# Microcompartment assembly around multicomponent fluid cargoes

Lev Tsidilkovski and Michael F Hagan*
*Martin Fisher School of Physics, Brandeis University, Waltham, MA, USA.*

This article describes dynamical simulations of the assembly of an icosahedral protein shell around a bicomponent fluid cargo. Our simulations are motivated by bacterial microcompartments, which are protein shells found in bacteria that assemble around a complex of enzymes and other components involved in certain metabolic processes. The simulations demonstrate that the relative interaction strengths among the different cargo species play a key role in determining the amount of each species that is encapsulated, their spatial organization, and the nature of the shell assembly pathways. However, the shell protein-protein and shell-cargo interactions that help drive assembly and encapsulation also influence cargo composition within certain parameter regimes. These behaviors are governed by combination of thermodynamic and kinetic effects. In addition to elucidating how natural microcompartments encapsulate multiple components involved within reaction cascades, these results have implications for synthetic biology efforts to colocalize alternative sets of molecules within microcompartments to accelerate reactions. More broadly, the results suggest that coupling between self-assembly and multicomponent liquid-liquid phase separation may play a role in organization of the cellular cytoplasm.

## I. INTRODUCTION

Compartmentalization is essential for many biological functions, including metabolism, cellular signaling, and genetic storage. While membrane-enveloped organelles are a prominent mode of compartmentalization in eukaryotic cells, it has become apparent that liquid-liquid phase separation [1–14] and proteinaceous organelles [15–17] play an important role in organizing the cytoplasm for all kingdoms of life. For example, in bacteria, certain metabolic pathways are enabled by bacterial microcompartments, which are organelles consisting of a large protein shell assembled around a dense complex of enzymes and reactants [18–25]. Other protein-shelled compartments include encapsulins [26, 27] and gas vesicles [26, 28] in bacteria and archea, and vault particles in eukaryotes [29]. From a practical perspective, there is great interest in exploiting principles and materials from biology to achieve similar nanoscale compositional control for synthetic biology and drug delivery applications. In particular, researchers have demonstrated the ability to target new molecules to microcompartment interiors and to transfect the systems into nonnative organisms, including bacterial and plant cells, suggesting a basis for designing microcompartments as customizable nanoreactors (*e.g.*, [22, 30–43]).

Many of these cellular functions and biotechnology applications require compartments that assemble around multiple cargo species. For example, metabolism relies on achieving high local concentrations of the enzymes and reactants involved in the reaction cascade, while the formation of cellular signaling complexes requires colocalizing signaling proteins and their regulatory ligands [11, 21, 44]. Similarly, synthetic biology applications require targeting multiple species to microcompartment interiors (e.g. [22, 24–26, 30, 31, 34–43, 45–52]). In particular, microcompartments could be exploited to increase the efficiency of arbitrary multi-enzyme cascades, if the species are encapsulated with controlled stoichiometry. Thus, understanding the factors that control the amount and composition of packaged cargo is essential for understanding both natural and reengineered functions of microcompartments.

Previous experimental and modeling studies of microcompartments and other shells provide an important starting point for understanding multicomponent encapsulation. The outer shells of microcompartments are roughly icosahedral, with diameters ranging from 40-400 nm, and assemble from pentameric, hexameric, and pseudo-hexameric (trimer-of-dimer) protein oligomers [18, 19, 26, 53–59]. Recent atomic-resolution structures of small empty microcompartment shells, as well as computational modeling [60–67] have elucidated mechanisms that control the structure and size of microcompartment shells. More broadly, previous modeling studies have elucidated the assembly of empty icosahedral shells [68–94] and templating effects of a cargo consisting of a nanoparticle or RNA molecule on shell assembly[75, 81–83, 95–119]. However, alternative models are needed for microcompartment encapsulation of multicomponent cargo complexes that do not have a specific size, structure, or composition.

Although previous experiments successfully targeted particular cargo species to microcompartment interiors [22, 30–43], the factors that control cargo coalescence and encapsulation remain incompletely understood. In some microcompartments, shell-cargo attractions are mediated by 'scaffold proteins' such as the CcmN protein in $\beta$-carboxysomes[120], while in other systems core enzymes have a short 'encapsulation peptide' sequence that binds to the shell inner surface[26, 31, 35–37, 39, 121]. Similarly, cargo-cargo attractions may be mediated by scaffolds, such as CcmN in $\beta$-carboxysomes[120] or potentially through direct pair interactions between cargo molecules [36, 37].

Previous experimental and modeling studies suggest that the packaged cargo of microcompartments undergoes phase separation, either prior to or during assembly of the outer protein shell [60–62, 65, 120, 122–126]. Furthermore, some microcompartments fulfill similar functions as liquid-liquid phase separated domains in cells — for example, the compo-

* hagan@brandeis.edu

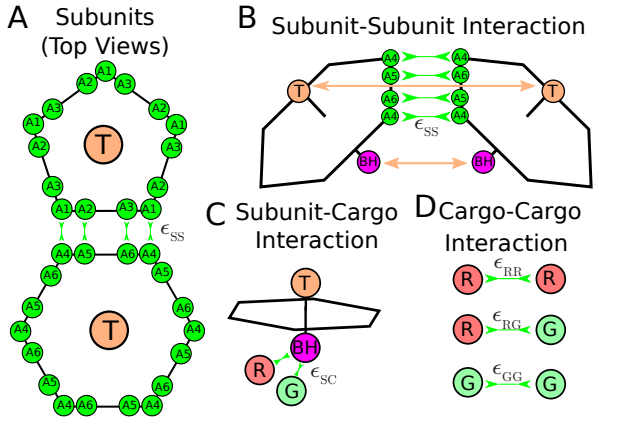


FIG. 1. **Description of the model. (A)** Each shell subunit contains 'Attractors' (green circles) on the perimeter, that define the shape of the subunit. **(B)** Interactions between Attractors drive subunit dimerization. Complementary pairs of attractors are indicated by green arrows in (A) for the pentamer-hexamer interface and in (B) for the hexamer-hexamer interface. A combination of Top-Top (T) and Bottom-Bottom (BH) repulsions controls the subunit-subunit angle in a complete shell. **(C)** Only hexamer Bottom psuedoatoms 'BH' bind cargo molecules (shown as R and G). **(D)** The cargo molecules have attractive interactions between each other that depend on pair type.

sition, structure, and function of the rubisco complex within carboxysomes has strong similarities to the pyrenoid, which is a liquid domain consisting of rubisco and other components that enables carbon fixation within plant cells [127–130]. It is therefore likely that some factors that control the composition of liquid-liquid phase separated domains in cells (e.g. [2, 6, 8, 10, 12–14, 131–145]) also affect the composition of microcompartment cargoes. However, in the case of microcompartment assembly, thermodynamic and kinetic factors resulting from coupling between liquid-liquid phase separation and assembly of the crystalline shell could have significant effects. Previous computational modeling of microcompartment assembly has focused on encapsulation of a single cargo species, driven either by direct cargo-cargo pair interactions [60, 61, 65] or mediated by scaffolds [62].

In this work, we build on these previous studies, by studying a minimal computational model for a microcompartment shell that assembles around a cargo containing two species. We investigate the amount and composition of the encapsulated cargo as a function of the affinities among these cargo species, as well as shell-cargo and shell-shell affinities. The simulation results suggest that the relative cargo-cargo affinities are the most important factor in determining the amount and composition of packaged cargo. However, within certain parameter ranges, varying the shell-cargo and shell-shell interactions provides an alternative means to sensitively control the packaged cargo. These effects are governed by a combination of thermodynamic and kinetic factors arising due to coupling between cargo coalescence and shell assembly.

## II. RESULTS AND DISCUSSION

We simulate the dynamics of microcompartment assembly in the presence of two cargo species, which we denote as 'R' and 'G'. There are numerous control parameters in the microcompartment system, many of which have been explored in previous works [60–67]. Here, our objective is to focus on the effect of interaction parameters on cargo encapsulation. To avoid complexities associated with assembly into different shell geometries[61, 62, 65, 66], we use a minimal model for the shell developed in Perlmutter *et al.* [60], for which the energy minimum corresponds to a $T$=3 icosahedral shell in the Caspar-Klug nomenclature [146], comprising 12 pentamers and 20 hexamers (Section IV A and Fig. 1). Although our model is motivated by bacterial microcompartments, it is sufficiently general to describe other proteinaceous shells, including viruses and microcompartments that are reengineered to encapsulate designer cargoes (*e.g.*, [22, 30–39, 41, 46–48, 50–52, 147–156]).

As shown previously for encapsulation of one cargo species [60], efficient assembly and cargo loading generically requires three classes of interactions: interactions between shell subunits that drive shell assembly, interactions between shell subunits and one or more cargo species, and interactions between cargo particles that drive cargo coalescence. While in some systems one or more of these interactions can be mediated by auxiliary proteins (*e.g.*, the CCmM and CCmN proteins in $\beta$-carboxysomes), for simplicity we model all interactions as direct pair interactions in this work. We have explored assembly dynamics over a range of the control parameters that most strongly affect the amount and composition of encapsulated cargo: the interaction strengths (binding affinities) among each type of cargo pair, $\varepsilon_{GG}$, $\varepsilon_{RR}$, and $\varepsilon_{RG}$, the shell-cargo affinity $\varepsilon_{SC}$, and the shell subunit-subunit affinity $\varepsilon_{SS}$. To minimize the number of parameters, we assume that both cargo species have the same binding affinity to shell subunits; i.e., $\varepsilon_{SC}$ is equal for R and G particles. For simplicity, we focus on parameter ranges that lead to formation of well-formed icosahedral shells; factors that lead to other shell morphologies have been investigated previously [60–62, 65, 66].

Throughout this article, all energy values are given in units of the thermal energy, $k_B T$, and all lengths are given in units of the cargo diameter, $r_*$.

There are multiple thermodynamic and kinetic factors which could influence the amount and composition of encapsulated cargo. Thermodynamic effects include the interplay between cargo-cargo and shell-cargo interactions, and the finite size of the encapsulated cargo globule. Kinetic effects arise from the fact that once a shell closes, even with relatively weak shell-shell interactions typical of productive assembly, exchange of shell subunits or cargo particles does not occur on experimentally relevant timescales. Thus, an assembling shell may trap its contents far out of equilibrium. To distinguish these effects, we have performed simulations in three different types of systems: *(i)* large simulations of cargo in the absence of shell subunits to approximate a bulk system (Fig. 2), *(ii)* simulations of cargo in the presence of assembled but permeable shells, allowing the encapsulated cargo to

reach equilibrium with the bulk (Fig. A1), and *(iii)* dynamical simulations of shell assembly around cargo. In the next section, we will briefly summarize the phase behavior of the bulk system *(i)* to serve as a reference point. Then, we will describe the results of dynamical shell assembly (system *(iii)*), with comparison to the results from systems *(i,ii)*.

