Microcompartment assembly around multicomponent fluid cargoes

Lev Tsidilkovski, Farzaneh Mohajerani, and Michael F Hagan
Martin A. Fisher School of Physics, Brandeis University, Waltham, MA, 02453

(*Electronic mail: hagan@brandeis.edu)

(Dated: 14 May 2022)

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-shell protein and shell protein-cargo component interactions that help drive assembly and encapsulation also influence cargo composition within certain parameter regimes. These behaviors are governed by a 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 efforts in synthetic biology to colocalize alternative sets of molecules within microcompartments to accelerate specific reactions. More broadly, the results suggest that coupling between self-assembly and multicomponent liquid-liquid phase separation may play a role in the 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 non-native 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 a reaction cascade, while the formation of cellular signaling complexes requires colocalizing signaling proteins and their regulatory ligands[11,21,44]. Similarly, applications in synthetic biology 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 re-engineered 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 shown 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 β-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 CcmM in β-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 composition, structure, and function of the RuBisCO complex within carboxysomes has strong similarities to the pyrenoid, which is

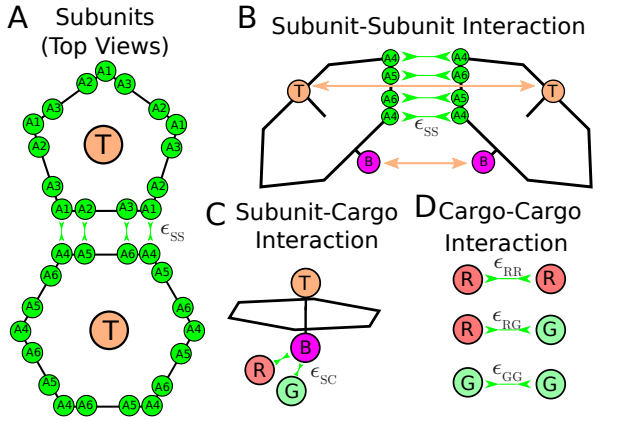

FIG. 1. **Description of the model. (A)** Each shell subunit contains 'Attractors' (green circles) on the perimeter, that define the shape of the subunit. **(B)** Attractive interactions between Attractors drive subunit dimerization. Complementary pairs of Attractors are indicated by green arrows in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

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 attractions[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, and of the 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 affinities 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.

shell-shell affinity parameter	ε_{SS}/k_BT	2.0	3.5
hex-hex affinity	g_{hh}/k_BT	4	-0.5
hex-hex dissociation constant	K_d^{hh}/μM	7×10^7	6×10^5
free energy per subunit in shell*	$\Delta g_{shell}/k_BT$	-15	-27
dissociation constant from shell*	K_d^{shell}/μM	0.4	1×10^{-6}

TABLE I. **The shell-shell binding affinity for the default affinity parameter values**. The binding free energy and dissociation constant for hexamer dimerization (g_{hh}, from Eq. (A1)), and from the shell-shell and shell-cargo interactions in a complete shell complex with $\varepsilon_{SC} = 6k_BT$ (from Eq. (A3)). The standard state concentration is 1 M. *Does not include cargo-cargo contributions, which will increase (decrease) the effective shell dissociation constant K_d^{shell} below (above) the threshold for cargo vapor-liquid coexistence (see Perlmutter et al.[60])

.

II. RESULTS AND DISCUSSION

A. Model overview

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]).

Interaction parameters. Efficient assembly and cargo loading generically require 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[60]. While in some systems one or more of these interactions can be mediated by auxiliary proteins (*e.g.* the CCmM and CCmN proteins in β-carboxysomes), for simplicity we model all interactions as direct pair interactions in this work.

We have explored assembly dynamics over a range of the simulation control parameters that most strongly affect the amount and composition of encapsulated cargo — these are the interaction potential well-depths that control the binding free energy (affinity) between each type of cargo pair, ε_{GG}, ε_{RR}, and ε_{RG}, as well as between shell-cargo and shell-shell pairs ε_{SC} and ε_{SS} (see Appendix B). To minimize the number of parameters, we assume that both cargo species have the

same binding affinity to shell subunits; i.e., ε_{SC} is equal for R and G particles. For convenience, throughout this article we refer to the well-depths as 'affinities'. We approximately relate ε_{SC} and ε_{SS} to the actual affinities in Appendix A and Table I. Throughout this article, all energy values are given in units of the thermal energy, k_BT, and all lengths are given in units of the cargo diameter, r_*, which translates to $r_* \approx 13\,\text{nm}$ for carboxysomes (see section IV B).

