$, specifically

$$\rho_S/\rho_L = \frac{e^{n_S \beta(\epsilon_L - \epsilon_S)}}{\rho_L^{(n_L - n_S)/n_L}} \qquad (29)$$

This condition shows that the larger entropy of smaller aggregates requires that $\rho_S/\rho_L > 1$ provided that the concentration of large aggregates remains sufficiently small, below $\rho_L < \rho_L^{**} = \left[e^{\beta(\epsilon_L - \epsilon_S)} \right]^{n_S n_L/(n_L - n_S)}$. Hence, when n_S and the differential in aggregation energy are small enough, small aggregates remain more populous than large aggregates up to total concentrations that exceed the first CAC to the small aggregate state, until the second CAC.

This simplified model illustrates a generic conclusion. Even if an aggregate state does not correspond to the *global* minimum of $\epsilon(n)$, it may exhibit an entropically-stabilized window of thermodynamic dominance at intermediate concentrations, provided its target size is sufficiently small and its energy is sufficiently close to the global minimum. Next, we illustrate this entropic stabilization of finite (compact) aggregates in models for which the competing states are *unlimited*.

Finite and unlimited aggregates: While polymorphic assembly into multiple finite-number aggregates occurs in some natural and biomimetic systems (Lutomski *et al.*, 2018; Sun *et al.*, 2007; Wingfield *et al.*, 1995) and may be desirable for nanomaterials applications, cases in which aggregates *change dimensionality* are more common. That is, aggregate structures that remain finite in at least one or more spatial directions, but undergo essentially unlimited growth in other directions. The most common examples are amphiphillic assemblies, which can either form spherical micelles (finite in all directions), cylindrical micelles (finite in two spatial dimensions, unlimited in one), or lamellar/layered assemblies (finite in one dimension, unlimited in two). In Sec. III.A.2 below we describe the molecular ingredients that lead to polymorphic transitions between aggregate dimensionality based on a model of surfactant assembly (Bergström, 2016; May and Ben-Shaul, 2001). In this section, we illustrate how the principles of secondary CAC behavior apply to models which can exhibit states of finite aggregation number (e.g. spherical) that can transition to states of 1D aggregation or a bulk (unlimited) morphology.

Since our primary interest is to describe conditions where ideal aggregation gives rise to concentration-dependent transitions in morphology, we consider a minimal description of a generic model including finite and unlimited aggregation states. As summarized in Fig.7A, this model considers three disconnected "branches" of assembly:

1) Finite (compact) aggregates, with respective target size and energy, n_F and ϵ_F

2) 1D aggregates, with energy per subunit

$$\epsilon(n) = \epsilon_{1D} + \Delta_0/n$$

where $\Delta_0 > 0$ characterizes the cost of finite "endcaps" and $\epsilon_{1D} < 0$ is the limiting $n \to \infty$ per subunit assembly energy in this morphology

3) Bulk aggregates, with energy density ϵ_{bulk}. Here, we consider only one macroscopic aggregate ($n \to \infty$) with negligible boundary energy.

Based on the foregoing analysis of the 2-finite assembly state model, it can be anticipated that secondary CAC behavior from finite to 1D aggregation takes places when the $n \to \infty$ energy density of 1D aggregates is lower than finite aggregates, but the energy gap is sufficiently small that the translational entropy associated with the compact aggregates can stabilize a window of n_F-mer aggregates. Likewise, when the energy density of the bulk state falls below these dimensionally limited states, we anticipate an upper limit to concentration (i.e. saturation) which can maintain equilibrium with dispersed aggregates.

With this in mind, we consider a simplified law of mass action for subunit populations

$$\Phi = \phi_1 + \Phi_F + \Phi_{1D} + \Phi_{bulk}, \qquad (30)$$
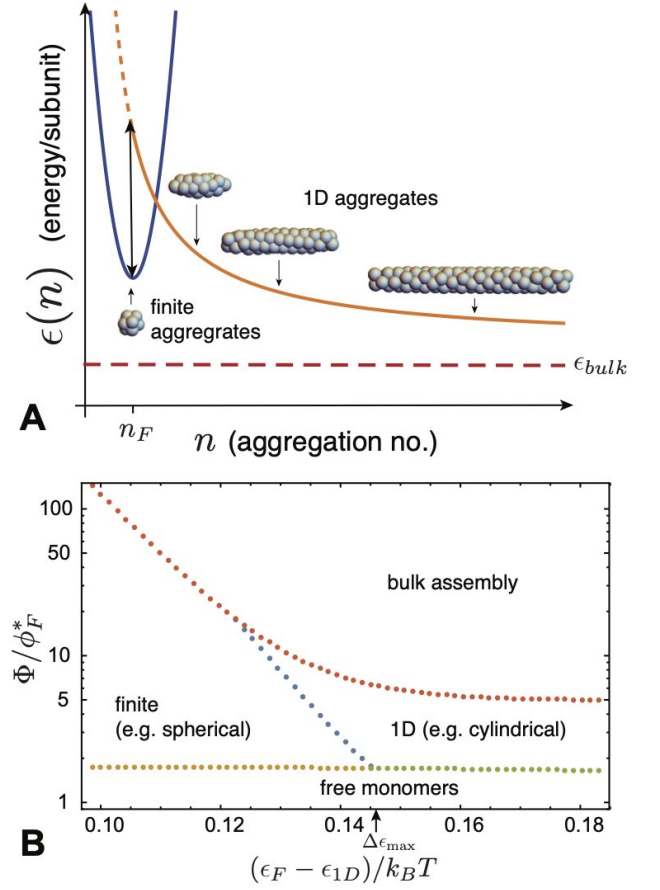

FIG. 7 (A) A schematic plot of polymorphic aggregation energetics with three competing branches of assembly: finite aggregates with a local minimum at n_F, 1D aggregates, and bulk aggregates ($n \to \infty$ energy shown as horizontal dashed line). In this case, the infinite 1D aggregate has a lower per subunit energy than finite aggregates, and there is a barrier (in total energy) δ that separates these states at $n = n_F$, i.e. the double arrow in (A) corresponds to δ/n_F. (B) Phase diagram for concentration-dependent size selection. The dominant aggregation state is shown for a system with coexistence among finite aggregates with $n_F = 100$ subunits, separated by an energy gap $\epsilon_F - \epsilon_{1D}$ and a barrier of δ, eq. (34), to 1D aggregates. There is an additional per subunit energy gap of $\epsilon_{1D} - \epsilon_{bulk} = 0.0005 k_B T$ between 1D and bulk aggregates. The horizontal axis gives the energy gap between spheres and cylinders, and the vertical axis gives the total concentration relative to the CAC for finite aggregates $\phi_F^* \simeq e^{\beta \epsilon_F}/n_F$. The boundaries between monomers, finite and 1D aggregates are determined by crossovers in the most populous aggregate type from eq. (30), while the point of bulk saturation is determined by the point when $\mu = \epsilon_{bulk}$. In this example, the energy per subunit in finite aggregates is fixed at $\epsilon_F = -10\ k_B T$ and the endcap energy of spherocylinders is $\Delta_0 = 20\ k_B T$. The maximum energy gap for which 2nd CAC behavior occurs ($\Delta \epsilon_{max} \approx 0.14\ k_B T$, eq. 35) is indicated on the x-axis.

where the terms respectively describe the populations of free monomers, subunits in a single finite aggregate size, 1D aggregates of various size, and bulk aggregation. The population of subunits in finite aggregates is given by (neglecting number fluctuations)

$$\Phi_{\mathrm{F}} = n_{\mathrm{F}}\big(\phi_1 e^{-\beta \epsilon_{\mathrm{F}}}\big)^{n_{\mathrm{F}}}, \qquad (31)$$

while the 1D aggregate population is given by

$$\Phi_{\mathrm{1D}} \simeq \sum_{n=n_{\mathrm{F}}}^{\infty} n e^{-\beta \Delta_0}\big(\phi_1 e^{-\beta \epsilon_{\mathrm{1D}}}\big)^n$$

$$\simeq n_{\mathrm{F}} e^{-\beta \Delta_0} \frac{\big(\phi_1 e^{-\beta \epsilon_{\mathrm{1D}}}\big)^{n_{\mathrm{F}}}}{\big(1 - \phi_1 e^{-\beta \epsilon_{\mathrm{1D}}}\big)}. \qquad (32)$$

In this final term, we made the additional assumptions that 1D assembly is not favorable below some aggregate size close to $n_{\mathrm{F}} > 1$, and that the monomer concentration remains well below the value $e^{\beta \epsilon_{\mathrm{1D}}}$ where Φ_{1D} diverges[7]. The results are not qualitatively sensitive to these approximations.

Based on these forms, it is straightforward to find the free monomer concentration ϕ_1^{**} where finite aggregates and 1D aggregates are equally populous, i.e. $\Phi_{\mathrm{F}}(\phi_1^{**}) = \Phi_{\mathrm{1D}}(\phi_1^{**})$,

$$\phi_1^{**} = e^{\beta \epsilon_{\mathrm{1D}}}\big(1 - e^{-\beta \delta}\big), \qquad (33)$$

where

$$\delta = \Delta_0 + n_{\mathrm{F}}(\epsilon_{\mathrm{1D}} - \epsilon_{\mathrm{F}}) \qquad (34)$$

is the energy difference, or "barrier", between an n_{F}-mer and a 1D aggregate of the same size (see Fig. 7A). As described above, a stable aggregate population requires at least a local minimum in the energy and hence a barrier necessarily separates aggregation states associated with distinct local maxima in population. Critically, the size of this barrier determines the window of secondary CAC transition behavior, as follows.

First, note that $\delta > 0$ implies that the energy of forming two "end caps" on the 1D aggregate exceeds that of the target n_{F}-mer. Second, the existence of a second CAC requires that this concentration exceeds the primary CAC to a n_{F}-mer dominated state, that is, the condition $\phi_1^{**} > \phi_{\mathrm{F}}^* \simeq e^{\beta \epsilon_{\mathrm{F}}}/n_{\mathrm{F}}^{1/n_{\mathrm{F}}}$. This gives an upper limit to the energy gap between n_{F}-mer aggregation and 1D assembly for second CAC behavior, $\epsilon_{\mathrm{F}} - \epsilon_{\mathrm{1D}} < \Delta \epsilon_{\mathrm{max}}$, with

$$\Delta \epsilon_{\mathrm{max}} \approx \frac{k_{\mathrm{B}}T}{n_{\mathrm{F}}} \ln n_{\mathrm{F}} + k_{\mathrm{B}}T \ln\big(1 - e^{-\delta}\big). \qquad (35)$$

, the stability window of the n_{F}-mer state expands, in terms of $\Delta \epsilon_{\mathrm{max}}$, both with *decreasing* target size and *increasing* energy barrier to 1D aggregation. For energy gaps larger than this limiting condition, the intermediate n_{F}-mer state disappears [8].

Last, note that monomers reach chemical equilibrium with the bulk state at a concentration $\phi_1 = e^{\beta \epsilon_{\mathrm{bulk}}}$, which sets an additional condition for second CAC behavior, $\phi_1^{**} < e^{\beta \epsilon_{\mathrm{bulk}}}$. This gives the upper limit to the energy gap between 1D and bulk assembly for second CAC behavior before saturation,

$$\epsilon_{\mathrm{1D}} - \epsilon_{\mathrm{bulk}} < -k_{\mathrm{B}}T \ln\big(1 - e^{-\beta \delta}\big) \qquad (36)$$

This condition shows that concentration range of the 1D aggregate state diminishes with increasing energy barrier between compact and 1D aggregates.

An example assembly state diagram for this model under conditions $\epsilon_{\mathrm{F}} > \epsilon_{\mathrm{1D}} > \epsilon_{\mathrm{bulk}}$ is shown in Fig. 7B. The concentration dependent state of aggregation is plotted versus energy gap between n_{F}-mers and 1D aggregates for variable ϵ_{F} with fixed n_{F}, Δ, ϵ_{1D}, and ϵ_{bulk}. Under these conditions, when the aggregation energy of (compact) n_{F}-mers is larger than, but sufficiently close to, that of (infinite) 1D aggregation, the system undergoes a sequence of concentration-driven transitions: first from monomers to finite-aggregations; then to 1D aggregates; and finally, to bulk (unlimited) assembly. In the following section, we revisit the possibility of multi-CAC behavior in the context of polymorphism of surfactant aggregates, focusing on its microscopic origin for this effect in terms of an underlying molecular model of aggregation.

III. MECHANISMS AND MODELS OF SELF-LIMITING ASSEMBLY

In Sec. II, we overviewed the basic thermodynamics of ideal aggregation, and described some of the generic ingredients and outcomes of finite-size equilibria. We showed that the key ingredient is a per subunit *aggregation free energy* which has one or more local minima as a function of aggregation number (for *finite-number* aggregates), or as a function of the size of one or more spatial dimensions of the aggregate (for *spatially finite* aggregates like quasi-cylindrical or planar structures). In this section, we review four broad mechanisms of self-limiting assembly, and a physical example of each mechanism. In each case, we first focus on the physical ingredients that give rise to the self-limiting aggregation energetics, and then illustrate some of the implications of the generic phenomonology overviewed in Sec. II.

---

[7] Specifically, we assume $n_{\mathrm{F}} > e^{\beta \delta}$. For $n_{\mathrm{F}} < e^{\beta \delta}$, $\Phi_{\mathrm{1D}} \simeq e^{-\beta \Delta}\big(\phi_1 e^{-\beta \epsilon_{\mathrm{1D}}}\big)^{n_{\mathrm{F}}+1}\big(1 - \phi_1 e^{-\beta \epsilon_{\mathrm{1D}}}\big)^{-2}$, $\phi_1^{**} \approx e^{\beta \epsilon_{\mathrm{1D}}}\big(1 - e^{-\beta \delta/2}/n_{\mathrm{F}}^{1/2}\big)$.

---

[8] Note that eq. (35) is an implicit relation for $\Delta \epsilon_{\mathrm{max}}$, since δ is a function of $(\epsilon_{\mathrm{1D}} - \epsilon_{\mathrm{F}})$. However, in the limit $\delta \gg 1$, the maximum gap is approximately $\Delta \epsilon_{\mathrm{max}} \approx (\epsilon_{\mathrm{1D}} + k_{\mathrm{B}}T \ln n_{\mathrm{F}})/n_{\mathrm{F}}$. Similarly, in the limit $n_{\mathrm{F}} < e^{\beta \delta}$, $\Delta \epsilon_{\mathrm{max}} \approx (\Delta_0 - 2k_{\mathrm{B}}T)/n_{\mathrm{F}}$ (see previous footnote7).

To organize the discussion, we divide these mechanisms into two broadly-delineated classes: *self-closing assembly* and *open-boundary assembly*. These classes are distinguished by the presence or absence of an open boundary and gradients in intra-aggregate stress in the target assembly. Self-closing assembly describes aggregates in which the subunits, by and large, share the same "cohesive environment" of neighboring subunits, and further adopt a common shape in the target assembly. In contrast, self-limitation of open-boundary assemblies requires *gradients* of the inter-subunit forces throughout the aggregate. Within the discussion, we highlight the distinct outcomes and potential "tradeoffs" between these different mechanisms in terms of size selection.

We divide the remainder of Sec. III into two main parts, focusing respectively on self-closing and open-boundary assemblies. Following the introduction of each broader class of SLA, we further subdivide each into two subclasses, representing four basic physical mechanisms of SLA. For each of the four basic mechanisms, we briefly overview the applicable physical systems and then introduce a simple example model that captures the emergence of self-limited assembly.

A. Self-closing assembly

We define *self-closing assembly* (SCA) as class of self-limiting assembly that achieves a finite target size, or finite target dimension, due to *anisotropic binding* between neighbors that leads to a preferred rotation of neighbor bonds. Such interactions generically arise when subunits are tapered or wedged-shaped, such that cohesive bonding leads to a relative rotation of the axes of neighbor units (see Figs. 2A and 9A). In combination with the relative displacement of subunit centers, this relative rotation, when built up over multiple subunits leads to a preferred intra-assembly curvature (along one or more principal directions). In the simplest case, this can be visualized as 1D "loops" of subunits, whose preferred curvature radius R_{close} leads the structure to close upon itself.

We include in the SCA class structures that close upon themselves in all directions of assembly and thus achieve a *finite number* of subunits, such as the spherical shell in Fig. 8A, as well as structures that close in one or more directions but remain unlimited in others, such as the tubule in Fig. 8B. In the latter case, structures have an unlimited number of subunits but achieve a *finite size* in the self-closing direction(s). For example, the tubule is unlimited in the axial direction, but has a well-defined radius and corresponding number of subunits in the circumferential direction. Notably, the underlying principles for size selection remain the same as for finite number; namely, the self-limiting size W of a self-closing direction is determined by the minimum of the energy per subunit with respect W. This can be readily understood by considering a system with nearly
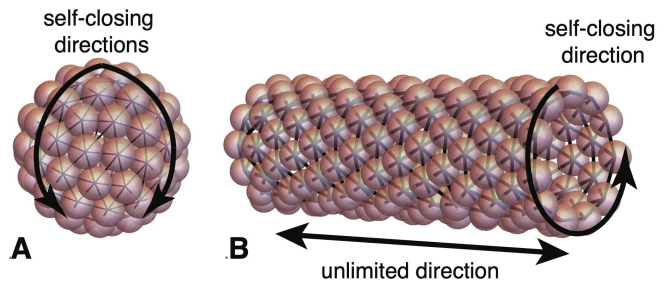

FIG. 8 (A) Schematic of spherical (shell) assembly with two independent self-closing directions of assembly. (B) Schematic of tubule assembly with one self-closing direction (circumferential) and one unlimited direction (axial) of assembly.

all of its N subunits assembled in aggregates of unlimited number, and correspondingly a negligible fraction of free monomers. To first approximation, the energetic costs of free edges and the translational entropy of the unlimited aggregates can be neglected, and the concentration of free monomers can be assumed a small contribution to the total free energy, $F \approx N\epsilon(W)$. Hence, the thermodynamic equilibrium at fixed N corresponds to the selection of the energy-minimizing size W_*, corresponding to $\partial_W \epsilon|_{W_*} = 0$. We describe similar considerations for self-limiting, open-boundary assemblies in Sec. III.B.1 below.

Strictly speaking, for SCA it need not be necessary to identify a continuously loop of bonds along the self-closing direction(s), nor that target curvatures are strictly uniform. We only require that there are one or more periodic directions on a representative 2D surface of the aggregate (e.g., the surface spanned by the subunit centers). Moreover, the preferred curvature does not need to select a perfectly commensurate number of subunits per cycle, since physical subunits generically possess some flexibility of shape and cohesion (bonding) that permits fluctuation in the inter-unit rotation. In the simplest case, e.g. with fluid-like intra-assembly order, a SCA may accommodate such strains through uniform deformation of subunits and their bonds. However, certain physical examples introduce extra geometric constraints (e.g. solid-like, spherical shells) which require at least some variable intra-assembly strains. Notwithstanding the possibility of such gradients, we categorize assembly as self-closing if its target size is selected through the curvature radius R_{close}, as opposed to the accumulation of stress-gradients, which is described as a distinct class of self-limiting assembly below.

Physical examples of SCA can be divided, roughly, into two groups according to the ratio of subunit size, characterized by some thickness d, and the target curvature radius R_{close}. When $R_{\mathrm{close}}/d \gg 1$ the target number of units per cycle is proportionately large. This case describes *tubules, shells and capsules*. The second case, $R_{\mathrm{close}} \approx d$ describes assemblies whose curvature (and thus self-limited size) is most often selected and regulated by the molecular dimension itself, which is characteristic

of *amphiphiles* and their micellar aggregates.

1. Shells, capsules, and tubules

We first review the case of tubule or shell-like assemblies. Examples of these are common in biology, where "tapered" protein subunits select a preferred radius of assembly curvature (Oosawa and Asakura, 1975). Quasi-cylindrical (tubular) examples include microtubules (Cheng *et al.*, 2012; Nogales, 2000) and the bacterial flagella (Namba and Vonderviszt, 1997), while quasi-spherical (shell and capsule) examples include clathrin cages (Bucher *et al.*, 2018; Giani *et al.*, 2017; Kirchhausen *et al.*, 2014; Mettlen *et al.*, 2018), viral capsids (Bruinsma and Klug, 2015; Hagan, 2014; Hagan and Zandi, 2016; Mateu, 2013; Perlmutter and Hagan, 2015; Twarock *et al.*, 2018; Zandi *et al.*, 2020; Zlotnick and Mukhopadhyay, 2011), bacterial microcompartments (Bobik *et al.*, 2015; Chowdhury *et al.*, 2014; Iancu *et al.*, 2007; Kerfeld *et al.*, 2010; Kerfeld and Melnicki, 2016; Rae *et al.*, 2013; Schmid *et al.*, 2006), and other protein-shell organelles (Nott *et al.*, 2015; Pfeifer, 2012; Sutter *et al.*, 2008; Zaslavsky *et al.*, 2018). Recently, engineered tapered or patchy colloids have also drawn interest for their ability to realize synthetic analogs of self-closing tubules and shells (Li *et al.*, 2011; Morphew and Chakrabarti, 2017), although many if not most such realizations to date are properly categorized as analogs to micelles where curvatures are comparable to colloidal dimension. Whatever the underlying subunit structure, the existence of 1D curvature only ensures equilibrium self-limitation along one of the two assembly dimensions in the tubular constructs, while the preferred positive Gaussian curvature of shells and capsules leads to equilibrium states with finite subunit *number*.

To illustrate the self-limiting thermodynamics of shells and capsules, where target curvature radii are much larger subunit dimensions, consider the following minimal model. Spherical fluid capsules, shown schematically in Fig. 9A, are composed of subunits with nominal area a_0 and a tapered shape that favors a preferred (target) spherical curvature radius R_T. Here, we restrict the analysis to cases in which the curvature preference sufficiently disfavors locally anisotropic curvature (Lázaro *et al.*, 2018a) to limit incomplete assembly to cap-like aggregation states with positive Gaussian curvature. Competition between incomplete assemblies with positive and zero Gaussian curvature was recently considered in Ref. (Mendoza and Reguera, 2020). We assume the cap covers an axisymmetric "cap" domain of radius R from its pole up to the *aperture angle* Θ (see Fig. 9A). While a closed capsule with a preferred curvature R_T has a target aggregation number $n_T = 4\pi R_T^2 / a_0$, a capsule may realize a different aggregation number $n = 2\pi R^2 (1 - \cos\Theta)$, provided that either it is open (i.e. $\Theta \neq \pi$) or is deformed from its preferred taper (i.e. $R \neq R_T$).

Taking the simplest possible model, we assume that the intra-shell order is fluid-like, such that the only elastic penalty derives from bending deformations away from target curvature, which we consider via a membrane-like bending energy

$$E_{\text{bend}} = \frac{B}{2} \int dA \Big(\frac{1}{R} - \frac{1}{R_T}\Big)^2, \tag{37}$$

where B is a bending modulus and the area integration is carried out over the incomplete shell. Additionally, we consider the line energy associated with open edge for an incomplete cap,

$$E_{\text{open}} = 2\pi R \sin\Theta \, \lambda, \tag{38}$$

where λ is the energy per unit length of the exposed edge, associated with the fewer cohesive bonds as well as a difference in solvation of subunits at the edge. The aggregation energy as a function of n then takes the form

$$\epsilon(n, \Theta) = -\epsilon_T + \frac{2\pi B}{n_T} \Big(\sqrt{\frac{n_T(1 - \cos\Theta)}{2n}} - 1\Big)^2 + \frac{\sqrt{2\pi a_0}\lambda}{n^{1/2}} \frac{\sin\Theta}{\sqrt{1 - \cos\Theta}} \tag{39}$$

where $-\epsilon_T$ is the "bare" aggregation energy for subunits in the bulk of undeformed capsules.

It is easy to see that the form of $\epsilon(n, \Theta)$ in eq. (39) has a global minimum for closed capsules of the target size (i.e. $n = n_T$ and $\Theta = \pi$). Nevertheless, the combination of bending elasticity and the edge energy of incomplete shells influences assembly for $n \neq n_T$. This can be seen by plotting the landscape of assembly energetics in the n/n_T - Θ plane, as shown in Fig. 9B. For very small aggregate sizes $n/n_T \ll 1$, caps lock into the preferred curvature, $R \to R_T$, described by the condition $\cos\Theta \to 1 - 2n/n_T$, shown as a dotted line in Fig. 9B. As n increases, the edge energy favors compression of the open caps to smaller curvature radii $R < R_T$. The amount of this shape compression grows up to a critical aggregation number n_S, beyond which the minimal energy capsule "snaps" discontinuously to a closed shell of suboptimal size, that is $\Theta \to \pi$ for $n_S < n < n_0$.

Hence, the minimal aggregation vs. n generically exhibits two branches (Fig. 9C): an open cap for $n < n_S$ and an (edge-free) closed shell for $n_S < n < n_T$. The transition between these two branches can be understood in terms of "nucleation" of an open pore in an over-curved shell, where $n_T - n_S$ corresponds to the size of the "critical nucleus". The inset of Fig.9C shows that the critical preclosure size n_S generically decreases with an increasing (dimensionless) ratio of edge energy and bending stiffness, $\bar{\lambda} \equiv (a_0 n_T / 8\pi)^{1/2} \lambda / B$.

Notwithstanding the generic existence of a transition between open-cap and preclosed branches of the energy landscape, the transition does not lead to stable partial shells (i.e. a minima in $\epsilon(n)$ for $n \neq n_T$). Hence, while the open cap branch and its transition to the closed shell
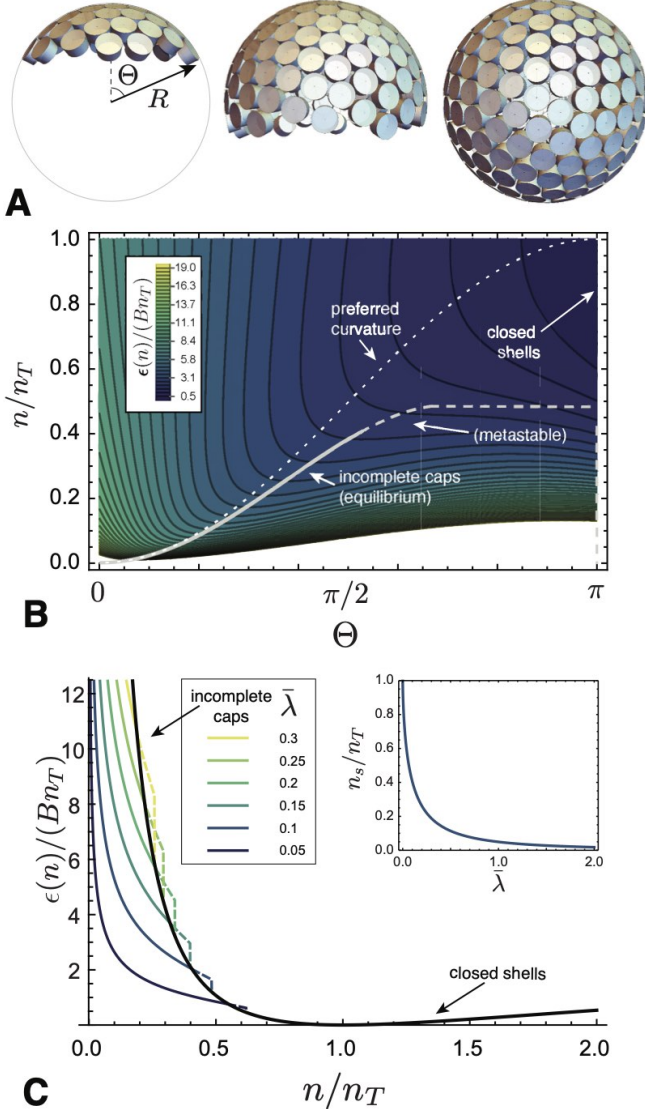
may have implications for assembly pathways and kinetics, the equilibrium distribution of self-limiting capsules is independent of the edge energy and generically governed by the energetics of the closed-shell branch. This fact has further generic consequences for the concentration dependence and dispersity of aggregate size, both of which are governed by the product of the convexity and cube of the target size, $\epsilon''(n_T)n_T^3/k_BT$, according to eqs. (19) and (20). The $\Theta \to \pi$ limit of (39) shows that $\epsilon''(n_T) = B\pi/n_T^3$. Hence, the decrease of convexity with target assembly size precisely cancels that entropic factor of n_T^3, such that the *relative* shift in mean aggregate size and *relative* dispersity, $(n_* - n_T)/n_T$ and $\langle\Delta n^2\rangle^{1/2}/n_T$, respectively, are limited only by the ratio of bending modulus to thermal energy, B/k_BT, and independent of self-closing target size. [9] Therefore, to regulate the self-limiting size of self-closing assemblies in *absolute* terms, the rigidity of subunits and their angular interactions must *grow with target size.*

The predicted growth of size fluctuations with target size would seem to contradict observations of the best studied example of self-closing shells, icosahedral virus capsids. At conditions of optimal assembly, size polydispersity of viruses is remarkably small. In fact, this high degree of monodispersity has been exploited by using 3D crystalline arrays of virus capsids for optical applications requiring precise spatial periodicity (Brillault *et al.*, 2017; Chen *et al.*, 2016; Dang *et al.*, 2011; Delalande *et al.*, 2016; Judd *et al.*, 2014; Malyutin *et al.*, 2015; Minten *et al.*, 2011; Park *et al.*, 2015; Rother *et al.*, 2016; Steinmetz *et al.*, 2011; Young *et al.*, 2008). While high-precision measurements of size-dispersity in capsids are challenging, electron microscopy structures that were obtained without the assumption of icosahedral symmetry show that as many as 40% of alphavirus nucleocapsid core particles exhibit defects (Wang *et al.*, 2018, 2015), and hence some dispersity in shape. Notably some non-icosahedral structures, like immature HIV caspids, exhibit variations in number of subunits (~ 1000 GAG protein subunits) that are on the order of the mean capsid size (Briggs *et al.*, 2009), although such effects may also be attributed to assembly kinetics (Dharmavaram *et al.*, 2019). More recently, size distributions of hepatitis B virus (HBV) capsids (and capsids of other viruses) have been achieved near or at single-subunit precision using resistive pulse sensing (Zhou *et al.*, 2018, 2011), mass spectrometry (Uetrecht *et al.*, 2011), and charge detection mass spectrometry (Lutomski *et al.*, 2018; Pierson *et al.*, 2016, 2014). Although metastable defective capsids are observed (Lutomski *et al.*, 2018; Pierson *et al.*, 2016), these measurements show that at long times (po-

FIG. 9 (A) Schematic geometry of a "fluid capsule" of tapered subunits, which assumes partial shell geometries of spherical caps with aperture angle $0 < \Theta \leq \pi$. (B) A contour map of the energy density landscape of the capsule model as a function of aperture angle and the ratio of subunits n to the preferred number in the ideal closed shell n_T. Low (high) values of $\epsilon(n)$ appear as purple (red). The dotted line shows a partial shell with the target curvature, while the solid line indicates the minimal energy cap, whose curvature radius is slightly compressed by the line tension of the boundary. Beyond a threshold cap size $n = n_S \simeq 0.4n_T$, this open cap becomes unstable to preclosure, and the minimal energy branch runs along $\Theta \to \pi$. This landscape corresponds to a dimensionless line tension $\bar{\lambda} = 0.1$. (C) Plots of the minimal energy branches of assembly: incomplete caps are shown as colored curves (metastable portions are dashed), and the closed shell ($\Theta \to \pi$) is shown as a black curve. The inset shows the size n_S corresponding to the preclosure, or "snap", transition between stable open caps and closed shells as a function of line tension.

[9] This continues until the asymptotic limit of zero spontaneous curvature ($n_T \to \infty$), at which point the free energy per subunit is independent of size and the size distribution becomes an unlimited exponential (Helfrich, W., 1986) as shown for 1D assemblies in section II.A.2.
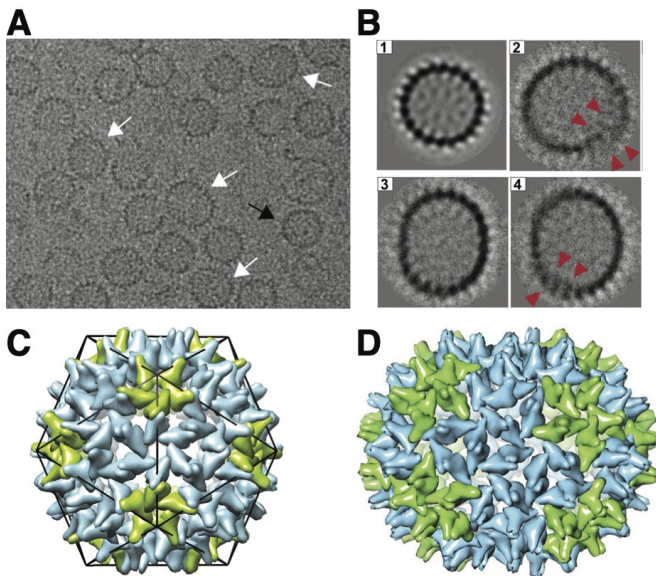
FIG. 10 Examples of well-formed and defective woodchuck hepatitis B virus (WHV) capsids, adapted from Ref. (Pierson *et al.*, 2016). (A) Portion of a representative cryo-EM image of empty WHV capsids assembled in vitro in the absence of RNA (from capsid proteins with the C-terminal RNA binding deleted, wCp149). The population of capsids is structurally heterogeneous; the black arrow indicates an example of a $T = 4$ icosahedral capsid, while the white arrows indicate examples of defective capsids. (B) 2D class averages of (1) icosahedral and (2-4) defective particles. Red arrows in (2) and (4) indicate locations where the capsid shell has overgrown and overlaps itself. (C) Schematic model of a $T = 4$ icosahedral HBV capsid, with the monomers forming the 12 five-fold vertices colored green, and others colored blue. (D) Hypothetical model of the structure of a capsid that is non-icosahedral but closed, with an elongated structure containing 150 protein dimers (the icosahedral capsid has 120 dimers). Here the green dimers are in pentamers or extend between pentamers and hexamers, and the blue dimers are in hexamers or extend between two hexamers. Images reprinted from (Pierson *et al.*, 2016).

tentially corresponding to a near equilibrium state) the population is dominated by icosahedral capsids with the native size [10].