### A. Cargo phase behavior without shell subunits

We begin by briefly summarizing the bulk phase behavior of the cargo in the absence of shell subunits, as a reference point from which to understand how shell assembly can change cargo coalescence. We focus on parameter regimes relevant to our shell assembly simulations; more comprehensive descriptions of the phase behavior of a binary Lennard-Jones system can be found in *e.g.*, [157–159]. In particular, we consider equal stoichiometry between R and G molecules and binding affinity values ($\varepsilon_{ij} < 1.8 \quad \text{for } ij = \text{RR}, \text{GG}, \text{RG}$) that maintain the system in vapor and/or liquid phases, but avoid binding affinities strong enough to drive crystallization. As a reference point, for a single species (e.g. R) at our simulated cargo volume fraction of 0.003, we would observe the vapor-fluid transition at $\varepsilon_{\text{RR}}^{\text{vf}} = 1.3$ and the crystallization transition at $\varepsilon_{\text{RR}}^{\text{fs}} \approx 3$. Within these limits, the phase behavior can be classified as follows: homogeneous demixed (no phase coexistence), phase separation of one cargo species (a dense phase rich in one cargo species in coexistence with a dilute phase containing an excess of the other), phase separation and mixing of both cargo species (coexistence between dilute and dense phases, with both species homogeneously mixed within both phases), and phase separation with demixing (coexistence between three phases — a dilute phase, a dense phase rich in R, and a dense phase rich in G). We consider three binding affinity values for the unlike cargo pair interaction $\varepsilon_{\text{RG}} = 1.0, 1.3, 1.6$, which respectively correspond to weak, moderate, and strong interactions (relative to $\varepsilon_{\text{RR}}^{\text{vf}}$). For each of these, Fig. 2 summarizes the bulk phase behaviors as a function of the like cargo pair interactions $\varepsilon_{\text{RR}}$ and $\varepsilon_{\text{GG}}$. Specifically, Figs. 2(A-C) compare the driving force for phase separation for each of the two species by showing the fraction of R particles $f_{\text{R}}$ in the high-density phase (hereafter referred to as 'globule'). Figs. 2 (D-F) quantify the extent of nonrandom compositional mixing within the high-density phase, $f_{\text{UL}}$. Here we define the mean fraction of unlike neighbors in the first solvation shell around each particle as: $\bar{f}_{\text{UL}} \equiv \frac{n_{\text{RG}}}{n_{\text{RR}}+n_{\text{GG}}n_{\text{RG}}}$, with $n_{ij}$ the number of cargo neighbor pairs of species $i$ and $j$ with a neighbor pair defined as two particles with separations of $r < 1.2$. For random mixing, we expect $f_{\text{rand}} = 2f_{\text{R}}(1-f_{\text{R}})$, so the plots indicate the difference between randomness and the results: $f_{\text{UL}} \equiv (\bar{f}_{\text{UL}} - f_{\text{rand}})$.

For weak unlike interactions ($\varepsilon_{\text{RG}} = 1.0$, Fig. 2A) the system exhibits either no phase separation, coexistence between a phase rich in one species and a dilute phase, or coexistence among a dilute phase and separate phases for each species. In the latter case, the R-rich and G-rich phases are attracted to each other by the weak but non-zero R-G interactions, leading to formation of one globule with an interface separating the two phases. The interface between these two phases becomes more diffuse with decreasing $\varepsilon_{\text{RR}}$ and $\varepsilon_{\text{GG}}$, and/or increasing $\varepsilon_{\text{RG}}$ because the surface tension between the two domains decreases. For moderate unlike interactions ($\varepsilon_{\text{RG}} = 1.3$, Fig. 2B) we observe no phase separation, coexistence between a phase rich in one species and a dilute phase, or coexistence between a dilute phase and a dense phase containing both species.

For strong unlike interactions ($\varepsilon_{\text{RG}} = 1.6$, Fig. 2C) we always observe mixing of both species within the dense phase, with the fraction of R or G particles depending on the relative magnitudes of $\varepsilon_{\text{RR}}$ and $\varepsilon_{\text{GG}}$. Since the primary driving force for phase separation is $\varepsilon_{\text{RG}}$ in this case, fluctuations in $f_{\text{R}}$ and $f_{\text{UL}}$ are smaller than for the other systems.

### B. Assembly pathways

We now consider how cargo and shell interactions affect the dynamics of assembly and cargo encapsulation. In previous work, simulations with a single cargo species showed that assembly pathways are strongly affected by the net strength of cargo-cargo interactions [60]. In particular, at our simulated volume fraction of 0.003, weak cargo-cargo interactions and/or strong shell-shell interactions ($\varepsilon_{\text{RR}} < 1.7$, $\varepsilon_{\text{SS}} > 2.5$) lead to one-step pathways in which shell assembly is concomitant with cargo coalescence, while strong cargo-cargo and weak shell-shell interactions ($\varepsilon_{\text{RR}} > 1.7$ and $\varepsilon_{\text{SS}} < 2.5$) allow two-step pathways in which a cargo globule coalesces, followed by disordered adsorption of shell subunits onto the globule surface and then cooperative rearrangement into an assembled shell. Other interaction combinations, such as high cargo-cargo and shell-shell affinities lead to pathways in between these two extremes.

The case of two cargo species exhibits qualitatively similar classes of assembly pathways. Fig. 3A,B show snapshots from typical simulations at relatively strong shell-shell interaction ($\varepsilon_{\text{SS}} = 3.5$), leading to one-step assembly pathways. In Fig. 3A the cargo-cargo affinities are uniform and relatively strong $\varepsilon_{\text{RR}} = \varepsilon_{\text{GG}} = \varepsilon_{\text{RG}} = 1.7$, leading to rapid and simultaneous coalescence of cargo and shell assembly, resulting in a uniform mixture of both cargo types within the complete shell. In Fig. 3B the R-R interactions ($\varepsilon_{\text{RR}} = 1.7$) are stronger than than the R-R and R-G interactions ($\varepsilon_{\text{GG}} = \varepsilon_{\text{RG}} = 1.3$), leading to simultaneous assembly and coalescence of a nearly pure R domain.

Fig. 3C-E show trajectories with a weaker shell-shell interaction of $\varepsilon_{\text{SS}} = 2.0$, for which shell nucleation is slow and we observe two-step assembly pathways. In Fig. 3C, the cargo-cargo interactions are strongest for R-R, and the assembly pathway closely resembles that for a single species — a domain of nearly pure R particles condenses, with subsequent assembly and encapsulation by the shell.

We observe a more interesting coupling between cargo composition and shell assembly for strong like-like interactions $\varepsilon_{\text{RR}} = \varepsilon_{\text{GG}} = 1.7$ and moderate unlike interactions $\varepsilon_{\text{RG}} = 1.3$ (Fig. 3D). The cargo initially coalesces into a globule that is mixed, albeit with significant compositional fluctuations, consistent with the bulk cargo phase behavior under

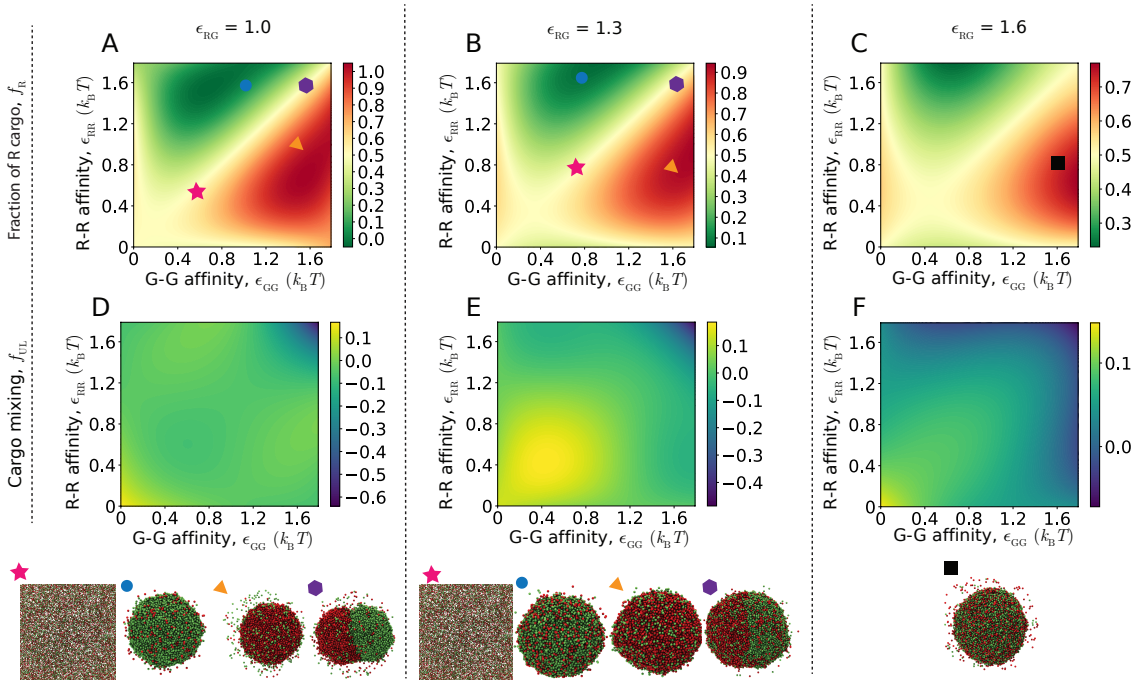


FIG. 2. **The bulk phase behavior of cargo particles (in the absence of shell subunits) as a function of their binding affinities. (A-C)** The fraction of R particles in the high-density phase, $f_R$, as a function of the like-like cargo affinities for different unlike affinities: (A) weak $\varepsilon_{RG} = 1.0$, (B) moderate $\varepsilon_{RG} = 1.3$ and (C) strong $\varepsilon_{RG} = 1.6$. All energies are given in units of the thermal energy, $k_BT$. The symbols on the plot identify corresponding snapshots below each plot that illustrate the phase behavior at the indicated value of $\varepsilon_{RG}$. **(D-F)** For the same parameters, the fraction of unlike cargo particles in the first solvation shell relative to random mixing, $f_{UL}$, defined in the text. Each simulation contains 46,938 total cargo particles, with equal composition of R and G.

these parameters (the ⬢ symbol in Fig. 2B). At these parameters, the domain is large enough to form two shells and shell closure couples to compositional fluctuations and occurs preferentially in the vicinity of an interface between R-rich and G-rich domains. In particular, we observe closure at the interface and consequently purified cargo encapsulation, meaning that the encapsulated cargo within an individual shell is nearly pure R or G. The origins of coupling between shell closure and compositional fluctuations can be understood from the competing forces that govern shell closure. As shown in Perlmutter et al. [60], closure requires that shell subunits drive formation of a neck within the cargo domain followed by scission. Since this process requires breaking cargo-cargo interactions, it is accompanied by a free energy barrier with a height that increases with cargo surface tension, and correspondingly increases with cargo-cargo affinity. In Fig. 3E, the R-R interactions are stronger ($\varepsilon_{RR} = 2.0$), resulting in a larger barrier for closure within the R-rich portion of the globule, and thus a complete shell forms only around G particles.

To quantify the extent to which a pathway is one-step or two-step, we developed an order parameter $\mathscr{O}_{pw}$ that measures the the maximum number of unassembled shell subunits adsorbed onto a globule at any point during a trajectory [60]. Fig. A2B shows $\mathscr{O}_{pw}$ as a function of the cargo-cargo affinities for relatively strong ($\varepsilon_{SS} = 3.5$) and weak ($\varepsilon_{SS} = 2.0$) shell-shell interactions. Given that a complete shell has 20 subunits, values of $\mathscr{O}_{pw} \gtrsim 15$ indicate clear two-step assembly pathways.

### C. Factors that control the composition of encapsulated cargo

#### 1. *Cargo-cargo affinity*

Under most conditions the cargo-cargo affinities have the strongest effect on the amount and composition of encapsulated cargo. We begin by considering one-step assembly pathways obtained with moderate shell-cargo interactions $\varepsilon_{SC} = 6.0$ and strong shell-shell interactions $\varepsilon_{SS} = 3.5$. These conditions allow assembly over a broad range of cargo-cargo affinities.

Fig. 4 shows observables that characterize cargo properties, as well as the assembly yield, quantified as the fraction of subunits within complete shells ($f_s$, Fig. 4D). Because the most interesting variations occur at moderate unlike affinities, we present most results in this figure with varying like-like affinities at fixed unlike affinity $\varepsilon_{RG} = 1.3$. First, Fig. 4C shows that the amount of packaged cargo increases monotonically with the net encapsulation driving force; *i.e.*, increasing either or both of $\varepsilon_{RR}$ and $\varepsilon_{GG}$ raises the number of encapsulated cargo particles $n_C$, which asymptotically approaches the value corresponding to crystalline density ($n_C \approx 150$)[60]. This demonstrates that cargo-cargo cohesive interactions are essential to obtain full shells [60].