We consider ranges of affinity parameters relevant to biological microcompartments (see section IV B for details). Except where mentioned otherwise, we consider two shell-shell affinities ($\varepsilon_{SS}/k_BT = 2, 3.5$), which are respectively below and above the transition between one-step and two-step assembly pathways (discussed in section II C). Both cases are below the threshold for assembly of empty shells (containing no cargo). Importantly, dimerization is very unfavorable, but the assembled shells are stable (see Table. I). The unfavorable dimerization affinities reflect the essential role that cargo plays in assembly at these shell-shell affinities, the cooperative nature of assembly, and the fact that a significant nucleation barrier is essential to avoid kinetic traps[60,80,89,157]. We primarily focus on a representative shell-cargo affinity $\varepsilon_{SC} = 6k_BT$, which corresponds to a free energy per shell subunit of $g_{SC} \approx -14k_BT$. Our simulations span the range of cargo-cargo affinities in the vicinity of vapor-liquid coexistence ($\varepsilon^{vl} \approx 1.3k_BT$ at our simulated cargo concentration), but remain below the threshold for cargo crystallization at $\varepsilon^{fs} \approx 3k_BT$. With ε^{vl} as a reference point, we denote cargo-cargo affinities as 'weak' for $\varepsilon_{RR} \lesssim 1.1k_BT$, 'moderate' for $1.2 \lesssim \varepsilon_{RR} \lesssim 1.4$, and 'strong' for $\varepsilon_{RR} \gtrsim 1.5k_BT$ (with analogous definitions for ε_{GG} and ε_{RG}).

Our shell subunit and cargo concentrations map to approximately 1 μM and 10 μM respectively, which are reasonable based on quantities found within bacterial cells.

Simulated systems. 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 affinities, 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 affinities 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 begin by describing large simulations of cargo in the absence of shell subunits to approximate a bulk system (Section II B). We then describe the dynamical assembly of shells and cargo particles (Sections II C and II D). We compare the results of these finite-time dynamical simulations against simulations with assembled but permeable shells, which allow the encapsulated cargo to equilibrate with the bulk.

B. 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.*[158–160]. In particular, we consider equal stoichiometry between R and G molecules and cargo-cargo affinities ($\varepsilon_{ij} < 1.8 \quad \text{for } ij = \text{RR}, \text{GG}, \text{RG}$) that maintain the system in vapor and/or liquid phases, but are below the crystallization threshold (the vapor-fluid and crystallization transitions for a single species occur at at $\varepsilon^{vl} = 1.3$ and $\varepsilon^{fs} \approx 3$). 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 de-mixing (coexistence between three phases — a dilute phase, a dense phase rich in R, and a dense phase rich in G). We consider weak, moderate, and strong R-G affinities: $\varepsilon_{RG} = 1.0, 1.3, 1.6$. For each of these, Fig. 2 summarizes the bulk phase behaviors as a function of the R-R and G-G affinities ε_{RR} and ε_{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_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_{UL}. Here we define the mean fraction of unlike neighbors in the first solvation shell around each particle as: $\bar{f}_{UL} \equiv \frac{n_{RG}}{n_{RR}+n_{GG}n_{RG}}$, 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 $f_{rand} = 2f_R(1-f_R)$, so the plots indicate the difference between randomness and the results: $f_{UL} \equiv (\bar{f}_{UL} - f_{rand})$.

For weak R-G affinities ($\varepsilon_{RG} = 1.0$, 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 affinities, leading to formation of one globule with an interface separating the two phases. The interface between these two phases becomes more diffuse with decreasing ε_{RR} and ε_{GG}, and/or increasing ε_{RG} because the surface tension between the two domains decreases. For moderate R-G affinities $\varepsilon_{RG} = 1.3$, 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 R-G affinities ($\varepsilon_{RG} = 1.6$, 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 ε_{RR} and ε_{GG}. Since the primary driving force for phase separation is ε_{RG} in this case, fluctuations in f_R and f_{UL} are smaller than for the other systems.