To place these measurements in the context of the results for the fluid shell model, we note that bending moduli for virus capsids have been estimated from the force-displacement curves measured in nanoindentation experiments in which virus capsids are compressed using an AFM tip [11]. Estimated bending modulus values vary from 10-200 $k_\mathrm{B}T$, but typical values fall in the higher

end of the range, 100-200 $k_\mathrm{B}T$. (For comparison, bending moduli of lipid bilayer membranes are typically in the range of $10 - 20$ $k_\mathrm{B}T$.) Using the result from above that $\langle \Delta n^2 \rangle^{1/2} \approx n_0/\sqrt{\pi B}$, with $n_0 = 120$ and an estimate of $B \sim 60 k_\mathrm{B}T$ for HBV[12] gives a root mean squared size fluctuation of about 9 dimers, considerably larger than the long-time experimental estimates.

This discrepancy highlights two physical ingredients neglected in the model described above: the discrete subunit size and the fact that most virus capsids, as well as most other protein shells, are crystalline rather than fluid. As discussed below in section V.B, the arrangement of proteins within icosahedral capsids can be mapped onto a triangular net. However, tiling a spherical topology with a triangular lattice requires the formation of 12 five-fold sites, often consider as "defects" in a hexagonal packing. A number of equilibrium calculations have shown that the elastic energy of the defects themselves and inter-defect elastic interactions significantly affect the energy landscape of such shells (Bruinsma *et al.*, 2003; Chen *et al.*, 2007; Li *et al.*, 2018, 2019; Mendoza and Reguera, 2020; Zandi *et al.*, 2004). These effects are reflected in local minima in the aggregate energy at certain 'magic numbers' of subunits (Zandi *et al.*, 2004). Notably, these minima correspond to shells with high degrees of symmetry, with the in-plane bonding energies corresponding to shells with icosahedral symmetry. Thus, when including the energetics of this bond-ordering a size fluctuation of even one subunit can incur a significant energy cost ($\gtrsim 10 k_\mathrm{B}T$), since it requires disruption of the low-energy geometry, for example, through the introduction of a pair of 5- and 7-fold defects. In fact, the metastable structures observed in HBV capsids are typically found at discrete intervals corresponding to deviations of multiple subunits from the native capsid size (Lutomski *et al.*, 2018; Pierson *et al.*, 2016, 2014), suggesting that typical fluctuations correspond to insertion or deletion of multi-subunit oligomers (e.g. hexamers of the capsid protein) which would minimize disruptions to the capsid symmetry. For example, Fig. 10A,B show cryo-EM images and examples of 2D class averages respectively of in vitro assembly products of woodchuck hepatitis B virus (WHV) capsid proteins, which exhibit heterogeneous structures including icosahedral capsids, elongated closed shells, and shells with overlapping edges. Fig. 10C shows a hypothetical model of a non-icosahedral closed shell, in which insertion of additional hexamers

---

[10] In fact HBV is dimorphic. Both in vitro and in vivo HBV capsid assembly yields mostly 120 protein dimer capsids with $T = 4$ icosahedral symmetry in the Caspar Klug nomenclature (see section V.B), but also a few percent of $T = 3$ icosahedral capsids. However, size fluctuations around the dominant $T = 4$ population were shown to be insignificant at long times.

[11] It should be noted that estimating elastic moduli from the force-displacement curve is sensitive to the value chosen for thickness of the capsid shell, and the relationship between the atomic structure and the effective mechanical thickness remains at least some-

what obscure (May *et al.*, 2011; May and Brooks, 2011).

[12] The bending modulus is calculated from the 3D Youngs modulus $E = 0.26$ GPa obtained from nanoindentation measurements in Roos et al.(Roos *et al.*, 2010), and assuming a thin shell model so that the 2D Youngs modulus and bending modulus are respectively given by $Y = Et$ and $B = Yt^3/(12(1 - \nu^2))$, with $t = 2.1$ nm the effective shell thickness (Roos *et al.*, 2010; Wynne *et al.*, 1999) and $\nu = 0.4$ the Poisson's ratio (Roos and Wuite, 2009; Uetrecht *et al.*, 2008).

leads to a prolate structure.

Computational models of icosahedral assembly have also identified ensembles of defective capsules in which additional hexamers were added to the icosahedral shell (Elrad and Hagan, 2010; Nguyen *et al.*, 2009), although these models and the corresponding defective structures differ from the HBV system. Hence, for more realistic models that incorporate both bending and "bond-network" elasticity, we expect that these corrugations in the energy landscape (i.e. due to communicability with icosahedral symmetry) versus n will be superposed on the smooth landscapes illustrated in Fig. 9 for the fluid shell model, which may account for size fluctuations to be restricted for a limited set of low-energy values of n at or near to high-symmetry, or magic number capsomer arrangements.

2. Amphiphillic aggregates

Arguably the most common and well-studied class of self-limiting assemblies is *amphiphiles*. In the broadest sense, these refer to subunits with chemically dissimilar ends, which consequently favor distinct solvent environments. For example, lipids and surfactants possess oily hydrocarbon tails attached to a polar or charged head group (Israelachvili, 2011), which imparts a respective hydrophobic and hydrophillic character to either end of the same molecule (e.g. schematic in Fig. 11A). Dispersing such amphiphiles in a solvent that has higher affinity to one end of the molecule generically drives them to form aggregates that partially hide, or sequester, the solvophobic portions while maintaining exposure of the solvophillic portions. Examples of such aggregrates, spherical and cylindrical micelles, and bilayer sheets, are shown schematically in Fig. 11B-C. Notably these structures curve upon themselves, but do so on a lengthscale that is limited by, and comparable to, the size of the amphiphile itself, e.g. the molecular tail length in Fig. 11A. The tendency to exclude unfavorable solvent from the "core" of the aggregate, in combination with the packing constraints of filling this region with the solvophobic portions, requires each amphiphillic subunit to "span" the entire thickness of the aggregate, which is fundamental to their self-limiting assembly.

In this section we describe a simple model to capture the self-limiting assembling of amphiphiles, and highlight, in particular, how thermodynamic considerations of changes in *aggregate thickness* shape the preferred aggregate curvature, but also give rise to polymorphism in the dimensionality of aggregates (e.g. spheres, cylinders, membranes). For illustration, we review a model for the thermodynamics of surfactant aggregation, of the type shown in Fig. 11A, capturing central ingredients of the well-known packing model developed by Israelachvili and coworkers (Israelachvili, 2011; Israelachvili *et al.*, 1976). While this model aims to capture molecular elements of low-molecular weight surfactants and lipids, the essential
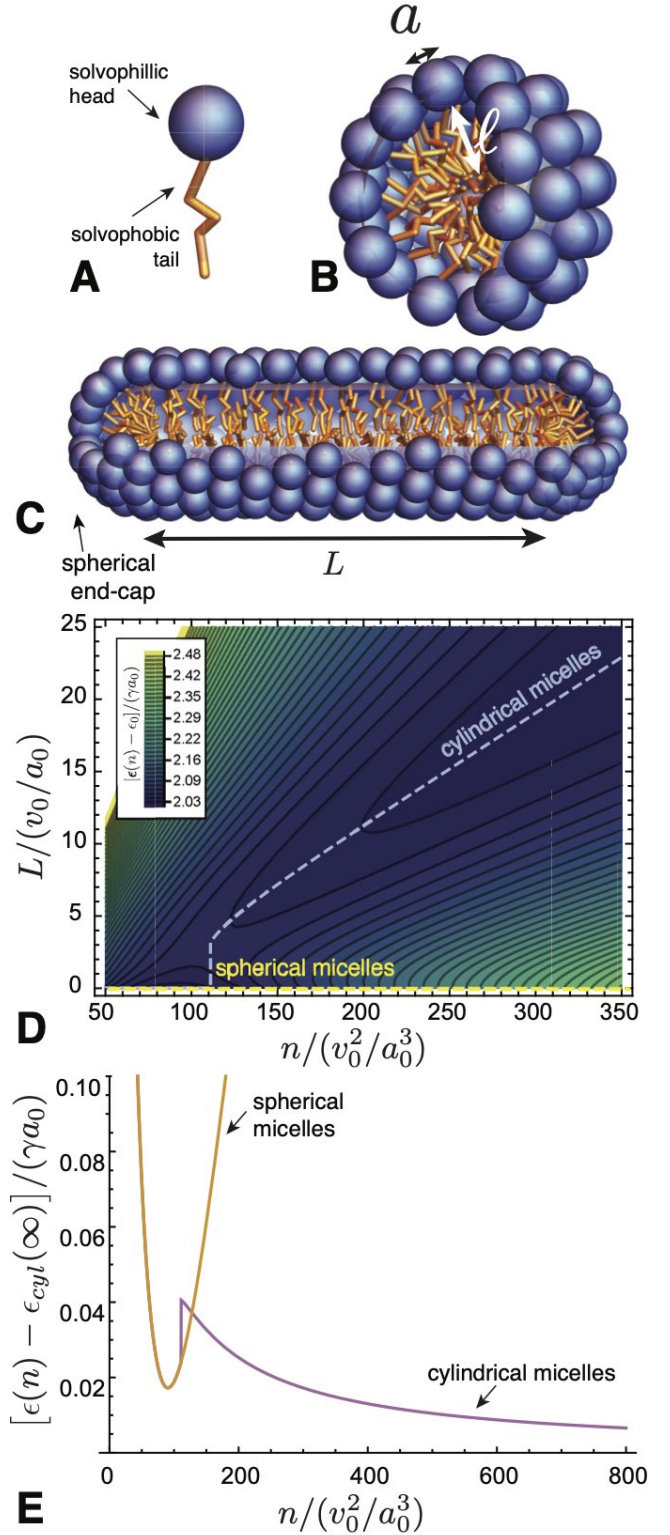

FIG. 11 (A) A cartoon model of a (single-tailed) surfactant molecule. (B) A schematic of a spherical micelle with the thickness of the solvophobic core (ℓ) and the area per head group (a) highlighted. A wedge-like portion of the micelle is cutaway to illustrate the interior packing of tails. (C) A spherocylinder model of a worm-like micelle, which has a cylindrical portion of length L capped by two hemispherical micelles of equal radius. (D) An energy landscape for spherocylinders described by the model in eq. (44) with $P^{-1} = 2.25$ and $\bar{k} = 0.1$. The dashed orange line shows the ($L = 0$) spherical micelle branch and the dashed pink line shows the ($L = 0$) worm-like micelle branch, whose corresponding energy as a function of reduced aggregation number is plotted in (E).

thermodynamic features carry over to other amphiphillic assemblies, such as block copolymers in selective solvents (Halperin *et al.*, 1992; Jain and Bates, 2003; Leibler *et al.*, 1983; Zhang and Eisenberg, 1996). The thermodynamics of amphiphile aggregation incorporates three ingredients: (i) the thermodynamics of area per solvophillic head group; (ii) the thermodynamics of molecular extension; and (iii) the constraint of uniform density in the solvophobic core, which links the first two elements.

A simple model to describe the head-group energetics considers the per subunit energy to form aggregates with an area a at the solvent/core interface (Israelachvili, 2011; Israelachvili *et al.*, 1976). The generic tendency to "hide" amphiphiles from the solvent is parameterized by a surface energy cost γa, where $\gamma > 0$ favors dense lateral packing of head groups. Competing against this lateral compression is the cost of inter-unit repulsive interactions, which in the simplest case are described by the two-body term in the virial series, giving a per subunit energy A_2/a, where $A_2 > 0$ (Tanford, 1974). These two terms can be combined into a single form

$$\epsilon_{\text{int}} = \gamma\Big(a + \frac{a_0^2}{a}\Big), \qquad (40)$$

where $a_0 = \sqrt{A_2/\gamma}$ is the *optimal head group area*.

In combination with the tendency to achieve optimal head group area are additional thermodynamics of tail packing in the core, and the costs to extend its length, ℓ. There are various models proposed for this effect, including a finite-maximum extension (Israelachvili, 2011; Israelachvili *et al.*, 1976) or instead treating the core as melt of flexible polymers (Ben-Shaul *et al.*, 1984; Dill and Flory, 1980; Nagarajan, 2002; Nagarajan and Ruckenstein, 1991). Here, we adopt a simplified model used by May and Ben-Shaul for the free energy of tail length ℓ that spans from the solvent/core interface into the center of the aggregate

$$\epsilon_{\text{stretch}} = \frac{k}{2}\big(\ell - \ell_0\big)^2, \qquad (41)$$

where k is an elastic constant for intra-subunit stretch and ℓ_0 is a preferred length, which parameterizes the free energy cost of deformations from a preferred conformation state of the short tail. Here, we consider k and ℓ_0 as a minimal description of the extensional thermodynamics, and like γ and a_0, these parameters can be varied through a combination of subunit structure and physical-chemical conditions, e.g. temperature and solvent properties.

Extensional energetics are linked by packing constraints associated with occupying the core with a fixed density of solvophobic portions of the subunits (Israelachvili, 2011; Israelachvili *et al.*, 1976). These constraints vary with the dimensionality of the limited directions in the aggregate: $d_{\text{L}} = 3$, spherical micelles; $d_{\text{L}} = 2$, cylindrical micelles; and $d_{\text{L}} = 1$ planar bilay-

ers [13]. By considering the ratio between the core volume and interfacial area of an aggregate of thickness (radius) ℓ, it is straightforward to show that the (solvophobic) volume per subunit satisfies

$$v_0 = \frac{a\ell}{d_{\text{L}}}. \qquad (42)$$

As aggregates change their shape and number, uniform density requires adjustment of a and ℓ to maintain constant v_0. Using this constraint, we rewrite the assembly energy in terms of a single *dimensionless thickness*,

$$r \equiv \frac{\ell}{(v_0/a_0)} = d_{\text{L}}\frac{a_0}{a}, \qquad (43)$$

giving

$$\frac{\epsilon(r, d_{\text{L}})}{\gamma a_0} = \Big(\frac{d_{\text{L}}}{r} + \frac{r}{d_{\text{L}}}\Big) + \frac{\bar{k}}{2}\big(r - P^{-1}\big)^2 + \epsilon_0, \qquad (44)$$

where $\bar{k} = k v_0^2/(\gamma a_0^3)$ is a scaled stretch modulus of the tail, and ϵ_0 parameterizes the negative energetic gain to assemble. The parameter P was introduced by Isrealachvili as the *packing parameter*, a measure of the commensurability of the preferred shape with accessible aggregate geometries,

$$P \equiv \frac{a_0 \ell_0}{v_0}. \qquad (45)$$

Written in this way, it is straightforward to understand how area and stretch thermodynamics compete to determine the optimal aggregate morphology. While area terms favor a thickness $r = d_{\text{L}}$, stretch thermodynamics favor a thickness $r = 1/P$. Only for particular preferred headgroup areas and subunit lengths do these two values coincide, i.e. when $P = 1/d_{\text{L}}$; otherwise there is at least some shape frustration between these terms. As a heuristic, we therefore expect aggregation to favor the dimensionality d_{L} closest to $1/P$, which corresponds to the "tapered" geometry that most closely approximates the favored areal and thickness packing at uniform density.

A more complete picture of the polymorphism of aggregates is given by considering assembly landscapes that allow for transitions of micellar dimension. Fig. 11C shows the structure of a "wormlike" micelle, modeled as a spherocylinder composed of a length L of cylindrical micelle capped by two equal-radius hemispherical micelle caps. As L increases, the fraction of the aggregate in the $d_{\text{L}} = 2$ (vs. $d_{\text{L}} = 3$) packing increases, and thus consideration of the energy as a function of both n and L illustrates the landscape of aggregates intermediate to a uniformly cylindrical or spherical geometry. An example

---

[13] Note that d_{L} refers to the number of *limited* directions, while we use d to refer to the dimensionality of the *unlimited* directions, e.g. $d = 1$ and $d = 2$ for cylindrical and lamellar aggregates.
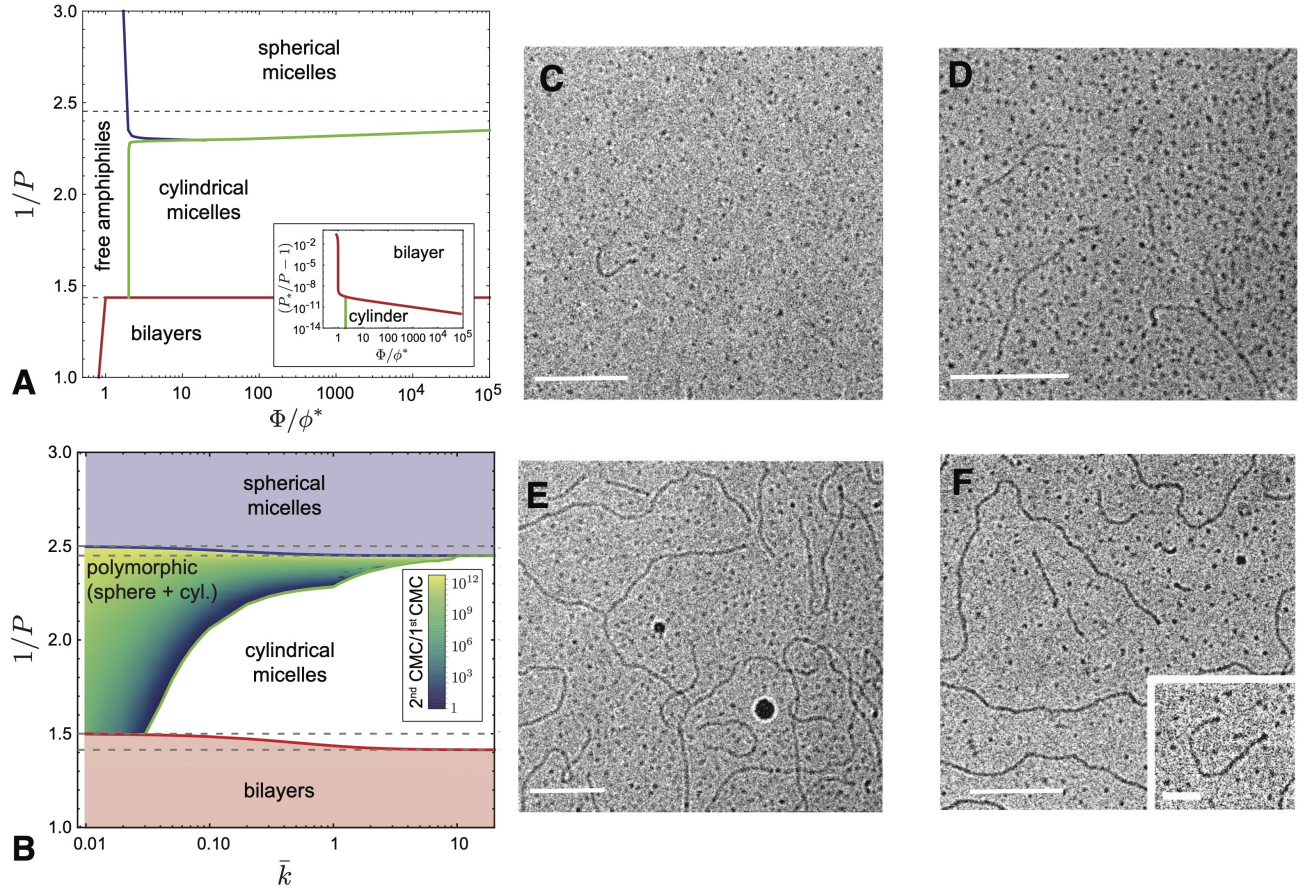
FIG. 12 (A) An assembly state phase diagram for the amphiphillic aggregate model in eq. (44) for dimensionless stiffness $\bar{k} = 1$ and $v_0^2/a_0^3 = 10$, where ϕ_* is the nominal CMC for cylinders. Solid lines mark the boundaries between the most populous aggregate type, and dashed lines indicate the boundaries in the infinite Φ limit. The inset shows a zoom in near the boundary between cylinders and bilayers, illustrating an extremely narrow window of secondary CMC behavior due to the large mean (finite) size of cylinders. (B) Summarizes the polymorphic assembly of the amphiphile model in the plane of stiffness $\bar{k}$ and inverse packing parameter P^{-1}, for $v_0^2/a_0^3 = 10$. Regimes of single CMC behavior are shown as solid red, white, and blue for bilayers, cylinders, and spheres, respectively. The regime of polymorphic concentration-driven sphere-to-spherocylinder transitions is colored on a blue-green scale according to the ratio of second CMC (spheres to spherocylinders) to the first CMC (monomers to spheres). (C-F) show cryo-transmission electron micrographs of micelles formed by dimeric (gemini) surfactancts, at 25° C at increasing weight percent: (C) 0.26%; (D) 0.5%; (E) 0.62%; and (F) 0.74% (scale bars equal 100 nm). The coexistence of cylindrical micelles of increasing total length for (D-F), and absence of lengths intermediate to spheres and the shortest cylinders is consistent with a concentration-dependent second-CMC. The inset of (F) shows the bulbous ends of a cylindrical micelle (scale bar, 25 nm), consistent with an energy barrier between spherical and (long) cylinders due to a mismatch in preferred radius. (C-F) are adapted from (Bernheim-Groswasser et al., 2000).

landscape is shown in Fig. 11D, for a packing parameter $P^{-1} = 2.25$ intermediate to spheres and cylinders, which exhibits two branches of local minima. In the spherical branch, $L = 0$ and changes in number are accommodated purely through changes in micelle radius. The second, cylindrical, branch appears only above a threshold aggregate number ($n \simeq 111 v_0^2/a_0^3$), beyond which further subunit addition is accommodated through increasing length. The per subunit aggregation energies for these two branches are shown in Fig. 11E. Notably, an energy barrier separates the convex minimum of the spherical branch from the cylindrical branch, which asymptotically approaches the global minimum of ϵ at $n \to \infty$ via the $1/n$ falloff characteristic of 1D assembly. The origin of this "barrier" between spherical and cylindrical

aggregates derives from the fact that the preferred radii of these two micelle types are different, and hence the confinement energy of the "endcaps" on the spherocylindrical micelle exceeds that of a minimal-energy (larger radius) sphere. We note that this spherocylinder geometry will overestimate the cost of endcaps relative to a more realistic model that includes, say, variation of the radial thickness along the micelle (May and Ben-Shaul, 2001).

As described in Sec. II.B.3, the existence of multiple low-energy branches of aggregation leads to a rich phase behavior, which can be analyzed according to the law of mass action eq. (5). As summarized in Appendix A, we analyze the polymorphic assembly of the model described by eq. (44), making several simplifying assump-

tions. Specifically, we adopt the continuum limit for n, calculate the optimal spherocylinder radius for each length L by minimizing the per-molecule energy eq. (44) as a function of r (i.e. neglecting radius fluctuations), and numerically calculate the mass fractions of spheres, spherocylinders, and membranes as functions of k, P, Φ from eq. (5).

Fig. 12A shows the phase diagram as a function of the packing parameter and total concentration Φ for a dimensionless stretching stiffness $\bar{k} = 1$. Analogous to the phase diagram of the generic model in Fig. 7B, with increasing packing parameter the system undergoes dimensional transitions from spheres to spherocylindrical assemblies and finally to bulk lamellar aggregates (corresponding to packing dimensionalities of $d_L = 3, 2, 1$). The infinite concentration limits of the phase boundaries (indicated by dashed lines in Fig. 12A) correspond to the packing parameter values where the bulk free energy per subunit of two aggregate geometries are equal: $\epsilon_3 = \epsilon_2$ at the sphere/spherocylinder boundary, and $\epsilon_2 = \epsilon_1$ at the spherocylinder/lamella boundary, with ϵ_{d_L} the optimal energy per subunit in the aggregate interior for aggregates with dimensionality d_L. Notably, for values of the packing parameter near each of these dimensional transitions, there is a concentration-dependent dimensional transition. I.e., for packing parameter values close to the infinite-concentration value for the spherocylinder-sphere transition $2.45 \gtrsim P^{-1} \gtrsim 2.3$, spheres are favored at low concentrations, with a transition to spherocylinders occurring above a threshold concentration that diverges exponentially as the packing parameter approaches the transition value, i.e. $P^{-1} \rightarrow 2.45$. An analogous behavior occurs at the spherocylinder/lamella transition (Fig. 12 inset), but the range of P^{-1} values is exceptionally narrow due to the very large (yet finite) mean size of spherocylinder aggregates (see below).

In Fig. 12B, we show an overview of the polymorphic assembly in terms of the two control parameters of the amphiphile aggregation model, P^{-1} and k. Boundaries in the $P^{-1} - \bar{k}$ space are shown for infinite-concentration sphere/spherocylinder and spherocylinder/lamella transitions, indicated by blue and red lines respectively, with gray dashed lines corresponding to the values of these transitions in the extensionally "floppy" ($\bar{k} \rightarrow 0$) and "stiff" ($\bar{k} \rightarrow \infty$) limits [14]. We also show the region in $P^{-1} - \bar{k}$ space for which there is a concentration-driven sphere/spherocylinder transition (a secondary CMC), with color indicating the width in concentration space of the transition. More specifically, the color scale indicates the ratio of pseudo-critical concentrations, $\Phi_{\text{cyl}}/\Phi_{\text{sph}}$ where spherocylinders or spheres, respectively, become the most populous subunit state (i.e. greater than 50%)

at a given value of stiffness and packing parameter.

The emergence of the polymorphic, concentration-driven transition between spherical and spherocylindrical micelles captured in Fig. 12 can be understood in terms of the three ingredients of the secondary CAC behavior encoded in eq. (35): the energy gap between minimal-energy spheres and infinite cylinders $\Delta\epsilon = \epsilon_3 - \epsilon_2$, the finite aggregation number in spherical micelles n_{sph}, and an energy barrier δ separating the spherical aggregates from spherocylinders (as in Fig. 11D). Below the blue curve in Fig. 12B, where $\Delta\epsilon > 0$, the physical origin of the intermediate concentration state of spheres is the higher (per subunit) entropy associated with their fewer subunits. This window of second CMC behavior widens in Φ as the gap between infinite cylinders and spheres vanishes to zero, which happens as the inverse packing parameter increases and approaches the blue curve in Fig. 12B. Likewise, from eq. (44) it is straightforward to see that the gap between $d_L = 3$ and $d_L = 2$ vanishes as $\bar{k} \rightarrow 0$, since the is no obstacle to achieving the optimal head group packing ($a \rightarrow a_0$) for any d_L in the absence of extensional stiffness. Hence, the ratio of the second CMC (spheres to spherocylinders) relative to the first (monomers to spheres) grows large in both of these regimes.

The regime of second CMC behavior is restricted to lower values of extensional stiffness, and disappears above a threshold value of $\bar{k}$, due to its effect on the energy barrier between spheres and spherocylinder micelles. In the limit of $k \rightarrow 0$, the thickness of spheres and cylinders is determined purely by head group packing, and hence $d_L = 3$ and $d_L = 2$ micelles have different radii, implying a finite frustration cost for the hemispherical endcaps of the spherocylindrical micelles. As summarized in eq. (35), the window of second CMC behavior is widened with increasing energy barrier between compact and 1D assemblies. With increasing $\bar{k}$, this barrier diminishes, ultimately vanishing the in the $\bar{k} \rightarrow \infty$ limit, because the high stiffness requires the micelle thickness to maintain $\ell = \ell_0$ independent of dimensionality. Hence, as is the case for the "ladder model" of cylindrical micelle thermodynamics (Missel *et al.*, 1980), in the absence of an energy barrier between spheres and elongated cylinders, there is only a single CMC to a state where mean aggregation number continuously increases with Φ.

Similar arguments apply to the spherocylinder/lamella transition, except that spherocylindrical aggregates have a very small translational entropy due to their large mean size, and thus are stabilized by entropy over a vanishingly narrow region of parameter space.

Evidence for the secondary CMC transition between spherical and cylindrical micelles has been reported for range of surfactant systems; see (Bergström, 2016) for a review. Many experiments report an indirect signature of an inflection point, and secondary upturn, in the mean aggregation number as a function of concentration. For example, the convex dependence of viscosity on concentration for certain ionic surfactants was inter-

---

[14] The transition values between d_L and $d_1 + 1$ are $P^{-1} = (d(d-1))^{1/2}$ and $P^{-1} = d_L - 1/2$ in the respectively stiff ($\bar{k} \rightarrow \infty$) and floppy ($\bar{k} \rightarrow 0$) limits.

preted by Porte in coworkers as a secondary CMC, and motivated their model with an energetic gap between spherical and cylindrical states for second CMC behavior (Porte *et al.*, 1984). Elsewhere, experimental imaging of the concentration-dependence micelle morphologies has been used to probe second CMC behavior. For example, Figs. 12 C-F show cryo-electron microscopy (cryo-EM) images of micelles formed by dimeric (gemini) surfactants (Bernheim-Groswasser *et al.*, 2000). Above the first CMC, but still below a second threshold concentration, cryo-EM shows monodisperse spherical micelles Fig. 12. With increased concentration (Fig. 12D-F), cryo-EM shows the appearance of cylindrical, or worm-like, micelles. While these grow in length with concentration, they also retain coexistence with a population of spherical micelles. This, along with an observed gap in micelle sizes intermediate to spheres and the shortest cylinders, is consistent with the secondary CMC transition induced by an energetic gap between spheres and cylinders. As described above, and consistent with the apparently "bulbous" ends of worm-like micelles (Fig. 12F), such an energetic gap is a natural consequence of the mismatch between radii of spherical and cylindrical micelle packings.

To conclude the overview of amphiphillic aggregation, we briefly return to the question of *convexity* of the energetics described by (40) - (44) for spherical micelles ($d_L = 3$). It is straightforward to consider two simple limits to esimtate the dependence of convexity on target size. When $\bar{k} \ll 1$, energetics are dominated by headgroup area terms and $\epsilon_T''(\bar{k} \ll 1) \propto \gamma v_0^{2/3}/n_T^{7/3}$. Whereas, the opposite limit is controlled by length elasticity, and $\epsilon_T''(\bar{k} \gg 1) \propto k v_0^{2/3}/n_T^{4/3}$. Notably, as with the fluid capsule model in the previous section, convexity generically decreases with target size, however, not nearly as quickly. In particular, for the amphiphile model the product $n_T^3 \epsilon_T''/kt$ is always *increasing* with n_T (i.e. as $n_T^{2/3}$ and $n_T^{5/3}$ for small and large $\bar{k}$, respectively). Hence, according to eq. (20) the relative number fluctuations of micelles *decrease* with mean size. This is in marked distinction with the predictions of the bending elasticity of the fluid shell model, for which $\epsilon_T'' \propto n_T^{-3}$ and size fluctuations grow in proportion to mean size. The origin of this relatively sharper minima of aggregation energetics, and correspondingly tighter control of aggregate size, can be traced to the fact the the closure radius is on the scale of the subunit thickness (i.e. $R_{\text{close}} \approx t$), or more specifically, the additional considerations of tail packing constant density in the micellar cores, which are absent for the membrane bending elasticity of fluid capsules.

B. Self-limited, open-boundary assembly

Here we describe a class of self-limiting assembly characterized by an *open boundary*, a surface that separates the aggregate interior from the solution of freely associating subunits. Unlike the SCA described above, in open-boundary assemblies (OBA) this free boundary is maintained in the target self-limiting states. Therefore, the target state has a finite surface energy associated with loss of short-range cohesion or differences in solvation at its exterior.

As described in Sec. III.A.1, a finite boundary energy alone generically favors unlimited aggregates. Hence, to be self-limited, OBA structures require additional interaction terms that grow with aggregate dimensions, and thus balance the generic tendency to minimize the boundary to interior ratio. We define this additional (non-surface) energetics as the *excess energy*, and its essential feature is a regime of *super-extensive growth*, meaning the total excess energy increases with size faster than the number of subunits n. Below we describe two example mechanisms that generate such a form of excess energy, but the key underlying feature is the existence of *gradients in stress* throughout the aggregate. Whereas SCA can realize finite target dimensions with uniform subunit shape and packing, in OBA long-range, gradient patterns of intra-aggregate stress are required for self-limitation and ultimately dictate the range of possible self-limiting sizes.

Before introducing these two physical mechanisms, which will illustrate the microscopic origins of excess energy accumulation, we begin with a generalized description of the thermodynamics of OBA.