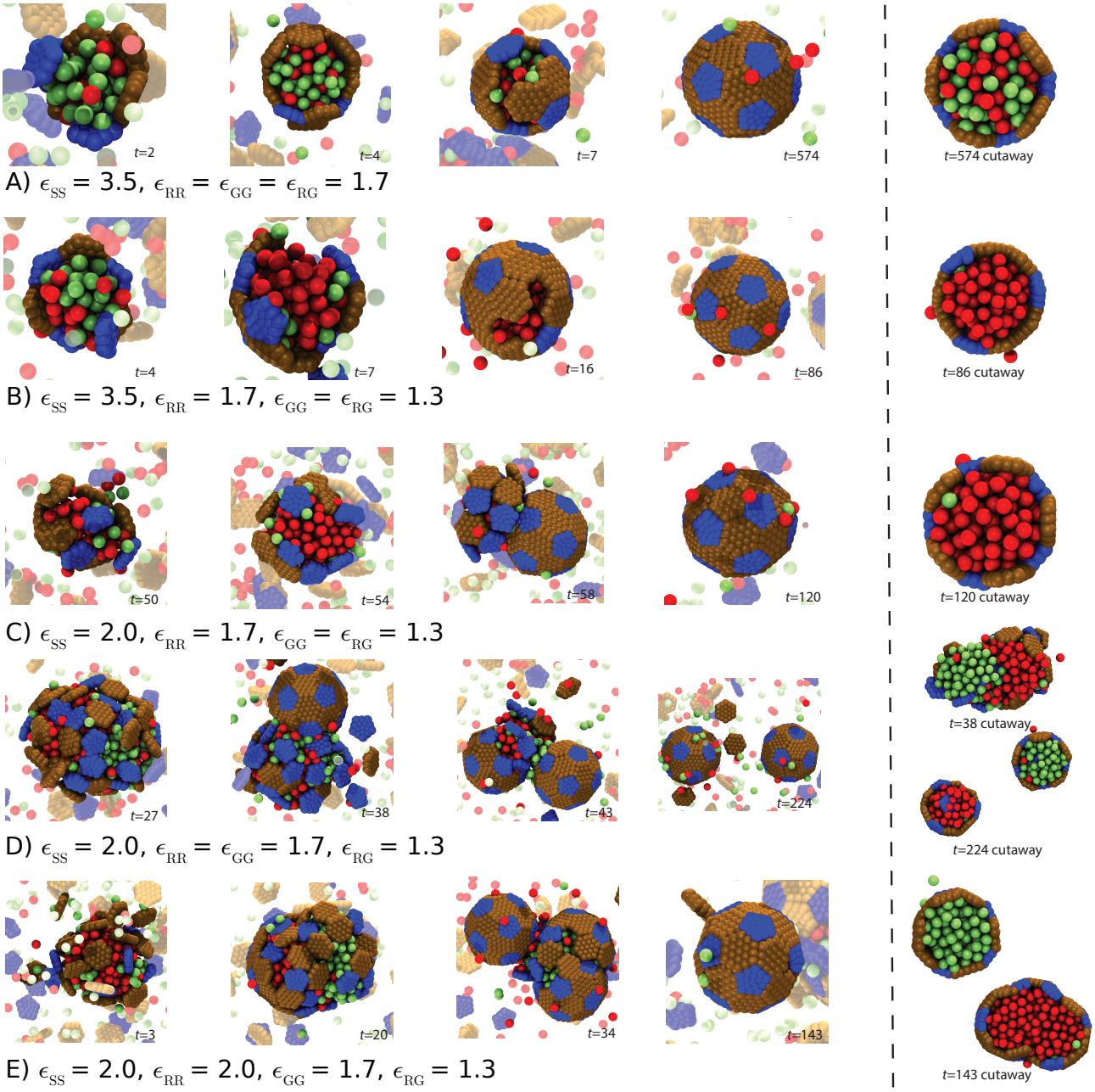


FIG. 3. **Assembly pathways.** Snapshots from simulation trajectories illustrate the classes of assembly pathways discussed in the text. In each row, the first four snapshots show frames at indicated time points (in units of the nondimensional time $\tau$), while the last snapshot is a cutaway view of the shell to show the encapsulated cargo. The parameters characterizing binding affinities are listed under each row (shell-shell, $\varepsilon_{SS}$; shell-cargo, $\varepsilon_{SC}$; like-like cargo pairs, $\varepsilon_{RR}$ and $\varepsilon_{GG}$; and unlike cargo pairs $\varepsilon_{RG}$. **(A)** All interactions are relatively high affinity, driving rapid and simultaneous coalescence of both cargo types and shell assembly (one-step assembly). The encapsulated cargo is uniformly mixed. **(B)** At high $\varepsilon_{SS}$, high $\varepsilon_{RR}$ but moderate $\varepsilon_{GG}$ and $\varepsilon_{RG}$ lead to one-step assembly and coalescence of R cargo, with G cargo essentially excluded from the shell. **(C)** At low $\varepsilon_{SS}$, strong $\varepsilon_{RR}$ interactions, but moderate $\varepsilon_{GG}$ and $\varepsilon_{RG}$ interactions, lead to two-step assembly around an almost pure R domain. **(D)** Strong $\varepsilon_{GG}$ and $\varepsilon_{RR}$ and moderate $\varepsilon_{RG}$ cargo interactions drive two-step assembly, with coupling between shell closure and cargo compositional fluctuations, leading to encapsulation of nearly pure R and G cargo domains in separate shells. The initial globule that becomes separated is shown in the top cutaway view, while the assembled shells are shown in the bottom view. **(E)** Very strong $\varepsilon_{RR}$, strong $\varepsilon_{GG}$, and moderate $\varepsilon_{RG}$ interactions lead to two-step assembly around an almost pure G domain. The strength of the $\varepsilon_{RR}$ interaction prevents the subunits from successfully encapsulating the globule of R cargo, while the G cargo is able to bud off similarly to **(D)** and form a properly assembled shell. The cutaway shows the state of the assembled shell and G cargo domain within as well as the R cargo globule that is unable to properly close. Supplemental videos, numbered corresponding to the respective trajectories above are included in the SI.

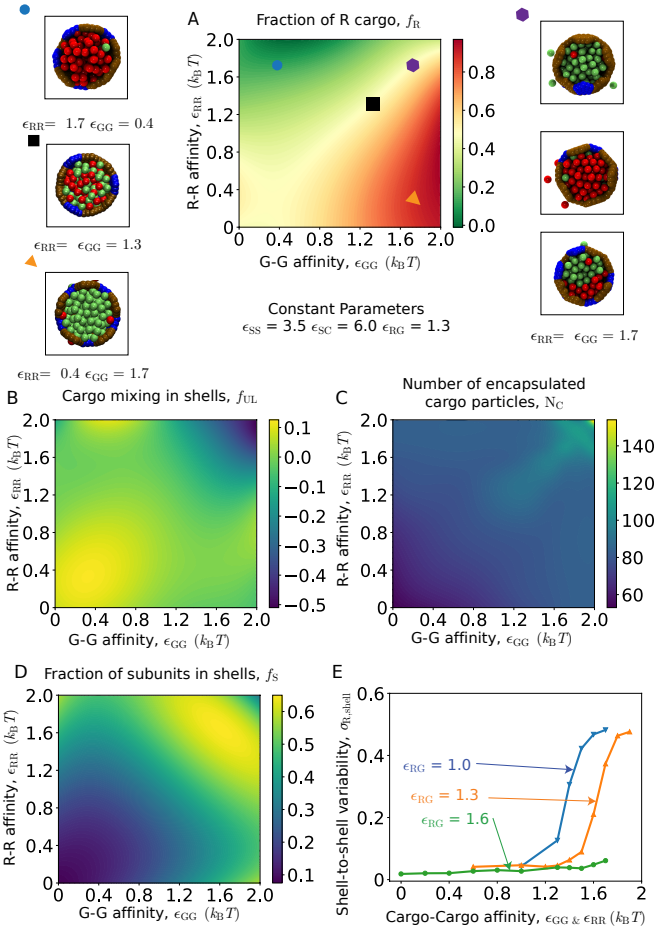


FIG. 4. **Effect of cargo-cargo affinities on encapsulation for one-step assembly pathways.** **(A-C)** Characterization of the encapsulated cargo as a function of the G-G and R-R affinities, for fixed unlike cargo affinity $\varepsilon_{RG} = 1.3$. **(A)** The fraction of R particles averaged over all shells $f_R$ to characterize the average composition of encapsulated cargo. Snapshots show cutaway views of assembled shells, to illustrate the cargo morphology at corresponding symbols on the plot. **(B)** The fraction of unlike cargo particles in first solvation shells, $f_{UL}$ averaged over all shells. **(C)** The mean number of encapsulated cargo particles per shell, to indicate cargo loading efficiency. **(D)** The fraction of subunits in complete shells, revealing the effect of cargo-cargo affinities on shell assembly. **(E)** The shell-to-shell variability of cargo composition, $\sigma_{R,shell}$, defined as the standard deviation of $f_R$ over shells, is shown as a function of the like-like affinities for three indicated values of the unlike affinity. High values of $\sigma_{R,shell} \gtrsim 0.4$ indicate that different shells have encapsulated pure domains of respectively R or G cargo species. To simplify the presentation, we only consider equal like-like affinities, $\varepsilon_{RR} = \varepsilon_{GG}$. For (A)-(E) other parameter values are $\varepsilon_{RG} = 1.3$, shell-cargo affinity $\varepsilon_{SC} = 6.0$, and shell-shell affinity $\varepsilon_{SS} = 3.5$.

Next, to determine the relative loading of the two cargo species, Fig. 4A shows the fraction of R species averaged over all shells, $f_R$, as a function of $\varepsilon_{GG}$ and $\varepsilon_{RR}$. Further, to characterize the arrangement of cargo particles within shells, Fig. 4B shows the fraction of unlike particles in the first solvation shell ($f_{UL}$). The snapshots surrounding Fig. 4A illustrate typical arrangements of the encapsulated cargo in each regime. Comparison of these results with the corresponding high-density cargo phase in the bulk system (Figs. 2B,E) shows that encapsulation within shells has a only small effect on the overall composition of encapsulated cargo.

Fig. 5 compares these results to two-step assembly pathways that occur for weaker shell-shell affinities, $\varepsilon_{SS} = 2.0$. Since we observe significant assembly only for a narrow range of strong cargo-cargo affinities (Fig. A2A, discussed next), we present these results as a function of $\varepsilon_{RR}$ for fixed $\varepsilon_{GG} = 1.7$. We see that average cargo loading and composition are qualitatively similar between the two cases, although there are quantitative differences revealing deviations from equilibrium, which we consider in more detail in Section II C 2.

However, the composition of cargo within individual shells can differ significantly from the mean. Fig. 4E shows the shell-to-shell variability of cargo composition, $\sigma_{R,shell}$, defined as the standard deviation of the fraction of R particles within each shell: $\sigma^2_{R,shell} = \frac{1}{N_{shell}} \sum_{n=1}^{N_{shell}} (f_{R,n} - f_R)^2$. Because compositional fluctuations depend strongly on the ratio of like-like/unlike affinities, we present the results as a function of like-like affinity (for $\varepsilon_{RR} = \varepsilon_{GG}$), for weak, moderate, and strong unlike affinities. In each case, as the like-like affinities exceed unlike affinities there is a transition from shells containing uniformly mixed cargo to nearly pure domains of R or G particles respectively within different shells, corresponding to large $\sigma_{R,shell}$.

Interestingly, this variability arises in different ways depending on the class of assembly pathways. To illustrate this difference, Fig. 5C,D show histograms of cargo compositions measured within individual shells for one-step and two-step pathways respectively. These parameters correspond to the maximum in Fig. 5A, where $f_R$ and $f_{UL}$ are approximately equal for the two pathways, but the mechanisms underlying the cargo separation differ. In one-step pathways, as shown in Fig. 3B, a compositional fluctuation in the initial cargo globule nucleus leads to preferential coalescence of the same cargo type as shell assembly proceeds, with the shell eventually closing around a nearly pure domain. However, for these parameters we do occasionally ($\sim 5\%$) observe shells that encapsulate two domains separated by an interface (see Fig. 5C). In contrast, two-step assembly pathways form purified shells via the compositional fluctuation mechanism discussed in Section II B and shown in Fig. 3D. This leads to highly purified cargo globules within individual shells (see Fig. 5D).