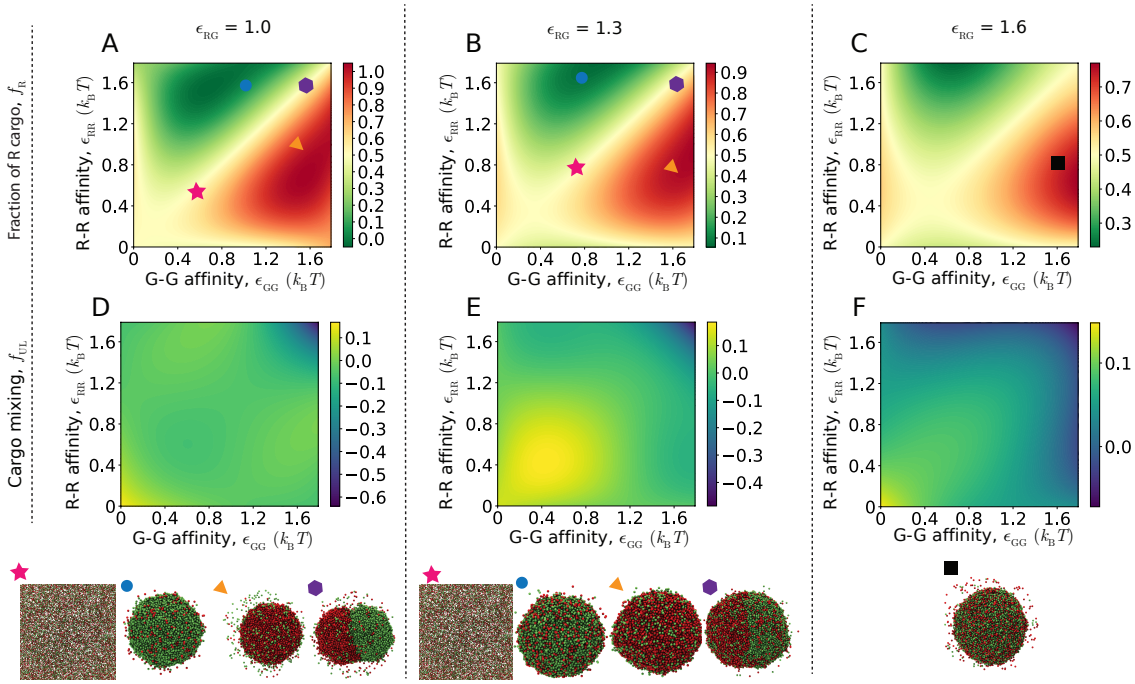

FIG. 2. **The bulk phase behavior of cargo particles (in the absence of shell subunits) as a function of their binding affinities. (A-C)** The fraction of R particles in the high-density phase, f_R, as a function of the R-R and G-G affinities for different R-G affinities: (A) weak $\varepsilon_{RG} = 1.0$, (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 ε_{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.

C. Assembly pathways

We now consider how cargo and shell affinities 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 affinities[60]. In particular, at our simulated volume fraction of 0.003, weak cargo-cargo affinities ($\varepsilon_{RR} < 1.7$) and/or the higher shell-shell affinities ($\varepsilon_{SS} > 2.5$) lead to one-step pathways in which shell assembly is concomitant with cargo coalescence, while strong cargo-cargo and lower shell-shell affinities ($\varepsilon_{RR} > 1.7$ and $\varepsilon_{SS} < 2.5$) 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 shell-shell affinity $\varepsilon_{SS} = 3.5$, leading to one-step assembly pathways. In Fig. 3A the cargo-cargo affinities are uniform and relatively strong $\varepsilon_{RR} = \varepsilon_{GG} = \varepsilon_{RG} = 1.7$, 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 affinities ($\varepsilon_{RR} = 1.7$) are stronger than the G-G and R-G affinities ($\varepsilon_{GG} = \varepsilon_{RG} = 1.3$), leading to simultaneous assembly and coalescence of a nearly pure R domain.

Fig. 3C-E show trajectories with a weaker shell-shell interaction of $\varepsilon_{SS} = 2.0$, for which shell nucleation is slow and we observe two-step assembly pathways. In Fig. 3C, the cargo-cargo affinities 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 (Multimedia view).