1. Limits of self-limitation

To describe the aggregation energetics of OBA, we consider a structure with an open boundary that can grow in D possible spatial dimensions (directions). We let d_L of these directions be potentially self-limiting, while in the remaining $d = D - d_L$ directions the structure undergoes unlimited growth. Denoting the limited and unlimited dimensions of the structure as W and L respectively, the scaling of aggregate volume gives $n \propto W^{d_L} L^d$. In the limit that $L \gg W$, the amount of the open boundary then grows as $A_b \sim W^{d_L-1} L^d$. [15] For example, consider subunits that can bind in all three spatial dimensions, i.e. $D = 3$. Quasi-cylindrical aggregates of such subunits accrue a surface energy cost deriving from their full boundaries, which are limited in two transverse spatial dimensions ($d_L = 2$) and unlimited in the axial direction ($d = 1$); whereas finite-thickness planar aggregates are limited along the normal direction ($d_L = 1$), but unlimited in the two in-plane directions ($d = 2$). Notably, such structures are self-limiting spatially, but do not necessarily have a finite or even well-defined peak ag-

----

[15] For $D = 2$, assemblies are sheet-like and the boundary corresponds to a 1D edge due to fewer (lateral) cohesive bonds, while for $D = 3$, the boundary corresponds to the entire 2D surface surrounding the aggregate.
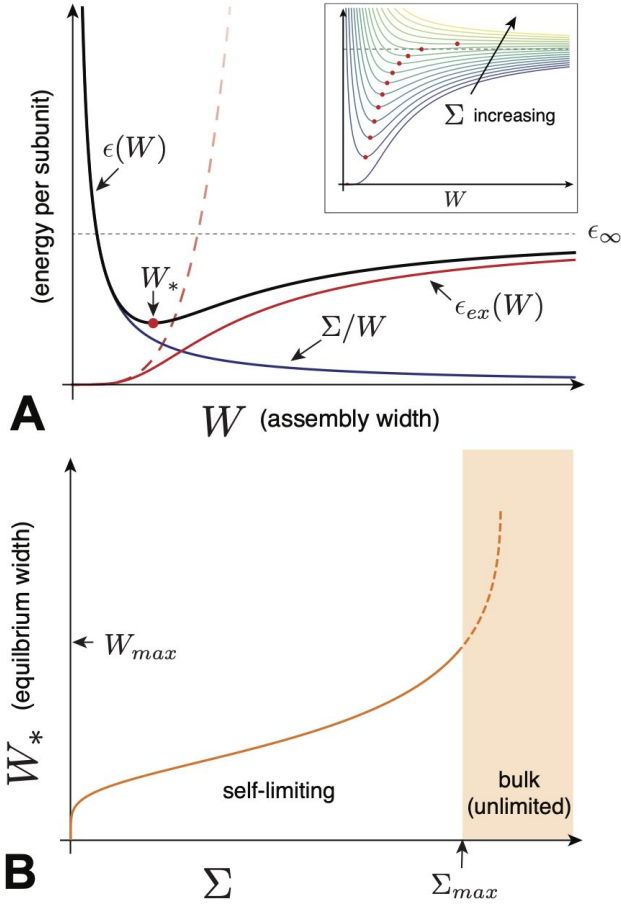
FIG. 13 (A) A schematic plot of the aggregation energetics (per subunit) for self-limiting open boundary assembly, where the blue, red, and black curves respectively show the surface energy density, excess energy, and total energy density. The dashed curve shows power-law growth of excess energy at small size. The red point highlights the equilibrium self-limiting width W_*, and the inset shows the variation of this equilibrium width with increasing surface energy Σ, assuming fixed $\epsilon_{ex}(W)$. (B) A schematic plot of the equilibrium width as a function of Σ, where the dashed portion indicates the possibility of a finite-width branch that is metastable relative to the bulk state $W \to \infty$. The boundary between equilibrium self-limiting and bulk states is marked by a maximum self-limiting size $W_{\max}$, denoted as the *escape size*.

gregation number [16]. Nevertheless, equilibrium in these cases (e.g. finite-thickness "filaments" and "slabs") derives from the optimal energy per subunit with respect to the self-limiting dimension(s) of the assembly.

Because surface subunits typically have fewer cohesive bonds with neighbor subunits and potentially more unfavorable contacts with surrounding solvent, an open boundary generically accrues a surface energy cost proportional to $A_{\rm b}$. Parameterizing this cost by the bound-

---

[16] Indeed, cylindrical aggregates have an exponential length distribution, while planar aggregates (e.g. plates and membranes) will correspond to a bulk state according to the analysis of Sec. II.A.2.

ary energy Σ and the "bare" aggregation energy for interior subunits $-\epsilon_0$, we may write a generic form for the per-subunit aggregation energy as a function of the self-limiting dimension W:

$$\epsilon(W) = -\epsilon_0 + \frac{\Sigma}{W} + \epsilon_{\rm ex}(W). \qquad (46)$$

For simplicity, we absorbed a geometric factor associated with the dimensionsality of the boundary into the definition of Σ. The first two terms describe the short-range cohesive interactions, with a constant bulk energy gain and a surface energy penalty, while the final term defines the *excess energy* relative to the short-range model. An example energy of this form is shown in Fig. 13A. As we illustrate below, such an excess energy arises from effects such as long-range inter-subunit repulsions or elastic stresses that increase with aggregate size. In self-limiting assemblies, these "cumulative" effects give rise to an excess energy (per subunit) that increases monotonically in size; hence, $\epsilon_{\rm ex}(W)$ captures energetic effects that grow *super-extensively* with aggregate size. In any physical system, this super-extensive energy growth will only persist up to some threshold assembly size, crossing over from convex (e.g. power-law) growth at small sizes to some asymptotically saturating energy density as $W \to \infty$. The large-W saturation of excess energy can occur for a variety of reasons. For example, long-range repulsions may be screened beyond some length scale. Alternatively, above a threshold excess energy cost, it will become energetically favorable for subunits to reorganize or deform to avoid further excess energy accumulation. Notwithstanding its microscopic origin, the effect of this saturating excess energy is to renormalize the per subunit energetics from its bare value to $-\epsilon_0 + \epsilon_\infty$.

In OBA, self-limitation follows directly from the balance between the accumulating cost of $\epsilon_{\rm ex}(W)$ and the generic decrease of surface energy with increasing W. It is then straightforward to show that equilibrium assemblies satisfy the following "equation of state" that links the finite equilibrium size W_* to the surface energy,

$$\Sigma = W_*^2 \epsilon'_{\rm ex}(W_*), \qquad (47)$$

with $\epsilon'_{\rm ex} = \partial_W \epsilon_{\rm ex}$. Stability criteria additionally require that $\epsilon''_{\rm ex} > 0$, but the basic results of the competition between surface energy and excess energy accumulation are shown schematically in Fig. 13. Since the surface energy always drives assembly toward larger sizes, the equilibrium finite size W_* generically increases with Σ given a fixed form of $\epsilon_{\rm ex}(W)$. Self-limitation can arise in two ways: either the equilibrium width can increase *continuously* with Σ to the bulk state (i.e. $W_* \to \infty$), or as illustrated in Fig. 13B, self-limitation will persist only up to a *maximal* finite size before a discontinuous transition to bulk assembly occurs. In the latter case, the energy density of the finite state eventually increases with surface energy, to the point where the energy densities of the finite and bulk (unlimited) states become equal; i.e. $\epsilon(W \to \infty) = -\epsilon_0 + \epsilon_\infty$. For surface energies above

this *maximal* value, the bulk (unlimited) state is favored. Hence, such systems can be characterized by a maximal self-limiting size, $W_{\max}$, and a maximal surface energy $\Sigma_{\max}$ below which equilibrium structures are finite (e.g. Fig. 13B).

For a general OBA, it is then useful to consider the following question: What are the range of possible self-limited equilibrium states that a given system can exhibit? As described in Sec. II.B.2, provided that the concentration is well above the aggregation threshold (CAC), the mean aggregate size is determined by the minimum of the per-subunit energy $\epsilon(n)$. Thus, the answer to this generic question depends only on the excess energy and its accumulation with width [17].

To see this, we reformulate the condition for equilibrium of the self-limiting state relative to bulk assembly, $\epsilon(W_*) < \epsilon(W \to \infty)$, in terms of the surplus of energy in the bulk relative to the finite state,

$$\Delta\epsilon(W_*) \equiv \epsilon(W \to \infty) - \left(\frac{\Sigma}{W_*} + \epsilon_{\mathrm{ex}}(W_*) - \epsilon_0\right) \quad (48)$$

$$= \frac{\partial}{\partial W_*}\Big(W_*\big[\epsilon_\infty - \epsilon_{\mathrm{ex}}(W_*)\big]\Big), \quad (49)$$

where we used the equation of state linking stable size to surface energy in eq. (47). The condition that $\Delta\epsilon(W_*) > 0$ is required for equilibrium finite (self-limited) states can then be simply formulated in terms of the first integral of surplus bulk energy

$$\mathcal{A}(W) \equiv W\big[\epsilon(W \to \infty) - \epsilon(W)\big] = W\big[\epsilon_\infty - \epsilon_{\mathrm{ex}}(W)\big] - \Sigma, \quad (50)$$

so the equilibrium of finite structures relative to bulk corresponds to the condition

$$\left(\frac{\partial\mathcal{A}}{\partial W}\right)_\Sigma > 0. \quad (51)$$

We refer to the function $\mathcal{A}(W)$ as the *accumulant*, and note that graphically it corresponds to the area of the rectangular regions of the plot of $\epsilon_{\mathrm{ex}}(W)$ versus W highlighted in Fig. 14A, for the model (1) described by the blue curve. Notably, because the fixed-Σ partial derivative in eq. (51) is independent of surface energy, for a given form of excess energy the accumulant may be constructed for any value of surface energy, and analyzed as a function of W to consider potential finite-size equilibria at *all values* of Σ.

According to eq. (51), finite-W equilibria correspond to the range of *increasing* $\mathcal{A}(W)$. Fig. 14 illustrates two models of excess energy, both of which are characterized by crossovers from power-law growth at small-W to
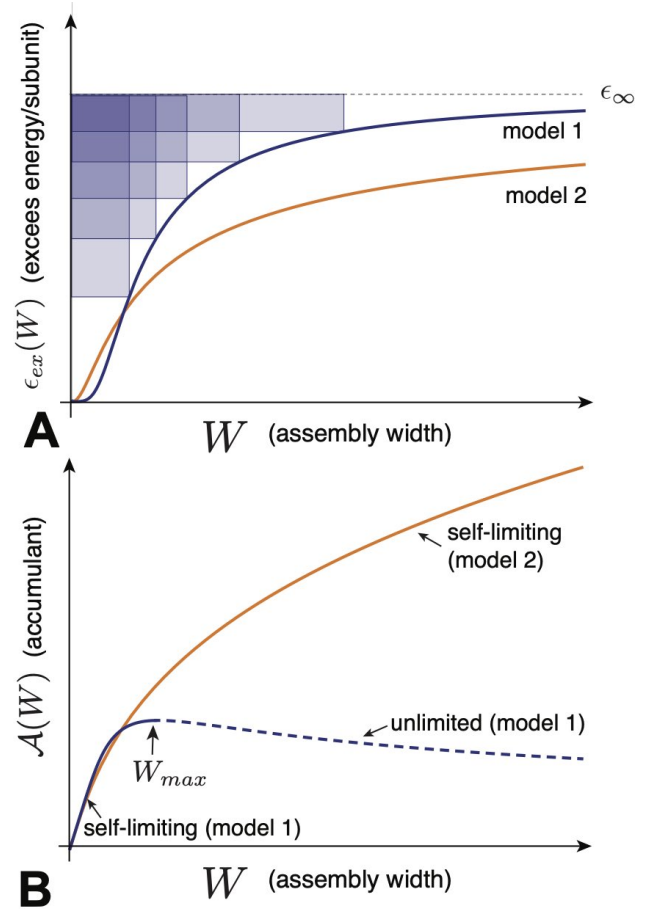

FIG. 14 (A) Schematic plots of excess energy versus width for two different models of OBA: model 1 (blue); and model 2 (orange). Both crossover from power-law growth at small size to saturating excess energy at large size, but do so with different functional forms (notably, the asymptotic approach to ϵ_∞ is slower in model 2 compared to model 1). The area of the blue rectangles graphically illustrates the definition of the *accumulant* in eq. (50). (B) Accumulant $\mathcal{A}(W)$ as a function of width for the two models shown in (A). Regions of increasing $\mathcal{A}(W)$ correspond to possible ranges of equilibrium self-limiting sizes for a given form of $\epsilon_{\mathrm{ex}}(W)$. That is, values of W for which $\mathcal{A}'(W) > 0$ correspond to energy minima of $\epsilon(W)$, which are lower in energy than the bulk state $\epsilon(W \to \infty)$, for a particular value of the surface energy Σ give by eq. (47). Hence, model 1 shows an upper limit for maximal self-limiting size $W_{\max}$, while model 2 exhibits stable self-limiting equilibria at all sizes.

asymptotic saturation to finite values of $\epsilon_{\mathrm{ex}}(W \to \infty) = \epsilon_\infty$, but do so via different functional dependencies on the finite size W.

For model 1 (blue), the range of increasing $\mathcal{A}(W)$ extends only up to a maximum, with corresponding width $W_{\max}$, indicating that a first-order transition between finite and bulk states occurs at $W_{\max}$ and the corresponding value of $\Sigma_{\max}$ [18]. For model 2 (orange), the monotonically increasing range of the accumulant extends to

---

[17] Eq. 19 shows a small concentration-dependent shift of the optimal size n^* below the size corresponding to the minimum of $\epsilon(n)$; the optimal size approaches the energy-minimizing size as $\Phi \to \infty$.

[18] Defining the accumulant in terms of the $\Sigma = 0$ energetics, i.e.

$W \to \infty$. This indicates that model 2 supports equilibrium self-limited states at all values of Σ, and only reaches bulk assembly in the limit of infinite surface cost. Hence, even if the excess energy saturates in the $W \to \infty$ limit, self-limitation may still, in principle, extend to all possible size ranges, depending on the nature of the asymptotic approach to the bulk energy.

Given the form of the accumulant defined in eq. (50) and the condition eq. (51) for its increase, the possibility for self-limited states that extend continuously up to the bulk state (i.e. $\lim_{\Sigma \to \infty} W_* \to \infty$) can be deduced from the asymptotic form of the residual energy $\Delta\epsilon(W) = \epsilon_{ex}(\infty) - \epsilon_{ex}(W)$ as $W \to \infty$. Following an argument made by (LeRoy, 2018; Terzi et al., 2020), we assume this residual vanishes as a power law $\epsilon_{ex}(\infty) - \epsilon_{ex}(W) \sim W^{-\nu}$. It is then straightforward to show that when $\nu > 1$, $\mathcal{A}(W)$ decreases as $W \to \infty$. Such cases correspond to the first-order type self-limitation exhibited by model 1 in Fig. 14. Alternatively, when $0 < \nu < 1$, indicating a slower saturation of excess energy, the accumulant continues to increase as $W \to \infty$, indicating the existence of self-limited equilibria extending up to the bulk state. The case of $\nu = 1$ is marginal, and can exhibit either first- or second-order like behavior. Below, we describe models in the context of geometrically frustrated assemblies that can illustrate both types of behavior, either a *continuous* or *discontinuous* transition between the finite and bulk states depending on the mechanisms underlying the accumulation of excess energy.

2. Short-range attractions, long-range repulsions

As described in the previous section, self-limitation in OBA requires superextensive growth of the excess energy. In this first class of examples, the accumulation of excess energy derives from long-range interactions between subunits, specifically, interactions characterized by *short-range attraction and long-range repulsion* (SALR) (Groenewold and Kegel, 2001; Sciortino et al., 2004). Models of such systems usually consider isotropic pair-potentials $u(r)$ that can be split into two parts,

$$u(r) = u_{SA}(r) + u_{LR}(r), \qquad (52)$$

where the short-ranged potential describes cohesive interactions $u_{SA}(r) < 0$ between neighboring subunits, which act on scales comparable to the subunit hard-core diameter d; i.e., $u_{SA}(r \gg d) \simeq 0$. Outside of this cohesive range, the potential is dominated by a long-range repulsion $u_{LR}(r) > 0$ that extends over sizes much larger than single particles. Even when the repulsive interactions

———————

$\mathcal{A}(W) \equiv W\big[\epsilon_\infty - \epsilon_{ex}(W)\big]$, as in Fig. 14B, it can be easily shown that $\Sigma_{max} = \mathcal{A}(W_{max})$.

between neighboring subunits are much weaker than the cohesion, the fact that repulsive interactions extend far beyond the first shell of neighbors can lead to superextensive growth of repulsive energy with increasing assembly size. A model of this form has been applied to explain finite-sized aggregate formation in a broad range of systems (Dinsmore et al., 2011), including protein complexes (Cardinaux et al., 2007; Foderà et al., 2013; Stradner et al., 2004), charged nanoparticles (Nguyen et al., 2007), colloidal particles in low-dielectric solvents (Sedgwick et al., 2004; Van Schooneveld et al., 2009), dipolar mesophases (Seul and Andelman, 1995), and even models of "nuclear matter" (Caplan and Horowitz, 2017).

To illustrate the mechanism of self-limitation in this class of systems, we consider the following specific model: Short-range cohesive interactions lead to $-u_0$ per neighbor contact, and aggregates maintain an (approximately) uniform density $\rho_0 \approx d^{-3}$. For repulsive interactions, we assume that subunits are isotropic and repel according to a screened (Yukawa) repulsion,

$$u_{LR}(r) = \frac{q^2}{r} e^{-\kappa r}. \qquad (53)$$

Here q is the electrostatic charge per subunit, assumed to be fixed, and κ is the screening length, which arises from Debye-Hückel screening by mobile ions in solution, and truncates the far-field repulsions for $r \gg \kappa^{-1}$. For simplicity, we illustrate a simplified version of the theory by Groenowold and Kegel (GK) (Groenewold and Kegel, 2001). In our presentation, we consider the case where charge per subunit is fixed, a point which we revisit below. Our primary purpose is to illuminate a model with the minimal ingredients for a self-limiting equilibrium, and due its finite screening length, the long-range Yukawa potential provides a convenient example. For purposes of illustration, we consider spherical aggregates with radius R, whose interaction (free) energy can be described by

$$\epsilon(R) = -u_0 \frac{\langle z \rangle}{2} + \frac{3\Sigma}{R} + \epsilon_{ex}(R) \qquad (54)$$

where $\langle z \rangle$ is the mean number of neighbor contacts in the bulk of the aggregate, and $\Sigma \approx u_0/d^2$ is the surface energy associated with fewer (short-ranged) cohesive contacts at the boundary. Here, the excess energy derives directly from the sum of long-range, pairwise repulsions
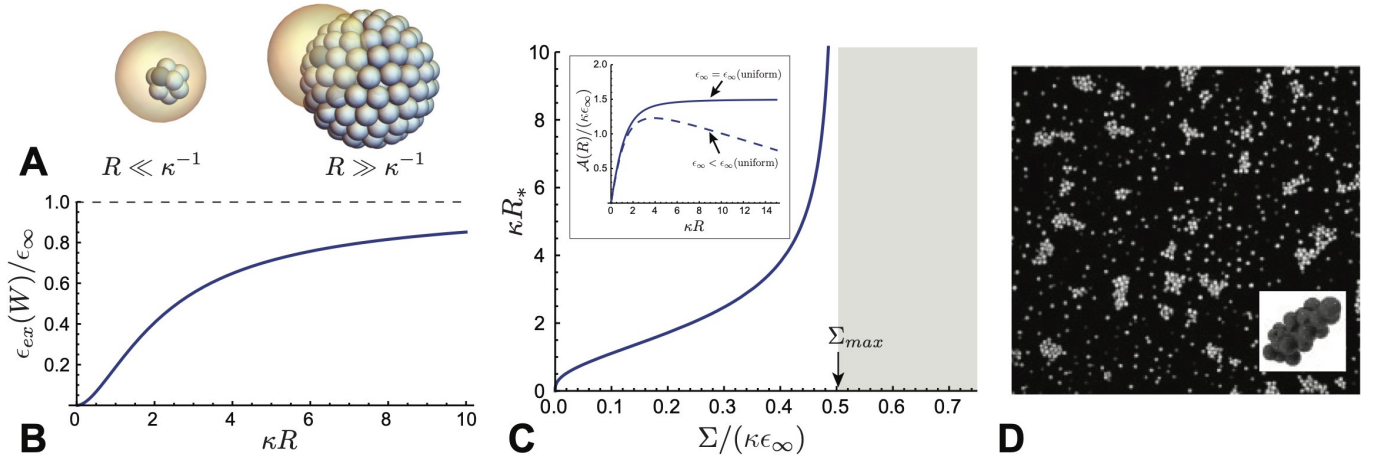
FIG. 15 (A) Schematics of aggregates of particles interacting via short-range (i.e. contact) interactions and long-range repulsions, with a screening length κ^{-1}, which can be large or small compared to the aggregate size R. (B) Excess energy of spherical aggregates versus radius, showing a crossover from power-law growth for $R \ll \kappa^{-1}$ to an asymptotically saturating bulk energy for $R \gg \kappa^{-1}$. (C) Plot of the equilibrium finite radius R_* of the SALR model versus surface energy Σ, showing a continuous divergence to the bulk at a finite $\Sigma = \Sigma_{\max}$. The second-order-like transition from finite-to-bulk states predicted by this model (solid curve) is consistent with the monotonically increasing form of the accumulant plotted in the inset, where this result assumes the bulk state has uniform density. Assuming instead, that the bulk state has a lower energy, non-uniform density (e.g. a periodic aggregate morphology (Zhang *et al.*, 2019)), the transition will be first-order, as illustrated by the non-monotonic accumulant shown as a dashed line. (D) shows a confocal microscope image of clusters of charged colloidal particles (radius 660 nm) adapted from (Stradner *et al.*, 2004). In these experiments short-range attractions are induced by depletion forces generated by inert polymers in suspension, while electrostatic repulsions are maintained at relative long range due to the low-concentration of mobile ions in host organic solvent (Stradner *et al.*, 2004).

in the aggregate volume $\mathcal{V}$ [19],

$$
\epsilon_{\text{ex}}(R) = \frac{q^2 \rho_0}{2\mathcal{V}} \int_{\mathcal{V}} d^3\mathbf{r}\, d^3\mathbf{r}'\, u_{LR}(|\mathbf{r} - \mathbf{r}'|)
$$
$$
= \frac{q^2 \rho_0}{2\kappa^2}\left\{ 1 - \frac{3(1+\kappa R)}{2(\kappa R)^3}(1 + e^{-2\kappa R})\left[\kappa R - \tanh(\kappa R)\right]\right\}.
$$
$$(55)$$

The behavior encoded in $\epsilon_{\text{ex}}(R)$ is plotted in Fig. 15B, and can be understood physically by considering the asymptotic limits of small and large aggregate size relative to the screening-length of the repulsive interactions, as shown schematically in Fig. 15A. For small aggregates,

$$
\epsilon_{\text{ex}}(R \ll \kappa^{-1}) \simeq \frac{q^2 \rho_0}{5} R^2, \qquad (56)
$$

which derives from the fact that, when repulsions extend over the entire aggregate, the per-subunit cost of the Coulomb self-energy of an aggregate of is roughly $Q^2/(nR) \propto R^2$, where the total aggregate charge $Q = qn \propto R^3$. In the opposite regime, where aggregate sizes

---

[19] The form of the self-energy can be readily calculated using Greens theorem, where the total repulsive energy can be written as $\int_{\mathcal{V}} d^3\mathbf{r}\, q\rho(\mathbf{r})\phi(\mathbf{r})$ with a potential that satisfies the linearized (Debye-Hückel) equation, $(\nabla^2 + \kappa^2)\phi(\mathbf{r}) = 4\pi q\rho(\mathbf{r})$. Solving this for spherically symmetric aggregates, $\rho(r \leq R) = \rho_0$ and $\rho(r > R)$, yields the explicit form of screened electrostatic energy.

far exceed the screening length,

$$
\epsilon_{\text{ex}}(R \gg \kappa^{-1}) \simeq \epsilon_\infty \left(1 - \frac{3}{2\kappa R} + \frac{3}{2(\kappa R)^3}\right), \qquad (57)
$$

where $\epsilon_\infty \equiv q^2 \rho_0 \kappa^{-2}/2$. This leading term $R \to \infty$ derives from the fact that each subunit in the bulk of the aggregate experiences repulsive interactions with roughly $\rho_0 \kappa^{-3}$ other subunits within a screening length, while the first correction accounts for the surface layer of thickness κ^{-1} with fewer neighbors within the screening length. Notably, the subleading $1/R^3$ term can be associated with a square curvature cost (per unit area) for deforming the boundary shape from planar, which alters the distribution of repulsive particles near the free surface of the aggregate.

Fig. 15C shows the predicted equation of state for the self-limiting radius $R*$ as a function of surface energy, $\Sigma(R_*) = R_*^2 \epsilon'_{\text{ex}}(R_*)/3$. For small Σ, the balance between the "charging energy" of aggregates and surface energy leads to a growth $R_* \sim \Sigma^{1/3}$, which proceeds until the optimal aggregate size grows beyond the screening length. In the large aggregate regime, the asymptotic approach to ϵ_∞ leads to an aggregate size that diverges continuously at a critical surface energy $\Sigma_{\max} = \kappa\epsilon_\infty/2$. The origin of the second-order-like transition to the bulk in this model can be traced to the $R \gg \kappa^{-1}$ form of the excess energy in eq. (57). The leading correction to $\epsilon_{\text{ex}}(R \to \infty) = \epsilon_\infty$ goes as $-1/R$, that is, in the notation of the foregoing Sec. III.B.1, $\nu = 1$. Hence, this leading correction at large R behaves like a negative contribution to the surface energy, $\Sigma_{\text{eff}} = \Sigma - \Sigma_{\max}$, due to the

reduced electrostatic repulsion within a screening length of the surface. When $\Sigma < \Sigma_{\max}$, the effective surface energy is *negative*, and the aggregate equilibrium derives from the balance between the $-1/R$ drive to create more surface and the sub-leading $+1/R^3$ term, to give

$$R_* \sim \sqrt{\frac{\epsilon_\infty \kappa^{-3}}{\Sigma_{\max} - \Sigma}}, \qquad (58)$$

which diverges (continuously) as $\Sigma \to \Sigma_{\max}$. This prediction, that there is a maximal cohesive surface energy for aggregates, but that their equilibrium self-limiting size extends to arbitrarily large values, is consistent with the plot in Fig. 15C(inset), which shows the accumulant of spherical aggregates to be monotonically increasing over the full range of R.

The result that self-limitation can extend up to arbitrarily large sizes (i.e. $W_{\max} \to \infty$) for this simplified model may be surprising, as it implies that the thermodynamics are sensitive to the finite size of aggregates over much larger size ranges than the finite interaction range, κ^{-1}. That is, a subunit whose interactions extend only κ^{-1} can still "sense" that it should join an aggregate with a radius smaller than $R_* \gg \kappa^{-1}$, but not a larger one. The resolution of this puzzle is that the physical term that restrains aggregate growth in this $R_* \gg \kappa^{-1}$ regime derives from the square curvature cost of the free boundary (contributing as $+1/R^3$ to the excess energy). Although repulsions are short-ranged compared to large aggregates (i.e. $R \gg \kappa^{-1}$), they are sufficiently non-local to sense the *curvature* of the boundary, and thereby, the global radius of the aggregate. That is, repulsions in this regime give rise to precisely the type of square-curvature energetics that selects for finite sizes in self-closing assemblies in Sec. III.A.1; i.e., with a preferred boundary curvature that vanishes as $\Sigma \to \Sigma_{\max}$ [20].

While self-limited assembly due to a competition between cohesive boundary costs and accumulation of long-range repulsion is generic, the specific features of this simplified model, including continuously diverging finite size for $R_* \gg \kappa^{-1}$, require several caveats. Foremost, the convexity of the assembly energetics decreases with aggregate size as $\epsilon_*'' \sim n_*^{-3}$ for $R_* \gg \kappa^{-1}$ for this model, similar to the case of self-closing shells in Sec. III.A.1. Thus, according to eq. (20), fluctuations in n grow with self-limiting size, and diverge at the threshold as $\langle (\Delta n)^2 \rangle^{1/2} \propto n_* \sim (\Sigma_c - \Sigma)^{-3/2}$. Hence, for all practical purposes the self-limitation to finite and well-defined sizes will not persist to arbitrarily large structures.

Beyond this, the model is oversimplified and fails to incorporate a number of physical ingredients, which can

influence the form of excess energy accumulation. For example, more realistic models would include: a finite compressibilty of the short-range cohesive forces, which allows for density variation with aggregate size and position within the aggregate; and for charged subunits in solvent, the effects of dielectric contrast between aggregates and the solvent, as well as variable degrees of charge condensation/dissociation as aggregates vary in size. Indeed, when initial applications of this GK model aimed to understand the aggregation of lysosyme proteins in low-salt aqueous solutions, comparison to scattering measurements of the aggregate size suggested that variable ionization of subunits could not be neglected (Zaccarelli, 2007). While to a first approximation, it was argued that subunit charging is independent of aggregation number (Groenewold and Kegel, 2001), the charge per subunit and the screening length, which varies with counterion concentration, both vary with *total concentration* of ionizable subunit species (Cardinaux *et al.*, 2007). Hence, accounting quantitatively for the concentration dependence of aggregation for these electrostatic systems requires consideration of subunit charge in a self-consistent fashion, leading potentially to aggregation-dependent renormalization of repulisive interactions (Nguyen *et al.*, 2007), not to mention additional effects associated with non-spherical aggregate shapes (Sciortino *et al.*, 2005). Notably, the non-spherical nature of aggregates of charged particles is evident in confocal images of a colloidal analog to the aggregation of nanoscopic proteins (Sedgwick *et al.*, 2004), as shown Fig. 15D.

Finally, even beyond specific physical considerations of electrostatic SALR systems, the prediction of the simplified Yukawa model of diverging finite aggregate size assumes a spatially uniform bulk state, whereas models of systems with short-range attractions and long-range repulsions have been shown to form periodically-modulated aggregate phases, e.g. stripes, layers and spheres, at high concentration (Sciortino *et al.*, 2004; Sear and Gelbart, 1999; Seul and Andelman, 1995; Zhuang and Charbonneau, 2016). That is, at sufficiently high densities, the uniform density bulk state is unstable to lower free energy periodic bulk states such that $\epsilon_\infty(\text{periodic}) = \epsilon_\infty(\text{uniform})$. If the stability of equilibrium self-limitation is reanalyzed in terms of a competition with a lower-energy, non-uniform bulk state, then the maximal size range of self-limiting equilibrium becomes finite (see e.g. the accumulant plot in Fig. 15D). The full thermodynamics of the transition between a dilute phase of self-limiting aggregates (above the CAC) to the long-range ordered bulk aggregate phases requires considerations beyond the ideal aggregation thermodynamics considered here.

---

[20] Consistent with this result, performing an analogous calculation for a planar slab geometry shows that the self-limited size diverges exponentially above a threshold value of Σ, due to the absence of such a curvature term.
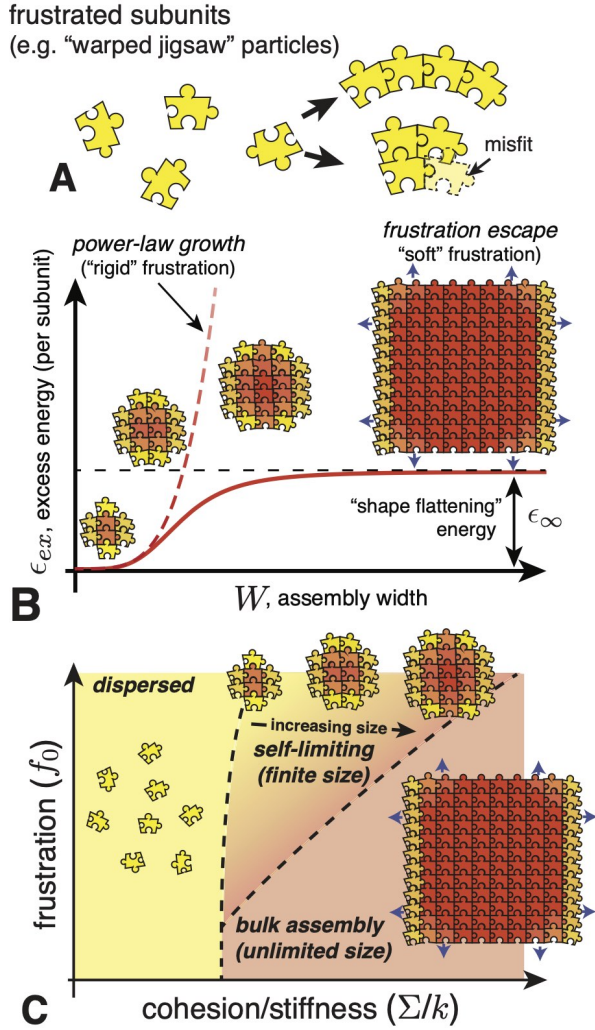
FIG. 16 (A) A heuristic "warped jigsaw" model for GFA, after (Grason, 2017), in which directional interactions promote curvature along rows of a locally preferred 2D "lattice". Assembly in both bonding directions leads to "misfits", as in the tetrameric aggregate. (B) A schematic plot of the excess energy (per subunit) for the warped jigsaw model, due to the superextensive build-up of elastic costs of misfits. Intra-assembly strains are illustrated via particle color, from unstrained (yellow) to highly-strained (red). The excess energy shows the characteristic crossover from power-law growth at small aggregates to an asymptotic approach to a strained bulk state (in this case envisioned as "shape flattening" of jigsaw particles). (C) A schematic phase diagram for a generic model GFA in eq. (59), considered (at fixed concentration and temperature) as a function of the ratio of surface energy to elastic stiffness and a measure of the "strength" of frustration, f_0.

3. Geometrically frustrated assembly (GFA)

The notion of *geometric frustration* (GF) originally emerged in the context of low-temperature, condensed matter systems (e.g. magnetic materials, spin models) (Vannimenus and Toulouse, 1977). It refers to the impossibility of propagating an energetically-preferred arrangement throughout space, due to global geometric constraints (Kléman, 1989; Sadoc and Mosseri, 2006).

For bulk, infinite systems, GF leads to a rich phenomenology: extensive arrays of topological defects thread through the highly-degenerate bulk ground states that populate a rough energy landscape.

Recently has it been recognized that GF gives rise to new behaviors in self-assembling materials (Grason, 2016), deriving from two key features. First, the constituent subunits (e.g. polymers, colloids, proteins) are relatively "soft" and held together by weak, non-covalent forces. Second, assemblies need not reach bulk states, and thus have additional degrees of freedom associated with the (potentially) finite-size and shape of the assembled domain. Unlike bulk or rigid systems where GF must be resolved by defects (Sadoc and Mosseri, 2006), in soft assemblies it can be tolerated, at least over some size range, by smooth gradients in the subunit shapes and packings. As an illustration, see the schematic of "warped jigsaw" particles in Fig. 16A, where the tapering of the particle shape favors curvature along one row of the lattice assembly (Grason, 2017). Provided particles or their interactions are sufficiently deformable, aggregates can accommodate frustration through strain gradients, leading to arrangements that are more (less) relaxed near to (far from) the free boundary of the aggregate.