***Cargo-cargo affinities also affect shell assembly.*** While our focus in this article is on factors that control cargo encapsulation, it is important to note that the driving force for cargo cohesion in turn affects the extent and robustness of shell assembly. Figs. 4D and Fig. A2A show the fraction of subunits in complete shells $f_s$ as a function of cargo-cargo affinities for strong ($\varepsilon_{SS} = 3.5$) and weak shell-shell interactions ($\varepsilon_{SS} = 2.0$) respectively. We observe that assembly occurs over a broad range of cargo-cargo affinity values for strong shell-shell interactions ($\varepsilon_{SS} = 3.5$, Fig. 4D), but for weak shell-shell interactions ($\varepsilon_{SS} = 2.0$, Fig. A2A) significant assembly requires at least one strong cargo-cargo affinity; *i.e.*,

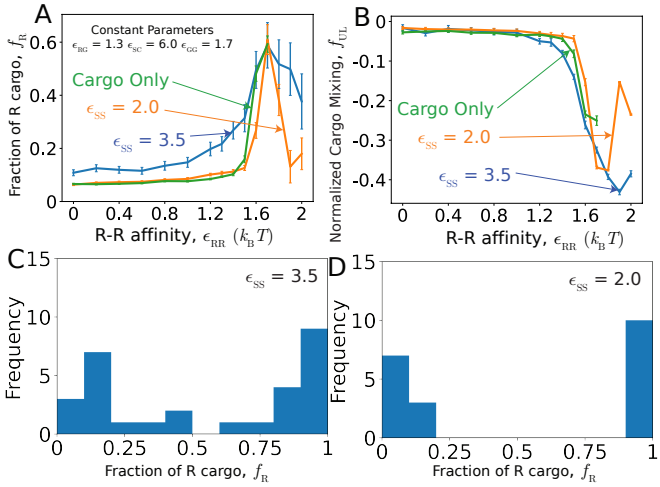


FIG. 5. **Comparison of cargo encapsulation for one-step and two-step pathways.** The fraction of R cargo in completed shells $f_R$ **(A)** and the fraction of unlike cargo particles in first solvation shells $f_{UL}$ **(B)** are shown as a function of $\varepsilon_{RR}$ for weak and strong shell-shell affinity values $\varepsilon_{SS} = 2.0$, and $\varepsilon_{SS} = 3.5$. Other parameters are $\varepsilon_{GG} = 1.7$, $\varepsilon_{RG} = 1.3$, and $\varepsilon_{SC} = 6.0$. **(C, D)** show histograms of cargo composition within individual shells to illustrate shell-to-shell variability, for the peak in panel (A) at $\varepsilon_{RR} = \varepsilon_{GG} = 1.7$ and $\varepsilon_{RG} = 1.7$, for (C) one-step and (D) two-step pathways.

$\varepsilon_{ij} \geq 1.7$ for at least one of $ij \in \{\text{RR, GG, RG}\}$. This result indicates that nucleation of shell assembly for $\varepsilon_{SS} \lesssim 2.5$ requires adsorption onto a cargo globule and thus requires that cargo-cargo affinities are sufficient to drive cargo phase separation. This observation is consistent with results from simulations with a single cargo species [60].

### 2. *Shell-shell and shell-cargo affinities affect cargo packaging through equilibrium and non-equilibrium mechanisms.*

Although the cargo-cargo affinities most strongly affect the amount and properties of encapsulated cargo, the shell protein affinities ($\varepsilon_{SC}$ and $\varepsilon_{SS}$) can also modulate cargo encapsulation. These effects depend on an interplay between thermodynamic and kinetic effects. The assembly trajectories in our simulations begin from an initial condition of dispersed subunits and cargo that are significantly out of equilibrium. Moreover, once a shell closes, the timescale for reconfiguration to another shell morphology or exchange of the encapsulated cargo is much longer than assembly timescales since all subunits have their maximum number of interactions. Therefore, even though the dynamics satisfy microscopic reversibility, trajectories are not guaranteed to approach equilibrium on the finite timescales of simulations and experiments [160].

To distinguish between equilibrium and kinetic effects, we compare the results of our assembly simulations to the alternative 'equilibrium' or permeable-shell system described above, in which shells are completely assembled but are made permeable to cargo particles, thus allowing the encapsulated cargo to fully equilibrate with cargo particles in solution (see section IV). The amount and composition of encapsulated cargo in these simulations is thus independent of assembly kinetics (and shell-shell interactions). Fig. A1 compares the results of assembly trajectories at parameters leading to one-step assembly (corresponding to Fig. 4) with the equilibrium simulations. Over much of parameter space the amount and compositions of encapsulated cargo are similar between the two systems, indicating that trajectories are near equilibrium during assembly. This can be understood since the diffusional exchange of cargo particles between the partially encapsulated cargo globule and bulk is rapid compared to shell growth under those conditions.

However, we do observe significant deviations from equilibrium cargo encapsulation under certain parameter ranges. These effects are strongly dependent on the shell-shell and shell-cargo interactions, since they are key determinants in the extent to which assembly is out of equilibrium.

For example, Fig. 5A shows the dependence of $f_R$ on $\varepsilon_{RR}$ for one-step ($\varepsilon_{SS} = 3.5$) and two-step ($\varepsilon_{SS} = 2.0$) assembly pathways. For moderate cargo-cargo affinities $\varepsilon_{RR} \lesssim 1.5$, the two-step assembly and equilibrium results closely agree, while the one-step simulations result in higher $f_R$. This difference arises because the initial globule is enriched in R particles (since globule nucleation is driven in part by shell-cargo interactions that are agnostic to particle type), and the rapid one-step assembly pathways do not allow the cargo globule time to fully equilibrate with the bulk. In contrast, the slow onset of nucleation and assembly in the two-step pathways allows globule equilibration. Interestingly, the situation reverses at high $\varepsilon_{RR} \gtrsim 1.6$, where there is an abrupt drop in $f_R$ for the two-step case. This arises because the strong R-R interactions favor segregation of the two cargo species and a high surface tension that inhibits closure of a shell around R particles. A comparable trajectory is shown in Fig. 3E. Notice that a shell successfully assembles around the cargo particles with weaker interactions (G), but only a malformed structure forms around the R globule because the shell subunits are unable to close around a commensurate-sized piece of the globule. Note that we do not show the permeable-shell results for such high $\varepsilon_{RR}$ because the strong cargo-cargo interactions result in a cargo globule that extends outside of the shell, which is an artifact of the permeable-shell condition imposed in those simulations.

We also observe that the extent of cargo mixing within shells (Fig. 5B) decreases with increasing $\varepsilon_{RR}$. This is partially an equilibrium affect due to the increasing drive for cargo segregation, especially as the like-like cargo affinities exceed the unlike affinity ($\varepsilon_{RG} = 1.3$). However, while the two-step results are close to equilibrium, mixing is further reduced for the one-step assembly trajectory. This effect can be attributed to the coupling between composition fluctuations and interfacial effects within the globule during the initial stage of the two-step pathway described in section II B and Fig. 3D. The tendency of budding to occur at the site of an interface enriches shell contents in one cargo species; this composition is then effectively locked in by shell closure.

Having compared the differences in cargo encapsulation between the two classes of assembly pathways, we now consider parameter values between these two extremes. Figs. 6A and

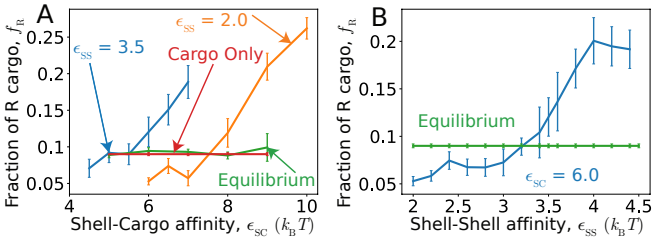


FIG. 6. **Effects of shell-cargo and shell-shell affinity on cargo composition reveal thermodynamic and kinetic influences on encapsulation.** **(A)** The fraction of encapsulated R particles $f_R$ as a function of shell-cargo affinity $\varepsilon_{SC}$, for weak and moderate shell-shell affinity values $\varepsilon_{SS} = 2.0$, and $\varepsilon_{SS} = 3.5$. To assess the importance of dynamics on these results, $f_R$ is also shown for the 'equilibrium' simulations in which cargo can exchange between shell interiors and the bulk. Other parameters are $\varepsilon_{GG} = 1.7$ and $\varepsilon_{RR} = \varepsilon_{RG} = 1.3$. **(B)** Cargo composition $f_R$ as a function of shell-shell affinity $\varepsilon_{SS}$, at a constant shell-cargo affinity of $\varepsilon_{SC} = 6.0$. Other parameters are $\varepsilon_{RR} = 1.3$, with $\varepsilon_{GG} = 1.7$ and $\varepsilon_{RG} = 1.3$. The 'equilibrium' results are also shown, as a straight line since they do not depend on $\varepsilon_{SS}$.

6B respectively show the effects of shell-cargo and shell-shell affinities on the composition of encapsulated cargo. Here, we have focused on the case for which cohesive interactions between pairs of G particles are stronger than for R-R and R-G, $\varepsilon_{GG} = 1.7$ and $\varepsilon_{RR} = \varepsilon_{RG} = 1.3$, for which the equilibrium and bulk composition of R particles is about 20%. Fig. 6A shows that for weak shell-cargo interactions, the encapsulated fraction of R particles is significantly smaller than the equilibrium value, but that it monotonically increases with $\varepsilon_{SC}$. The trend is the same for both strong and weak shell-shell interactions, except that higher shell-cargo affinities are required to observe assembly for weaker shell-shell interactions. Similarly, Fig. 6B shows that $f_R$ increases monotonically with shell-shell affinity. These trends can be understood as follows. At low $\varepsilon_{SC}$, the cargo-cargo affinities are the primary driving force for encapsulation, which at these parameters favors encapsulation of G cargo and thus small $f_R$. The fact that $f_R$ in the dynamical assembly simulations is smaller than that observed in the bulk cargo and equilibrium simulations suggests that the initial globule that nucleates is enriched in G, and that the shell closes before there is sufficient time for the globule composition to equilibrate with the bulk. However, as $\varepsilon_{SC}$ or $\varepsilon_{SS}$ increases, the strong shell-cargo interactions enhance nucleation, allowing a cargo globule to form with a composition closer to the bulk composition, $f_R = 0.5$.

## III. CONCLUSIONS

We have performed equilibrium and dynamical simulations to investigate the encapsulation of multicomponent fluid cargoes by self-assembling microcompartments. The results from these simulations show that the binding affinities among the different fluid components are the strongest determinant of the amount, composition, and spatial organization of cargo components within shells.

In certain regimes, the shell-cargo and shell-shell binding affinities can also be manipulated to control cargo encapsulation. A combination of thermodynamic and kinetic mechanisms underlie these effects. In particular, we see a marked difference between the properties of the encapsulated cargo in shells that assemble by one-step or two-step pathways. When cargo-cargo affinities are asymmetric leading to preferential encapsulation of one species, weaker shell-cargo and shell-shell interactions and two-step assembly pathways tend to enhance the degree of preferential encapsulation leading to more purified cargo. This behavior reflects the fact that the cargo globule that nucleates in the first step tends to be enriched in the species with higher affinities, and, if assembly occurs more rapidly than cargo exchange with bulk, shell closure locks in this out-of-equilibrium composition. Similarly, when the set of cargo-cargo affinities places the system close to ternary phase coexistence (a dilute phase and two dense phases respectively rich in either cargo species), two-step assembly pathways can enhance separation of the cargo species. This result can be explained at least in part by the tendency of assembling shells to close in the vicinity of transient interfaces between domains of the two cargo species.