We observe a more interesting coupling between cargo composition and shell assembly for strong R-R and G-G affinities ($\varepsilon_{RR} = \varepsilon_{GG} = 1.7$) and moderate R-G affinities $\varepsilon_{RG} = 1.3$ (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 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. For the shell to close, curvature of the assembling shell must first generate the formation

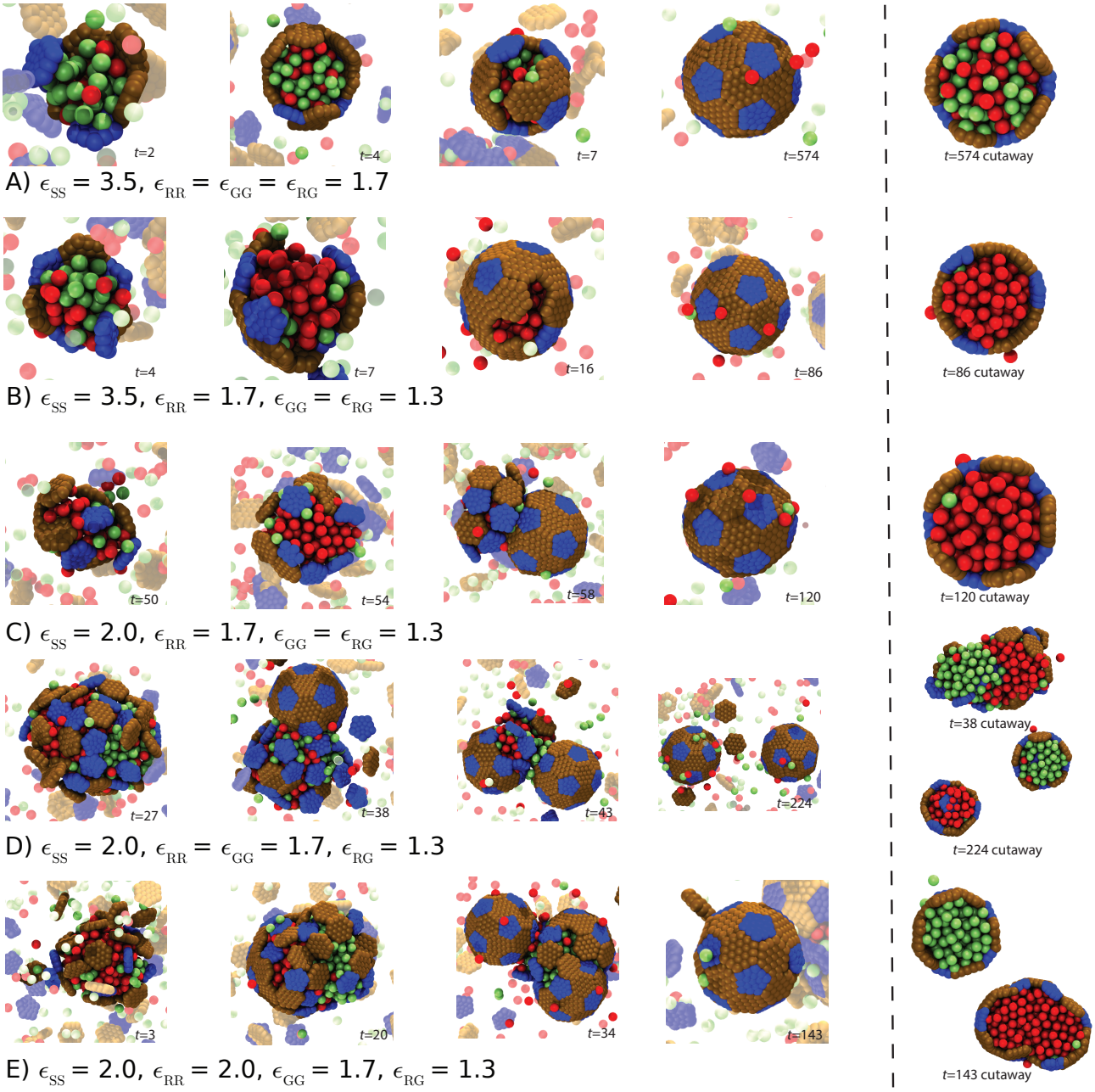

FIG. 3. **Assembly pathways.** Snapshots from simulation trajectories illustrate the classes of assembly pathways discussed in the text. In each row, the first four snapshots show frames at indicated time points (in units of the nondimensional time τ), while the last snapshot is a cutaway view of the shell to show the encapsulated cargo. The parameters characterizing binding affinities are listed under each row (shell-shell, ε_{SS}; shell-cargo, ε_{SC}; R-R and G-G, ε_{RR} and ε_{GG}; and R-G ε_{RG}. **(A)** All affinities are relatively strong, driving rapid and simultaneous coalescence of both cargo types and shell assembly (one-step assembly). The encapsulated cargo is uniformly mixed. **(B)** At $\varepsilon_{SS} = 3.5 k_B T$, high ε_{RR} but moderate ε_{GG} and ε_{RG} lead to one-step assembly and coalescence of R cargo, with G cargo essentially excluded from the shell. **(C)** At $\varepsilon_{SS} = 2 k_B T$, strong ε_{RR} but moderate ε_{GG} and ε_{RG}, lead to two-step assembly around an almost pure R domain. **(D)** Strong ε_{GG} and ε_{RR} and moderate ε_{RG} 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 ε_{RR}, strong ε_{GG}, and moderate ε_{RG} lead to two-step assembly around an almost pure G domain. The strength of the ε_{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, corresponding to each of the respective trajectories are included in the SI (Multimedia view).