This *self-organization of long-range stress gradients* is the defining characteristic of geometrically frustrated assembly (GFA), giving rise to the form of accumulating excess energy that can limit assembly size (Sec. III.B.1). The balance between the surface energy and the (superextensive) cost of GF can select equilibrium domain sizes that are finite, and in principle, arbitrarily larger than the subunits themselves.

To date, GFA has been implicated in the emergent structures of various soft matter systems, including self-twisting protein bundles (Aggeli *et al.*, 2001; Brown *et al.*, 2014; Cameron *et al.*, 2018; Grason and Bruinsma, 2007; Hall *et al.*, 2016; Turner *et al.*, 2003; Yang *et al.*, 2010), twisted molecular crystals (Haddad *et al.*, 2019; Li *et al.*, 2020), chiral smectics (Hough *et al.*, 2009; Matsumoto *et al.*, 2009) and membranes (Armon *et al.*, 2014; Ghafouri and Bruinsma, 2005; Gibaud *et al.*, 2012; Kang and Lubensky, 2017; Sakhardande *et al.*, 2017; Selinger *et al.*, 2004; Sharma *et al.*, 2014; Zhang *et al.*, 2019), particle-coated droplets (Bausch, 2003; Irvine *et al.*, 2010; Meng *et al.*, 2014; Yu *et al.*, 2016), curved protein shells (Li *et al.*, 2018; Zandi *et al.*, 2004), and phase-separated lipid vesicles (Schneider and Gompper, 2005). Consideration of GF in these systems has primarily stemmed from experimental observations of assemblies that (A) terminate at finite size and/or (B) exhibit "defect-ordered" morphologies. Continuum models that consider the interplay between the elastic costs of "misfit" and domain formation have been developed to address a range of distinct frustration mechanisms, including frustration of 2D liquid crystalline/crystalline order on non-Euclidean manifolds (Bowick and Giomi, 2009; Nelson and Peliti, 1987), metric and orientational frustration of chiral fibers (Brown *et al.*, 2014; Grason, 2015;

Haddad *et al.*, 2019), shape-frustration in stacking assemblies of curved layers (Achard *et al.*, 2005; DiDonna and Kamien, 2003; Matsumoto *et al.*, 2009), chirality frustration in crystalline (Armon *et al.*, 2014; Ghafouri and Bruinsma, 2005) and liquid crystalline membranes (Kang and Lubensky, 2017; Selinger *et al.*, 2001), and the assembly of non-tiling polygonal particles (Lenz and Witten, 2017).

We first describe a heuristic model that highlights the common thermodynamic features of these apparently diverse realizations of GFA, which are encoded in a simplified form of the excess energy,

$$\epsilon_{\text{ex}}(W) \approx \frac{k}{2}\delta^2(f, W) + \frac{C}{2}(s - s_0)^2. \qquad (59)$$

Here, k is a generalized elastic parameter for straining inter-element packing, as measured by a generalized (mean) strain $\delta(f, W)$, which itself varies with the finite size W of the domain and a parameter f that measures the "strength of frustration". To be clear, local strains vary with position throughout the assembly, as we describe for a specific example below, but for simplicity we focus here on how the magnitude of strain varies with size and frustration. While frustration mechanisms vary considerably among distinct GFA systems, they all exhibit power-law growth of strain with domain size W. This can be modeled heuristically as

$$\delta \approx f(s)W^\eta \qquad (60)$$

where η and f vary for different cases of GFA. For example, the orientational strains in 2D liquid crystal domains grow linearly with domain size ($\eta = 1$), whereas positional strains in frustrated 2D crystals grow quadratically ($\eta = 2$) (Grason, 2016; Niv and Efrati, 2018). The definition of "frustration strength" f depends on the specific GF mechanism; however, it can generally be expressed as a function $f(s)$ of the local shape of inter-subunit packing (e.g. inter-subunit bend or twist), which we denote generically with the shape parameter s. In many cases, the frustration strength can be expressed as a simple power law of shape,

$$f(s) \approx s^\mu \qquad (61)$$

where μ is a positive exponent. For the example of a crystalline cap on a spherical surface (Grason, 2016), f corresponds to the Gaussian curvature, which is the square of the 1D curvature in this geometry and thus corresponds to a shape parameter with $\mu = 2$. In the form of eq. (61), the strength of frustration generically vanishes, intuitively, in the limit that the shape "flattens" to $s \to 0$. The second term in eq. (59) describes generic costs for deformations away from an ideal, misfitting shape with $s = s_0 \neq 0$, which incur elastic penalties described by the "shape modulus" C.

This basic form of eq. (59) implies a generic size dependence for the excess energy shown schematically in Fig. 16B. For small sizes, assemblies retain their preferred, misfitting shape ($s \simeq s_0$) leading to a power-law growth of excess energy, $\epsilon_{\text{ex}}(W \to 0) \simeq kf_0^2W^{2\eta}/2$, where $f_0 = f(s_0) = s_0^\mu$ is the frustration strength of the preferred shape. If power-law growth of $\epsilon_{\text{ex}}(W)$ extended to all size scales, then the compromise between costs of GF and surface energy would select a finite equilibrium size $W_* \sim f_0^{-2/(2\eta - 1)}$ for any surface energy Σ. However, in any physical system, the excess energy can only accumulate up to some maximal size scale, beyond which the assemblies *escape frustration* through one of a number of competing morphological "modes" (Hall *et al.*, 2016; Hall and Grason, 2017). These include, for example, the formation of *topological defects* that screen far-field frustration stresses (Grason, 2012; Li *et al.*, 2019), as well as *shape flattening*, which refers to the smooth deformation of an incompatible (i.e. misfit) shape to a uniformly strained, compatible one (Grason, 2020). Because subunits and their interactions are generically soft, the excess energy required to escape frustration must be finite. Thus, as shown in Fig. 16B, $\epsilon_{\text{ex}}(W)$ generically crosses over from power-law accumulation at small W to an asymptotic approach to this finite energy $\epsilon_{\text{ex}}(W \to \infty) = \epsilon_\infty$. In the minimal description of eq. (59), the cost of shape flattening is simply $\epsilon_\infty = Cs_0^2/2$ per subunit.

Comparing power-law accumulation at small sizes to the asymptotic shape-flattened energy, one expects a crossover between these regimes at a characteristic "flattening" size $W_{\text{flat}} \approx \left[(C/k)/s_0^{2(\mu - 1)}\right]^{1/\eta}$. This size scale defines the maximum size for which assemblies will tolerate the accumulating frustration cost, beyond which deformation to an unfrustrated shape ($s \to 0$) becomes energetically favorable. Intuitively, this length scale also sets a bound on the *escape size*, $W_{\text{max}} \leq W_{\text{flat}}$, since for $W \gg W_{\text{flat}}$ the bulk energy is simply renormalized by the cost of flattening. In general, the escape size is set by the lowest-energy mode of relaxing frustration in the bulk state, which may be also involve "Wigner lattice" states of defects that neutralize the long-range cost of frustration (Li *et al.*, 2019). Understanding the practical limit of self-limitation then requires considering all possible competing modes of relaxing frustration, and determining which of these has the lowest energy for a particular regime of assembly. Notably, this distinction between power-law growth of energy density with W vs. asymptotic approach to a constant value has been proposed, by Meiri and Efrati, to classify distinct regimes of GF itself, as respectively "cumulative" and "non-cumulative" frustration (Meiri and Efrati, 2021). The analysis of the accumulant in Sec. III.B.1 translates that criteria into directly into their thermodynamics consequences for equilibrium SLA.

Notwithstanding which mode facilitates escape to the bulk, the heuristic picture of GFA implies a common phase diagram (Fig. 16C), spanned by the bare frustration strength f_0 on one axis, and the ratio of cohesion to intra-assembly stiffness Σ/k on the other (say for fixed
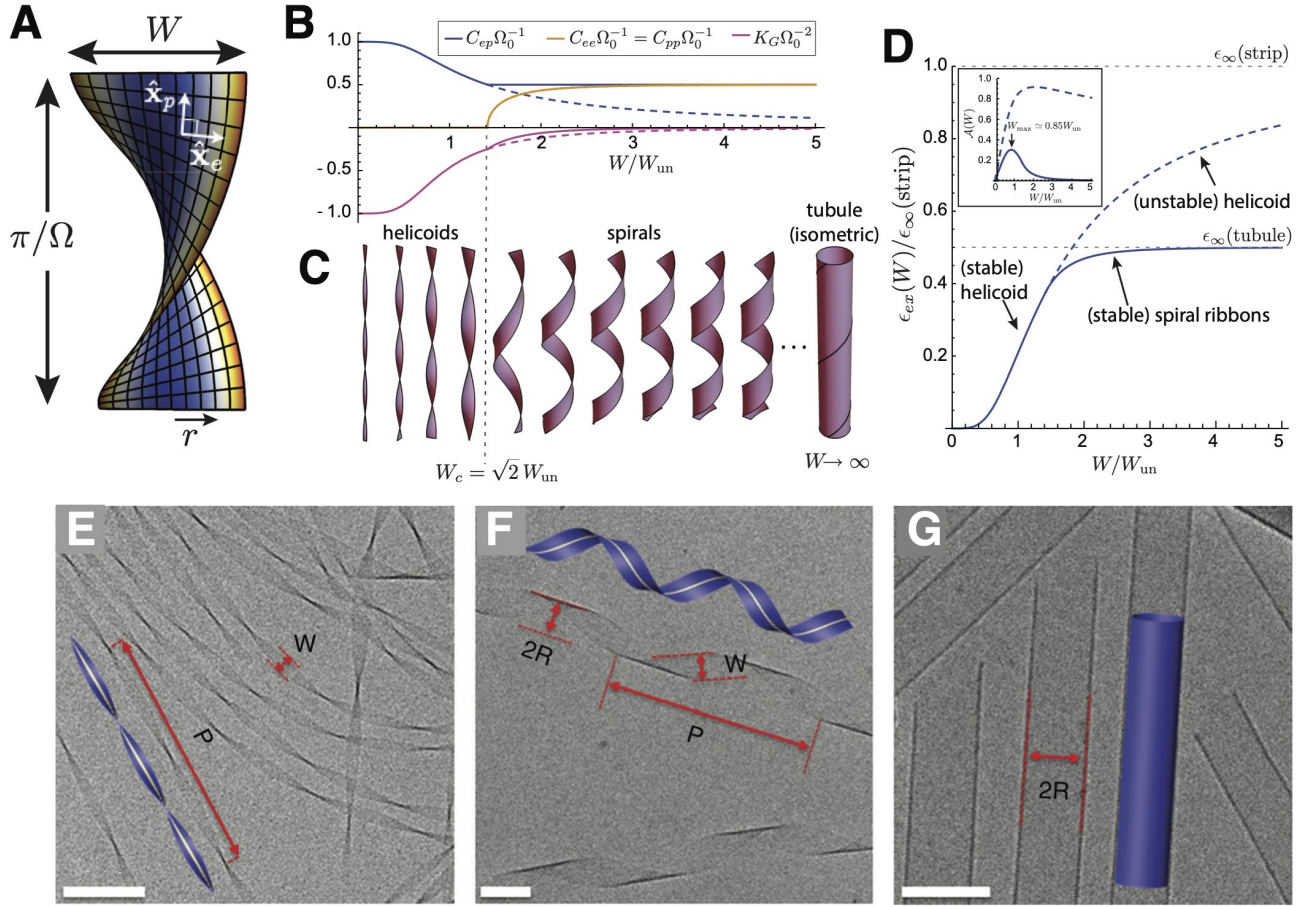
FIG. 17 (A) A schematic of a chiral, crystalline ribbon, where colors indicate the local extensional strains required by negative Gaussian curvature, from low (blue) to high (red). (B) Plots of the equilibrium shape relaxation of the narrow ribbon model of (Ghafouri and Bruinsma, 2005) as a function of increasing ribbon width, for curvature components C_{ij} and Gaussian curvature K_G, which is approximately uniform over the ribbon. The solid branches show the minimal energy configurations, while the dashed line indicates the (unstable) helicoidal equilibrium. Widths are scaled by the characteristic length W_{un}, defined in Eq. (63). (C) shows schematics of the shape equilibrium, in particular the shape transition from helicoids to spiral ribbons at a critical width W_c. (D) and (E) plot the respective excess energies for helicoidal (dashed) and helicoid-spiral shape branches (solid). While both shape modes predict an asymptotically relaxed frustration energy and a finite self-limiting width, spiral ribbons expel K_G at much faster rate, leading to a narrower range of self-limitation. (E-G) show transmission electron microscopy images of ribbons assembled from chiral bola-amphiphiles adapted from (Zhang *et al.*, 2019). The panels illustrate morphologies evolving with assembly time (scale bars 100 nm): in (E), helicoids are observed after 24 hours; in (F) spiral ribbons are observed after one week; and finally, in (G), closed tubules are observed after 5 months.

subunit concentration and temperature). Above a critical frustration strength, there is a regime of self-limiting aggregates between the dispersed state (below the CAC) and the bulk state. Within the self-limited regime, the equilibrium domain size increases with Σ/k but decreases with frustration strength f_0.

As a specific example of the origin and implications of GFA, we briefly discuss a well-studied model for crystalline ribbons frustrated by chirality (Armon *et al.*, 2014; Ghafouri and Bruinsma, 2005; Selinger *et al.*, 2004). This model has been developed to understand the polymorphic assembly of chiral surfactant bilayers that adopt a variety of quasi-1D structures (Oda *et al.*, 1999; Selinger *et al.*, 2001; Zhang *et al.*, 2019; Ziserman *et al.*, 2011); i.e., they form 2D sheets that are much narrower in width than length, $W \ll L$. Ghafouri and Bruinsma first de-

scribed this mechanism of frustration (Ghafouri and Bruinsma, 2005), posing it as a competition of the bending cost of a chiral (anisotropic) membrane against the elastic costs of in-plane stretching for a 2D crystal with Gaussian curvature. More recently, this model has been extended (Armon *et al.*, 2014; Grossman *et al.*, 2016) to describe experimental observations of bola-amphiphile ribbons by Sharon and coworkers (Zhang *et al.*, 2019).

For the purposes of this discussion, we present here only a simplified picture of the frustration and its effects in chiral ribbons, and provide a more detailed summary in Appendix B. As shown in Fig. 17A, we consider assemblies of ribbons of width W and much longer (unlimited) length $L \gg W$. For the case of sufficiently narrow ribbons, it can be shown that the favored morphology is that of a *helicoid*: a strip whose width axis twists around

a straight central ribbon at a uniform rate Ω. In this regime, the excess energy of the helicoidal chiral ribbon takes the following form (see Appendix B),

$$\epsilon_{\text{ex}}(W) = \frac{Ya_0}{1440}\Omega_0^4 W^4 + Ba_0(\Omega - \Omega_0)^2, \qquad (62)$$

where a_0 is the area per subunit in the membrane, and Y and B are the respective in-plane stretching and out-of-plane bending moduli for the assembled membrane. Here Ω plays the role of a shape parameter in the heuristic model introduced above, and the second term has the form of a *chiral bending energy* that favors uniform twist Ω_0 (Ghafouri and Bruinsma, 2005; Helfrich and Prost, 1988). The first term derives from the elastic cost of in-plane deformation of the 2D lattice of the ribbon, and describes the accumulating elastic cost of frustration. It is intuitive to understand the growth of stretching energy with ribbon width by considering the contour length per helical turn of the longitudinal strip of the ribbon at a distance r from the center, $(2\pi/\Omega)\sqrt{1+(\Omega r)^2}$. Due to the resistance to shearing and stretching of local elements in the 2D solid, this motif leads to strains that grow with relative longitudinal extension compared to the midline, roughly $\approx (\Omega r)^2$ to lowest order, and hence, it generates an elastic energy density of $\sim Y(\Omega^2 W^2)^2$. It can be shown that the underlying source for stress gradients in such an assembly derives from the incompatibility of a 2D planar metric with the non-zero (negative) Gaussian curvature of the membrane, $K_G \simeq -\Omega^2$ (Armon *et al.*, 2014; Ghafouri and Bruinsma, 2005). In the language of the heuristic picture introduced above, we can identify the strength of frustration with K_G, and hence $\eta = \mu = 2$, due to the quartic growth of frustration cost with shape misfit Ω and width W.

An instructive, albeit oversimplified, analysis of the size-dependence of frustration in chiral ribbons considers only the relaxation of helicoidal twist for variable width. For small widths, the ribbon adopts a preferred twist, $\Omega_*(W \to 0) = \Omega_0$, due to the vanishing cost of frustration as $W \to 0$, and $\epsilon_{\text{ex}} \approx Y\Omega_0^4 W^4$. For large widths $W \to \infty$, the balance between stretching and bending favors unwinding of the pitch, $\Omega_* \simeq \Omega_0(W_{\text{un}}/W)^{4/3}$, where

$$W_{\text{un}} \equiv \left(\frac{720B}{Y\Omega_0^2}\right)^{1/4}, \qquad (63)$$

defines a characteristic *unwinding size*. That is, for $W \ll W_{\text{un}}$ the ribbon retains roughly the twist preferred by chirality and the power-law ($\sim W^4$) accumulation in excess energy, while for much larger ribbons, the prohibitive cost of frustration causes the ribbon to unwind, expelling Gaussian curvature. Fig. 17D (dashed curve) shows the characteristic crossover in the excess energy of helicoidal ribbons, with an asymptotic flattening of the excess energy characterized by the exponent $\nu = 4/3$, which implies self-limiting widths are not stable in the unwinding region (see section III.B.1). In terms of the accumulant analysis (inset Fig. 17D), the escape of frus-

tration by helicoid unwinding would imply a maximum self-limiting size of $W_{\text{max}}(\text{helicoid}) \simeq 2.1 W_{\text{un}}$.

How do we understand physical parameters that determine the range of possible self-limiting widths? Here, we note that it arises from the combination of two physical lengths. One of these derives from the ratio of bending to stretching moduli, $\sqrt{B/Y} \equiv t$, which is typically of order of the thickness an elastic membrane. In other words, in can be expected that t is of order of the molecular ($\approx$ nm) size of the constituent amphiphiles. The second length scale is the preferred pitch $P_0 = 2\pi/|\Omega_0|$, which derives from the chiral preference for local skew packing of neighbor amphiphiles (Zhang *et al.*, 2019). Unlike t, the size range of P_0 is mesoscopic, on the order of 100s of nm. With these definitions, we see that the unwinding size scale, and hence the maximum self-limiting width, are the geometric mean of these two length scales $W_{\text{un}} \simeq 11\sqrt{tP_0}$, one molecular and one mesoscopic. It can then be understood that frustration can limit the size of helicoids up to a size range intermediate to these molecular and mesoscopic sizes, that is, ribbons of order of 10s of nm in width.

As we describe in more detail in Appendix B, frustration escape in chiral ribbons is more complex than suggested by considering only helicoidal shapes. As chiral ribbons grow beyond a critical width $W_c = \sqrt{2}W_{\text{un}}$ (Fig. 17B-C), they become mechanically unstable to a new class of shape equilibria, *spiral ribbons* (Armon *et al.*, 2014; Ghafouri and Bruinsma, 2005). With increasing width, this class of shapes approaches an isometric strip that is wound helically around a cylinder. Hence, unlike the helicoid, this more complex shape relaxation allows the assembly to retain some residual chiral twist while at the same time expelling the Gaussian curvature that includes in-plane stresses. Therefore, these spiral shapes facilitate a more "efficient" escape of frustration than helicoidal unwinding (see Fig. 17D).

Notwithstanding the quantitative frustration relaxation for large widths, it can be shown that the mechanical instability does not change the basic conclusion, that the scale of the self-limiting width is set by the geometric mean of the t and P_0. That is, according to the accumulant analysis of the spiral ribbon branch (inset of Fig. 17D), $W_{\text{max}} \simeq 0.85 W_{\text{un}}$, and the instability only reduces the numerical prefactor of the unwinding scale. This is consistent with a recent experimental study of ribbon morphologies of bola-amphiphiles (Zhang *et al.*, 2019) (Figs. 17E-G). In this system, the molecular size suggests $t \approx 3 - 4$ nm, while skewed packing of chiral neighbors in the crystal leads to much larger pitches of 200 nm. The observation of helicoidal ribbons up to ≈ 50 nm (consistent with the geometric mean of molecular size and helical pitch) demonstrates that frustration stress can propagate far beyond molecular dimensions. Notably, when ribbons grow beyond this size, they do not grow to infinitely wide and untwisted sheets, but instead, transition to a second mechanism of self-limitation, forming closed tubules of finite diameter proportional to

the pitch.

The preceding discussion has neglected relaxation of frustration by defects. For 2D solid chiral ribbons this can be justified because the critical width for defect formation in helicoids [21] far exceeds the transition to a shape-flattening state, whose lower energy is facilitated by the soft bending modes of thin solids. This conclusion does not hold in general, even when the same morphologies arise, if the underlying mechanism of frustration differs. For example, chiral ribbons with only liquid-crystalline (LC) in-plane order (e.g. hexatic) would be described by the same shape (bending) elasticity, but with a frustration cost that arises from *orientational strains* from Gaussian curvature (Mbanga *et al.*, 2012; Vitelli and Turner, 2004). Such angular strains grow with a weaker power-law $\sim K_{\mathrm{G}}W$ than positional strain (Niv and Efrati, 2018), and hence, are an example of $\eta = 1$ strain growth. This "softer" growth of frustration energetics implies a shape flattening transition that takes place at a much larger size, proportional to the mesoscopic pitch $|\Omega_0|^{-1}$. This is a size scale at which disclinations may also be expected to lower the ribbon energy, implying that frustration escape for LC ribbons likely falls into a different class than solid ribbons, one that may mix both smooth (shape flattening) and singular (defect-mediated) modes.

IV. KINETIC PATHWAYS TOWARDS SELF-LIMITING EQUILIBRIUM

As the primary focus of this review is the equilibrium ingredients and thermodynamics of SLA, the foregoing sections have not considered the non-equilibrium pathways by which such systems, starting from an out-of-equilibrium initial condition, may arrive at a self-limited (equilibrium) distribution. However, the kinetics of assembly can have significant influence over the size distributions and resulting morphologies formed in experimental systems of SLA systems, since they are necessarily limited to observations at finite times. Hence, interpretation of practically all experiments must allow for the possibility of non-equilibrium effects. In this section, we provide a basic introduction to a few of the key principles for understanding and modeling kinetics of SLA. Rather than an exhaustive review, our purpose is largely to illustrate how the language and formalism introduced in the foregoing discussions of equilibrium SLA "translates" to the basic conceptual and theoretical frameworks used for analyzing these same systems out of equilibrium. A reader will find a far more thorough and general description of the interplay between kinetics and assembly products, in both bulk and SLA processes, in several

reviews (Agarwal and Peters, 2014; Hagan, 2014; Sear, 2007; Whitelam and Jack, 2015).

Theory, computation, and experiments have shown that the ability of a self-assembling system to approach equilibrium within practical timescales depends on a delicate balance between thermodynamic and kinetic effects (Ceres and Zlotnick, 2002a,b; Cheng *et al.*, 2012; Endres and Zlotnick, 2002; Grant and Jack, 2012; Grant *et al.*, 2011; Hagan, 2014; Hagan and Chandler, 2006; Hagan *et al.*, 2011; Jack *et al.*, 2007; Klotsa and Jack, 2011; Nguyen *et al.*, 2007; Rapaport, 2008; Whitelam *et al.*, 2009; Whitelam and Jack, 2015; Whitesides and Grzybowski, 2002; Wilber *et al.*, 2009, 2007; Zlotnick, 2003; Zlotnick *et al.*, 1999). In a broad variety of systems, yields are nonmonotonic with the strength of cohesive interactions that drive assembly. Optimal assembly occurs when the cohesive free energy is on the order of $5-10k_{\mathrm{B}}T$ per subunit-subunit contact depending on the initial subunit concentration and valency of subunit contacts. To understand this behavior, we require a consideration of the kinetics of assembly.

To begin, we consider a typical bulk (in vitro) assembly kinetics experiment, such as, for example, solution assembly of reconstituted proteins into multi-unit structures. To prepare an initial state, the system is equilibrated under conditions that do not lead to assembly; i.e., the total subunit concentration is below the CAC ($\Phi < \phi^*$) and thus the equilibrium aggregate distribution is peaked at monomers (see Fig. 4). In practice, this is accomplished by setting a low subunit concentration, a solution pH and salt concentration that ensure very weak subunit-subunit attractions, or sufficiently high temperature that the translational entropy of unassembled subunits dominates over inter-subunit attractions. The system is then rapidly quenched into a condition that favors assembly ($\Phi > \phi^*$), either by increasing the subunit concentration Φ, or decreasing the CAC above the current concentration, by changing temperature or physicochemical changes that increase subunit-subunit interactions (e.g. changes to pH and salt concentration). That is, the new equilibrium state (post-quench) corresponds to a population of aggregates coexisting with monomers as described in Sec. II. However, the pre-quench distribution is out-of-equilibrium, consisting predominantly of unassembled monomers, and as such gradients in the system free energy will drive assembly toward lower free energy states with assembled clusters. Note that this scenario can equally well describe experiments of both unlimited (e.g. bulk crystals) and self-limited (e.g. capsules) assembly.

The study of assembly kinetics is concerned with the timescale required to approach the equilibrium state, as well as the structures and lifetimes of long-lived metastable states which may occur along the way. To understand the influence of kinetics on practical applications or biological function, these timescales must be compared to those that are accessible to an experimenter or a biological organism.

---

[21] This can be conjectured to be $\propto |\Omega_0|^{-1}$ based on standard arguments of Gaussian curvature "screening", see e.g. (Bowick and Giomi, 2009).

A. Classical nucleation theory and assembly timescales

A useful starting point to understand the dependence of assembly timescales on control parameters is given by the framework of classical nucleation theory; see e.g. (Agarwal and Peters, 2014; Becker and Döring, 1935; Binder and Stauffer, 1976; De Yoreo and Vekilov, 2003; Hagan, 2014; Hagan and Elrad, 2010; Oxtoby, 1992; Whitelam and Jack, 2015; Zandi *et al.*, 2006) and (Oxtoby, 1992) for a review. In this approach, the assembly of a cluster is broken into two phases: *nucleation* and *growth* (often called 'elongation' in the context of finite structures). Nucleation refers to the process of overcoming a free energy barrier to form a small but relatively stable aggregate, while the second phase describes growth of such an aggregate to its final (optimal) size. We will see below that for most assembly reactions to be productive (i.e. observable on experimentally realistic timescales), nucleation must be the rate limiting process, and thus the assembly timescale can be estimated by calculating the nucleation rate.

1. Nucleation kinetics

Consider a general form of aggregation free energy for aggregates that describes assembly driven by short-range cohesive interactions, with possible additional terms to describe higher-order effects such as those that give rise to SLA:

$$\epsilon(n) = -\epsilon_{\min} + \frac{\Delta_0}{n^{1/d}} + \epsilon_{\mathrm{ex}}(n) \tag{64}$$

where $\epsilon_{\min}$ is the bulk energy per subunit in the aggregate with optimal size, Δ_0 accounts for surface energy (as in Eq. (6)) and following Sec. III.B.1 we define $\epsilon_{\mathrm{ex}}(n)$ as the excess energy relative to the bulk and surface effects, including effects that are super-extensive in n [22]. Note that this description may apply equally to either self-limiting (i.e. with a minimum at a finite $n = n_{\mathrm{T}}$ and $\epsilon_{\mathrm{T}} = -\epsilon_{\min}$ from Sec. II.B) or unlimited (i.e. where $\epsilon(n)$ is minimal for $n \to \infty$ and $\epsilon_{\mathrm{T}} = -\epsilon_{\min}$ per Sec. II.A.2) assembly. We begin the discussion by considering nucleation for the simplest case $\epsilon_{\mathrm{ex}}(n) = 0$, i.e. unlimited assembly. We will see shortly that this analysis also qualitatively applies to most SLA examples, since nucleation occurs at small sizes, where the size dependence of the excess energy, $\epsilon_{\mathrm{ex}}(n)$ is much smaller than that of the surface terms, whose large values at small sizes constitute the generic origin of the nucleation barrier.

As noted in section II.A.2, small aggregates generically have a smaller cohesive energy than large aggregates, because the fraction of subunits at the aggregate surface

with unsatisfied interactions (accounted for by the second term Δ_0 in Eq. 6) decreases with aggregate size. In contrast, the bulk energy ($\epsilon_{\min}$) dominates over surface terms for large aggregates. We define the 'critical nucleus size' n_{nuc} as the crossover between these two regimes. Below n_{nuc} disassembly is favored over assembly because the bulk cohesive energy driving assembly is outcompeted by this unfavorable surface energy and the greater translational entropy of unassembled monomers. Above the critical nucleus size, on the other hand, the bulk cohesion dominates and assembly is favored [23].

Because forward assembly of pre-nuclei aggregates is unfavorable, growth of an aggregate to the critical nucleus size is improbable and nucleation is a rare event. In particular, we will see that productive assembly requires that nucleation is a rare event on the timescale of typical subunit-subunit association. This gives rise to a separation of timescales — pre-nucleated aggregates rapidly reach a quasi-equilibrium on timescales much shorter than the overall timescale required for the assembly process to approach equilibrium. Thus, based on this assumption, the aggregation distribution for sizes below the critical nucleus size n_{nuc} can be modeled by a variant of the law of mass action

$$\phi_n = n\left(\phi_1 e^{-\beta\epsilon(n)}\right)^n = ne^{-\beta\Omega(n)} \qquad \text{for } n < n_{\mathrm{nuc}} \tag{65}$$

where $\Omega(n) = n\big[\epsilon(n) - \mu\big]$ with $\mu = -k_{\mathrm{B}}T \ln \phi_1$ is the size-dependent *grand free energy* of n-mers (sometimes referred to as the 'excess' free energy) that accounts for the inter-aggregate interactions as well as the entropy cost incurred by subunits joining an aggregate. It is important to point out that this nonequilibrium description is a departure from from the thermodynamic one introduced in Sec. II, in which μ and ϕ_1 are thermodynamically defined by the total concentration and temperature. Here, the chemical potential is used to define only the partial equilibrium of pre-nuclei with free monomer. Hence, in this usage, ϕ_1 and hence μ and Ω should be understood as time-dependent quantities according to the depletion of free monomers as assembly proceeds. However, the quasi-equilibrium approximation assumes that these quantities vary slowly in comparison to the timescale required for the pre-nuclei aggregate distribution to reach this form.

By substituting Eq. (6) for the aggregate energy into Eq. (65), we see that (for $d > 1$) there will be a maximum in the grand free energy $\Omega(n)$ at a size $n_{\mathrm{nuc}} = \left(\frac{d-1}{d}\frac{\Delta_0}{-\epsilon_{\min}-\mu}\right)^d$, owing to the competing drives of (negative) bulk assembly and (positive) surface growth. This

---

[22] Here, the notion of excess energy is shifted by an unimportant constant relative to eq. (46).

[23] Specifically, critical nucleus denotes a structure for which either complete disassembly or growth to a large aggregate are equally probable. In general, for a particular system there will be an ensemble of critical nuclei that have different structures and (if size is not a complete reaction coordinate) different sizes, see e.g. (Pan and Chandler, 2004)

corresponds to the critical nucleus size, since aggregate growth beyond n_{nuc} will decrease the free energy. Here, the factor $\frac{\Delta_0}{-\epsilon_{\min}-\mu}$ gives a ratio of the unfavorable surface energy Δ_0 that impedes assembly to the net thermodynamic driving force for assembly $-\epsilon_{\min}-\mu$. This provides a natural measure for how far out of equilibrium are initial conditions

$$\Delta\mu \equiv -\epsilon_{\min} - \mu = k_{\mathrm{B}}T \ln\left(\frac{\Phi}{\Phi_{\mathrm{s}}}\right) \geq 0 \qquad (66)$$

where we have taken $\phi_1 \simeq \Phi$ since we are considering the initial conditions where nearly all subunits are free. We have defined $\Phi_{\mathrm{s}} \equiv e^{\beta\epsilon_{\min}}$ so the that ratio Φ/Φ_{s} approximately measures how far the total subunit concentration exceeds the CAC and is often referred to as the *supersaturation*.

The key argument of classical nucleation theory is that because aggregates of the critical nucleus size are rarefied, the nucleation timescale grows exponentially with the barrier height

$$\tau_{\mathrm{nuc}} \sim e^{-\beta\Omega(n_{\mathrm{nuc}})}. \qquad (67)$$

There are several important points to make here. First, this is the *initial* nucleation timescale at the inception of the assembly; the nucleation timescale increases as assembly proceeds because free subunits are depleted and thus μ decreases, in turn, increasing $\Omega(n_{\mathrm{nuc}})$. Second, Eq. (65) only applies up to n_{nuc}; above the critical nucleus size the decreasing free energy implies that growth is relatively rapid and thus post-nuclei aggregates do not reach a quasi-equilibrium. Instead, there is a predominant flux of monomers (via association to intermediates) from the population of pre-nuclei toward larger aggregates, either toward system-sized aggregates in the case of unlimited assembly or toward a population of target-sized structures in SLA. Third, if the barrier height becomes too small (e.g. $\Omega(n_{\mathrm{nuc}}) \lesssim 10k_{\mathrm{B}}T$ depending on the relevant timescales), the separation of timescales that enabled the quasi-equilibrium approximation breaks down, and Eq. (67) will under-predict nucleation timescales since free subunits are rapidly depleted. Fourth, for SLA, additional corrections to this classical nucleation picture will arise (such as an apparent size-dependent surface tension) if the critical nucleus size approaches the finite system size due to the higher-order effects captured by the excess energy (Alder and Wainwright, 1962; Mayer and Wood, 1965; Reguera *et al.*, 2003; Thompson *et al.*, 1984). Fifth, this analysis has assumed that aggregate size n is a good 'reaction coordinate', meaning that it accounts for all relevant slow degrees of freedom and thus the dynamics and probability of an aggregate successfully nucleating can be determined as a function of n. In practice, a complete reaction coordinate must include other aggregate characteristics such as its surface area (Pan and Chandler, 2004). Sixth, a variety of other extensions to classical nucleation theory have been investigated, but it remains highly challenging to quantitatively

predict nucleation rates (Auer and Frenkel, 2001; Jacobson *et al.*, 2010; Joswiak *et al.*, 2013; Knott *et al.*, 2012; Loeffler *et al.*, 2013; Mcgraw *et al.*, 1997; Peters, 2009; Prestipino *et al.*, 2012; Statt *et al.*, 2015; Zimmermann *et al.*, 2015).