The latter predictions can be tested in microcompartment systems by performing mutagenesis of shell protein-protein binding interfaces (to alter shell-shell affinities) or 'encapsulation peptides' that mediate shell-cargo interactions [161–164]. Cargo-cargo interactions are less well understood at present. Depending on the system, cargo-cargo affinities may arise from direct pair interactions, scaffold-mediated interactions, or a combination of the two [31, 36, 39, 121, 122, 126, 165, 166]. Thus, it will be interesting to extend the model considered here to explicitly incorporate scaffold-mediated cargo-cargo and shell-cargo interactions, for example as Ref. [62] did for a single-component cargo. Such a system is inherently more complex since its behavior will depend on not only cargo-cargo and cargo-scaffold affinities for each species, but also the properties of the scaffold itself, such as its flexibility and self interactions [62]. Further, it will be important to allow for shell assembly into different geometries, to study coupling between shell size and the composition and spatial organization of the different cargo components.

## IV. METHODS

### A. Computational model

We consider a minimal model for microcompartment assembly and encapsulation of two species of cargo particles, denoted as 'R' and 'G'. Microcompartment shells assemble from pentameric, hexameric, and pseudo-hexameric (trimeric) protein oligomers (*e.g.*, Fig. 3A in Ref. [58] and Refs. [18, 19, 167]). Experiments suggest that these oligomers are the basic assembly units, meaning that smaller protein complexes do not contribute significantly to the assembly process [19, 168]. To focus on factors that control the composition of encapsulated cargo, we consider a minimal shell model developed in Perlmutter et al. [60], with

only pentamer and hexamer subunits that have interactions designed to drive assembly into a $T$=3 shells (containing 12 pentamers and 20 hexamers in a truncated icosahedron geometry). Our model builds on previous simulations of microcompartment assembly with no cargo [66] or one cargo species [60, 61, 65], as well as previous models for virus assembly [76, 79, 80, 83, 102, 169–174].

**Shell-shell interactions.** In our model, shell subunits interact through both repulsive and attractive forces. The repulsions consist of 'Top' pseudoatoms that exist above the plane of the hexamer and 'Bottom' pseudoatoms that exist below the plane. We denote the top pseudoatoms as 'TH' and 'TP' respectively for hexamer and pentamer subunits, and similarly the bottom pseudoatoms as 'BH' and 'BP'. All pairs of Top and Bottom pseudoatoms (excepting those on the same subunit) interact via a repulsive Lennard-Jones potential (Eq. A3). The Top-Top diameters and repulsion strengths set the shell spontaneous curvature and bending modulus (see section A).

Microcompartment protein shell assembly is primarily driven by interactions along the edges of the hexameric and pentameric subunits [58]. We represent these interactions by 'Attractors' on the perimeter of each shell subunit. Complementary Attractors on nearby subunits have short-range interactions modeled by a Morse potential (Eq. (A4) in section ). Attractors which are not complementary do not interact. The arrangement of Attractors on subunit edges is shown in Fig. 1, with pairs of complementary Attractors indicated by cyan double-headed arrows. The shell-shell binding affinity is proportional to the well-depth of the Morse potential between complementary Attractors, $\varepsilon_{\mathrm{HH}}$.

To focus on the role of interaction strengths in determining cargo encapsulation, we consider shell subunits that preferentially form only T=3 shells. To enforce this restriction, only three of the 6 edges of a hexamer have attractive interactions with pentamers, and there are no pentamer-pentamer attractive interactions (see Fig. 1).

**Shell-cargo interactions.** We model attractive interactions between hexamer subunits and the cargo particles by a Morse Potential between cargo particles and hexamer 'BH' pseudoatoms, with well-depth $\varepsilon_{\mathrm{SC}}$. These interactions represent the effect of 'encapsulation peptides' that target cargo to microcompartment interiors by mediating cargo-hexamer interactions [31, 39, 121, 122, 165]. To minimize the number of parameters, we keep $\varepsilon_{\mathrm{SC}}$ the same for both cargo species. We also add a layer of 'Excluders' in the plane of the 'Top' pseudoatoms, which represent shell-cargo excluded volume interactions. Since the shell-shell interaction geometries are already controlled by the Attractor, Top, and Bottom pseudoatoms, we do not consider Excluder-Excluder interactions.

**Cargo-cargo interactions.** In natural microcompartment systems, the interior cargo undergoes phase separation prior to or during shell assembly. The attractions between cargo particles that drive phase separation may be mediated by microcompartment scaffolding proteins (e.g. [120, 122, 126]) and/or direct pair interactions between cargo particles [36]. Similarly, in synthetic systems it is possible to engineer direct pair or scaffold-mediated attractions cargo molecules. In our model, we represent these scenarios in a minimal manner. We model both species of cargo particles with spherically symmetric excluded volume and direct pair attractions, implemented via an attractive Lennard-Jones (LJ) potential, with well-depth values $\varepsilon_{\mathrm{GG}}$, $\varepsilon_{\mathrm{RR}}$, and $\varepsilon_{\mathrm{RG}}$ for pairs of G-G, R-R, and R-G particles.

### B. Simulations and systems

We performed simulations using the Langevin dynamics algorithm in HOOMD (which uses GPUs to accelerate molecular dynamics simulations[175]), and periodic boundary conditions in a cubic box to represent a bulk system. The subunits were modeled as rigid bodies [176]. Each simulation was performed in the NVT ensemble, using the HOOMD fundamental units [177], with the unit length scale $r_*$ defined as the circumradius of the pentagonal subunit (the cargo diameter is also set to $r_*$), and energies given in units of the thermal energy, $k_{\mathrm{B}}T$. The simulation timestep was 0.005 in dimensionless time units for all systems.

**Systems.** We simulated several systems as follows:

*Shell assembly and cargo encapsulation.* For dynamical simulations of shell assembly in the presence of cargo, each simulation contained enough subunits to form four complete microcompartments (48 pentamers and 80 hexamers), in a cubic simulation box with sidelength $40r_*$. Each simulation also contained 854 cargo particles, with composition 50% 'R' and 'G' respectively. We performed each simulation for $2 \times 10^6$ timesteps. Simulations in the two-step assembly regime were performed for longer, $2 \times 10^7$ timesteps, as they assemble more slowly.

*Bulk simulations without shell subunits.* We also performed large simulations of cargo without shell subunits, to determine the cargo phase behavior of the cargo as a function of the cargo-cargo interactions. Each of these simulations contained 46,938 cargo particles (50% 'R' and 'G' respectively) in a box with sidelength 160, to maintain the same cargo concentration as in the shell assembly simulations. We performed each simulation for $2 \times 10^7$ timesteps.

*Equilibrium cargo encapsulation.* To identify effects of out-of-equilibrium dynamics of shell assembly on the properties of the encapsulated cargo, we performed additional simulations in which we simulated cargo dynamics in the presence of an assembled shell that was made permeable to cargo. In particular, these simulations consisted of a fully assembled immobile shell placed in the center of the box, with one missing pentamer to allow permeability to occur with the surrounding cargo. While the shell could not move, the $\varepsilon_{\mathrm{SC}}$ was still active. Each simulation contained 854 cargo particles with box sidelength $40r_*$ to maintain the same cargo concentration and composition as in the other systems. We performed each simulation for $2 \times 10^6$ timesteps, which we determined was sufficient to achieve equilibration of cargo inside of the shell by monitoring the amount, composition, and mixing of the encapsulated cargo.

**Initial conditions.** Except for the equilibrium cargo encapsulation systems, all simulations were initialized with random positions for cargo particles and shell subunits (except

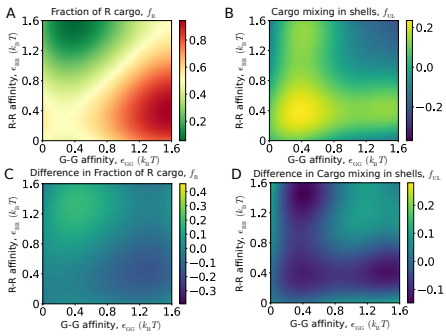


FIG. A1. **Comparison of encapsulated cargo between equilibrium and dynamical simulations. (A,B)** The fraction of R particles **(A)** and mixing **(B)** are shown as a function of the affinities for like-like cargo interactions. Other parameters are $\varepsilon_{RG} = 1.3$ and $\varepsilon_{SC} = 6.0$, as in Fig. 4. **(C,D)** Difference between cargo encapsulation in dynamical assembly simulations and equilibrium. The plots show the differences between the quantities $f_R$ and $f_{UL}$ for the dynamical simulations Fig. 4 (with $\varepsilon_{SS} = 3.5$) and equilibrium simulations (panels (A,B) of this figure).

that significant overlaps between pseudoatoms were forbidden). The equilibrium cargo encapsulation simulations were initialized with an assembled microcompartment shell whose subunits were fixed throughout the simulation, and the cargo particles were initialized with random locations with overlaps forbidden.

**Statistical Significance.** For most systems, we performed 10 independent trials at each parameter set. For most parameter sets at which we report statistical errors, we performed 20 independent trials. For parameter sets that resulted in low probability events ($\leq 5\%$), we performed 60 trials to ensure low statistical error.

## ACKNOWLEDGEMENTS

This work was supported by Award Number R01GM108021 from the National Institute Of General Medical Sciences and the Brandeis Center for Bioinspired Soft Materials, an NSF MRSEC, DMR-2011846. Computational resources were provided by NSF XSEDE computing resources (Maverick, Bridges, and Comet) and the Brandeis HPCC which is partially supported by DMR-2011846.

## Appendix A: Model details

In our model, all potentials can be decomposed into pairwise interactions. Potentials involving shell subunits further decompose into pairwise interactions between their constituent building blocks – the excluders, attractors, 'Top', and 'Bottom' pseudoatoms. It is convenient to state the total energy of the system as the sum of three terms, involving shell-shell ($U_{SS}$), cargo-cargo ($U_{CC}$), and shell-cargo ($U_{SC}$) interactions, each summed over all pairs of the appropriate type:

$$U = \sum_{\text{shell } i} \sum_{\text{shell } j<i} U_{SS} + \sum_{\text{cargo } i} \sum_{\text{cargo } j<i} U_{CC} + \sum_{\text{shell } i} \sum_{\text{cargo } j} U_{SC} \tag{A1}$$

where $\sum_{\text{shell } i} \sum_{\text{shell } j<i}$ is the sum over all distinct pairs of shell subunits in the system, $\sum_{\text{shell } i} \sum_{\text{cargo } j}$ is the sum over all shell-cargo particle pairs, etc.

**Shell-shell interaction potentials.** The shell-shell potential $U_{SS}$ is the sum of the attractive interactions between complementary attractors, and geometry guiding repulsive interactions between 'Top' - 'Top', 'Bottom' - 'Bottom', and 'Top' - 'Bottom' pairs. There are no interactions between members of the same rigid body. Thus, for notational clarity, we index rigid bodies and non-rigid pseudoatoms in Roman, while the pseudoatoms comprising a particular rigid body are indexed in Greek. For subunit $i$ we denote its attractor positions as $\{\mathbf{a}_{i\alpha}\}$ with the set comprising all attractors $\alpha$, its 'Top' position $\mathbf{t}_i$, and its 'Bottom' position $\mathbf{b}_i$.