of a 'neck' of cargo, i.e. a narrow region of cargo particles that connects a shell assembly to the remainder of the globule or to other assemblies (*e.g.* see Fig. 3D at $t = 43$ and Fig. 3E at $t = 34$). Then, additional shell subunits must assemble to close the shell, causing the neck to break and the completed shell to separate from the remainder of the globule (Fig. 3D at $t = 224$ and Fig. 3E at $t = 143$). The processes of forming and then breaking the neck increase the total cargo interfacial area and correspondingly reduce the number of cargo-cargo interactions. Thus, they are 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 affinities 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.

Shell closure locks the encapsulated cargo composition out of equilibrium. After a shell closes, fluctuations of subunits large enough to allow cargo particles to escape are extremely rare because all subunits have their maximum number of interactions. Thus, the cargo composition is essentially fixed once the shell completes. However, as noted in Section II D the cargo remains fluid in the range of cargo-cargo affinities that we focus on, and thus the cargo can spatially reorganize inside of the shell even after its completion.

D. 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 affinities $\varepsilon_{SC} = 6.0$ and the higher shell-shell affinities $\varepsilon_{SS} = 3.5$. 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 R-G affinities, we present most results in this figure with varying R-R and G-G affinities at fixed $\varepsilon_{RG} = 1.3$. 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 ε_{RR} and ε_{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]. Importantly though, N_C does not reach crystalline density until cargo-cargo affinities nearly reach the bulk crystallization threshold $\varepsilon^{fs} \approx 3\,k_BT$. Ref. [60] found that well-formed shells cannot assemble when the cargo crystallizes, because the solid cargo cannot be deformed by assembly subunits, thus preventing shell closure. They also measured radial distribution functions for a single cargo species