While the discussion so far has considered an aggregate energy form that drives unlimited assembly, in general the results are qualitatively similar for SLA with large target structures, since the interaction terms that eventually limit assembly grow superextensively and thus are small for small aggregates. As an example, here and throughout the remainder of this section, we consider the fluid capsid model of section III.A.1 for which self-closing leads to a finite assembly size n_{T}. To simplify the presentation, we assume the limit of high bending modulus $B \to \infty$, so the curvature radius of the assembling shell is fixed to $R_{\mathrm{T}} = \sqrt{a_0 n_{\mathrm{T}}/4\pi}$, and the free energy becomes

$$\epsilon(n) = -\epsilon_{\min} + \tilde{\lambda}\sqrt{\frac{n_{\mathrm{T}} - n}{n}} \qquad (68)$$

with the effective line tension associated with the boundaries of incomplete shells (which gives rise to the nucleation barrier) given by $\tilde{\lambda} = \sqrt{4\pi a_0/n_{\mathrm{T}}}\lambda$, with λ the bare line energy. Notice that in the limit of small n, Eq. (68) reduces to a bulk cohesive energy $\epsilon_{\min}$ and a surface term $\propto \sqrt{n}$ as anticipated above.

This specific ($B \to \infty$) model has been considered in the context of viral capsid assembly in Refs (Hagan and Elrad, 2010; Zandi *et al.*, 2006). The grand free energy that corresponds to Eq. (68) is given by

$$\Omega(n) = \Delta\mu n + \tilde{\lambda}\sqrt{n(n_{\mathrm{T}} - n)}. \qquad (69)$$

which corresponds to a barrier height (Zandi *et al.*, 2006) of

$$\Omega(n_{\mathrm{nuc}}) = \frac{n_{\mathrm{T}}\tilde{\lambda}}{2}\left(\sqrt{\Gamma^2 + 1} - \Gamma\right) \qquad (70)$$

where $\Gamma = \Delta\mu/\tilde{\lambda}$ defines a measure of the dimensionless 'quench depth'; i.e., the driving force for assembly in the initial state compared to line energy that impedes assembly (Zandi *et al.*, 2006).

Fig. 18a shows the grand free energy as a function of aggregate size for several values of the target shell size n_{T}, as well as for unlimited assembly into a flat disk, $n_{\mathrm{T}} \to \infty$. Here we have plotted the portions of the free energy profiles below(above) the critical nucleus size as solid(dashed) lines to emphasize that the quasi-equilibrium assumption only applies below the critical nucleus. That is, $\phi(n) \propto e^{-\beta\Omega(n)}$ for $n \leq n_{\mathrm{nuc}}$, but for $n > n_{\mathrm{nuc}}$ there is a constant flux of subunits toward complete shells and the nonequilibrium distribution of intermediate concentrations does not follow the law of mass action.

Notice that the free energy profiles for the self-limited cases are qualitatively similar to that of the unlimited
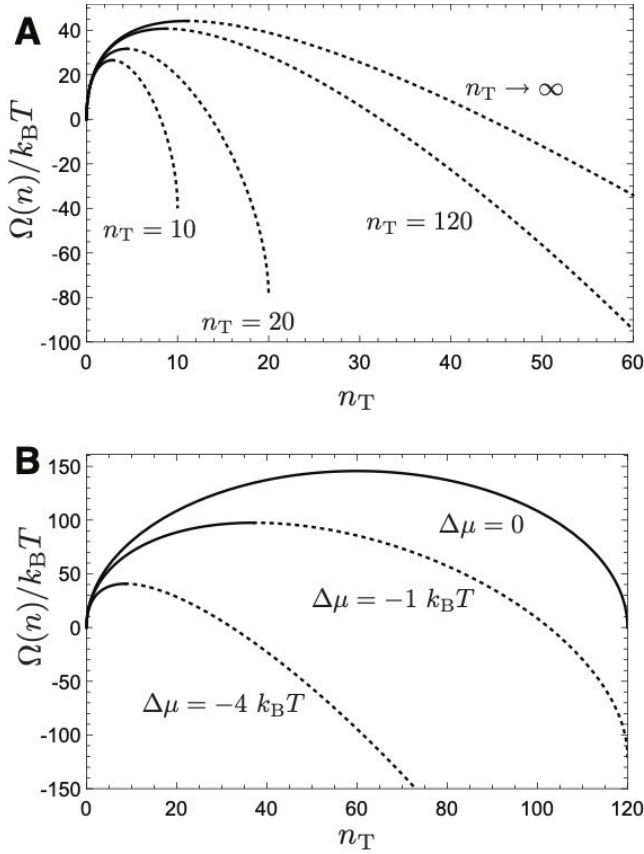
FIG. 18 The grand free energy Ω for the capsid model with fixed curvature R_T (Eq. (70)), as a function of partial capsid size n. **(A)** shows $\Omega(n)$ for indicated complete capsid sizes n_T, with the limit $n_T \to \infty$ corresponding to unlimited assembly (a flat disk). The calculation is performed for a chemical potential difference $\Delta\mu = -\epsilon_T - \mu = -4k_BT$, and a per-subunit binding free energy $\epsilon_T = -15k_BT$, corresponding to a subunit-subunit contact energy of $7.5k_BT$ and tetravalent subunits (e.g. (Ceres and Zlotnick, 2002b; Zlotnick *et al.*, 2000)). We set the line energy to $\lambda = \epsilon_T/2a_0^{1/2}$, corresponding to one unsatisfied contact per subunit on the partial shell rim. **(B)** shows $\Omega(n)$ as a function of the chemical potential difference $\Delta\mu$ for $\epsilon_T = -15k_BT$ and complete capsid size $n_T = 120$. In (A) and (B), the solid lines correspond to sizes for which the quasi-equilibrium approximation described in the text applies (i.e., $n \leq n_{\text{nuc}}$) and thus the concentration of intermediates is approximately given by $\phi(n) \propto e^{-\beta\Omega(n)}$. The dashed lines correspond to sizes $n > n_{\text{nuc}}$ for which this assumption is not valid and the intermediate concentrations cannot be described by a quasi-equilibrium (except approximately for the case $\Delta\mu = 0$).

case at small aggregate sizes, consistent with the analysis above. However, the shell geometry that gives rise to SLA does have quantitative effects — the shell curvature causes the length of the unfavorable free boundary to decrease relative to a flat disk (i.e. of equal area). Thus, the critical nucleus size and barrier height increase with target size, asymptotically approaching the flat disk limit.

Note that, while in this case the self-closing physics giving rise to SLA decreases the nucleation barrier relative to the unlimited case, SLA effects can in general shift it in either direction. For example, the strain energy that gives rise to GFA in section III.B will increase the barrier beyond the surface costs. We reiterate though that the effects of SLA on nucleation are typically quantitative rather than qualitative.

Fig. 18b shows the grand free energy as a function of aggregate size for several values of the supersaturation, at fixed target size and line tension. We see that as the supersaturation decreases, the critical nucleus size and corresponding barrier height increase, since the translational entropy of free monomers increases and thus the net assembly driving force decreases. Note that the different curves corresponding to different supersaturation levels in Fig. 18a can be viewed from two perspectives. On one hand, each supersaturation level can correspond to the initial condition of a separate experiment with a different total subunit concentration Φ, with each curve corresponding to the aggregate free energy at the beginning of the experiment. In this interpretation, Eq. 67 describes the *initial* nucleation rate and its dependence on quench conditions. On the other hand, as a single experiment proceeds, the supersaturation is continually decreasing as free monomers are depleted by assembly, and curves at decreasing supersaturation levels reveal the instantaneous nucleation rate as the experiment proceeds (and subunits are effectively removed from the pre-nuclei pool). Notice, then, that the barrier height increases as the reaction proceeds, and thus the nucleation rate decreases over time. For large target sizes $n_T \gg 1$, as the reaction proceeds the nucleation barrier eventually becomes so large in comparison to the thermal energy that assembly ceases on relevant timescales; thus, the reaction only asymptotically approaches equilibrium.

Finally, the curve corresponding to no supersaturation, $\Delta\mu = 0$, corresponds to the equilibrium state with coexistence of shells and free monomers (hence the entire curve is plotted with a solid line). In this case the critical nucleus size is given by $n_{\text{nuc}} = n_T/2$, corresponding to a half shell. Notice that at equilibrium, intermediates are higher in free energy than free monomers or complete shells due to the unfavorable line energy at the boundaries and are thus present only at low concentrations (see also section II.B.2).

2. Growth

We define growth as the process by which a critical nucleus assembles to its final state, commensurate with the optimal aggregation size. In contrast to unlimited assembly (e.g. bulk crystal growth), for SLA there is a well-defined mean timescale for growth of the aggregate, since about $n_T - n_{\text{nuc}} \approx n_T$ subunits must associate to

reach the optimal size[24].

The physical process of growth differs fundamentally from nucleation. Nucleation is a highly cooperative process — from Eq. (70), we see that the nucleation timescale decreases exponentially with subunit interaction strength $\epsilon_{\min}$ and decreases with subunit concentration according to $\phi_1^{n_{\text{nuc}}}$. The latter condition reflects the fact that n_{nuc} subunits must come together within a short timescale in order to create a stable nucleus. In contrast, because post-nucleus intermediates are relatively stable, the growth phase can proceed through independent additions of individual subunits or small oligomers. Thus, it can generally be expected that growth timescale τ_{grow} depends only weakly on $\epsilon_{\min}$. Moreover, it can be expected that τ_{grow} varies inversely with subunit concentration $\tau_{\text{grow}} \propto \phi_1^{-1}$, since the rate of addition should increase in proportion to concentration unless growth requires overcoming any secondary nucleation barriers. This latter prediction has been tested and confirmed experimentally for Hepatitis B capsid assembly by Selzer et al (Selzer *et al.*, 2014).

In general, the mean growth timescale can be estimated as

$$\tau_{\text{grow}} = \frac{n_{\text{T}}^{\alpha}}{\phi_1 f_{\text{assem}}} \tag{71}$$

where we have assumed $n_{\text{T}} \gg n_{\text{nuc}}$ so that $n_{\text{T}} - n_{\text{nuc}} \approx n_{\text{T}}$. The quantity f_{assem} is the association rate constant for subunit addition, averaged over the growth phase since it may vary with aggregate size. The factor in the numerator indicates that the growth timescale generically *increases with optimal aggregate size* (i.e. $\alpha > 0$) since $\mathcal{O}(n_{\text{T}})$ independent subunit additions must occur. The value of the exponent α will depend on factors such as the dimensionality, the aggregate geometry, and the relative stability of intermediates. For example, for growth of a globular aggregate in 3D (i.e. a crystal or the SALR system of section III.B.2), we expect $\alpha \approx 1/3$ if growth is proportional to the diffusion limited rate. For the capsid example that we consider in this section, we expect $1/2 \leq \alpha \leq 2$. In particular, if assembly is strongly biased over disassembly during growth and the assembly rate is proportional to the perimeter of the free boundary, we obtain $\alpha = 1/2$ (Hagan and Elrad, 2010). For moderately biased assembly, growth tends to occur along a single point on the perimeter, giving $\alpha = 1$, while for weakly biased assembly (near the reversible limit), the growth timescale approaches that of a random walk, giving $\alpha = 2$.

----

[24] However, even for a system with no minimum in its aggregation free energy, one can define a growth timescale to reach a given finite-size, which can be analyzed as described here.

3. Beyond nucleation and growth

It is important to note that not all assembly processes can be adequately described by the nucleation and growth mechanism. In some systems there are additional timescales which may become rate-limiting. These include subunit conformational changes, or cooperative global rearrangements required to achieve the optimal self-limited structure. For example, strains must propagate across scales on the order of the size of the structure in the frustrated open-boundary assemblies described in section III.B. Similar behaviors may occur in self-closing aggregates; for example recent evidence suggests that assembly of empty HBV capsids proceeds by nucleation, growth into large but defective or disordered intermediates, followed by a 'completion phase' in which the intermediate rearranges into the icosahedral capsid structure (Chevreuil *et al.*, 2020). Additional multi-step mechanisms are possible if we move beyond the scope of this review (SLA from a single species). For example, computation and experiments have shown that assembly around a substrate or template can proceed by an alternative 'en masse' pathway (Elrad and Hagan, 2010; Garmann *et al.*, 2014, 2016; Hagan, 2008; Panahandeh *et al.*, 2020; Perlmutter *et al.*, 2014; Tsvetkova *et al.*, 2012). In this process, subunits rapidly adsorb on a substrate in a disordered manner, and then cooperatively rearrange to form an ordered aggregate. While that process involves condensation and assembly of individual substrates, it is also possible for the assembling components to undergo bulk phase separation into a metastable liquid phase prior to assembly. For example in many virus families the host cell undergoes liquid-liquid phase separation to form a domain that is concentrated in viral proteins and nucleic acids, within which the nucleocapsid (capsid assembled around the viral RNA) assembles, e.g. (Brocca *et al.*, 2020; Carlson *et al.*, 2020; Fernández de Castro *et al.*, 2020; Guseva *et al.*, 2020; Kieser *et al.*, 2020; Nikolic *et al.*, 2017; Savastano *et al.*, 2020; Schoelz and Leisner, 2017). A related mechanism occurs for unlimited assembly — it has been shown that crystallization can proceed by a two-step mechanism in which subunits first condense into a metastable liquid phase, followed by formation of an ordered crystal, see e.g. (Basios *et al.*, 2008; Fortini *et al.*, 2008; Gliko *et al.*, 2005; Nicolis and Nicolis, 2003; Schubert *et al.*, 2017; Sear, 2009; Veesler *et al.*, 2006; Whitelam, 2010; ten Wolde and Frenkel, 1997).

B. Interplay between thermodynamic stability, assembly rates and kinetic traps

Achieving productive assembly requires thermodynamic stability of the target structure, which implies that subunit interactions must be strong enough to overcome the translational and rotational entropy losses incurred by the subunits forming an aggregate. Achieving productive assembly in finite time places even more restrictive

conditions on interactions, based on the kinetics of assembly. Interactions must be strong enough to ensure that the nucleation timescale discussed above falls within relevant timescales. However, overly strong subunit interactions lead to *kinetic traps*, or metastable states which evolve toward equilibrium very slowly. These kinetic traps can be broadly classified into two categories. We discuss each category and its effects in turn, followed by a brief survey of open questions for optimizing assembly kinetics.

1. Overnucleation (i.e. monomer starvation)

First, the *monomer starvation* trap arises when nucleation timescales are short in comparison to *growth* timescales, or the time required for a nucleated aggregate to grow to its equilibrium size. In this situation, so many nuclei form that the system becomes depleted of monomers before most nuclei grow to completion. Subsequent evolution to equilibrium requires either redistribution of subunits from smaller to larger aggregates (Ostwald ripening), which incurs significant free energy barriers, or coalescence of large intermediates, which is rare and frequently leads to mis-assembled structures.

In the context of SLA, this condition becomes more stringent as the target assembly size increases, since about n_T monomers will eventually be depleted during the growth of each nucleus, and the growth timescale typically increases with target size since the critical nucleus size depends at most weakly on n_T. The parameter regimes that give rise to the monomer starvation trap can be understood from the different dependence of nucleation and growth timescales on subunit-subunit interaction strengths and subunit concentrations (Endres and Zlotnick, 2002; Hagan and Elrad, 2010; Zlotnick *et al.*, 1999). In particular, from Eq. (67) the *initial* nucleation timescale for the capsid model is $\tau_{\mathrm{nuc}}(\Phi) \sim \exp\left[-\beta\Omega(n_{\mathrm{nuc}}(\Phi))\right]$ However, as noted above the nucleation timescale increases as the reaction proceeds due to monomer depletion. By evolving aggregates according to the kinetics described by Eq. (C1), we can integrate the cumulative depletion of monomers as a function of time, from which one can obtain median assembly time $\tau_{1/2}$ (defined as the time required for half of the subunits to be assembled, $\phi_1(\tau_{1/2}) = \Phi/2$ (Hagan and Elrad, 2010) (see Appendix C for details):

$$\tau_{1/2} \sim \frac{\tau_{\mathrm{nuc}}(\Phi)}{n_T}. \tag{72}$$

The boundary between productive assembly (which we define as having initial nucleation times shorter than, say, 1 day) and monomer starvation can be estimated by the locus in parameter space at which the median assembly time and growth time are equal. Fig. 19 shows this and boundaries between other kinetic regimes as a function of supersaturation and n_T for the capsid model described above. For any optimal size, we see that as the
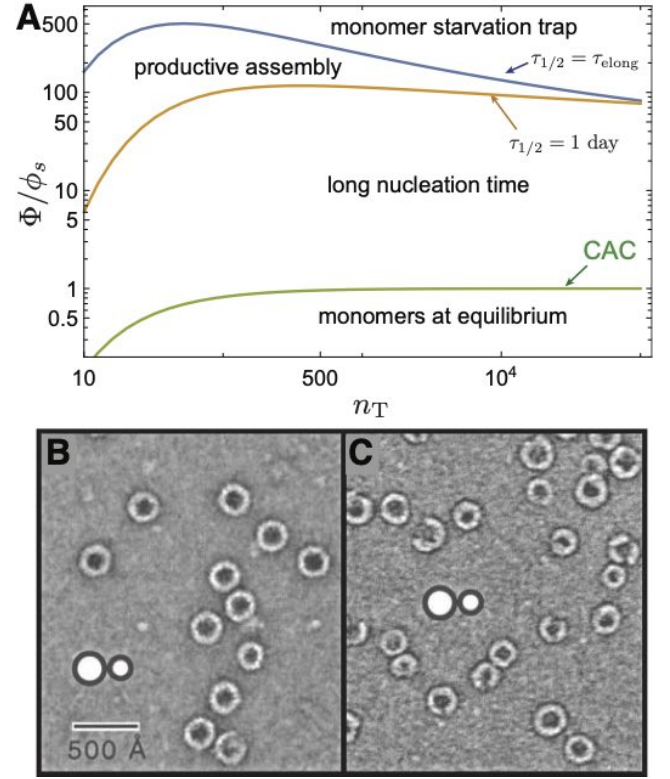

FIG. 19 **(A)** Assembly 'phase diagram' for the capsid model ($B \to \infty$ model of the fluid capsule model). The boundaries between the different kinetics regimes discussed in the text are shown as a function of capsid size and supersaturation $\Phi/\Phi_s = e^{-\beta\Delta\mu}$. The calculation was performed using Eqs. 68, 5, 71 with $\alpha = 1$, and C2. We set $\epsilon_{\mathrm{min}} = -15k_BT$ and $\lambda = \epsilon_T/2a_0^{1/2}$ as in Fig. 18. **(B), (C)** TEM images of in vitro assembly of empty capsids from CCMV capsid proteins. (B) corresponds to productive assembly, while (C) corresponds to assembly of long-lived partial shells (the monomer starvation trap) that occurs under stronger subunit-subunit interactions. Images in (B) and (C) reproduced from (Zlotnick *et al.*, 2000).

total subunit concentration increases, the system transitions from a monomer-dominated equilibrium phase, to a state in which the monomer rich phase is only metastable with respect to aggregates, but assembly does not occur on relevant timescales due to a large nucleation barrier, to a window of productive assembly, and finally to the monomer starvation regime. Notice that the concentration at which assembly becomes kinetically accessible can significantly exceed the CAC, showing that accurately inferring the CAC from experimental measurements can be challenging. Moreover, the region of productive assembly between the nucleation threshold and the monomer starvation trap narrows with increasing target size, due to the different dependencies of the nucleation and elongation timescales on n_T. In particular, the median assembly time increases with n_T a small sizes due to rim curvature as discussed above, but decreases as $1/n_T$ at large sizes because the nucleation time saturates while

each nucleus consumes n_T subunits (see Eq. (72)). Further, the growth time increases with n_T (Eq. (71) with $\alpha = 1$)[25].

2. Malformed assemblies

In the second class of kinetic traps, mis-assembled structures arise when incorrectly bound subunits do not have time to anneal before becoming trapped within the aggregate by binding of additional subunits. That is, subunit interactions must be reversible on the timescale of subunit association, even in the growth phase of aggregate assembly. This condition depends on both the cohesive interaction strength and initial monomer concentration, since the annealing time increases exponentially with interaction strength while the time interval between subunit association events decreases inversely with monomer concentration. Moreover, in the case of open-boundary assembly (and possibly some instances of curvature-controlled assembly), subunit association must occur sufficiently slowly and close to reversibility that strains due to frustration have time to propagate across the structure. Otherwise, defects or cracks that allow the system to escape frustration may occur. The conditions leading to such escape scenarios remain an open question in open-boundary assembly.

A similar trade-off between thermodynamic and kinetic considerations applies to the orientational specificity of subunit interactions. While some degree of orientational specificity is required to stabilize the target structure over other competing morphologies and to avoid mis-assembly, overly high specificity (e.g. extremely precise lock-and-key interactions) leads to small kinetic cross-sections for subunit association and thus low assembly rates (Whitelam *et al.*, 2009).

3. Non-equilibrium protocols for optimal kinetics

Finally, we briefly note that it may be possible to achieve faster assembly and/or higher assembly yields of SLA over a larger region of parameter space by using non-equilibrium protocols in which assembly driving forces are varied over time, for example by varying the temperature, solution conditions, or concentration. An intuitive approach is to use strong subunit interactions for sufficient duration to rapidly form a desired number of nuclei, and then to reduce subunit interactions to a level where post-nuclei intermediates can undergo growth but further nucleation is suppressed by a large nucleation barrier. Several approaches have been developed to use

---

[25] Note that a similar analysis can be performed for crystallization in a finite-sized system, to tune nucleation and growth rates so that a single nucleus forms on accessible timescales but grows to system-size before additional nuclei arise.

feedforward for feedback control to optimize such time varying protocols, e.g. (Green *et al.*, 2019; Grover *et al.*, 2019; Klotsa and Jack, 2013; Pineros *et al.*, 2018; Tang *et al.*, 2016b, 2017). A simpler but highly effective strategy is to seed nucleation of particular structures, e.g. (Mohammed and Schulman, 2013). This is related to the more general context of harnessing non-equilibrium assembly pathways to achieve size-controlled aggregates *out of equilibrium*, which we return to in detail in the next section.

V. FINITE SIZES BY OTHER MEANS

In Secs. II.B and III, we overviewed the statistical thermodynamics of and known physical mechanisms for identical subunits to achieve equilibrium states with well-defined and finite dimensions. Here, we briefly survey two broad classes of mechanisms for achieving size-controlled assemblies that fall outside of this basic paradigm: non-equilibrium size control and programmable assemblies multiple species.

A. Non-equilibrium mechanisms of size-controlled assembly

While Sec. IV showed that kinetic effects introduce limitations to the practical ability to achieve equilibrium SLA, in this section we consider how non-equilibrium effects can be exploited to achieve size-control of finite assemblies even when the equilibrium states are non self-limiting.

Such nonequilibrium mechanisms leading to size-controlled assembly distributions can be further classified into two categories. (i) In 'kinetically-controlled' assembly reactions, kinetic effects drive a system to a well-defined metastable state which is either sufficiently long-lived for practical applications or can be subsequently stabilized by additional reactions. Examples of such kinetically-controlled reactions include polymer/particle synthesis processes (de Pablo *et al.*, 2019), nonequilibrium formation of finite size droplets in microemulsions (Woltornist *et al.*, 2015, 2017), flow-driven aggregate breakup (Conchúir and Zaccone, 2013) and kinetically arrested coarsening (Siggia, 1979). (ii) The second class, typically referred to as 'non-equilibrium assembly' mechanisms, requires continual energy input into the system to stabilize the self-limited size distribution. These mechanisms use energy consumption to modify the aggregate-size-dependence of subunit association and/or dissociation rates. In both equilibrium and nonequilibrium mechanisms, a stable assemblage requires that association and dissociation rates are equal at a finite aggregate size. The resulting fixed point must also be stable, requiring that dissociation rates exceed association rates at larger aggregate sizes, and association rates are larger for smaller aggregates. These conditions are guaranteed in an equilibrium system by the criteria for self-limited

assembly sizes discussed in section II.B.2. However, additional nonequilibrium mechanisms to modify association or dissociation rates can also lead to stable finite sizes. For example, a number of biological mechanisms have been proposed in which assembly sizes are regulated by energy-consuming processes, such as active assembly and disassembly by molecular machines, or coupling between protein conformational states and phosphotransfer reactions (e.g. changes in a protein's phosphorylation state or hydrolysis of a bound, nucleotide triphosphate). Similarly, assembly/disassembly rates can be modulated by externally controlled gradients in monomer concentrations or nucleation factors, as occurs during embryogenesis (Briscoe and Small, 2015).

In a well-studied example, the *lengths* of microtubules or actin filaments are regulated by a suite of accessory proteins that modulate assembly and disassembly rates at filament ends, as well as molecular motors that actively remove subunits from filament ends or sever filaments in their interiors (Desai and Mitchison, 1997; Mohapatra *et al.*, 2016; Pollard, 2016). Moreover, the subunits themselves undergo conformational changes powered by hydrolysis of nucleotide triphosphates (ATP or GTP). The hydrolysis event shifts subunits into geometries that are incompatible with the global filament structure, thus inhibiting further assembly and/or weakening the existing structure. These nonequilibrium processes not only allow stable finite-length distributions, but also structural dynamics, such as actin treadmilling and microtubule dynamical instability, that allow the cellular cytoskeleton to rapidly respond and reconfigure to environmental cues. Using models ranging from idealized one-dimensional filaments to geometrically realistic particle-based dynamical simulations (Bollinger and Stevens, 2018; Fai *et al.*, 2019; Hemmat and Odde, 2020; Mohapatra *et al.*, 2016; Tong and Voth, 2020), researchers have identified multiple mechanisms by which active filament assembly/disassembly processes and energy-driven subunit conformational changes can lead to 1D filaments which exhibit dynamical instabilities and/or with well-defined stable sizes. In contrast, recall from section II.A.2 that equilibrium 1D filaments generically exhibit exponential length distributions.

Other biological structures thought to be subject to non-equilibrium size regulation include COP protein bound vesicles in the eukaryotic secretory system (Foret and Sens, 2008), neuronal synapses (Broadhead *et al.*, 2016; Burlakov *et al.*, 2012; Lisman and Raghavachari, 2006, 2015; Liu *et al.*, 2017; Miermans *et al.*, 2017; Shomar *et al.*, 2017; Tang *et al.*, 2016a), transcriptional regulatory complexes (enhancers) (Cho *et al.*, 2018; Chong *et al.*, 2018; Hnisz *et al.*, 2017; Sabari *et al.*, 2018), and other phase-separated liquid domains (Weber *et al.*, 2019; Zwicker *et al.*, 2014, 2015). The key characteristic of all of these systems is that the finite size of the assembled structure depends on continual energy consumption; e.g. through the replacement of 'inactive' subunits with 'active' ones by the disassembly of GDP bound tubulin subunits and reassembly of GTP bound tubulin at a microtubule end, or the continual dephosphorylation and re-phosphorylation of subunits by phosphatases and kinases within a neuronal synapse (Lisman and Raghavachari, 2006, 2015).

More broadly, it has been known since Turing's seminal paper (Turing, 1952) that combining imbalances in diffusion rates with interconverting molecular species (e.g. through chemical reactions) can lead to compositional inhomogeneities with well-defined, steady state sizes, e.g. (Halatek and Frey, 2018; Haselwandter *et al.*, 2011, 2015). More recently, advances in stochastic thermodynamics (Seifert, 2008, 2012) have demonstrated that active processes can play important roles in regulating the structures and functions of assembly and self-organization (Marsland and England, 2018; Nguyen and Vaikuntanathan, 2016). Moreover, theoretical and experimental studies suggest that spatiotemporal patterns with well-defined domain sizes can occur in some active matter systems, whose constituent components consume energy at the particle scale to drive motion, e.g. (Bär *et al.*, 2020; Bechinger *et al.*, 2016; Doostmohammadi *et al.*, 2018; Marchetti *et al.*, 2013; Needleman and Dogic, 2017; Shaebani *et al.*, 2020).

B. Addressable assembly of programmable subunit mixtures

In Sec. III above, we reviewed two broad categories of self-limiting assemblies, both of which achieve equilibrium finite-size assemblies from a single species of subunit. Here, we briefly describe an emerging class of self-assembling systems that also realizes equilibrium finite-size assembly, but which falls (at least partly) outside of these two categories.

We refer to this class of systems as *addressable assemblies* (AAs) following the terminology introduced in (Jacobs and Frenkel, 2016; Jacobs *et al.*, 2015). AAs are formed by mixtures of multiple assembling subunit species (say species A, B, C, D *etc.*), each with specific interactions that selectively bind to a subset of all subunit species (e.g. A binds selectively only to D, while B binds to itself as well as C). The core design principle of AAs is: one can "program" the matrix of species interactions to match the 3D adjacency matrix of a desired structure, perhaps uniquely, such that this target structure becomes the equilibrium assembly state in mixtures of controlled subunit stoichiometry (Hormoz and Brenner, 2011). That is, each particle has an "address" (or set of addresses) where it sits in the 3D target assembly.

Examples of AAs include colloids and nanoparticles functionallized by single-stranded DNA tethers that mediate interactions via complementary base pairing (Jones *et al.*, 2015), as well as "DNA bricks" assembled by mixtures of oligomeric DNA strands whose sequences are designed to interleave ends into 3D patterns via complementary base pairing (Ke *et al.*, 2012). Each of these systems has been considered and studied for its poten-

tial to selectively design and assemble target superstructures that terminate at specifically predetermined dimensions (Zeravcic *et al.*, 2014). In the sense that the self-limiting target structure can be designed to be stress-free with each cohesive bond ideally satisfied, finite AAs can be thought of as multi-species analogs to self-closing structures, albeit with more complex bond networks. As an example, rectangular beams with precisely defined and finite cross-sectional dimensions were assembled using the programmable DNA bricks (Ke *et al.*, 2014). The different values of target finite dimension each required mixtures composed of different numbers of distinct oligomeric species. For example, finite width beams of 6×6, 8×8 and 10×10 dsDNA helices across were assembled, respectively, from mixtures containing 60, 112 and 180 distinct oligomers. Generically, equilibrium termination at specific size via AA requires a number of distinct subunit species n_S that grows with the finite target size W_* (Ong *et al.*, 2017), presumably with some power law $n_S \sim W_*^{\beta}$ (e.g. $\beta = 2$ if distinct subunits are required in every 2D cross-section of a beam).

The unbounded growth of the number of subunit species with target size for such an implementation of AA would seem to limit its practical applicability, due to the cost (in terms of design, synthesis, and processing) and limited scalability of the programmable mixture. This raises a basic question about AA: Are there optimal strategies that minimize the complexity (e.g. number of distinct species) of a subunit mixture needed to achieve self-limiting AAs of a given topology, and how do these scale in the limit of large target size? The self-assembly of viral shells, or capsids, may provide clues for how to approach this question.

Capsids are "crystalline" shells composed of protein subunits (capsomers) that self-assemble to enclose the viral genome. Since the enclosed genome has to code for the capsomers themselves, it has been understood since Crick and Watson's seminal paper (Crick and Watson, 1956) that such assemblies should be *economical*. I.e., a viral capsid is under selective pressure to enclose the largest possible volume using the minimal number of distinct capsomer types. Caspar and Klug (CK) proposed that viruses achieve this optimization by exploiting symmetry principles (Caspar and Klug, 1962). In their well-known construction, quasi-spherical capsid structures are mapped to high-symmetry triangulations of the sphere, in which each triangle is constituted by 3 capsomeric subunits (Prasad and Schmid, 2012). Based on this reasoning, CK conjectured that "optimal" capsids correspond to sub-triangulations of the icosahedron. Each CK structure can be classified by the number $T = h^2 + k^2 + hk$ of sub-triangles per each of the 20 triangular faces of the icosahedron, where h and k are positive integers. Since regular triangulations of the sphere are not possible beyond those corresponding to Platonic solids, CK noted that higher triangulation numbers require capsomers to accommodate different local environments, corresponding to variations in neighbor spacing, orientation, and

the number of neighbors. In CK capsids, the triangulation number T is the number of symmetry-inequivalent capsomer positions.

In this way, we might view CK capsids as a highly-symmetric and economical limit of AA. While the complete shell could be assembled from $60T$ distinct and specifically interacting capsomeric subunits, the large number of symmetry elements of the icosahedral net implies that there are many redundancies in such a design, and in fact, the same unique target structure could be realized from specific interactions of only T distinct units. This approach has been adopted by the protein design community, enabling researchers to engineer proteins that assemble into icosahedral shells of various sizes (Bale *et al.*, 2016; Butterfield *et al.*, 2017; King *et al.*, 2014; Lai *et al.*, 2014; Mosayebi *et al.*, 2017). Recently, the CK design principles have been repurposed for the de novo design of triangular DNA origami particles with precisely defined geometry and edge interactions that selectively assemble into T icoshedral shells (Sigl *et al.*, 2020). Notably they show such capsids can be uniquely programmed and assembled from even a fewer number of distinct subunit species, $\lceil T/3 \rceil$, because each triangle can have three inequivalent edges.