The shell-shell interaction potential between two subunits $i$ and $j$ is then defined as:

$$\begin{aligned} U_{SS}(\{\mathbf{a}_{i\alpha}\}, \mathbf{t}_i, \mathbf{a}_j, \mathbf{t}_j) = {} & \varepsilon_{SS} Ł\left(\left|\mathbf{t}_i - \mathbf{t}_j\right|,\ \sigma_{t,ij}\right) \\ & + \varepsilon_{SS} Ł\left(\left|\mathbf{b}_i - \mathbf{b}_j\right|,\ \sigma_b\right) \\ & + \varepsilon_{SS} Ł\left(\left|\mathbf{b}_i - \mathbf{t}_j\right|,\ \sigma_{tb}\right) \\ & + \sum_{\alpha,\beta}^{N_{ai}, N_{aj}} \varepsilon_{SS} \mathscr{M}\left(\left|\mathbf{a}_{i\alpha} - \mathbf{a}_{j\beta}\right|,\ r_0, \rho, r_{\text{cut}}^{\text{att}}\right) \end{aligned} \tag{A2}$$

The function Ł is defined as the repulsive component of the Lennard-Jones potential shifted to zero at the interaction diameter:

$$Ł(x, \sigma) \equiv \theta(\sigma - x)\left[\left(\frac{\sigma}{x}\right)^{12} - 1\right] \tag{A3}$$

with $\theta(x)$ the Heaviside function. The function $\mathscr{M}$ is a Morse potential:

$$\begin{aligned} \mathscr{M}(x, r_0, \rho, r_{\text{cut}}) = {} & \theta(r_{\text{cut}} - x) \times \\ & \left[\left(e^{\rho\left(1 - \frac{x}{r_0}\right)} - 2\right) e^{\rho\left(1 - \frac{x}{r_0}\right)} - V_{\text{shift}}(r_{\text{cut}})\right] \end{aligned} \tag{A4}$$

with $V_{\text{shift}}(r_{\text{cut}})$ the value of the (unshifted) potential at $r_{\text{cut}}$.

The parameter $\varepsilon_{SS}$ sets the strength of the shell-shell attraction at each attractor site, $N_{ai}$ is the number of attractor pseudoatoms in subunit $i$, and $\varepsilon_{\text{angle}}$ scales the repulsive interactions that enforce the geometry.

**Shell-shell interaction parameter values.** *Attractors:* The strength of attractive interactions is parameterized by the well-depth $\varepsilon_{SS}$ for a pair of attractors on hexamers as follows. Hexamer-Hexamer edge attractor pairs (A2-A6, A3-A5, and A5-A6) have a well-depth of $\varepsilon_{SS}$. Because vertex attractors (A1, A4) have multiple partners in an assembled structure,

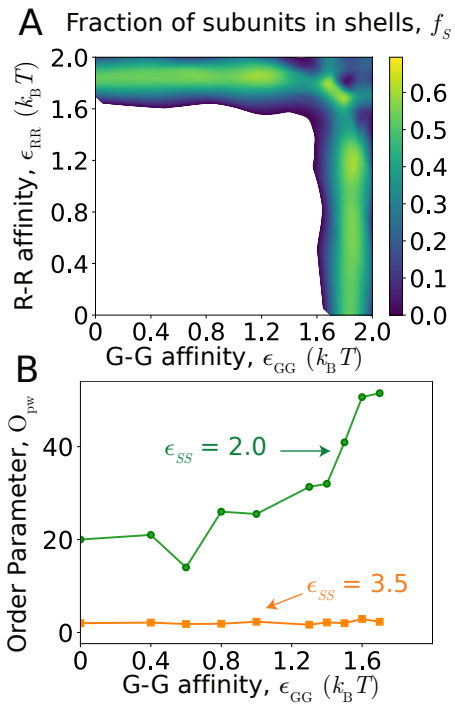


FIG. A2. **(A)** Fraction of shell subunits in complete shells in the two-step assembly pathway regime, $\varepsilon_{SS} = 2.0$ and $\varepsilon_{RG} = 1.3$. **(B)** The pathway order parameter $\mathscr{O}_{pw}$ as a function of the cargo-cargo affinity $\varepsilon_{GG}$ for one-step and two-step pathways, with fixed $\varepsilon_{RR} = 1.7$ and $\varepsilon_{RG} = 1.3$.

whereas edge attractors have only one, the well-depth for the vertex pairs (A1-A4 and A4-A4) is set to $0.5\varepsilon_{SS}$. Similarly, for pentamer-hexamer interactions, the well-depth for edge attractor pairs (A2-A5, A3-A6) is $\varepsilon_{PH}$, while the vertex interaction pairs (A1-A4 and A4-A4) have $0.5\varepsilon_{PH}$.

*Repulsive interactions:* The 'Top' and 'Bottom' heights, or distance out of the attractor plane, are set to $h = 1/2r_b$, with $r_b = 1$ the distance between a vertex attractor and the center of the pentagon. For all simulations shells have $T$=3 preferred curvature and,$\sigma_{tb} = 1.8r_b$ is the diameter of the 'Top' - 'Bottom' interaction (this prevents subunits from binding in inverted configurations [76]), and $\sigma_b = 1.5r_b$ is the diameter of the 'Bottom' - 'Bottom' interaction. In contrast to the latter parameters, $\sigma_{t,ij}$ the effective diameter of the 'Top' - 'Top' interaction, depends on the species of subunits $i$ and $j$; denoting a pentagonal or hexagonal subunit as 'p' or 'h' respectively, $\sigma_{t,pp} = 2.1r_b$, $\sigma_{t,hh} = 2.4r_b$, and $\sigma_{t,ph} = 2.2r_b$. The parameter $r_0$ is the minimum energy attractor distance, set to $0.2r_b$, $\rho = 4r_b$ determines the width of the attractive interaction, and $r_{cut}^{att} = 2.0r_b$ is the cutoff distance for the attractor potential. Since the interactions just described are sufficient to describe assembly of the shell subunits, we included no excluder-excluder interactions.

**Cargo-cargo interactions.** The interaction between cargo particles is given by a sum over all interacting pairs:

$$U_{CC}(\{\mathbf{l}_i\},\{\mathbf{l}_j\}) = \sum_{z=RR,GG,RG} \sum_{i<j}^{N_z} \varepsilon_z \mathscr{L}\left(\left|\mathbf{l}_i - \mathbf{l}_j\right|,\ \sigma_C, r_{cut}^C\right) \tag{A5}$$

with $\mathscr{L}$ the full Lennard-Jones interaction:

$$\mathscr{L}(x,\sigma,r_{cut}) = \theta(x - r_{cut}) \times \left\{4\left[\left(\frac{x}{\sigma}\right)^{12} - \left(\frac{x}{\sigma}\right)^{6}\right] - V_{shift}(r_{cut})\right\} \tag{A6}$$

and $N_{RR}$, $N_{GG}$, and $N_{RG}$ as the number of R-R, G-G, and R-G pairs in the system, and the cargo diameter $\sigma_C = r_b$ and cutoff $r_{cut}^C = 3\sigma_C$ are the same for all cargo species.

**Shell-cargo interactions.** The shell-cargo interaction is modeled by a short-range repulsion between cargo-excluder and cargo-'Top' pairs representing the excluded volume, plus an attractive interaction between pairs of cargo particles and hexamer 'Bottom' pseudoatoms. (We do not consider pentamer-cargo attractions because there is no experimental evidence for them.) For subunit $i$ with excluder positions $\{\mathbf{x}_{i\alpha}\}$ and 'Bottom' psuedoatom $\mathbf{b}_i$, and cargo particle $j$ with position $\mathbf{R}_j$, the potential is:

$$U_{SC}(\{\mathbf{x}_{i\alpha}\},\mathbf{R}_j) = \sum_{\alpha}^{N_x} Ł\left(|\mathbf{x}_{i\alpha} - \mathbf{R}_j|, \sigma_{ex}\right) \tag{A7}$$

$$+ \sum_{\alpha}^{N_t} Ł\left(|\mathbf{t}_{i\alpha} - \mathbf{R}_j|, \sigma_t\right) \tag{A8}$$

$$+ \sum_{\alpha}^{Nb} \varepsilon_{SC} \mathscr{M}\left(\left|\mathbf{c}_{i\alpha} - \mathbf{R}_j\right|,\ r_0, \rho^{SC}, r_{cut}^{SC}\right) \tag{A9}$$

where $\varepsilon_{SC}$ parameterizes the shell-cargo interaction strength, $N_x$, $N_t$, and $N_b$ are the numbers of excluders, 'Top', and 'Bottom' pseudoatoms on a shell subunit, $\sigma_{ex} = 0.5r_b$ and $\sigma_t = 0.5r_b$ are the effective diameters of the Excluder - cargo and 'Top' - cargo repulsions, $r_0^{SC} = 0.5r_b$ is the minimum energy attractor distance, the width parameter is $\rho^{SC} = 2.5r_b$, and the cutoff is set to $r_{cut}^{SC} = 3.0r_b$.


[1] C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher, and A. A. Hyman, Science **324**, 1729 (2009).
[2] C. Brangwynne, P. Tompa, and R. Pappu, Nat. Phys. **11**, 899 (2015).
[3] C. F. Lee, C. P. Brangwynne, J. Gharakhani, A. A. Hyman, and F. Julicher, Phys Rev Lett **111**, 088101 (2013).
[4] C. P. Brangwynne, J. Cell Biol. **203**, 875 (2013).
[5] S. C. Weber and C. P. Brangwynne, Cell **149**, 1188 (2012).
[6] A. A. Hyman and C. P. Brangwynne, Dev. Cell **21**, 14 (2011).
[7] A. A. Hyman and K. Simons, Science **337**, 1047 (2012).
[8] L. P. Bergeron-Sandoval, N. Safaee, and S. W. Michnick, Cell

**165**, 1067 (2016).
[9] M. Feric, N. Vaidya, T. S. Harmon, D. M. Mitrea, L. Zhu, T. M. Richardson, R. W. Kriwacki, R. V. Pappu, and C. P. Brangwynne, Cell **165**, 1686 (2016).
[10] D. Hnisz, K. Shrinivas, R. A. Young, A. K. Chakraborty, and P. A. Sharp, Cell **169**, 13 (2017).
[11] M. B. Stone, S. A. Shelby, M. F. Nunez, K. Wisser, and S. L. Veatch, Elife **6** (2017), 10.7554/eLife.19891.
[12] L. Zhu and C. P. Brangwynne, Curr Opin Cell Biol **34**, 23 (2015).
[13] S. F. Banani, H. O. Lee, A. A. Hyman, and M. K. Rosen, Nat Rev Mol Cell Biol **18**, 285 (2017).
[14] J. A. Joseph, J. R. Espinosa, I. Sanchez-Burgos, A. Garaizar, D. Frenkel, and R. Collepardo-Guevara, Biophys J **120**, 1219 (2021).
[15] T. O. Yeates, C. A. Kerfeld, S. Heinhorst, G. C. Cannon, and J. M. Shively, Nat. Rev. Microbiol. **6**, 681 (2008).
[16] T. J. Nott, 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.*, Mol. Cell **57**, 936 (2015).
[17] B. Y. Zaslavsky, L. A. Ferreira, A. L. Darling, and V. N. Uversky, Int. J. Biol. Macromol. **117**, 1224 (2018).
[18] C. A. Kerfeld, S. Heinhorst, and G. C. Cannon, Microbiology+ **64**, 391 (2010).
[19] S. Tanaka, C. A. Kerfeld, M. R. Sawaya, F. Cai, S. Heinhorst, G. C. Cannon, and T. O. Yeates, Science **319**, 1083 (2008), http://www.sciencemag.org/content/319/5866/1083.full.pdf.
[20] B. D. Rae, B. M. Long, M. R. Badger, and G. D. Price, Microbiol. Mol. Biol. Rev. **77**, 357 (2013), http://mmbr.asm.org/content/77/3/357.full.pdf+html.
[21] T. A. Bobik, B. P. Lehman, and T. O. Yeates, Mol. Microbiol. **98**, 193 (2015).
[22] C. Chowdhury, S. Sinha, S. Chun, T. O. Yeates, and T. A. Bobik, Microbiol. Mol. Biol. Rev. **78**, 438 (2014).
[23] C. A. Kerfeld and M. R. Melnicki, Curr. Opin. Plant Biol. **31**, 66 (2016).
[24] J. K. Polka, S. G. Hays, and P. A. Silver, Cold Spring Harb Perspect Biol **8** (2016), 10.1101/cshperspect.a024018.
[25] C. A. Kerfeld, C. Aussignargues, J. Zarzycki, F. Cai, and M. Sutter, Nat. Rev. Microbiol. **16**, 277 (2018).
[26] M. Sutter, D. Boehringer, S. Gutmann, S. Gunther, D. Prangishvili, M. J. Loessner, K. O. Stetter, E. Weber-Ban, and N. Ban, Nat. Struct. Mol. Biol. **15**, 939 (2008).
[27] R. J. Nichols, B. LaFrance, N. R. Phillips, L. M. Oltrogge, L. E. Valentin-Alvarado, A. J. Bischoff, E. Nogales, and D. Savage, bioRxiv (2020).
[28] F. Pfeifer, Nat Rev Micro **10**, 705 (2012).
[29] V. A. Kickhoefer, K. S. Rajavel, G. L. Scheffer, W. S. Dalton, R. J. Scheper, and L. H. Rome, J. Biol. Chem. **273**, 8971 (1998), http://www.jbc.org/content/273/15/8971.full.pdf+html.
[30] A. Chessher, R. Breitling, and E. Takano, ACS Biomaterials Science & Engineering **1**, 345 (2015).
[31] B. P. Lehman, C. Chowdhury, and T. A. Bobik, J. Bacteriol. **199** (2017), 10.1128/JB.00785-16.
[32] B. M. Long, M. R. Badger, S. M. Whitney, and G. D. Price, J. Biol. Chem. **282**, 29323 (2007).
[33] B. M. Long, L. Tucker, M. R. Badger, and G. D. Price, Plant Physiol. **153**, 285 (2010).
[34] M. J. Mayer, R. Juodeikis, I. R. Brown, S. Frank, D. J. Palmer, E. Deery, D. M. Beal, W. F. Xue, and M. J. Warren, Sci. Rep. **6**, 36899 (2016).
[35] M. B. Quin, S. A. Perdue, S. Y. Hsu, and C. Schmidt-Dannert, Appl. Microbiol. Biotechnol. **100**, 9187 (2016).
[36] M. Slininger Lee and D. Tullman-Ercek, Current Opinion in Systems Biology (2017), 10.1016/j.coisb.2017.05.017.
[37] M. F. Slininger Lee, C. M. Jakobson, and D. Tullman-Ercek, ACS Synth. Biol. (2017), 10.1021/acssynbio.7b00042.
[38] A. Hagen, M. Sutter, N. Sloan, and C. A. Kerfeld, Nat. Comm. **9**, 1 (2018).
[39] J. N. Kinney, A. Salmeen, F. Cai, and C. a. Kerfeld, J. Biol. Chem. **287**, 17729 (2012).
[40] B. Ferlez, M. Sutter, and C. A. Kerfeld, Metab. Eng. **54**, 286 (2019).
[41] G. Kalnins, E.-E. Cesle, J. Jansons, J. Liepins, A. Filimonenko, and K. Tars, Nat. Comm. **11**, 1 (2020).
[42] R. Juodeikis, M. J. Lee, M. Mayer, J. Mantell, I. R. Brown, P. Verkade, D. N. Woolfson, M. B. Prentice, S. Frank, and M. J. Warren, MicrobiologyOpen **9**, e1010 (2020).
[43] M. Künzle, J. Mangler, M. Lach, and T. Beck, Nanoscale **10**, 22917 (2018).
[44] B. Alberts, *Essential cell biology* (Garland Science, New York; London, 2010).
[45] H. M. Huttanus and X. Feng, Biotechnol. J. (2017), 10.1002/biot.201700052.
[46] D. P. Patterson, P. E. Prevelige, and T. Douglas, ACS Nano **6**, 5000 (2012), http://dx.doi.org/10.1021/nn300545z.
[47] D. P. Patterson, B. Schwarz, R. S. Waters, T. Gedeon, and T. Douglas, ACS Chem. Biol. **9**, 359 (2014), http://dx.doi.org/10.1021/cb4006529.
[48] B. Worsdorfer, Z. Pianowski, and D. Hilvert, J. Am. Chem. Soc. **134**, 909 (2012), http://dx.doi.org/10.1021/ja211011k.
[49] F. Cai, S. L. Bernstein, S. C. Wilson, and C. A. Kerfeld, Plant Physiol. **170**, 1868 (2016).
[50] S. M. Lee, N. Bond, C. Callaway, B. Clark, E. Farmer, M. Mallard, and S. S. Jang, RSC Adv. **8**, 37866 (2018).
[51] S. Chakraborti, A. Korpi, M. Kumar, P. Stepien, M. A. Kostiainen, and J. G. Heddle, Nano Lett **19**, 3918 (2019).
[52] T. G. W. Edwardson and D. Hilvert, J Am Chem Soc **141**, 9432 (2019).
[53] G. D. Price and M. R. Badger, Can. J. Bot. **69**, 963 (1991).
[54] J. M. Shively, F. Ball, D. H. Brown, and R. E. Saunders, Science **182**, 584 (1973).
[55] J. M. Shively, F. L. Ball, and B. W. Kline, J. Bacteriol. **116**, 1405 (1973).
[56] C. V. Iancu, H. J. Ding, D. M. Morris, D. P. Dias, A. D. Gonzales, A. Martino, and G. J. Jensen, J. Mol. Biol. **372**, 764 (2007).
[57] C. V. Iancu, D. M. Morris, Z. Dou, S. Heinhorst, G. C. Cannon, and G. J. Jensen, J. Mol. Biol. **396**, 105 (2010).
[58] M. Sutter, B. Greber, C. Aussignargues, and C. A. Kerfeld, Science **356**, 1293 (2017).
[59] M. Sutter, T. G. Laughlin, N. B. Sloan, D. Serwas, K. M. Davies, and C. A. Kerfeld, Plant Physiol. **181**, 1050 (2019).
[60] J. D. Perlmutter, F. Mohajerani, and M. F. Hagan, eLife **5**, e14078 (2016).
[61] F. Mohajerani and M. F. Hagan, PLoS Comput. Biol. **14**, e1006351 (2018).
[62] F. Mohajerani, E. Sayer, C. Neil, K. Inlow, and M. F. Hagan, ACS Nano **15**, 4197 (2021), pMID: 33683101, https://doi.org/10.1021/acsnano.0c05715.
[63] J. P. Mahalik, K. A. Brown, X. Cheng, and M. Fuentes-Cabrera, ACS Nano **10**, 5751 (2016).
[64] E. J. Young, R. Burton, J. P. Mahalik, B. G. Sumpter, M. Fuentes-Cabrera, C. A. Kerfeld, and D. C. Ducat, Front. Microbiol. **8** (2017), 10.3389/fmicb.2017.01441.
[65] G. M. Rotskoff and P. L. Geissler, Proc. Natl. Acad. Sci. **115**, 6341 (2018).

[66] Y. Li, N. W. Kennedy, S. Li, C. E. Mills, D. Tullman-Ercek, and M. Olvera de la Cruz, ACS Cent Sci **7**, 658 (2021).
[67] L. Garcia-Alles, E. Lesniewska, K. Root, N. Aubry, N. Pocholle, C. Mendoza, E. Bourillot, K. Barylyuk, D. Pompon, R. Zenobi, D. Reguera, and G. Truan (2017).
[68] T. Chen and S. C. Glotzer, Physical Review E. **75**, 051504 (2007).
[69] R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, Phys. Rev. Lett. **90**, 248101 (2003).
[70] R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, and J. Rudnick, Proc. Natl. Acad. Sci. U. S. A. **101**, 15556 (2004).
[71] B. Berger, P. W. Shor, L. Tuckerkellogg, and J. King, Proc. Natl. Acad. Sci. U. S. A. **91**, 7732 (1994).
[72] R. Schwartz, R. L. Garcea, and B. Berger, Virology **268**, 461 (2000).
[73] D. Rapaport, Phys. Rev. E. **70**, 051905 (2004).
[74] H. D. Nguyen, V. S. Reddy, and C. L. Brooks, J. Am. Chem. Soc. **131**, 2606 (2009).
[75] O. M. Elrad and M. F. Hagan, Nano Lett. **8**, 3850 (2008).
[76] I. G. Johnston, A. A. Louis, and J. P. K. Doye, J. Phys.: Condens. Matter **22**, 104101 (2010).
[77] D. C. Rapaport, J. Phys.: Condens. Matter **22**, 104115 (2010).
[78] H. Nguyen and C. Brooks, Nano Lett. **8**, 4574 (2008).
[79] M. G. Mateu, Arch. Biochem. Biophys. **531**, 65 (2013).
[80] M. F. Hagan, Adv. Chem. Phys. **155**, 1 (2014).
[81] J. D. Perlmutter and M. F. Hagan, Annu. Rev. Phys. Chem. **66**, 217 (2015), http://dx.doi.org/10.1146/annurev-physchem-040214-121637.
[82] R. F. Bruinsma and W. S. Klug, Annu. Rev. Condens. Matter Phys. **6**, 245 (2015), https://doi.org/10.1146/annurev-conmatphys-031214-014325.
[83] M. F. Hagan and R. Zandi, Curr. Opin. Vir. **18**, 36 (2016).
[84] D. C. Rapaport, Phys. Rev. E **86**, 051917 (2012).
[85] D. C. Rapaport, J. E. Johnson, and J. Skolnick, Comput. Phys. Commun. **122**, 231 (1999).
[86] C. I. Mendoza and D. Reguera, eLife **9**, e52525 (2020).
[87] J. Wagner and R. Zandi, Biophys. J. **109**, 956 (2015).
[88] R. Schwartz, P. W. Shor, P. E. Prevelige, and B. Berger, Biophys. J. **75**, 2626 (1998).
[89] M. F. Hagan and D. Chandler, Biophys. J. **91**, 42 (2006).
[90] A. W. Wilber, J. P. K. Doye, A. A. Louis, E. G. Noya, M. A. Miller, and P. Wong, J. Chem. Phys. **127**, 085106 (2007).
[91] J. E. Baschek, H. C. R. Klein, and U. S. Schwarz, Bmc Biophysics **5** (2012), 10.1186/2046-1682-5-22.
[92] M. Castelnovo, T. Verdier, and L. Foret, Epl-europhys Lett **105** (2014), 10.1209/0295-5075/105/28006.
[93] M. Castelnovo, D. Muriaux, and C. Faivre-Moskalenko, New J. Phys. **15** (2013), 10.1088/1367-2630/15/3/035028, castelnovo, M. Muriaux, D. Faivre-Moskalenko, C.
[94] M. A. Boettcher, H. C. R. Klein, and U. S. Schwarz, Phys. Biol. **12** (2015), 10.1088/1478-3975/12/1/016014.
[95] S. N. Fejer, D. Chakrabarti, and D. J. Wales, ACS Nano **4**, 219 (2010).
[96] R. Kusters, H.-K. Lin, R. Zandi, I. Tsvetkova, B. Dragnea, and P. van der Schoot, J. Phys. Chem. B **119**, 1869 (2015).
[97] R. Zandi and P. van der Schoot, Biophys. J. **96**, 9 (2009).
[98] M. F. Hagan, Phys. Rev. E **77**, 051904 (2008).
[99] M. F. Hagan, J. Chem. Phys. **130**, 114902 (2009).
[100] V. A. Belyi and M. Muthukumar, Proc. Natl. Acad. Sci. U. S. A. **103**, 17174 (2006).
[101] R. Twarock, R. J. Bingham, E. C. Dykeman, and P. G. Stockley, Curr. Opin. Vir. **31**, 74 (2018).
[102] R. Zandi, B. Dragnea, A. Travesset, and R. Podgornik, Phys. Rep. (2020).
[103] S. Li, P. Roy, A. Travesset, and R. Zandi, Proc. Natl. Acad. Sci. U. S. A. **115**, 10971 (2018).
[104] S. Li, R. Zandi, A. Travesset, and G. M. Grason, Phys. Rev. Lett. (2019), 10.1103/PhysRevLett.123.145501, arXiv:1906.03301.
[105] P. van der Schoot and R. Zandi, J. Biol. Phys. **39**, 289 (2013).
[106] G. Erdemci-Tandogan, J. Wagner, P. van der Schoot, R. Podgornik, and R. Zandi, Phys. Rev. E **89**, 032707 (2014).
[107] A. Siber, R. Zandi, and R. Podgornik, Phys. Rev. E **81**, 051919 (2010).
[108] E. C. Dykeman, P. G. Stockley, and R. Twarock, J. Mol. Biol. **425**, 3235 (2013).
[109] N. Patel, E. C. Dykeman, R. H. A. Coutts, G. P. Lomonossoff, D. J. Rowlands, S. E. V. Phillips, N. Ranson, R. Twarock, R. Tuma, and P. G. Stockley, Proc. Natl. Acad. Sci. U.S.A. (2015), 10.1073/pnas.1420812112.
[110] E. C. Dykeman, P. G. Stockley, and R. Twarock, Proc. Natl. Acad. Sci. U. S. A. **111**, 5361 (2014).
[111] E. C. Dykeman, P. G. Stockley, and R. Twarock, Phys Rev E **87**, 022717 (2013).
[112] V. L. Morton, E. C. Dykeman, N. J. Stonehouse, A. E. Ashcroft, R. Twarock, and P. G. Stockley, J. Mol. Biol. **401**, 298 (2010).
[113] R. Zhang, E. Wernersson, and P. Linse, RSC Adv. **3**, 25258 (2013).
[114] R. Zhang and P. Linse, J. Chem. Phys. **140** (2014), 10.1063/1.4883056, zhang, Ran Linse, Per 1089-7690.
[115] R. Zhang and P. Linse, J. Chem. Phys. **138** (2013), 10.1063/1.4799243, zhang, Ran Linse, Per 1089-7690.
[116] D. G. Angelescu, J. Stenhammar, and P. Linse, J. Phys. Chem. B **111**, 8477 (2007).
[117] D. G. Angelescu and P. Linse, Phys. Rev. E **75**, 051905 (2007).
[118] S. Panahandeh, S. Li, and R. Zandi, Nanoscale **10**, 22802 (2018).
[119] S. Panahandeh, S. Li, L. Marichal, R. Leite Rubim, G. Tresset, and R. Zandi, ACS Nano **14**, 3170 (2020).
[120] F. Cai, Z. Dou, S. L. Bernstein, R. Leverenz, E. B. Williams, S. Heinhorst, J. Shively, G. C. Cannon, and C. A. Kerfeld, Life **5**, 1141 (2015).
[121] C. Fan, S. Cheng, Y. Liu, C. M. Escobar, C. S. Crowley, R. E. Jefferson, T. O. Yeates, and T. a. Bobik, Proc. Natl. Acad. Sci. U. S. A. **107**, 7509 (2010).
[122] J. C. Cameron, S. C. Wilson, S. L. Bernstein, and C. A. Kerfeld, Cell **155**, 1131 (2013).
[123] A. H. Chen, A. Robinson-Mosher, D. F. Savage, P. A. Silver, and J. K. Polka, PLoS One **8**, e76127 (2013).
[124] T. Wunder, S. L. H. Cheng, S.-K. Lai, H.-Y. Li, and O. Mueller-Cajar, Nat. Comm. **9**, 1 (2018).
[125] T. Chaijarasphong, R. J. Nichols, K. E. Kortright, C. F. Nixon, P. K. Teng, L. M. Oltrogge, and D. F. Savage, J. Mol. Biol. **428**, 153 (2016).
[126] L. M. Oltrogge, T. Chaijarasphong, A. W. Chen, E. R. Bolin, S. Marqusee, and D. F. Savage, Nature Structural & Molecular Biology , 1 (2020).
[127] J.-D. Rochaix, Cell **171**, 28 (2017).
[128] H. Wang, X. Yan, H. Aigner, A. Bracher, N. D. Nguyen, W. Y. Hee, B. Long, G. D. Price, F. Hartl, and M. Hayer-Hartl, Nature **566**, 131 (2019).
[129] S. He, H.-T. Chou, D. Matthies, T. Wunder, M. T. Meyer, N. Atkinson, A. Martinez-Sanchez, P. D. Jeffrey, S. A. Port, and W. Patena, Nat. Plants , 1 (2020).
[130] M. T. Meyer, A. K. Itakura, W. Patena, L. Wang, S. He, T. Emrich-Mills, C. S. Lau, G. Yates, L. C. Mackinder, and M. C. Jonikas, Science advances **6**, eabd2408 (2020).