In its simplest implementation, a CK capsid can be assembled by synthesizing a distinct subunit species for each of the T symmetry-equivalent positions. While some small viruses follow this approach, it becomes increasingly impractical as the target capsid size grows, since the number of distinct species scales linearly with the capsid area. In practice, there are two mechanisms to reduce the number of distinct subunits that need to be synthesized by a virus (or by any other manufacturer). In many viruses, the capsid protein interconverts between different 'quasi-equivalent' conformations with slightly different interaction geometries that accommodate the different local symmetry environments in the capsid, thus enabling assembly of capsids with $T > 1$ from a single subunit species (Caspar and Klug, 1962; Johnson and Speir, 1997). More broadly, it has been shown that such high symmetry constructions correspond to free energy minima of assembled structures with spherical topologies with relatively generic types of short-range cohesive subunit-subunit interactions, although the free energy minimum symmetry depends on the size of the assembled structure (Chen *et al.*, 2007; Zandi *et al.*, 2004). Correspondingly, CK-like capsids emerge naturally from the assembly of elastic structures from subunits resembling the tapered subunits discussed in section III.A.2 (Chen *et al.*, 2007; Fejer *et al.*, 2010; Lázaro *et al.*, 2018a,b; Reguera *et al.*, 2019), and many of the principles discussed in that section can be extended to systems that form CK capsids. However, the assembly dynamics of such structures remains an open question — how does the 'right' subunit conformation end up in the appropriate location within an assembling capsid (Berger *et al.*, 1994; Elrad and Hagan, 2008; Li *et al.*, 2018; Morton *et al.*, 2010; Panahandeh *et al.*, 2020, 2018; Perkett *et al.*,

2016; Perlmutter and Hagan, 2015; Stockley *et al.*, 2013; Twarock *et al.*, 2018; Zandi *et al.*, 2020)? More recently, the CK construction has been extended to account for other viruses in which capsomers accommodate more extreme differences in local environments and other symmetry classes (Luque and Reguera, 2010; Twarock, 2004; Twarock and Luque, 2019). In a second strategy, finite shells are assembled from a single (or few) subunit types which do not adopt explicitly distinct conformations, but are sufficiently deformable to accommodate different local symmetry environments by the formation of inhomogeneous strains within the capsid structure. It is easy to imagine that these two strategies (different conformations with specific interactions or subunits with deformable interaction geometries) could be combined to extend the size and complexity of a shell that can be assembled by the CK mechanism.

Returning to the context of AA, CK constructions and their extensions can be viewed as limiting cases of AA in which one uses symmetry to putatively minimize the number of distinct subunits types need to enclose a given volume. The CK framework suggests that there is a trade-off between the complexity of the mixture of distinct subunits and the complexity, or asymmetry, of the target assembly (i.e. fewer symmetry elements in the target structure implies more subunit species). This idea can be extended more broadly to designing self-limiting AAs that target other, non-spherical topologies. In this context, a more precise, and potentially useful notion of "economy" may be to consider the ratio of target finite size to number of subunit types W_*/n_S, and ask what are the analogs to CK designs for arbitrary topologies that maximize this ratio in the limit of increasing W_*.

VI. CONCLUDING REMARKS: SELF-LIMITING ASSEMBLY BY DISCOVERY AND DESIGN

In this review we have attempted to provide a unified theoretical perspective on assembly processes that, while occurring in chemically and physically diverse systems, share the common thread of autonomously terminating at a well-defined equilibrium finite size. Starting from the framework of ideal aggregation theory, we showed that the necessary conditions for such self-limited assembly are linked to the existence of a minimum in the size-dependent aggregation energetics $\epsilon(n)$. In contrast, systems which do not meet these conditions generically assemble bulk structures such as crystals. We saw that properties of self-limited assembly reactions can be identified from the functional form of $\epsilon(n)$, including the onset of aggregation with increasing subunit concentration and the size of fluctuations around the optimal aggregate size. Systems with multiple local minima in $\epsilon(n)$ can exhibit polymorphic self-limited assembly, with concentration-dependent transitions (secondary CACs) between different aggregation states. We saw that the small, but non-zero, translational entropy of the aggregates plays a key role in driving secondary CACs.

The existence of local minima at non-trivial sizes $n_T \gg 1$ implies that the energetics of assembly, as mediated by the shape and interactions between subunits themselves, requires the ability to "sense" the aggregate size on scales comparable to the optimal size. Surveying known examples of equilibrium SLA from identical subunits, we argued that there are two broad classes of physical mechanisms that achieve this. Either the subunit interaction geometries encode a target assembly geometry that "returns to itself" after a characteristic number of subunits, or instead, there is a source of intra-assembly stress gradients that can propagate up to the finite-size scale of the assembly. An important feature of the latter mechanism is that it enables finite-sized aggregates that have open boundaries. We presented generic descriptions for understanding each of these mechanisms, including how to assess the limits to the finite size that can be achieved. Toward that end, we considered mechanisms by which the system can "escape" finite size, resulting in bulk assembly products. For example, a system can curtail the accumulation of stress gradients with assembly size by expelling strain to its periphery or by locally relaxing strain through defects. We presented distinct physical mechanisms that can give rise to accumulating stress gradients: the interplay between short-range attractions and long-range repulsions, or an incompatibility between the preferred local subunit packing and the large-scale assembly geometry. We also presented example experimental systems that may correspond to each of these SLA mechanisms.

In part, our purpose in spotlighting the relatively rarefied conditions required for equilibrium self-limitation is to reframe broadly open challenges in understanding and engineering SLA. One such challenge could be described as the experimental inference of SLA. That is, for a given set of experimental observations of assembly, is it possible to determine if finite aggregates are the result of equilibrium self-limitation? This is a particularly vexing issue for experimental observations of both synthetic and biological assemblies that appear to be finite and well-defined. While it is often desirable to link the observations to specific microscopic models that recapitulate aspects of finite assembly *a posteriori*, such models often require assumptions about the interactions and energetics of complex subunits that are poorly understood. Thus, it is challenging to rule out alternative mechanisms of kinetic trapping of assembly in such systems. As one example, in living tissues collagen forms fibrillar assemblies that appear to have well-defined diameters reaching up to microns, well beyond the nm-scale width of a single pro-collagen molecule (Ottani *et al.*, 2002). Moreover, the mean diameter varies considerably between tissue types. Fibers in tendons have mean diameters in excess of $\gtrsim 1~\mu m$, while those found in corneal tissue have a tighter distribution around $\lesssim 50$nm (Wess, 2008). Such observations combined with the functional needs of these different tissues (Meek, 2009), high stiffness vs. optical trans-

parency respectively, suggest a need to regulate the assembly of the same subunits to form architectures of *tunable* finite size. Indeed, physical models have proposed mechanisms of geometric frustration deriving from the chiral organization within fibers as a means of imposing a self-limited diameter (Brown *et al.*, 2014; Grason and Bruinsma, 2007; Turner *et al.*, 2003; Yang *et al.*, 2010). While these explanations are plausible and broadly consistent with observations of finite-diameter collagen fibers and other fibrous biofilament structures, these physical models require knowledge of parameters describing chiral inter-molecular forces and inter-molecular mechanics, which are difficult to predict under conditions relevant to assembly (Grason, 2020). Without direct knowledge of these inter- and intra-molecular parameters, not to mention the non-equilibrium conditions of assembly, at best, such models can plausibly explain the finite size of fiber diameters. Developing generalizable experimental methodologies that strictly prove (or disprove) mechanisms of equilibrium self-limitation, particularly from complex biomolecular subunits for which the intermolecular aggregation energetics is poorly understood, remains a more distant and unmet goal.

A related challenge is to use advances in synthetic techniques to design and engineer self-limiting assemblies that target *a priori* finite dimensions. Great advances have been made in designing shape-controlled particles (Glotzer and Solomon, 2007; Sacanna and Pine, 2011) whose symmetries and interactions direct assembly to targeted structures. However, assembly targets have thus far been largely restricted to various bulk structures (albeit with complex unit cells), or 1D or 2D aggregates of uncontrolled ultimate size. Alternately, the field of "supramolecular chemistry" has leveraged chemical synthesis of an outstanding variety of architecturally- and compositionally-defined macromolecules to direct their assembly. The chemical control over these "precision amphiphiles" has significantly increased the ability to rationally design and form periodic mesophases, thermotropic supramolecular crystals, or liquid crystals (Su *et al.*, 2020). However, while the symmetries of these phases have become increasingly complex, they remain bulk structures. Synthetic advances in amphiphile assembly have largely focused on imbuing micellar assemblies with functional properties, such as controlled uptake and release of drugs (Geng *et al.*, 2007; Oltra *et al.*, 2014). While micellar assemblies are finite-size in terms of diameter, as in the case of traditional surfactants, the finite size remains limited to the size of the molecules that span the aggregate core. Thus, notwithstanding tremendous advances in synthesizing shape- and interaction-controlled subunits, controlling the finite-size of target superstructures, particularly on size scales much larger than the subunits themselves, remains a relatively unexplored aspect of engineered assemblies. While recent advances in methods such as DNA nanotechnology seem to pave the way to geometric control of subunits needed to realize bioinspired capsules and tubules (Benson *et al.*,

2015; Rothemund *et al.*, 2004; Sigl *et al.*, 2020; Tian *et al.*, 2014), it remains to be explored what the experimentally realizable upper limits to finite sizes are, and what mechanisms of self-limitation are needed to reach this limit.

A further challenge is to design supramolecular structures that do not have a single finite size but, rather, can exist in multiple different sizes, all of which are stable. Such classes of structures enable essential functions in biology. For example, recent evidence suggests that a neuronal synapse changes size during long-term memory storage, but then must remain stable at that size over the lifetime of a memory (Lisman and Raghavachari, 2006; Tang *et al.*, 2016a). Designing such variable-size stable structures is also becoming of interest to nanomaterials science, since materials capable of "learning" or "remembering" multiple stable configurations (Murugan *et al.*, 2015; Zhong *et al.*, 2017) could adapt their structures to store information, self-heal, or respond to environmental cues. Despite this interest and insights from biology, the principles underlying such variable-size stable structures remain far from clear. In this context, it would be of interest to extend the considerations in section II.B.3 of secondary CACs, to understand more broadly how the interplay between aggregate translational entropy and interaction energies can lead to controllable transitions between structures with different finite number sizes and/or dimensionalities. Similarly, can these principles be combined with the concepts of nonequilibrium assembly to design subunits that are preprogrammed to organize into nanoscale machines capable of autonomously manipulating matter or performing other functions currently found only in living organisms?

References

Achard, M. F., M. Kleman, Y. A. Nastishin, and H. T. Nguyen, 2005, European Physical Journal E **16**(1), 37, ISSN 12928941.

Agarwal, V., and B. Peters, 2014, *Solute Precipitate Nucleation: A Review of Theory and Simulation Advances* (John Wiley & Sons, Ltd), chapter 3, ISBN 9781118755815, pp. 97–160, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118755815.ch03.

Aggeli, A., I. A. Nyrkova, M. Bell, R. Harding, L. Carrick, T. C. B. Mcleish, A. N. Semenov, and N. Boden, 2001, **98**(21), 11857.

Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, 2002, *Molecular Biology of the Cell* (Garland Science, New York), fourth edition.

Alder, B. J., and T. E. Wainwright, 1962, Phys. Rev. **127**, 359, URL http://link.aps.org/doi/10.1103/PhysRev.127.359.

Armon, S., H. Aharoni, M. Moshe, and E. Sharon, 2014, Soft Matter **10**(16), 2733, ISSN 1744-683X, URL http://www.ncbi.nlm.nih.gov/pubmed/24668301http://xlink.rsc.org/?DOI=c3sm52313f.

Auer, S., and D. Frenkel, 2001, Nature **409**(6823), 1020, ISSN 0028-0836, 10.1038/35059035, URL http://dx.doi.org/10.1038/35059035.

Baker, T. S., N. H. Olson, and S. D. Fuller, 2000, Microbiology and Molecular Biology Reviews **64**(1), 237, ISSN 1092-2172.

Bale, J. B., S. Gonen, Y. Liu, W. Sheffler, D. Ellis, C. Thomas, D. Cascio, T. O. Yeates, T. Gonen, N. P. King, and D. Baker, 2016, Science **353**(6297), 389, ISSN 1095-9203 (Electronic) 0036-8075 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/27463675.

Bär, M., R. Grossmann, S. Heidenreich, and F. Peruani, 2020, Annual Review of Condensed Matter Physics **11**(1), 441, ISSN 1947-5454 1947-5462.

Basios, V., J. Lutsko, G. Nicolis, D. Maes, and C. Kirschhock, 2008, Microgravity Science and Technology **21**(1-2), 47, ISSN 0938-0108 1875-0494.

Bausch, A. R., 2003, Science **299**(5613), 1716, ISSN 00368075, URL https://www.sciencemag.org/lookup/doi/10.1126/science.1081160.

Bechinger, C., R. Di Leonardo, H. Löwen, C. Reichhardt, G. Volpe, and G. Volpe, 2016, Reviews of Modern Physics **88**(4), 045006, URL https://link.aps.org/doi/10.1103/RevModPhys.88.045006.

Becker, R., and W. Döring, 1935, Ann. Phys. (Leipzig) **416**, 719.

Ben-Shaul, A., I. Szleifer, and W. M. Gelbart, 1984, Proceedings of the National Academy of Sciences **81**(14), 4601, ISSN 0027-8424, URL http://www.pnas.org/cgi/doi/10.1073/pnas.81.14.4601.

Benson, E., A. Mohammed, J. Gardell, S. Masich, E. Czeizler, P. Orponen, and B. Högberg, 2015, Nature **523**(7561), 441, ISSN 0028-0836, URL http://www.nature.com/articles/nature14586.

Berger, B., P. W. Shor, L. Tuckerkellogg, and J. King, 1994, Proc. Natl. Acad. Sci. U. S. A. **91**(16), 7732.

Bergström, L. M., 2016, Current Opinion in Colloid & Interface Science **22**, 46, ISSN 13590294, URL http://dx.doi.org/10.1016/j.cocis.2016.02.008https://linkinghub.elsevier.com/retrieve/pii/S1359029416300127.

Bernheim-Groswasser, A., R. Zana, and Y. Talmon, 2000, The Journal of Physical Chemistry B **104**(17), 4005, ISSN 1520-6106, URL https://pubs.acs.org/doi/10.1021/jp994301a.

Binder, K., and D. Stauffer, 1976, Adv. Phys. **25**(4), 343.

Bobik, T. A., B. P. Lehman, and T. O. Yeates, 2015, Mol Microbiol **98**(2), 193, ISSN 1365-2958 (Electronic) 0950-382X (Linking), URL http://www.ncbi.nlm.nih.gov/pubmed/26148529.

Boles, M. A., M. Engel, and D. V. Talapin, 2016, Self-assembly of colloidal nanocrystals: From intricate structures to functional materials.

Bollinger, J. A., and M. J. Stevens, 2018, Soft Matter **14**, 1748, URL http://dx.doi.org/10.1039/C7SM02033C.

Bowick, M. J., and L. Giomi, 2009, Advances in Physics **58**(5), 449, ISSN 0001-8732, URL http://www.tandfonline.com/doi/abs/10.1080/00018730903043166.

Briggs, J. A. G., J. D. Riches, B. Glass, V. Bartonova, G. Zanetti, and H.-G. Krausslich, 2009, Proceedings of the National Academy of Sciences **106**(27), 11090, ISSN 0027-8424, URL http://www.pnas.org/cgi/doi/10.1073/pnas.0903535106.

Brillault, L., P. V. Jutras, N. Dashti, E. C. Thuenemann, G. Morgan, G. P. Lomonossoff, M. J. Landsberg, and F. Sainsbury, 2017, ACS Nano **11**(4), 3476, ISSN 1936-0851, URL http://pubs.acs.org/doi/abs/10.1021/acsnano.6b07747.

Briscoe, J., and S. Small, 2015, Development (Cambridge, England) **142**(26628090), 3996, ISSN 0950-1991, URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4712844/.

Broadhead, M. J., M. H. Horrocks, F. Zhu, L. Muresan, R. Benavides-Piccione, J. DeFelipe, D. Fricker, M. V. Kopanitsa, R. R. Duncan, D. Klenerman, N. H. Komiyama, S. F. Lee, *et al.*, 2016, Scientific Reports **6**, 24626.

Brocca, S., R. Grandori, S. Longhi, and V. Uversky, 2020, Int J Mol Sci **21**(23), ISSN 1422-0067 (Electronic) 1422-0067 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/33260713.

Brown, A. I., L. Kreplak, and A. D. Rutenberg, 2014, Soft Matter **10**(42), 8500, ISSN 1744-683X, URL http://xlink.rsc.org/?DOI=C4SM01359J.

Bruinsma, R. F., W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, 2003, Physical Review Letters **90**(24), 248101, ISSN 0031-9007, URL http://link.aps.org/doi/10.1103/PhysRevLett.90.248101https://link.aps.org/doi/10.1103/PhysRevLett.90.248101.

Bruinsma, R. F., and W. S. Klug, 2015, Annual Review of Condensed Matter Physics **6**(1), 245, URL https://doi.org/10.1146/annurev-conmatphys-031214-014325.

Bucher, D., F. Frey, K. A. Sochacki, S. Kummer, J. P. Bergeest, W. J. Godinez, H. G. Krausslich, K. Rohr, J. W. Taraska, U. S. Schwarz, and S. Boulant, 2018, Nat Commun **9**(1), 1109, ISSN 2041-1723 (Electronic) 2041-1723 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/29549258.

Burlakov, V. M., N. Emptage, A. Goriely, and P. C. Bressloff, 2012, Physical Review Letters **108**(2), ISSN 0031-9007, URL <GotoISI>://WOS:000298990900008.

Butterfield, G. L., M. J. Lajoie, H. H. Gustafson, D. L. Sellers, U. Nattermann, D. Ellis, J. B. Bale, S. Ke, G. H. Lenz, A. Yehdego, R. Ravichandran, S. H. Pun, *et al.*, 2017, Nature **552**(7685), 415, ISSN 1476-4687 (Electronic) 0028-0836 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/29236688.

Cameron, S., L. Kreplak, and A. D. Rutenberg, 2018, Soft Matter **14**(23), 4772, ISSN 1744-683X, URL http://xlink.rsc.org/?DOI=C8SM00377G.

Caplan, M. E., and C. J. Horowitz, 2017, Reviews of Modern Physics **89**(4), 041002, ISSN 0034-6861, URL https://link.aps.org/doi/10.1103/RevModPhys.89.041002.

Cardinaux, F., A. Stradner, P. Schurtenberger, F. Sciortino, and E. Zaccarelli, 2007, Europhysics Letters (EPL) **77**(4), 48004, ISSN 0295-5075, URL https://iopscience.iop.

org/article/10.1209/0295-5075/77/48004.

Carlson, C. R., J. B. Asfaha, C. M. Ghent, C. J. Howard, N. Hartooni, and D. O. Morgan, 2020, bioRxiv URL https://www.ncbi.nlm.nih.gov/pubmed/32637943.

Caspar, D. L. D., and A. Klug, 1962, Cold Spring Harbor Symposia on Quantitative Biology **27**, 1, ISSN 0091-7451, URL http://symposium.cshlp.org/cgi/doi/10.1101/SQB.1962.027.001.005.

Fernández de Castro, I., R. Tenorio, and C. Risco, 2020, *Virus Factories*, ISBN 9780128096338.

Ceres, P., and A. Zlotnick, 2002a, Society **4**(13), 11525.

Ceres, P., and A. Zlotnick, 2002b, Biochemistry **41**(39), 11525.

Chen, T., Z. L. Zhang, and S. C. Glotzer, 2007, Proc. Natl. Acad. Sci. U. S. A. **104**(3), 717.

Chen, Z., N. Li, S. Li, M. Dharmarwardana, A. Schlimme, and J. J. Gassensmith, 2016, Wiley Interdiscip. Rev.: Nanomed. Nanobiotechnol. **8**(4), 512, ISSN 19395116, URL http://doi.wiley.com/10.1002/wnan.1379.

Cheng, S., A. Aggarwal, and M. J. Stevens, 2012, Soft Matter **8**(20), 5666, ISSN 1744-683X, URL http://xlink.rsc.org/?DOI=c2sm25068c.

Chevreuil, M., L. Lecoq, S. Wang, L. Gargowitsch, N. Nhiri, E. Jacquet, T. Zinn, S. Fieulaine, S. Bressanelli, and G. Tresset, 2020, J. Phys. Chem. B **124**(45), 9987, ISSN 1520-6106, URL https://doi.org/10.1021/acs.jpcb.0c05024.

Cho, W.-K., J.-H. Spille, M. Hecht, C. Lee, C. Li, V. Grube, and I. I. Cisse, 2018, Science **361**(6400), 412, URL https://science.sciencemag.org/content/sci/361/6400/412.full.pdf.

Chong, S., C. Dugast-Darzacq, Z. Liu, P. Dong, G. M. Dailey, C. Cattoglio, A. Heckert, S. Banala, L. Lavis, X. Darzacq, and R. Tjian, 2018, Science **361**(6400), ISSN 1095-9203 (Electronic) 0036-8075 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/29930090.

Chowdhury, C., S. Sinha, S. Chun, T. O. Yeates, and T. A. Bobik, 2014, Microbiol Mol Biol Rev **78**(3), 438, ISSN 1098-5557 (Electronic) 1092-2172 (Linking), URL http://www.ncbi.nlm.nih.gov/pubmed/25184561.

Conchúir, B. O., and A. Zaccone, 2013, Phys. Rev. E **87**, 032310, URL https://link.aps.org/doi/10.1103/PhysRevE.87.032310.

Cozzoli, P. D., T. Pellegrino, and L. Manna, 2006, Chemical Society Reviews **35**(11), 1195, ISSN 0306-0012, URL http://xlink.rsc.org/?DOI=b517790c.

Crick, F. H. C., and J. D. Watson, 1956, Nature **177**(4506), 473.

Dang, X., H. Yi, M.-H. Ham, J. Qi, D. S. Yun, R. Ladewski, M. S. Strano, P. T. Hammond, and A. M. Belcher, 2011, Nat. Nanotechnol. **6**(6), 377.

De Yoreo, J. J., and P. G. Vekilov, 2003, Reviews in Mineralogy and Geochemistry **54**(1), 57, ISSN 1529-6466, URL https://doi.org/10.2113/0540057.

Debye, P., 1949, Annals of the New York Academy of Sciences **51**(4), 575, ISSN 00778923, URL http://doi.wiley.com/10.1111/j.1749-6632.1949.tb27293.x.

Delalande, L., I. B. Tsvetkova, C. Zeng, K. Bond, M. F. Jarrold, and B. Dragnea, 2016, Nanoscale **6**, 205, ISSN 2040-3364, URL http://xlink.rsc.org/?DOI=C6NR04469G.

Desai, A., and T. J. Mitchison, 1997, Annual Review of Cell and Developmental Biology **13**(1), 83, pMID: 9442869, URL https://doi.org/10.1146/annurev.cellbio.13.1.83.

Dharmavaram, S., S. B. She, G. Lázaro, M. F. Hagan, and R. Bruinsma, 2019, PLOS Computational Biology **15**(8), e1006602, URL https://doi.org/10.1371/journal.pcbi.1006602.

DiDonna, B. A., and R. D. Kamien, 2003, Construction , 15ISSN 1539-3755, eprint 0307149, URL http://arxiv.org/abs/cond-mat/0307149.

Dill, K. A., and P. J. Flory, 1980, Proceedings of the National Academy of Sciences **77**(6), 3115, ISSN 0027-8424, URL http://www.pnas.org/cgi/doi/10.1073/pnas.77.6.3115.

Dinsmore, A., P. Dubin, and G. Grason, 2011, The Journal of Physical Chemistry B **115**(22), 7173, ISSN 1520-6106, URL https://pubs.acs.org/doi/10.1021/jp202724b.

Doostmohammadi, A., J. Ignes-Mullol, J. M. Yeomans, and F. Sagues, 2018, Nat Commun **9**(1), 3246, ISSN 2041-1723 (Electronic) 2041-1723 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/30131558.

Dufresne, E. R., H. Noh, V. Saranathan, S. G. J. Mochrie, H. Cao, and R. O. Prum, 2009, Soft Matter **5**(9), 1792, ISSN 1744-683X, URL http://xlink.rsc.org/?DOI=b902775k.

Elrad, O., and M. F. Hagan, 2010, Phys. Biol. **7**, 045003.

Elrad, O. M., and M. F. Hagan, 2008, Nano Lett. **8**(11), 3850.

Endres, D., and A. Zlotnick, 2002, Biophysical Journal **83**(2), 1217, ISSN 00063495, URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1302223{&}tool=pmcentrez{&}rendertype=abstracthttp://linkinghub.elsevier.com/retrieve/pii/S0006349502752454.

Fai, T. G., L. Mohapatra, P. Kar, J. Kondev, A. Amir, F. Jülicher, and N. Barkai, 2019, eLife **8**, e42599, ISSN 2050-084X, URL https://doi.org/10.7554/eLife.42599.

Fejer, S. N., D. Chakrabarti, and D. J. Wales, 2010, ACS Nano **4**(1), 219.

Fisher, M. E., 1984, Journal of Statistical Physics **34**(5-6), 667, ISSN 0022-4715, URL http://link.springer.com/10.1007/BF01009436.

Foderà, V., A. Zaccone, M. Lattuada, and A. M. Donald, 2013, Phys. Rev. Lett. **111**, 108105, URL https://link.aps.org/doi/10.1103/PhysRevLett.111.108105.

Foret, L., and P. Sens, 2008, Proc. Nat. Acad. Sci. USA **105**(39), 14763.

Fortini, A., E. Sanz, and M. Dijkstra, 2008, Phys Rev E Stat Nonlin Soft Matter Phys **78**(4 Pt 1), 041402, ISSN 1539-3755 (Print) 1539-3755 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/18999422.

Fox, J. M., G. J. Wang, J. A. Speir, N. H. Olson, J. E. Johnson, T. S. Baker, and M. J. Young, 1998, Virology **244**(1), 212.

Fraenkel-Conrat, H., and R. C. Williams, 1955, Proc. Natl. Acad. Sci. U. S. A. **41**(10), 690.

Fratzl, P., 2003, Cellulose and collagen: From fibres to tissues.

Garmann, R. F., M. Comas-Garcia, A. Gopal, C. M. Knobler, and W. M. Gelbart, 2014, Journal of Molecular Biology **426**(5), 1050, ISSN 00222836, URL http://www.ncbi.nlm.nih.gov/pubmed/24148696http://linkinghub.elsevier.com/retrieve/pii/S0022283613006621.

Garmann, R. F., M. Comas-Garcia, C. M. Knobler, and W. M. Gelbart, 2016, Acc. Chem. Res. **49**(1), 48.

Gelbart, W., A. Ben-Shaul, and D. Roux, 1994, *Micelles, Membranes, Microemulsions and Monolayers* (Springer, New York).

Geng, Y., P. Dalhaimer, S. Cai, R. Tsai, M. Tewari, T. Minko,

and D. E. Discher, 2007, Nature Nanotechnology **2**(4), 249, ISSN 1748-3387, URL http://www.nature.com/articles/nnano.2007.70.

Ghafouri, R., and R. Bruinsma, 2005, Physical Review Letters **94**(13), 138101, ISSN 0031-9007, URL http://link.aps.org/doi/10.1103/PhysRevLett.94.138101https://link.aps.org/doi/10.1103/PhysRevLett.94.138101.

Giani, M., W. K. D. Otter, W. J. Briels, M. Giani, W. K. D. Otter, and W. J. Briels, 2017, J. Chem. Phys. **146**, 155102.

Gibaud, T., E. Barry, M. J. Zakhary, M. Henglin, A. Ward, Y. Yang, C. Berciu, R. Oldenbourg, M. F. Hagan, D. Nicastro, R. B. Meyer, and Z. Dogic, 2012, Nature **481**(7381), 348, ISSN 0028-0836, URL http://www.nature.com/doifinder/10.1038/nature10769.

Gliko, O., N. Neumaier, W. Pan, I. Haase, M. Fischer, A. Bacher, S. Weinkauf, and P. G. Vekilov, 2005, Journal of the American Chemical Society **127**(10), 3433, ISSN 0002-7863, URL https://doi.org/10.1021/ja043218k.

Glotzer, S. C., and M. J. Solomon, 2007, Nature Materials ISSN 14764660.

Grant, J., and R. L. Jack, 2012, Phys. Rev. E **85**(2).

Grant, J., R. L. Jack, and S. Whitelam, 2011, The Journal of Chemical Physics **135**(21), 214505, ISSN 0021-9606, URL http://aip.scitation.org/doi/10.1063/1.3662140.

Grason, G. M., 2012, Physical Review E **85**(3), 031603, ISSN 1539-3755, URL http://link.aps.org/doi/10.1103/PhysRevE.85.031603https://link.aps.org/doi/10.1103/PhysRevE.85.031603.

Grason, G. M., 2015, Reviews of Modern Physics **87**(2), 401, ISSN 0034-6861, URL https://link.aps.org/doi/10.1103/RevModPhys.87.401.

Grason, G. M., 2016, The Journal of Chemical Physics **145**(11), 110901, ISSN 0021-9606, URL http://scitation.aip.org/content/aip/journal/jcp/145/11/10.1063/1.4962629http://aip.scitation.org/doi/10.1063/1.4962629.

Grason, G. M., 2017, Nature Physics **13**(12), 1149, ISSN 1745-2473, URL http://www.nature.com/articles/nphys4201.

Grason, G. M., 2020, Soft Matter **16**(4), 1102, ISSN 1744-683X, URL http://xlink.rsc.org/?DOI=C9SM01840A.

Grason, G. M., and R. F. Bruinsma, 2007, Physical Review Letters **99**(9), 098101, ISSN 0031-9007, URL http://link.aps.org/doi/10.1103/PhysRevLett.99.098101.

Green, L. N., H. K. K. Subramanian, V. Mardanlou, J. Kim, R. F. Hariadi, and E. Franco, 2019, Nature Chemistry **11**(6), 510, ISSN 1755-4349, URL https://doi.org/10.1038/s41557-019-0251-8.

Groenewold, J., and W. K. Kegel, 2001, Journal of Physical Chemistry B **105**(47), 11702, ISSN 10895647.

Grossman, D., E. Sharon, and H. Diamant, 2016, Physical Review Letters **116**(25), 258105, ISSN 0031-9007, URL https://link.aps.org/doi/10.1103/PhysRevLett.116.258105.

Grover, M. A., D. J. Griffin, and X. Tang, 2019, IFAC-PapersOnLine **52**(1), 1, ISSN 2405-8963, URL https://www.sciencedirect.com/science/article/pii/S2405896319301144.

Guseva, S., S. Milles, M. R. Jensen, N. Salvi, J.-P. Kleman, D. Maurin, R. W. H. Ruigrok, and M. Blackledge, 2020, Science Advances **6**(14), eaaz7095, URL http://advances.sciencemag.org/content/6/14/eaaz7095.abstract.

Haddad, A., H. Aharoni, E. Sharon, A. G. Shtuken-berg, B. Kahr, and E. Efrati, 2019, Soft Matter **15**(1), 116, ISSN 1744-683X, URL http://xlink.rsc.org/?DOI=C8SM01290C.

Hagan, M. F., 2008, Phys. Rev. E **77**, 051904.

Hagan, M. F., 2014, Adv. Chem. Phys. **155**, 1.

Hagan, M. F., and D. Chandler, 2006, Biophysical Journal **91**(1), 42, ISSN 00063495, URL http://linkinghub.elsevier.com/retrieve/pii/S0006349506717064.

Hagan, M. F., and O. M. Elrad, 2010, Biophysical Journal **98**(6), 1065, ISSN 00063495, URL http://www.ncbi.nlm.nih.gov/pubmed/20303864http://linkinghub.elsevier.com/retrieve/pii/S0006349509017524.

Hagan, M. F., O. M. Elrad, and R. L. Jack, 2011, J. Chem. Phys. **135**, 104115.

Hagan, M. F., and R. Zandi, 2016, Current opinion in virology **18**, 36.

Halatek, J., and E. Frey, 2018, Nature Physics **14**(5), 507, ISSN 17452481, URL http://dx.doi.org/10.1038/s41567-017-0040-5.

Hall, D. M., I. R. Bruss, J. R. Barone, and G. M. Grason, 2016, Nature Materials **15**(7), 727, ISSN 1476-1122, URL http://www.nature.com/articles/nmat4598.

Hall, D. M., and G. M. Grason, 2017, Interface Focus **7**(4), 20160140, ISSN 2042-8898, URL http://rsfs.royalsocietypublishing.org/lookup/doi/10.1098/rsfs.2016.0140.

Halperin, A., M. Tirrell, and T. P. Lodge, 1992, in *Macromolecules: Synthesis, Order and Advanced Properties* (Springer-Verlag, Berlin/Heidelberg), pp. 31–71, URL http://link.springer.com/10.1007/BFb0051635.

Hamley, I. W., 2003, Angewandte Chemie International Edition **42**(15), 1692, ISSN 14337851, URL http://doi.wiley.com/10.1002/anie.200200546.

Harris, A. B., R. D. Kamien, and T. C. Lubensky, 1999, Reviews of Modern Physics **71**(5), 1745, ISSN 0034-6861, URL https://link.aps.org/doi/10.1103/RevModPhys.71.1745.

Haselwandter, C. A., M. Calamai, M. Kardar, A. Triller, and R. A. da Silveira, 2011, Physical review letters **106**(23), 238104.

Haselwandter, C. A., M. Kardar, A. Triller, and R. A. da Silveira, 2015, Physical Review E **92**(3), 032705.