[131] C. P. Brangwynne, Soft Matter **7**, 3052 (2011).
[132] R. W. Style, T. Sai, N. Fanelli, M. Ijavi, K. Smith-Mannschott, Q. Xu, L. A. Wilen, and E. R. Dufresne, Phys. Rev. X **8** (2018), 10.1103/PhysRevX.8.011028.
[133] W. M. Jacobs and D. Frenkel, Biophys. J. **112**, 683 (2017).
[134] W. M. Jacobs, Phys. Rev. Lett. **126**, 258101 (2021).
[135] S. L. Perry, Current Opinion in Colloid & Interface Science **39**, 86 (2019).
[136] I. Sanchez-Burgos, J. R. Espinosa, J. A. Joseph, and R. Collepardo-Guevara, Biomolecules **11** (2021), 10.3390/biom11020278.
[137] J. Zhou, P. Kondylis, D. G. Haywood, Z. D. Harms, L. S. Lee, A. Zlotnick, and S. C. Jacobson, Anal. Chem. **90**, 7267 (2018).
[138] R. B. Best, Curr Opin Struct Biol **42**, 147 (2017).
[139] G. L. Dignon, W. Zheng, R. B. Best, Y. C. Kim, and J. Mittal, Proc Natl Acad Sci U S A **115**, 9929 (2018).
[140] G. L. Dignon, W. Zheng, Y. C. Kim, R. B. Best, and J. Mittal, PLoS Comput Biol **14**, e1005941 (2018).
[141] G. L. Dignon, W. Zheng, and J. Mittal, Curr Opin Chem Eng **23**, 92 (2019).
[142] A. Ianiro, H. Wu, M. M. J. van Rijt, M. P. Vena, A. D. A. Keizer, A. C. C. Esteves, R. Tuinier, H. Friedrich, N. Sommerdijk, and J. P. Patterson, Nat Chem **11**, 320 (2019).
[143] A. E. Posey, A. S. Holehouse, and R. V. Pappu, Methods Enzymol **611**, 1 (2018).
[144] K. M. Ruff, R. V. Pappu, and A. S. Holehouse, Curr Opin Struct Biol **56**, 1 (2018).
[145] S. Alberti, A. Gladfelter, and T. Mittag, Cell **176**, 419 (2019).
[146] D. L. D. Caspar and A. Klug, Cold Spring Harbor Symp. Quant. Biol. **27**, 1 (1962).
[147] T. Douglas and M. Young, Nature **393**, 152 (1998).
[148] T. Douglas and M. Young, Science **312**, 873(3) (2006).
[149] C. P. Goodall, B. Schwarz, E. Selivanovitch, J. Avera, J. Wang, H. Miettinen, and T. Douglas, ACS Applied Bio Materials **4**, 8205 (2021), https://doi.org/10.1021/acsabm.1c00718.
[150] S. A. Robinson, E. C. Hartman, B. C. Ikwuagwu, M. B. Francis, and D. Tullman-Ercek, Biomacromolecules **21**, 4194 (2020), pMID: 32880435, https://doi.org/10.1021/acs.biomac.0c00987.
[151] I. J. Minten, L. J. A. Hendriks, R. J. M. Nolte, and J. J. L. M. Cornelissen, J. Am. Chem. Soc. **131**, 17771 (2009), http://dx.doi.org/10.1021/ja907843s.
[152] J.-K. Rhee, M. Hovlid, J. D. Fiedler, S. D. Brown, F. Manzenrieder, H. Kitagishi, C. Nycholat, J. C. Paulson, and M. G. Finn, Biomacromolecules **12**, 3977 (2011), http://dx.doi.org/10.1021/bm200983k.
[153] W. F. Rurup, J. Snijder, M. S. T. Koay, A. J. R. Heck, and J. J. L. M. Cornelissen, J. Am. Chem. Soc. **136**, 3828 (2014), http://dx.doi.org/10.1021/ja410891c.
[154] W. F. Rurup, F. Verbij, M. S. T. Koay, C. Blum, V. Subramaniam, and J. J. L. M. Cornelissen, Biomacromolecules **15**, 558 (2014), http://dx.doi.org/10.1021/bm4015792.
[155] N. H. Dashti, R. S. Abidin, and F. Sainsbury, ACS Nano **12**, 4615 (2018).
[156] M. Faulkner, L. S. Zhao, S. Barrett, and L. N. Liu, Nanoscale Res Lett **14**, 54 (2019).
[157] B. Smit, J. Chem. Phys. **96**, 8639 (1992), https://doi.org/10.1063/1.462271.
[158] J. N. C. LOPES, Mol. Phys. **96**, 1649 (1999), https://doi.org/10.1080/00268979909483108.
[159] V. Shen and J. Errington, J. Chem. Phys. (2005).
[160] S. Whitelam and R. L. Jack, Ann Rev Phys Chem **66**, 143 (2015).
[161] J. B. Parsons, S. Frank, D. Bhella, M. Liang, M. B. Prentice, D. P. Mulvihill, and M. J. Warren, Mol. Cell **38**, 305 (2010).
[162] S. Choudhary, M. B. Quin, M. A. Sanders, E. T. Johnson, and C. Schmidt-Dannert, PLoS One **7**, e33342 (2012).
[163] J. K. Lassila, S. L. Bernstein, J. N. Kinney, S. D. Axen, and C. A. Kerfeld, J. Mol. Biol. **426**, 2217 (2014).
[164] G. N. Condezo and C. San Martin, PLoS Pathog. **13** (2017).
[165] C. Aussignargues, B. C. Paasch, R. Gonzalez-Esquer, O. Erbilgin, and C. A. Kerfeld, Communicative & integrative biology **8**, e1039755 (2015).
[166] P. Ryan, T. J. Forrester, C. Wroblewski, T. M. Kenney, E. N. Kitova, J. S. Klassen, and M. S. Kimber, J. Biol. Chem. **294**, 2593 (2019).
[167] C. a. Kerfeld and O. Erbilgin, Trends Microbiol. **23**, 22 (2015).
[168] M. Sutter, M. Faulkner, C. Aussignargues, B. C. Paasch, S. Barrett, C. A. Kerfeld, and L.-N. Liu, Nano Lett. **16**, 1590 (2016), pMID: 26617073, http://dx.doi.org/10.1021/acs.nanolett.5b04259.
[169] J. D. Perlmutter, C. Qiao, and M. F. Hagan, eLife **2**, e00632 (2013).
[170] J. D. Perlmutter, M. R. Perkett, and M. F. Hagan, J. Mol. Biol. (2014), 10.1016/j.jmb.2014.07.004.
[171] J. D. Perlmutter and M. F. Hagan, J. Mol. Biol. (2015), 10.1016/j.jmb.2015.05.008.
[172] D. J. Wales, Phil. Trans. R. Soc. A **363**, 357 (2005).
[173] S. N. Fejer, T. R. James, J. Hernandez-Rojas, and D. J. Wales, Phys. Chem. Chem. Phys. **11**, 2098 (2009).
[174] T. Ruiz-Herrero and M. F. Hagan, Biophys. J. , 1 (2015).
[175] J. A. Anderson, C. D. Lorenz, and A. Travesset, J. Comput. Phys. **227**, 5342 (2008).
[176] T. D. Nguyen, C. L. Phillips, J. A. Anderson, and S. C. Glotzer, Comput. Phys. Commun. **182**, 2307 (2011).
[177] J. Anderson, (URL), http://codeblue.umich.edu/hoomd-blue/doc/page_units.html.