Helfrich, W., and J. Prost, 1988, Physical Review A **38**(6), 3065, ISSN 0556-2791, URL https://link.aps.org/doi/10.1103/PhysRevA.38.3065.

Helfrich, W., 1986, J. Phys. France **47**(2), 321, URL https://doi.org/10.1051/jphys:01986004702032100.

Hemmat, M., and D. J. Odde, 2020, bioRxiv URL https://www.biorxiv.org/content/early/2020/01/08/2020.01.07.897439.

Hiemenz, P. C., and T. P. Lodge, 2007, *Polymer Chemistry* (CRC Press, Baton Rouge), second edition, ISBN 1-57444-779-3.

Hnisz, D., K. Shrinivas, R. A. Young, A. K. Chakraborty, and P. A. Sharp, 2017, Cell **169**(1), 13, ISSN 1097-4172 (Electronic) 0092-8674 (Linking), URL http://www.ncbi.nlm.nih.gov/pubmed/28340338.

Hormoz, S., and M. P. Brenner, 2011, Proceedings of the National Academy of Sciences **108**(13), 5193, ISSN 0027-8424, URL http://www.pnas.org/cgi/doi/10.1073/pnas.1014094108.

Hough, L. E., H. T. Jung, D. Krüerke, M. S. Heberling, M. Nakata, C. D. Jones, D. Chen, D. R. Link, J. Zasadzin-ski, G. Heppke, J. P. Rabe, W. Stocker, *et al.*, 2009, Sci-

ence **325**(5939), 456, ISSN 0036-8075, URL https://www.sciencemag.org/lookup/doi/10.1126/science.1170027.

Iancu, C. V., H. J. Ding, D. M. Morris, D. P. Dias, A. D. Gonzales, A. Martino, and G. J. Jensen, 2007, J. Mol. Biol. **372**(3), 764, ISSN 00222836.

Irvine, W. T. M., V. Vitelli, and P. M. Chaikin, 2010, Nature **468**(7326), 947, ISSN 0028-0836, URL http://www.nature.com/articles/nature09620.

Israelachvili, J., 2011, *Intermolecular and Surface Forces* (Elsevier), ISBN 9780123751829, eprint arXiv:1011.1669v3, URL https://linkinghub.elsevier.com/retrieve/pii/C20090215601.

Israelachvili, J. N., D. J. Mitchell, and B. W. Ninham, 1976, Journal of the Chemical Society, Faraday Transactions 2 **72**, 1525, ISSN 0300-9238, URL http://xlink.rsc.org/?DOI=f29767201525.

Jack, R. L., M. F. Hagan, and D. Chandler, 2007, Phys. Rev. E **76**, 021119.

Jacobs, W. M., and D. Frenkel, 2016, Journal of the American Chemical Society **138**(8), 2457, ISSN 0002-7863, URL https://pubs.acs.org/doi/10.1021/jacs.5b11918.

Jacobs, W. M., A. Reinhardt, and D. Frenkel, 2015, Proceedings of the National Academy of Sciences **112**(20), 6313, ISSN 0027-8424, URL http://www.pnas.org/lookup/doi/10.1073/pnas.1502210112.

Jacobson, L. C., W. Hujo, and V. Molinero, 2010, J. Am. Chem. Soc. **132**(33), 11806, ISSN 0002-7863, URL https://doi.org/10.1021/ja1051445.

Jain, S., and F. S. Bates, 2003, Science **300**(5618), 460, ISSN 00368075, URL https://www.sciencemag.org/lookup/doi/10.1126/science.1082193.

Johnson, J. E., and J. a. Speir, 1997, Journal of Molecular Biology **269**(5), 665, ISSN 00222836, URL http://www.ncbi.nlm.nih.gov/pubmed/9223631http://linkinghub.elsevier.com/retrieve/pii/S0022283697910688.

Jones, M. R., N. C. Seeman, and C. A. Mirkin, 2015, Science **347**(6224), 1260901, ISSN 0036-8075, URL https://www.sciencemag.org/lookup/doi/10.1126/science.1260901.

Joswiak, M. N., N. Duff, M. F. Doherty, and B. Peters, 2013, J. Phys. Chem. Lett. **4**(24), 4267, URL https://doi.org/10.1021/jz402226p.

Judd, J., M. L. Ho, A. Tiwari, E. J. Gomez, C. Dempsey, K. Van Vliet, O. A. Igoshin, J. J. Silberg, M. Agbandje-McKenna, and J. Suh, 2014, ACS Nano **8**(5), 4740, ISSN 1936-086X, URL http://apps.webofknowledge.com/full{_}record.do?product=UA{&}search{_}mode=OneClickSearch{&}qid=22{&}SID=1ApSBJR93ZEtQWSFNbp{&}page=1{&}doc=5.

Kang, L., and T. C. Lubensky, 2017, Proceedings of the National Academy of Sciences of the United States of America ISSN 10916490, eprint 1608.07331.

Ke, Y., L. L. Ong, W. M. Shih, and P. Yin, 2012, Science **338**(6111), 1177, ISSN 0036-8075, URL https://www.sciencemag.org/lookup/doi/10.1126/science.1227268.

Ke, Y., L. L. Ong, W. Sun, J. Song, M. Dong, W. M. Shih, and P. Yin, 2014, Nature Chemistry **6**(11), 994, ISSN 1755-4330, URL http://www.nature.com/articles/nchem.2083.

Kerfeld, C. A., S. Heinhorst, and G. C. Cannon, 2010, Microbiology **64**(1), 391.

Kerfeld, C. A., and M. R. Melnicki, 2016, Curr Opin Plant Biol **31**, 66, ISSN 1879-0356 (Electronic) 1369-5266 (Linking), URL http://www.ncbi.nlm.nih.gov/pubmed/27060669.

Kieser, Q., R. S. Noyce, M. Shenouda, Y. J. Lin, and D. H. Evans, 2020, PLoS One **15**(1), e0228028, ISSN 1932-6203 (Electronic) 1932-6203 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/31945138.

King, N. P., J. B. Bale, W. Sheffler, D. E. McNamara, S. Gonen, T. Gonen, T. O. Yeates, and D. Baker, 2014, Nature **510**(7503), 103, ISSN 1476-4687 (Electronic) 0028-0836 (Linking), URL http://www.ncbi.nlm.nih.gov/pubmed/24870237.

Kirchhausen, T., D. Owen, and S. C. Harrison, 2014, Cold Spring Harbor perspectives in biology **6**(5), a016725, ISSN 1943-0264, URL https://pubmed.ncbi.nlm.nih.gov/24789820https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3996469/.

Kléman, M., 1989, Advances in Physics **38**(6), 605, ISSN 0001-8732, URL http://www.tandfonline.com/doi/abs/10.1080/00018738900101152.

Klok, H.-A., and S. Lecommandoux, 2001, Advanced Materials **13**(16), 1217, ISSN 09359648, URL http://doi.wiley.com/10.1002/1521-4095{%}28200108{%}2913{%}3A16{%}3C1217{%}3A{%}3AAID-ADM/0.CO{%}3B2-D.

Klotsa, D., and R. L. Jack, 2011, Soft Matter **6**, 6294.

Klotsa, D., and R. L. Jack, 2013, Journal of Chemical Physics **138**(9), ISSN 0021-9606, URL <GotoISI>://WOS:000315874200039.

Knott, B. C., V. Molinero, M. F. Doherty, and B. Peters, 2012, J. Am. Chem. Soc. **134**(48), 19544, ISSN 0002-7863, URL https://doi.org/10.1021/ja309117d.

Lai, Y. T., E. Reading, G. L. Hura, K. L. Tsai, A. Laganowsky, F. J. Asturias, J. A. Tainer, C. V. Robinson, and T. O. Yeates, 2014, Nat Chem **6**(12), 1065, ISSN 1755-4349 (Electronic) 1755-4330 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/25411884.

Lázaro, G. R., B. Dragnea, and M. F. Hagan, 2018a, Soft Matter **14**(28), 5728, ISSN 1744-683X, URL http://dx.doi.org/10.1039/C8SM00129D.

Lázaro, G. R., S. Mukhopadhyay, and M. F. Hagan, 2018b, Biophys. J. **114**(3), 619, ISSN 0006-3495, URL http://dx.doi.org/10.1016/j.bpj.2017.11.3782.

Leibler, L., H. Orland, and J. C. Wheeler, 1983, The Journal of Chemical Physics **79**(7), 3550, ISSN 0021-9606, URL http://aip.scitation.org/doi/10.1063/1.446209.

Lenz, M., and T. A. Witten, 2017, Nature Physics **13**(11), 1100, ISSN 17452481.

LeRoy, H., 2018, Master's thesis, Université Paris-Saclay, France.

Li, C., A. G. Shtukenberg, L. Vogt-Maranto, E. Efrati, P. Raiteri, J. D. Gale, A. L. Rohl, and B. Kahr, 2020, The Journal of Physical Chemistry C , acs.jpcc.0c04258ISSN 1932-7447, URL https://pubs.acs.org/doi/10.1021/acs.jpcc.0c04258.

Li, F., D. P. Josephson, and A. Stein, 2011, Angewandte Chemie International Edition **50**(2), 360, ISSN 14337851, URL http://doi.wiley.com/10.1002/anie.201001451.

Li, S., P. Roy, A. Travesset, and R. Zandi, 2018, Proceedings of the National Academy of Sciences of the United States of America ISSN 10916490.

Li, S., R. Zandi, A. Travesset, and G. M. Grason, 2019, Physical Review Letters ISSN 10797114, eprint 1906.03301.

Lisman, J., and S. Raghavachari, 2006, Sci. STKE **2006**(356), re11.

Lisman, J., and S. Raghavachari, 2015, Brain research **1621**, 51.

Liu, K. K. L., M. F. Hagan, and J. E. Lisman, 2017, Philosophical transactions of the Royal Society of London. Series B, Biological sciences **372**(1715), 20160328.

Loeffler, T. D., D. E. Henderson, B. Chen, T. D. Loeffler, D. E. Henderson, and B. Chen, 2013, J. Chem. Phys. **137**, 19430.

Luque, A., and D. Reguera, 2010, Biophysical Journal **98**, 2993.

Lutomski, C., N. A. Lyktey, E. E. Pierson, Z. Zhao, A. Zlotnick, and M. F. Jarrold, 2018, J. Am. Chem. Soc. **140**(17), 5784, ISSN 0002-7863, URL https://doi.org/10.1021/jacs.8b01804.

Malyutin, A. G., R. Easterday, Y. Lozovyy, A. Spilotros, H. Cheng, O. R. Sanchez-Felix, B. D. Stein, D. G. Morgan, D. I. Svergun, B. Dragnea, and L. M. Bronstein, 2015, Chem. Mater. **27**(1), 327, ISSN 0897-4756, URL http://apps.webofknowledge.com/full{_}record.do?product=UA{&}search{_}mode=GeneralSearch{&}qid=1{&}SID=1DPMSQiX7vdAIESPdPm{&}page=1{&}doc=2.

Manoharan, V. N., 2015, Science **349**(6251), 1253751, ISSN 0036-8075.

Marchetti, M., J. Joanny, S. Ramaswamy, T. Liverpool, J. Prost, M. Rao, and R. A. Simha, 2013, Reviews of Modern Physics **85**(3), 1143.

Marsland, R., and J. L. England, 2018, Physical Review E **98**(2-1), 022411.

Mateu, M. G., 2013, Arch. Biochem. Biophys. **531**(1–2), 65, ISSN 0003-9861, URL http://www.sciencedirect.com/science/article/pii/S0003986112003839.

Matsumoto, E. A., G. P. Alexander, and R. D. Kamien, 2009, Physical Review Letters **103**(25), ISSN 00319007.

May, E. R., A. Aggarwal, W. S. Klug, and C. L. Brooks, 2011, Biophys. J. **100**(11), L59.

May, E. R., and C. L. Brooks, 2011, Phys. Rev. Lett. **106**(18), 188101.

May, S., and A. Ben-Shaul, 2001, The Journal of Physical Chemistry B **105**(3), 630, ISSN 1520-6106, URL https://pubs.acs.org/doi/10.1021/jp003021o.

Mayer, J. E., and W. W. Wood, 1965, J. Chem. Phys. **42**, 4268.

Mbanga, B. L., G. M. Grason, and C. D. Santangelo, 2012, Physical Review Letters **108**(1), 017801, ISSN 0031-9007, URL https://link.aps.org/doi/10.1103/PhysRevLett.108.017801.

Mcgraw, R., A. Laaksonen, and R. Mcgraw, 1997, J. Chem. Phys. **106**, 5284.

McPhedran, R. C., and A. R. Parker, 2015, Physics Today **68**(6), 32, ISSN 0031-9228, URL https://doi.org/10.1063/PT.3.2816.

Meek, K. M., 2009, Biophysical Reviews **1**(2), 83, ISSN 1867-2450, URL http://link.springer.com/10.1007/s12551-009-0011-x.

Meiri, S., and E. Efrati, 2021, arxiv/2101.09816 URL https://arxiv.org/abs/2101.09816.

Mendoza, C. I., and D. Reguera, 2020, eLife **9**, ISSN 2050-084X, URL https://elifesciences.org/articles/52525.

Meng, G., J. Paulose, D. R. Nelson, and V. N. Manoharan, 2014, Science **343**(6171), 634, ISSN 10959203.

Mettlen, M., P. H. Chen, S. Srinivasan, G. Danuser, and S. L. Schmid, 2018, Annu Rev Biochem **87**, 871, ISSN 1545-4509 (Electronic) 0066-4154 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/29661000.

Miermans, C. A., R. P. T. Kusters, C. C. Hoogenraad, and C. Storm, 2017, PLOS ONE **12**(2), 1, URL https://doi.org/10.1371/journal.pone.0170113.

Minten, I. J., V. I. Claessen, K. Blank, A. E. Rowan, R. J. M. Nolte, and J. J. L. M. Cornelissen, 2011, Chem. Sci. **2**(2), 358.

Missel, P. J., N. A. Mazer, G. B. Benedek, C. Y. Young, and M. C. Carey, 1980, The Journal of Physical Chemistry ISSN 00223654.

Mohammed, A. M., and R. Schulman, 2013, Nano Lett **13**(9), 4006, ISSN 1530-6992 (Electronic) 1530-6984 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/23919535.

Mohapatra, L., B. L. Goode, P. Jelenkovic, R. Phillips, and J. Kondev, 2016, Annu Rev Biophys **45**, 85, ISSN 1936-1238 (Electronic) 1936-122X (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/27145876.

Morphew, D., and D. Chakrabarti, 2017, Current Opinion in Colloid & Interface Science **30**, 70, ISSN 13590294, URL https://linkinghub.elsevier.com/retrieve/pii/S1359029417300432.

Morphew, D., J. Shaw, C. Avins, and D. Chakrabarti, 2018, ACS Nano **12**(3), 2355, ISSN 1936-0851, URL https://pubs.acs.org/doi/10.1021/acsnano.7b07633.

Morton, V. L., E. C. Dykeman, N. J. Stonehouse, A. E. Ashcroft, R. Twarock, and P. G. Stockley, 2010, J. Mol. Biol. **401**(2), 298.

Mosayebi, M., D. K. Shoemark, J. M. Fletcher, R. B. Sessions, N. Linden, D. N. Woolfson, and T. B. Liverpool, 2017, Proceedings of the National Academy of Sciences **114**(34), 9014, URL https://www.pnas.org/content/pnas/114/34/9014.full.pdf.

Murugan, A., J. Zou, and M. P. Brenner, 2015, Nature Communications **6**(1), 6203, ISSN 2041-1723, URL http://www.ncbi.nlm.nih.gov/pubmed/25669898http://www.nature.com/articles/ncomms7203.

Nagarajan, R., 2002, Langmuir **18**(1), 31, ISSN 0743-7463, URL https://pubs.acs.org/doi/10.1021/la010831y.

Nagarajan, R., and E. Ruckenstein, 1991, Langmuir **7**(12), 2934, ISSN 0743-7463, URL https://pubs.acs.org/doi/abs/10.1021/la00060a012.

Namba, K., and F. Vonderviszt, 1997, Quarterly Reviews of Biophysics **30**(1), 1, ISSN 0033-5835, URL https://www.cambridge.org/core/product/identifier/S0033583596003319/type/journal{_}article.

Needleman, D., and Z. Dogic, 2017, Nature Reviews Materials **2**(9), 17048, ISSN 2058-8437.

Nelson, D., and L. Peliti, 1987, Journal de Physique **48**(7), 1085, ISSN 0302-0738, URL http://jphys.journaldephysique.org/index.php?option=com{_}article{&}access=doi{&}doi=10.1051/jphys:019870048070108500{&}Itemid=129http://www.edpsciences.org/10.1051/jphys:019870048070108500.

Neville, A. C., 1993, *Biology of Fibrous Composites* (Cambridge University Press, Cambridge), ISBN 9780521410519, URL http://ebooks.cambridge.org/ref/id/CBO9780511601101https://www.cambridge.org/core/product/identifier/9780511601101/type/book.

Nguyen, H. D., V. S. Reddy, and C. L. Brooks, 2007, Nano Letters **7**(2), 338, ISSN 1530-6984, URL http://www.ncbi.nlm.nih.gov/pubmed/17297998http://pubs.acs.org/doi/abs/10.1021/nl062449h.

Nguyen, H. D., V. S. Reddy, and C. L. Brooks, 2009, J. Am. Chem. Soc. **131**(7), 2606.

Nguyen, M., and S. Vaikuntanathan, 2016, Proceedings of the National Academy of Sciences of the United States of

America **113**(50), 14231.

Nicolis, G., and C. Nicolis, 2003, Physica A: Statistical Mechanics and its Applications **323**, 139, ISSN 03784371.

Nikolic, J., R. Le Bars, Z. Lama, N. Scrima, C. Lagaudriere-Gesbert, Y. Gaudin, and D. Blondel, 2017, Nat Commun **8**(1), 58, ISSN 2041-1723 (Electronic) 2041-1723 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/28680096.

Niv, I., and E. Efrati, 2018, Soft Matter **14**(3), 424, ISSN 1744-683X, URL http://xlink.rsc.org/?DOI=C7SM01672G.

Nogales, E., 2000, Annual Review of Biochemistry **69**(1), 277, ISSN 0066-4154, URL http://www.annualreviews.org/doi/10.1146/annurev.biochem.69.1.277.

Nott, T. J., E. Petsalaki, P. Farber, D. Jervis, E. Fussner, A. Plochowietz, T. D. Craggs, D. P. Bazett-Jones, T. Pawson, J. D. Forman-Kay, *et al.*, 2015, Molecular cell **57**(5), 936.

O'Brien, M. N., M. R. Jones, and C. A. Mirkin, 2016, Proceedings of the National Academy of Sciences **113**(42), 11717, ISSN 0027-8424, URL http://www.pnas.org/lookup/doi/10.1073/pnas.1605289113.

Oda, R., I. Huc, M. Schmutz, S. J. Candau, and F. C. MacKintosh, 1999, Nature **399**(6736), 566, ISSN 0028-0836, URL http://www.nature.com/articles/21154.

Oltra, N. S., P. Nair, and D. E. Discher, 2014, Annual Review of Chemical and Biomolecular Engineering **5**(1), 281, ISSN 1947-5438, URL http://www.annualreviews.org/doi/10.1146/annurev-chembioeng-060713-040447.

Ong, L. L., N. Hanikel, O. K. Yaghi, C. Grun, M. T. Strauss, P. Bron, J. Lai-Kee-Him, F. Schueder, B. Wang, P. Wang, J. Y. Kishi, C. Myhrvold, *et al.*, 2017, Nature **552**(7683), 72, ISSN 0028-0836, URL http://www.nature.com/articles/nature24648.

Oosawa, F., and S. Asakura, 1975, *Thermodynamics of the Polymerization of Protein* (Academic Press, London).

Ottani, V., D. Martini, M. Franchi, A. Ruggeri, and M. Raspanti, 2002, Micron **33**(7-8), 587, ISSN 09684328, URL https://linkinghub.elsevier.com/retrieve/pii/S0968432802000331.

Oxtoby, D. W., 1992, Journal of Physics: Condensed Matter **4**(38), 7627, URL https://doi.org/10.1088%2F0953-8984%2F4%2F38%2F001.

de Pablo, J. J., N. E. Jackson, M. A. Webb, L.-Q. Chen, J. E. Moore, D. Morgan, R. Jacobs, T. Pollock, D. G. Schlom, E. S. Toberer, J. Analytis, I. Dabo, *et al.*, 2019, npj Computational Materials **5**(1), 41, ISSN 2057-3960, URL http://www.nature.com/articles/s41524-019-0173-4.

Pan, A. C., and D. Chandler, 2004, J. Phys. Chem. B **108**(51), 19681, ISSN 1520-6106, URL https://doi.org/10.1021/jp0471249.

Panahandeh, S., S. Li, L. Marichal, R. Leite Rubim, G. Tresset, and R. Zandi, 2020, ACS Nano **14**(3), 3170, ISSN 1936-0851, URL https://doi.org/10.1021/acsnano.9b08354.

Panahandeh, S., S. Li, and R. Zandi, 2018, Nanoscale **10**, 22802, URL http://dx.doi.org/10.1039/C8NR07202G.

Park, H., N. Heldman, P. Rebentrost, L. Abbondanza, A. Iagatti, A. Alessi, B. Patrizi, M. Salvalaggio, L. Bussotti, M. Mohseni, F. Caruso, H. C. Johnsen, *et al.*, 2015, Nat. Mater. **15**(2), 211, ISSN 1476-1122, URL http://www.nature.com/doifinder/10.1038/nmat4448.

Perkett, M. R., D. T. Mirijanian, and M. F. Hagan, 2016, J Chem Phys **145**(3), 035101, ISSN 1089-7690 (Electronic) 0021-9606 (Linking), URL http://www.ncbi.nlm.nih.gov/pubmed/27448905.

Perlmutter, J. D., and M. F. Hagan, 2015, Annu. Rev. Phys. Chem. **66**(1), 217, URL http://dx.doi.org/10.1146/annurev-physchem-040214-121637.

Perlmutter, J. D., M. R. Perkett, and M. F. Hagan, 2014, J. Mol. Biol. .

Peters, B., 2009, The Journal of Chemical Physics **131**(24), 244103, URL https://doi.org/10.1063/1.3271024.

Pfeifer, F., 2012, Nat Rev Micro **10**(10), 705, ISSN 1740-1526, URL http://dx.doi.org/10.1038/nrmicro2834.

Pierson, E. E., D. Z. Keifer, A. A. Kukreja, J. C. Wang, A. Zlotnick, and M. F. Jarrold, 2016, Journal of Molecular Biology **428**(2), 292, ISSN 0022-2836, URL http://dx.doi.org/10.1016/j.jmb.2015.06.019.

Pierson, E. E., D. Z. Keifer, L. Selzer, L. S. Lee, N. C. Contino, J. C. Y. Wang, A. Zlotnick, and M. F. Jarrold, 2014, J. Am. Chem. Soc. **136**(9), 3536, ISSN 0002-7863.

Pineros, W. D., B. A. Lindquist, R. B. Jadrich, and T. M. Truskett, 2018, J Chem Phys **148**(10), 104509, ISSN 1089-7690 (Electronic) 0021-9606 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/29544270.

Pollard, T. D., 2016, Cold Spring Harbor perspectives in biology **8**(26988969), a018226, ISSN 1943-0264, URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4968159/.

Popp, D., and R. C. Robinson, 2012, Cytoskeleton **69**(2), 71, ISSN 1949-3584, URL https://doi.org/10.1002/cm.21006.

Porte, G., Y. Poggi, J. Appell, and G. Maret, 1984, The Journal of Physical Chemistry **88**(23), 5713, ISSN 0022-3654, URL https://pubs.acs.org/doi/abs/10.1021/j150667a051.

Prasad, B. V. V., and M. F. Schmid, 2012, in *Viral Molecular Machines*, edited by M. G. Rossmann and V. B. Rao (Springer US, Boston, MA), ISBN 978-1-4614-0980-9, pp. 17–47, URL https://doi.org/10.1007/978-1-4614-0980-9_3.

Prestipino, S., A. Laio, and E. Tosatti, 2012, Phys. Rev. Lett. **108**, 225701, URL http://link.aps.org/doi/10.1103/PhysRevLett.108.225701.

Prum, R. O., E. R. Dufresne, T. Quinn, and K. Waters, 2009, Journal of The Royal Society Interface **6**(suppl_2), ISSN 1742-5689, URL https://royalsocietypublishing.org/doi/10.1098/rsif.2008.0466.focus.

Rae, B. D., B. M. Long, M. R. Badger, and G. D. Price, 2013, Microbiol. Mol. Biol. Rev. **77**(3), 357, URL http://mmbr.asm.org/content/77/3/357.abstract.

Rapaport, D., 2008, Phys. Rev. Lett. **101**, 186101.

Reguera, D., R. K. Bowles, Y. Djikaev, and H. Reiss, 2003, The Journal of Chemical Physics **118**(1), 340, URL https://doi.org/10.1063/1.1524192.

Reguera, D., J. Hernández-Rojas, and J. M. Gomez Llorente, 2019, Soft Matter **15**, 7166, URL http://dx.doi.org/10.1039/C9SM01593K.

Roos, W. H., R. Bruinsma, and G. J. L. Wuite, 2010, Nature Physics **6**(10), 733, ISSN 1745-2473, URL http://dx.doi.org/10.1038/nphys1797{%}5Cnhttp://www.nature.com/doifinder/10.1038/nphys1797.

Roos, W. H., and G. L. Wuite, 2009, Advanced Materials **21**(10-11), 1187.

Rothemund, P. W. K., A. Ekani-Nkodo, N. Papadakis, A. Kumar, D. K. Fygenson, and E. Winfree, 2004, Journal of the American Chemical Society **126**(50), 16344, ISSN 0002-7863, URL https://pubs.acs.org/doi/10.1021/ja0443191.

Rother, M., M. G. Nussbaumer, K. Renggli, and N. Bruns,

2016, Chem. Soc. Rev. **45**(22), 6213, ISSN 0306-0012, URL http://xlink.rsc.org/?DOI=C6CS00177G.

Royle, S. J., 2012, Journal of Cell Science **125**(1), 19, ISSN 0021-9533, URL http://jcs.biologists.org/cgi/doi/10.1242/jcs.094607.

Ruan, L., J. A. Hadden, and A. Zlotnick, 2018, Journal of Virology **92**(20), ISSN 0022-538X, URL https://jvi.asm.org/content/92/20/e01082-18.

Russel, W. B., D. A. Saville, and W. R. Schowalter, 1989, *Colloidal Dispersions* (Cambridge University Press), ISBN 9780521426008, URL https://www.cambridge.org/core/product/identifier/9780511608810/type/book.

Sabari, B. R., A. Dall'Agnese, A. Boija, I. A. Klein, E. L. Coffey, K. Shrinivas, B. J. Abraham, N. M. Hannett, A. V. Zamudio, J. C. Manteiga, C. H. Li, Y. E. Guo, *et al.*, 2018, Science **361**(6400), ISSN 1095-9203 (Electronic) 0036-8075 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/29930091.

Sacanna, S., and D. J. Pine, 2011, Current Opinion in Colloid & Interface Science **16**(2), 96, ISSN 13590294, URL https://linkinghub.elsevier.com/retrieve/pii/S1359029411000069.

Sadoc, J.-F., and R. Mosseri, 2006, *Geometrical Frustration* (Cambridge University Press, Cambridge), ISBN 0 521 44198 6.

Safran, S., 1994, *Statistical Thermodynamics of Surfaces, Interfaces, and Membranes* (Addison-Wesley Pub.).

Sakhardande, R., S. Stanojevlea, A. Baskaran, A. Baskaran, M. F. Hagan, and B. Chakraborty, 2017, Physical Review E **96**(1), 012704, URL https://link.aps.org/doi/10.1103/PhysRevE.96.012704.

Saranathan, V., J. D. Forster, H. Noh, S.-F. Liew, S. G. J. Mochrie, H. Cao, E. R. Dufresne, and R. O. Prum, 2012, Journal of The Royal Society Interface **9**(75), 2563, URL https://doi.org/10.1098/rsif.2012.0191.

Savastano, A., A. Ibanez de Opakua, M. Rankovic, and M. Zweckstetter, 2020, Nat Commun **11**(1), 6041, ISSN 2041-1723 (Electronic) 2041-1723 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/33247108.

Schmid, M. F., A. M. Paredes, H. A. Khant, F. Soyer, H. C. Aldrich, W. Chiu, and J. M. Shively, 2006, Journal of Molecular Biology **364**(3), 526, ISSN 0022-2836, URL http://www.sciencedirect.com/science/article/pii/S0022283606012150.

Schneider, S., and G. Gompper, 2005, Europhysics Letters (EPL) **70**(1), 136, ISSN 0295-5075, URL http://stacks.iop.org/0295-5075/70/i=1/a=136?key=crossref.d31bcf32083bb357bec592501d698c94https://iopscience.iop.org/article/10.1209/epl/i2004-10464-2.

Schoelz, J. E., and S. Leisner, 2017, Front Plant Sci **8**, 1832, ISSN 1664-462X (Print) 1664-462X (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/29163571.

Schubert, R., A. Meyer, D. Baitan, K. Dierks, M. Perbandt, and C. Betzel, 2017, Crystal Growth & Design **17**(3), 954, ISSN 1528-7483, URL https://doi.org/10.1021/acs.cgd.6b01826.

Sciortino, F., S. Mossa, E. Zaccarelli, and P. Tartaglia, 2004, Physical Review Letters ISSN 00319007, eprint 0312161.

Sciortino, F., P. Tartaglia, and E. Zaccarelli, 2005, The Journal of Physical Chemistry B **109**(46), 21942, ISSN 1520-6106, URL https://pubs.acs.org/doi/10.1021/jp052683g.

Sear, R. P., 2007, Journal of Physics: Condensed Matter **19**(3), 033101, URL https://doi.org/10.1088%2F0953-8984%2F19%2F3%2F033101.

Sear, R. P., 2009, J Chem Phys **131**(7), 074702, ISSN 1089-7690 (Electronic) 0021-9606 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/19708753.

Sear, R. P., and W. M. Gelbart, 1999, The Journal of Chemical Physics **110**(9), 4582, ISSN 0021-9606, URL http://aip.scitation.org/doi/10.1063/1.478338.

Sedgwick, H., S. U. Egelhaaf, and W. C. K. Poon, 2004, Journal of Physics: Condensed Matter **16**(42), S4913, ISSN 0953-8984, URL https://iopscience.iop.org/article/10.1088/0953-8984/16/42/009.

Seifert, U., 2008, The European Physical Journal B **64**(3-4), 423.

Seifert, U., 2012, Reports on progress in physics. Physical Society (Great Britain) **75**(12), 126001.

Selinger, J. V., M. S. Spector, and J. M. Schnur, 2001, The Journal of Physical Chemistry B **105**(30), 7157, ISSN 1520-6106, URL https://pubs.acs.org/doi/10.1021/jp010452d.

Selinger, R. L. B., J. V. Selinger, A. P. Malanoski, and J. M. Schnur, 2004, Physical Review Letters **93**(15), 158103, ISSN 0031-9007, URL https://link.aps.org/doi/10.1103/PhysRevLett.93.158103.

Selzer, L., S. P. Katen, and A. Zlotnick, 2014, Biochemistry **53**(34), 5496, ISSN 0006-2960, selzer, Lisa Katen, Sarah P. Zlotnick, Adam, URL <GotoISI>://WOS:000341229800004.

Seul, M., and D. Andelman, 1995, Science **267**(5197), 476, ISSN 0036-8075, URL https://www.sciencemag.org/lookup/doi/10.1126/science.267.5197.476.

Seung, H. S., and D. R. Nelson, 1988, Physical Review A **38**(2), 1005, ISSN 0556-2791, URL https://link.aps.org/doi/10.1103/PhysRevA.38.1005.

Shaebani, M. R., A. Wysocki, R. G. Winkler, G. Gompper, and H. Rieger, 2020, Nature Reviews Physics **2**(4), 181, ISSN 2522-5820.

Sharma, P., A. Ward, T. Gibaud, M. F. Hagan, and Z. Dogic, 2014, Nature **513**(7516), 77, ISSN 0028-0836, URL http://www.nature.com/articles/nature13694.

Shomar, A., L. Geyrhofer, N. E. Ziv, and N. Brenner, 2017, PLOS Comput Biol **13**(7), e1005668.

Siggia, E. D., 1979, Phys. Rev. A **20**, 595, URL https://link.aps.org/doi/10.1103/PhysRevA.20.595.

Sigl, C., E. M. Willne, W. Engelen, K. Sachenbacher, A. Liedl, F. Kolbe, F. Wilsch, S. A. Aghvami, U. Protzer, M. F. Hagan, S. Fraden, and H. Dietz, 2020, Submitted .

Statt, A., P. Virnau, and K. Binder, 2015, Phys. Rev. Lett. **114**(2), 26101, URL http://link.aps.org/doi/10.1103/PhysRevLett.114.026101.

Steinmetz, N. F., C.-F. Cho, A. Ablack, J. D. Lewis, and M. Manchester, 2011, Nanomedicine (London, U. K.) **6**(2), 351.

Stockley, P. G., N. A. Ranson, and R. Twarock, 2013, Future Virology **8**(6), 531.

Stradner, A., H. Sedgwick, F. Cardinaux, W. C. K. Poon, S. U. Egelhaaf, and P. Schurtenberger, 2004, Nature **432**(7016), 492, ISSN 0028-0836, URL http://www.nature.com/articles/nature03109.

Stupp, S. I., and L. C. Palmer, 2014, Chemistry of Materials **26**(1), 507, ISSN 0897-4756, URL https://pubs.acs.org/doi/10.1021/cm403028b.

Su, Z., R. Zhang, X.-Y. Yan, Q.-Y. Guo, J. Huang, W. Shan, Y. Liu, T. Liu, M. Huang, and S. Z. Cheng, 2020,

Progress in Polymer Science **103**, 101230, ISSN 00796700, URL https://linkinghub.elsevier.com/retrieve/pii/S007967002030023X.

Sun, J., C. DuFort, M.-C. Daniel, A. Murali, C. Chen, K. Gopinath, B. Stein, M. De, V. M. Rotello, A. Holzenburg, C. C. Kao, and B. Dragnea, 2007, Proceedings of the National Academy of Sciences **104**(4), 1354, ISSN 0027-8424, URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1783121{&}tool=pmcentrez{&}rendertype=abstracthttp://www.pnas.org/cgi/doi/10.1073/pnas.0610542104.

Sutter, M., D. Boehringer, S. Gutmann, S. Gunther, D. Prangishvili, M. J. Loessner, K. O. Stetter, E. Weber-Ban, and N. Ban, 2008, Nat. Struct. Mol. Biol. **15**(9), 939, ISSN 1545-9993, URL http://dx.doi.org/10.1038/nsmb.1473.

Tanaka, S., C. A. Kerfeld, M. R. Sawaya, F. Cai, S. Heinhorst, G. C. Cannon, and T. O. Yeates, 2008, Science **319**(5866), 1083, URL http://www.sciencemag.org/content/319/5866/1083.abstract.

Tanford, C., 1974, The Journal of Physical Chemistry **78**(24), 2469, ISSN 0022-3654, URL https://pubs.acs.org/doi/abs/10.1021/j100617a012.

Tang, A.-H., H. Chen, T. P. Li, S. R. Metzbower, H. D. MacGillavry, and T. A. Blanpied, 2016a, Nature **536**(7615), 210.

Tang, X., B. Rupp, Y. Yang, T. D. Edwards, M. A. Grover, and M. A. Bevan, 2016b, ACS Nano **10**(7), 6791, ISSN 1936-0851 1936-086X.

Tang, X., J. Zhang, M. A. Bevan, and M. A. Grover, 2017, Journal of Process Control **60**, 141, ISSN 09591524.

Terzi, M., H. LeRoy, and M. Lenz, 2020, to be published.

Thompson, S. M., K. E. Gubbins, J. P. R. B. Walton, R. A. R. Chantry, and J. S. Rowlinson, 1984, J. Chem. Phys. **81**, 530.

Tian, C., X. Li, Z. Liu, W. Jiang, G. Wang, and C. Mao, 2014, Angewandte Chemie International Edition **53**(31), 8041, ISSN 14337851, URL http://doi.wiley.com/10.1002/anie.201400377.

Tong, D., and G. A. Voth, 2020, Biophysical Journal **118**(12), 2938, ISSN 0006-3495, URL http://www.sciencedirect.com/science/article/pii/S0006349520303726.

Tsvetkova, I., C. Chen, S. Rana, C. C. Kao, V. M. Rotello, and B. Dragnea, 2012, Soft Matter **8**(17), 4571.

Turing, A. M., 1952, Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **237**(641), 37, ISSN 0962-8436, URL <GotoISI>://WOS:A1952YF90500001.

Turner, M. S., R. W. Briehl, F. A. Ferrone, and R. Josephs, 2003, Physical Review Letters **90**(12), 128103, ISSN 0031-9007, URL https://link.aps.org/doi/10.1103/PhysRevLett.90.128103.

Twarock, R., 2004, J. Theor. Biol. **226**(4), 477.

Twarock, R., R. J. Bingham, E. C. Dykeman, and P. G. Stockley, 2018, Current opinion in virology **31**(30078702), 74, ISSN 1879-6257, URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6281560/.

Twarock, R., and A. Luque, 2019, Nat Commun **10**(1), 4414, ISSN 2041-1723 (Electronic) 2041-1723 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/31562316.

Uetrecht, C., I. M. Barbu, G. K. Shoemaker, E. van Duijn, and J. R. Heck Albert, 2011, Nat Chem **3**(2), 126.

Uetrecht, C., C. Versluis, N. R. Watts, W. H. Roos, G. J. L. Wuite, P. T. Wingfield, A. C. Steven, and A. J. R. Heck, 2008, Proc. Natl. Acad. Sci. (USA) **105**(27), 9216.

Van Schooneveld, M. M., V. W. De Villeneuve, R. P. Dullens, D. G. Aarts, M. E. Leunissen, and W. K. Kegel, 2009, Journal of Physical Chemistry B ISSN 15206106.

Vannimenus, J., and G. Toulouse, 1977, Journal of Physics C: Solid State Physics **10**(18), L537, ISSN 0022-3719, URL http://stacks.iop.org/0022-3719/10/i=18/a=008?key=crossref.41b3904df55d3368826e81b583b12522.

Veesler, S., E. Revalor, O. Bottini, and C. Hoff, 2006, Organic Process Research & Development **10**(4), 841, ISSN 1083-6160, URL https://doi.org/10.1021/op060085+.

Vitelli, V., and A. M. Turner, 2004, Physical Review Letters **93**(21), 215301, ISSN 0031-9007, URL https://link.aps.org/doi/10.1103/PhysRevLett.93.215301.

Wang, C. J., S. Mukhopadhyay, and A. Zlotnick, 2018, Viruses **10**(1), ISSN 1999-4915.

Wang, J., V. Rayaprolu, S. Mukhopadhyay, and Z. A., 2015, ACS Nano **9**(9), 8898.

Weber, C. A., D. Zwicker, F. Julicher, and C. F. Lee, 2019, Rep Prog Phys **82**(6), 064601, ISSN 1361-6633 (Electronic) 0034-4885 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/30731446.

Weisel, J. W., 2004, Biophysical Chemistry **112**(2-3), 267, ISSN 03014622, URL https://linkinghub.elsevier.com/retrieve/pii/S030146220400198X.

Wess, T., 2008, in *Collagen* (Springer US, Boston, MA), ISBN 9780387739052, pp. 49–80, URL http://link.springer.com/10.1007/978-0-387-73906-9{_}3.

Whitelam, S., 2010, Phys Rev Lett **105**(8), 088102, ISSN 1079-7114 (Electronic) 0031-9007 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/20868132.

Whitelam, S., E. H. Feng, M. F. Hagan, and P. L. Geissler, 2009, Soft Matter **5**(6), 1251.

Whitelam, S., and R. L. Jack, 2015, Annual Review of Physical Chemistry **66**(1), 143, ISSN 0066-426X, URL http://www.annualreviews.org/doi/10.1146/annurev-physchem-040214-121215.

Whitesides, G. M., and B. Grzybowski, 2002, Science **295**(5564), 2418.

Wilber, A. W., J. P. K. Doye, A. A. Louis, and A. C. F. Lewis, 2009, J. Chem. Phys. **131**(17), 175102.

Wilber, A. W., J. P. K. Doye, A. a. Louis, E. G. Noya, M. a. Miller, and P. Wong, 2007, The Journal of Chemical Physics **127**(8), 085106, ISSN 0021-9606, URL http://www.ncbi.nlm.nih.gov/pubmed/17764305http://aip.scitation.org/doi/10.1063/1.2759922.

Wingfield, P. T., S. J. Stahl, R. W. Williams, and A. C. Steven, 1995, Biochemistry **34**(15), 4919.

ten Wolde, P. R., and D. Frenkel, 1997, Science **277**(5334), 1975, ISSN 0036-8075 (Print) 0036-8075 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/9302288.

Woltornist, S. J., J.-M. Y. Carrillo, T. O. Xu, A. V. Dobrynin, and D. H. Adamson, 2015, Macromolecules **48**(3), 687, ISSN 0024-9297, URL https://doi.org/10.1021/ma5024236.

Woltornist, S. J., D. Varghese, D. Massucci, Z. Cao, A. V. Dobrynin, and D. H. Adamson, 2017, Advanced Materials **29**(18), 1604947, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201604947.

Wynne, S. A., R. A. Crowther, and A. G. W. Leslie, 1999, Mol. Cell **3**(6), 771.

Yang, Y., R. B. Meyer, and M. F. Hagan, 2010, Physical Review Letters **104**(25), 258102, ISSN 0031-9007, URL https://link.aps.org/doi/10.1103/PhysRevLett.

104.258102.

Yin, Y., and A. P. Alivisatos, 2005, Nature **437**, 644, ISSN 00280836.

Young, M., W. Debbie, M. Uchida, and T. Douglas, 2008, Annu. Rev. Phytopathol. **46**(1), 361, ISSN 0066-4286, URL http://www.annualreviews.org/doi/10.1146/annurev.phyto.032508.131939.

Yu, N., A. Ghosh, and M. F. Hagan, 2016, Soft Matter **12**, 8990.

Zaccarelli, E., 2007, Journal of Physics Condensed Matter ISSN 09538984, eprint 0705.3418, URL http://arxiv.org/abs/0705.3418http://dx.doi.org/10.1088/0953-8984/19/32/323101.

Zandi, R., B. Dragnea, A. Travesset, and R. Podgornik, 2020, Physics Reports **847**, 1, ISSN 0370-1573, URL http://www.sciencedirect.com/science/article/pii/S0370157320300016.

Zandi, R., D. Reguera, R. F. Bruinsma, W. M. Gelbart, and J. Rudnick, 2004, Proceedings of the National Academy of Sciences **101**(44), 15556, ISSN 0027-8424, URL http://www.pnas.org/cgi/doi/10.1073/pnas.0405844101.

Zandi, R., P. van der Schoot, D. Reguera, W. Kegel, and H. Reiss, 2006, Biophysical Journal **90**(6), 1939, ISSN 00063495, URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1386774{&}tool=pmcentrez{&}rendertype=abstracthttp://linkinghub.elsevier.com/retrieve/pii/S0006349506723803.

Zaslavsky, B. Y., L. A. Ferreira, A. L. Darling, and V. N. Uversky, 2018, International journal of biological macromolecules **117**, 1224.

Zeravcic, Z., V. N. Manoharan, and M. P. Brenner, 2014, Proceedings of the National Academy of Sciences **111**(45), 15918, ISSN 0027-8424, URL http://www.ncbi.nlm.nih.gov/pubmed/25349380http://www.pnas.org/lookup/doi/10.1073/pnas.1411765111.

Zeravcic, Z., V. N. Manoharan, and M. P. Brenner, 2017, Reviews of Modern Physics **89**(3), 031001, ISSN 0034-6861, URL https://link.aps.org/doi/10.1103/RevModPhys.89.031001.

Zhang, L., and A. Eisenberg, 1996, Journal of the American Chemical Society **118**(13), 3168, ISSN 0002-7863, URL https://pubs.acs.org/doi/10.1021/ja953709s.

Zhang, M., D. Grossman, D. Danino, and E. Sharon, 2019, Nature Communications **10**(1), 3565, ISSN 2041-1723, URL http://dx.doi.org/10.1038/s41467-019-11473-6http://www.nature.com/articles/s41467-019-11473-6.

Zhong, W., D. J. Schwab, and A. Murugan, 2017, Journal of Statistical Physics ISSN 00224715, eprint 1701.01769.

Zhou, J., P. Kondylis, D. G. Haywood, Z. D. Harms, L. S. Lee, A. Zlotnick, and S. C. Jacobson, 2018, Anal Chem **90**(12), 7267, ISSN 1520-6882 (Electronic) 0003-2700 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/29708733.

Zhou, K., L. Li, Z. Tan, A. Zlotnick, and S. C. Jacobson, 2011, J. Am. Chem. Soc. **133**(6), 1618.

Zhuang, Y., and P. Charbonneau, 2016, The Journal of Physical Chemistry B **120**(32), 7775, ISSN 1520-6106, URL http://pubs.acs.org/doi/10.1021/acs.jpcb.6b05471.

Zimmermann, N. E. R., B. Vorselaars, D. Quigley, and B. Peters, 2015, J. Am. Chem. Soc. **137**(41), 13352, ISSN 0002-7863, URL https://doi.org/10.1021/jacs.5b08098.

Ziserman, L., A. Mor, D. Harries, and D. Danino, 2011, Physical Review Letters **106**(23), 238105, ISSN 0031-9007, URL https://link.aps.org/doi/10.1103/PhysRevLett.106.238105.

Zlotnick, A., 2003, Virology **315**(2), 269, ISSN 00426822, URL http://linkinghub.elsevier.com/retrieve/pii/S0042682203005865.

Zlotnick, A., R. Aldrich, J. M. Johnson, P. Ceres, and M. J. Young, 2000, Virology **277**(2), 450.

Zlotnick, a., J. M. Johnson, P. W. Wingfield, S. J. Stahl, and D. Endres, 1999, Biochemistry **38**(44), 14644, ISSN 0006-2960, URL http://www.ncbi.nlm.nih.gov/pubmed/10545189.

Zlotnick, A., and S. Mukhopadhyay, 2011, Trends Microbiol. **19**(1), 14.

Zwicker, D., M. Decker, S. Jaensch, A. A. Hyman, and F. Julicher, 2014, Proc Natl Acad Sci U S A **111**(26), E2636, ISSN 1091-6490 (Electronic) 0027-8424 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/24979791.

Zwicker, D., A. A. Hyman, and F. Julicher, 2015, Phys Rev E Stat Nonlin Soft Matter Phys **92**(1), 012317, ISSN 1550-2376 (Electronic) 1539-3755 (Linking), URL https://www.ncbi.nlm.nih.gov/pubmed/26274171.

Appendix A: Polymorphic amphiphile assembly phase diagram

Here we summarize the calculation of the polymorphic assembly phase diagram for the amphiphile model in Fig. 12. The aggregate energy for each of the dimensionalities is given by Eq. (44). To determine the phase diagram, we calculate the following law of mass action for subunit populations:

$$\Phi = \phi_1 + \Phi_3 + \Phi_2 + \Phi_1 \qquad (A1)$$

with $\Phi_{d_{\mathrm{L}}}$ the mass fraction of subunits in spherical, cylindrical, or planar aggregates for $d_{\mathrm{L}} = 3, 2, 1$ respectively, and ϕ_1 the free monomer population

Adopting the continuum limit, the mass fraction of subunits in spheres is given by

$$\Phi_{3,\mathrm{cont.}} = \int_0^\infty dr 4\pi r^2 n_0 \phi_3(r)$$
$$\phi_3(r) = n_{\mathrm{sph}}(r) \exp\left[-\left(\epsilon_{\mathrm{sph}}(r) - \mu\right) n_{\mathrm{sph}}(r)\right] \quad (A2)$$

with $\epsilon_{\mathrm{sph}}(r) = \epsilon(r, 3)$ from Eq. (44), $n_{\mathrm{sph}}(r) = \frac{4}{3}\pi r^3 n_0$, and $n_0 = v_0^2/a_0^3$ a dimensionless number density.

Likewise, the mass fraction in spherocylinders is

$$\Phi_2 = \int_{L_{\mathrm{B}}}^\infty dL \int_0^\infty dr 2\pi r n_0 \phi_2(r, L)$$
$$\phi_2(r, L) = (n_{\mathrm{sph}}(r) + n_{\mathrm{cyl}}(r, l)) \times$$
$$\exp\left[-\left(\epsilon_{\mathrm{SC}}(r, L) - \mu\right) (n_{\mathrm{sph}}(r) + n_{\mathrm{cyl}}(r, L))\right] \qquad (A3)$$

with L_{B} set by the minimal length of the stable spherocylinder branch (calculated below), $n_{\mathrm{cyl}}(r, L) = n_0 \pi r^2 L$, and

$$\epsilon_{\mathrm{SC}}(r, L) = \frac{n_{\mathrm{sph}}(r)\epsilon_{\mathrm{sph}}(r) + n_{\mathrm{cyl}}(r, l)\epsilon_{\mathrm{cyl}}(r)}{n_{\mathrm{sph}}(r) + n_{\mathrm{cyl}}(r, L)} \qquad (A4)$$

with $\epsilon_{\rm cyl}(r) = \epsilon(r, 2)$ from Eq. (44). Note that the first term in the numerator of Eq. (A3) corresponds to the "endcap energy" arising from the hemispherical cap at either end of the spherocylinder. We make the simplest assumption, that the radius of the hemispherical cap is equal to that of the cylindrical portion of the micelle, so that solvophobic tails remain shielded from solvent contact at the cylinder-endcap connection. More realistic models consider lower energy shapes that smoothly connect "bulbous" ends to cylindrical cores (May and Ben-Shaul, 2001).

Finally, Φ_1 is the mass fraction in layers, and is calculated below.

To proceed, recall from section II.B.2 that for the case of a minimum in the aggregate size around an optimal size $n_{\rm T}$, fluctuations vanish in the limit of large $n_{\rm T}$ or $\epsilon''|_{n_{\rm T}}$. Specifically, consider a generic aggregate energy function $\epsilon(n)$ with a minimum ϵ_* at the optimal size $n_{\rm T}$. The mass fraction of subunits in aggregates is then given by

$$\Phi_{\rm T} = \int_0^\infty dn\, n \exp[\mu - \epsilon(n)]. \qquad (A5)$$

Performing a saddle point as in section II.B.2 then results in

$$R_{\rm fluc} = \frac{\Phi_{\rm T}}{\Phi_{\rm T}(n_{\rm T})} = \sqrt{\frac{2\pi}{\epsilon''|_{n_{\rm T}} n_{\rm T}}} \qquad (A6)$$

with $\Phi_{\rm T}(n_{\rm T}) = n_{\rm T} e^{-(\epsilon_* - \mu)n_{\rm T}}$ the mass of subunits in aggregates if fluctuations are neglected. Thus, we see that when $\epsilon''|_{n_{\rm T}} n_{\rm T} \gtrsim 1$ the width of the distribution smaller than a subunit, and fluctuations are negligible. For spherocylinders, we see that the contribution due to fluctuations in the radial direction r diminishes with length as $R_{\rm fluc} \sim L^{-1/2}$. Thus, even for narrow spherocylinders, fluctuations become negligible in the large-length limit.

With this in mind, we account for polydispersity in micelles as follows. We first calculate the optimal radius for spherical aggregates, $\bar{r}_{\rm sph}$ by minimizing $\epsilon_{\rm sph}(r)$ for given values of k and P. We then numerically calculate the location of the barrier between the spherical and spherocylindrical branches of Eq. (A4), as $L_{\rm B} = \arg\max_L \left(\min_r \epsilon_{\rm SC}(r, L) \right)$; i.e., the length at which a spherocylinder of optimal radius has a maximum energy per particle. The number of particles at the barrier is then given by $n_{\rm B} = n_{\rm sph}(r_{\rm B}) + n_{\rm cyl}(r_{\rm B})$ with $r_{\rm B} = \arg\max_r \epsilon_{\rm SC}(r, L_{\rm B})$ the optimal radius at the barrier.

To calculate the mass of spherical micelles, we numerically integrate the expression in Eq. (A2). To maintain the assumptions of section III.A.2, we perform the integral over the range $n \in [\bar{n}_{\rm sph}, n_{\rm B}]$ with $\bar{n}_{\rm sph} = n_{\rm sph}(\bar{r}_{\rm sph})$, although the result is largely insensitive to increasing these integration bounds. We then include fluctuations only when they exceed the size of a single subunit by

setting

$$\Phi_3 = \max\left[\Phi_{3,{\rm cont.}}, \bar{n}_{\rm sph} e^{-(\epsilon_{\rm sph}(\bar{n}_{\rm sph}) - \mu)\bar{n}_{\rm sph}} \right]. \qquad (A7)$$

where the second argument is simply the concentration of micelles at the optimal size.

For spherocylinders we make the simplifying assumption that radial fluctuations can be neglected at all lengths, and take the optimal radius $\bar{r}_{\rm SC}(L) = \arg\max_r \epsilon_{\rm SC}(r, L)$ as a function of spherocylinder length L. In practice, we found that the numerics are more tractable if the integral is performed over particle number n rather than spherocylinder length L, so we calculate the optimal spherocylinder length $L_{\rm SC}(n) = \arg\max_L \epsilon_{\rm SC}(\bar{r}_{\rm SC}(n, L), L)$ with $\bar{r}_{\rm SC}(\hat{n}, L)$ determined from the volume of the spherocylinder: $n_{\rm sph}(\bar{r}_{\rm SC}) + n_{\rm cyl}(\bar{r}_{\rm SC}, L) = \hat{n}$. To make the integral numerically tractable, we perform the integral to a predefined (large) size $n_{\rm max}$, beyond which we assume that the effect of changing radius from the hemispherical caps is negligible, so that the optimal radius is given by the minimum of the cylinder energy, $\bar{r}_{\rm SC} \approx \bar{r}_{\rm cyl} = \arg\max_r \epsilon_{\rm cyl}(r)$, and the energetics becomes simply the 1D energetics of the form in sec. II.A.2:

$$\Phi_2 \approx \int_{n_{\rm B}}^{n_{\rm max}} dn\, \Phi_2(\bar{r}_{\rm SC}(n, L_{\rm SC}(n)), L_{\rm SC}(n)) + \Phi_\infty(n_{\rm max}) \qquad (A8)$$

with the contribution from the spherocylinders with sizes larger than $n_{\rm max}$ given by

$$\Phi_\infty(n_{\rm max}) = e^{-\epsilon_{\rm cap}} e^{(\mu - \bar{\epsilon}_{\rm cyl})n_{\rm max}} \frac{1 + (\bar{\epsilon}_{\rm cyl} - \mu)n_{\rm max}}{(\bar{\epsilon}_{\rm cyl} - \mu)^2} \qquad (A9)$$

with $\bar{\epsilon}_{\rm cyl} = \epsilon_{\rm cyl}(\bar{r}_{\rm cyl})$ the energy per particle within the cylindrical region and $\epsilon_{\rm cap} = n_{\rm sph}(\bar{r}_{\rm cyl}) \left(\epsilon_{\rm sph}(\bar{r}_{\rm cyl}) - \bar{\epsilon}_{\rm cyl} \right)$ the total *extra* energy that arises due to the unfavorable hemispherical caps.

Finally, the fraction of subunits in layers Φ_1 is calculated by noting that the free subunit chemical potential can never exceed the chemical potential of a subunit in a sheet, $\phi_1 \leq e^{\bar{\epsilon}_{\rm layer}}$ with $\bar{\epsilon}_{\rm layer} = \min_r \epsilon_{\rm layer}(r)$, where $\epsilon_{\rm layer}(r) = \epsilon(r, 1)$ from eq. (44). Thus, the amount of subunits in layers and the corresponding free subunit concentration are given by mass conservation as

$$\Phi_1 = \max\left[\Phi - \Phi_3 - \Phi_2 - e^{\bar{\epsilon}_{\rm layer}}, 0 \right]$$
$$\phi_1 = \Phi - \Phi_3 - \Phi_2 - \Phi_1, \qquad (A10)$$

effectively treating layers as an unlimited bulk phase, with negligible edge energy.

Phase boundaries in Fig. 12A are calculated from Eq. (A2) - Eq. (A10) by determining the total subunit concentration at which the fraction of subunits in an aggregate of a given dimensionality exceeds 50%.

That is, the concentrations corresponding to transitions between monomers/layers, layers/spherocylinders, and spherocylinders/spheres are calculated as the lowest total concentration at which $\Phi\left(\Phi_1 = 0.5\right)$, $\Phi\left(\Phi_2 = 0.5\right)$, and $\Phi\left(\Phi_3 = 0.5\right)$ respectively. The concentrations in Fig. 12 are normalized by the CAC for cylinders, $\phi^* = e^{\bar{\epsilon}_{\rm cyl1}}$. Note that ϕ^* is the only result within this section that depends on the cohesive energy strength ϵ_0; the relative concentrations corresponding to the transitions depend only on the elastic and boundary energy terms.

The CAC ratio shown within the spheres/spherocylinders coexistence region in Fig. 12B is computed as $\Phi\left(\Phi_3 = 0.5\right)/\Phi\left(\Phi_2 = 0.5\right)$. The infinite-concentration transitions, shown as solid red and blue lines in Fig. 12B, are calculated respectively from $\bar{\epsilon}_{\rm layer}(\bar{k}, P) = \bar{\epsilon}_{\rm cyl}(\bar{k}, P)$ and $\bar{\epsilon}_{\rm cyl}(\bar{k}, P) = \bar{\epsilon}_{\rm sph}(\bar{k}, P)$.

Finally, the boundary of the spheres/spherocylinders coexistence region (solid green line in Fig. 12B) is estimated as the minimum concentration at which either the height of the barrier between the sphere and spherocylinder branches goes to zero (i.e. corresponding to $\delta = 0$ in Eq. (34) of section II.B.3) or the point at which the transition concentration from spheres to spherocylinders becomes equal to the threshold concentration for assembling cylinders in the absence of other aggregates Eq. (35).

Appendix B: Continuum elastic theory of frustrated chiral ribbons

Here we present details of the "narrow ribbon" theory of frustrated chiral ribbons. As the model has been described elsewhere (Armon *et al.*, 2014; Ghafouri and Bruinsma, 2005; Grossman *et al.*, 2016), our primary aim is to present provide more details on the physical ingredients of the model, and further to describe how the elastic instability of wide helicoids quantitatively alters the picture of "frustration escape" presented in the Sec. III.B.3.

Following the approach of (Ghafouri and Bruinsma, 2005), we consider a simplified theory that describes the shape of ribbons in terms of the surface curvature tensor C_{ij} along the mid-line of ribbons, written in terms of coordinate directions $\hat{x}_e$ and $\hat{x}_p$ that point, respectively, along and perpendicular to the wide direction of the ribbon (see e.g. Fig. 17A). Specifically, this assumes that the in-plane curvatures vary little away from the mid-line of the ribbon, which is strictly valid when the ribbon widths are narrow with respect to their curvature radii. Since the ribbons effective "flatten" in shape as they grow wider, this narrow-ribbon approximation provides at least a qualitative picture of the ribbons' thermodynamics over the entire range of the widths.

The excess energy derives from two elastic contributions,

$$E_{\rm elast} = E_{\rm intrinsic} + E_{\rm extrinsic}, \qquad (B1)$$

where the first term depends on the intrinsic geometry,

or metric distortions away from a planar 2D lattice, while the second term describes the elasticity of the extrinsic geometry of the ribbon, i.e. , a generalized form of its bending energy expressed as quadratically in terms of curvature elements C_{ij}. The former term is captured by a 2D elastic energy,

$$E_{\rm intrinsic} = \int dA \; \sigma_{ij} u_{ij}, \qquad (B2)$$

with u_{ij} and $\sigma_{ij} \approx Y u_{ij}$ as the respective in-plane stress and strain of the 2D crystal ribbon order, and Y as the 2D Youngs modulus of the crystal (Seung and Nelson, 1988). Assuming that the crystalline packing favors uniform inter-subunit spacing, there is a geometrical and mechanical coupling between in-plane stress and out-of-plane deflections described by the so-called compatibility equation,

$$\nabla_\perp^2 \sigma_{ii} = -Y K_{\rm G}, \qquad (B3)$$

where $\sigma_{ii} = \sigma_{pp} + \sigma_{ee}$ and $K_{\rm G} \simeq C_{ee} C_{pp} - C_{ep}^2$ is the Gaussian curvature (neglecting variations of $K_{\rm G}$ across the ribbon width). Assuming approximately uniform negative curvature, it is straightforward to show that $E_{\rm intrinsic}/(WL) = Y K_{\rm G}^2 W^4/1440$ (Ghafouri and Bruinsma, 2005; Grason, 2016).

The extrinsic energy takes the form of the generalized bending energy

$$E_{\rm extrinsic} = \frac{1}{2}\int dA\left[B_{pp}C_{pp}^2 + B_{ee}C_{ee}^2 + 2B_{ep}(C_{ep}-\Omega_0)^2\right] \qquad (B4)$$

where B_{ij} are bending coefficients for different curvature elements. Here we consider the case of $B = B_{pp} = B_{ee} = B_{ep}$; relaxing this restriction does not alter the qualitative behavior. Symmetry considerations (i.e. lack of inversion symmetry) argue that chirality at the subunit scale generates a linear coupling to the off-diagonal curvature component, which we define as the ribbon twist,

$$\Omega \equiv C_{ep} \qquad (B5)$$

and hence Ω_0 can be associated with the preferred rotation, or *twist*, of the tangent plane along the edge or pitch axis (Helfrich and Prost, 1988). At a subunit scale, this preference for mesoscopic twist derives form a energetic preference for locally skewed packing of molecules in the membrane (Zhang *et al.*, 2019), although a predictive understanding of the relationship between preferred pitch and structure of constituent chiral molecules is a notoriously complex and long-standing issue in and of itself, see e.g. (Harris *et al.*, 1999).

Combining these together and dividing by the number of subunits per ribbon $n = WL/a_0$ gives the excess energy

$$\epsilon_{\rm ex}(W) = \frac{Y a_0}{1440}\left(C_{pp}C_{ee} - C_{ep}^2\right)^2 W^4$$
$$+ \frac{B a_0}{2}\left[C_{pp}^2 + C_{ee}^2 + 2(C_{ep} - \Omega_0)^2\right], \quad (B6)$$

where again, the curvature components in this expression are taken to correspond in to the ribbon shape at the mid-line, and further assumed to be constant along the length of the ribbon. Reconsidering the comparison to the heuristic description in Eq. (59), inspection of the chirality-frustrated ribbon energy in Eq. (B6) shows that Gaussian curvature plays the role of frustration strength (i.e. $f \to K_G \simeq \det(C_{ij})$), while the shape parameter can be captured by the curvature tensor (i.e. $s \to C_{ij}$) with a preferred (tensorial) shape component that is non-zero only along the off-diagonal elements.

The shape equilibrium of the ribbon varies with width from the roots of the equations

$$\frac{\partial \epsilon_{\text{ex}}}{\partial C_{ee}} = \frac{\partial \epsilon_{\text{ex}}}{\partial C_{pp}} = \frac{\partial \epsilon_{\text{ex}}}{\partial C_{ep}} = 0 \qquad \text{(B7)}$$

for fixed W. The shape equilibrium is characterized by two branches. The first is the *helicoidal* branch,

$$\left.\begin{array}{c} C_{ee} = C_{pp} = 0 \\[2mm] C_{ep} + \frac{YW^4}{720B}C_{ep}^3 = \Omega_0 \end{array}\right\} \quad \text{(helicoid).} \qquad \text{(B8)}$$

This is the branch of equilibria discussed in the main text. Helicoid twist tends to its preferred value for narrow ribbons $C_{ep}(W \ll W_{\text{un}}) \simeq \Omega_0$; and the helicoid unwinds in the wide limit as $C_{ep}(W \gg W_{\text{un}}) \simeq \Omega_0(W_{\text{un}}/W)^{4/3}$, where the *unwinding size* $W_{\text{un}} \equiv \left(720B/Y\Omega_0^2\right)^{1/4}$, defined in eq. (63, characterizes the crossover width between the these two regimes.

The second branch corresponds to a symmetry-breaking transition to a *spiral ribbon* shape,

$$\left.\begin{array}{c} C_{ee} = C_{pp} = \pm\sqrt{C_{ep}^2 - \Omega_0^2(W_{\text{un}}/W)^4} \\[2mm] C_{ep} = \Omega_0/2 \end{array}\right\} \quad \text{(spiral ribbon).}$$
$$\text{(B9)}$$

Note that this branch only exists above a critical width $W_c = \sqrt{2}W_{\text{un}}$, for which C_{ee} and C_{pp} are real.

The critical value corresponds to an elastic instability. For $W < W_c$ the helicoid branch is stable, and no spiral equilibrium exists. For $W \geq W_c$, the helicoid branch becomes unstable, and the stable branches become the two degenerate spiral states, which differ by signs of C_{ee} and C_{pp}. The shape equilibria and excess energy of both branches are plotted in Fig. 17B and C respectively.

Notice that for wide ribbons, the Gaussian curvature of both branches vanishes: for (unstable) helicoids, $K_G = -C_{ep}^2(W \gg W_{\text{un}}) \approx -\Omega_0^2(W_{\text{un}}/W)^{2/3}$; and for (stable) spiral ribbons, $K_G(W \geq W_c) = -\Omega_0^2(W_{\text{un}}/W)^4$. The difference in power law suggests a much more rapid expulsion with Gaussian curvature with increasing width of spirals. Notwithstanding the faster "frustration escape" of spiral ribbons, the accumulant analysis shown in the inset of Fig. 17C predicts that the maximum self-limiting size preempts the mechanical stability with $W_{\text{max}} = 0.85W_{\text{un}} < W_c = \sqrt{2}W_{\text{un}}$. Hence, a generic prediction of this narrow ribbon model is that self-limiting ribbons can only be helicoidal in shape (Armon *et al.*, 2014; Ghafouri and Bruinsma, 2005).

Appendix C: Assembly timescales and kinetic traps

From Eq. (67) the initial nucleation timescale for the capsid model is

$$\tau_{\text{nuc}}(\phi) \approx \frac{\exp\left[-\beta\Omega(n_{\text{nuc}}(\phi))\right]}{Z\phi f_{\text{assem}}(n_{\text{nuc}}(\phi))}. \qquad \text{(C1)}$$

with $Z = \sqrt{\frac{\beta\bar{\lambda}}{\pi n_T}}\left(1 + \Gamma^2\right)^{3/4}$ the Zeldovich factor that accounts for the time the system spends in the vicinity of the critical nucleus and $f_{\text{assem}}(n_{\text{nuc}}) \approx f_{\text{assem}}$.

The median assembly time is then given by using Eq. (C1) to integrate the cumulative depletion of monomers as a function of time, approximately accounting for reversibility of the reaction as in (Hagan and Elrad, 2010), resulting in

$$\tau_{1/2} \approx 2^{n_{\text{nuc}}(\Phi)-1}/\left(n_{\text{nuc}}(\Phi) - 1\right) f_T \frac{\tau_{\text{nuc}}(\Phi)}{n_T} \qquad \text{(C2)}$$

in which we made the approximation that the critical nucleus size remains constant over time, $n_{\text{nuc}}(\phi_1(t)) \approx n_{\text{nuc}}(\Phi) \equiv n_{\text{nuc}}(\Phi)$.