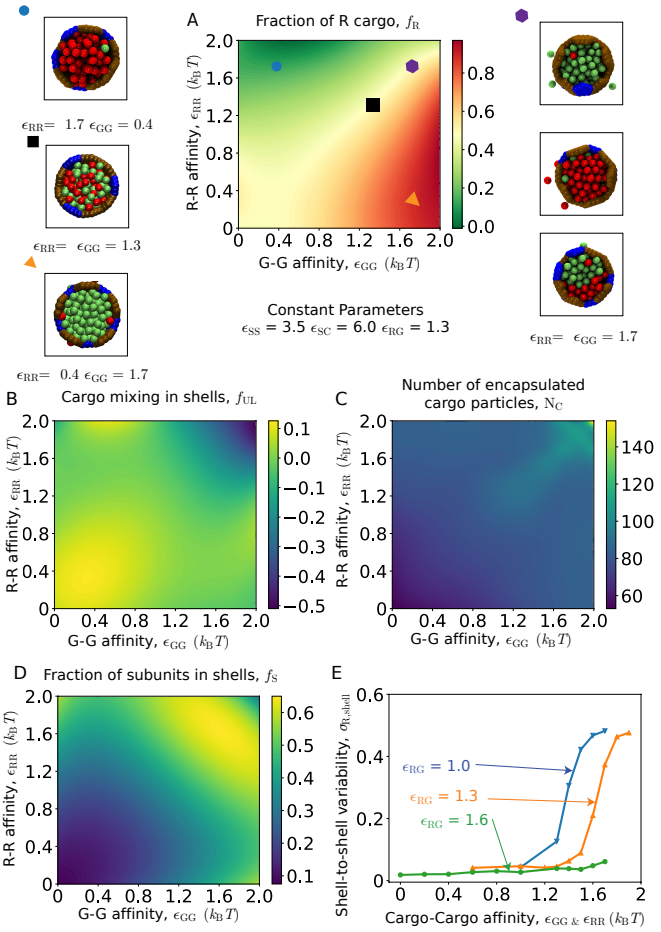

FIG. 4. **Effect of cargo-cargo affinities on encapsulation for one-step assembly pathways. (A-C)** Characterization of the encapsulated cargo as a function of the G-G and R-R affinities, for fixed R-G affinity $\varepsilon_{RG} = 1.3$. **(A)** The fraction of R particles averaged over all shells f_R 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 R-R and G-G affinities for three indicated values of ε_{RG}. 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 R-R and G-G 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 $\varepsilon_{SS} = 3.5$.

system, which showed that confinement by the shell imposes layering on the encapsulated cargo due to steric effects, but that the cargo remains fluid. The same effect is present for the bicomponent system studied in the present work, since the two cargo species differ only in their cargo-cargo affinities.

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 ε_{GG} and ε_{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. Comparing Figs. 4A,B with results from the bulk cargo system in the absence of microcompartment shells (Figs. 2B,E), shows that the composition and mixing of the two cargo species within shells is similar to the case of bulk cargo coalescence for this parameter regime. However, we will see below that shell assembly has a more significant effect on cargo properties in other parameter regimes.

Fig. 5 compares these results to two-step assembly pathways that occur for lower shell-shell affinities, $\varepsilon_{SS} = 2.0$. Since we observe significant assembly only for a narrow range of strong cargo-cargo affinities (Fig. 8, discussed next), we present these results as a function of ε_{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 D 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 ε_{RR} and ε_{GG} to ε_{RG}, we present the results as a function of equal R-R and G-G interactions ($\varepsilon_{RR} = \varepsilon_{GG}$), for weak, moderate, and strong ε_{RG}. In each case, as the R-R and G-G affinities exceed ε_{RG} 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 C 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. 8A show the fraction of subunits

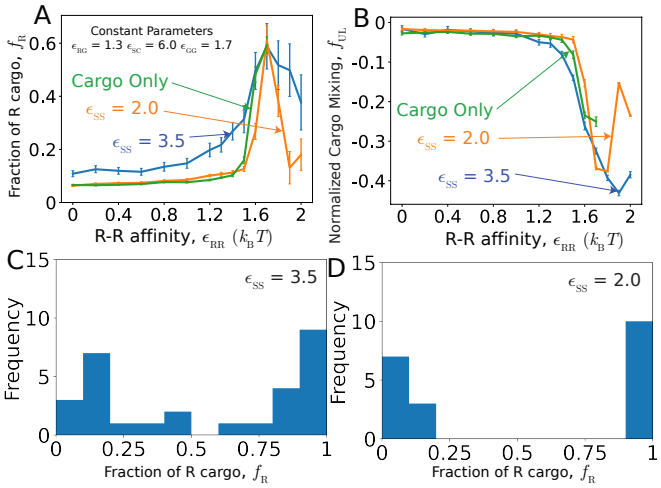

FIG. 5. **Comparison of cargo encapsulation for one-step and two-step pathways.** The fraction of R cargo in completed shells f_R **(A)** and the fraction of unlike cargo particles in first solvation shells f_{UL} **(B)** are shown as a function of ε_{RR} for the lower and higher shell-shell affinities $\varepsilon_{SS} = 2.0$, 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.

in complete shells f_s as a function of cargo-cargo affinities for $\varepsilon_{SS} = 2.0$ and $\varepsilon_{SS} = 3.5$ respectively. We observe that assembly occurs over a broad range of cargo-cargo affinity values for the higher shell-shell affinity ($\varepsilon_{SS} = 3.5$, Fig. 4D), but for the lower shell-shell affinity ($\varepsilon_{SS} = 2.0$, Fig. 8) significant assembly requires at least one strong cargo-cargo affinity; *i.e.*, $\varepsilon_{ij} \geq 1.7$ for at least one of $ij \in \{RR, GG, RG\}$. 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 (ε_{SC} and ε_{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[157].

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 affinities). Fig. 7 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 affinities, 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 ε_{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 attractive 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 affinities 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 affinities (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 ε_{RR} because the strong cargo-cargo affinities 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 ε_{RR}. This is partially an equilibrium effect due to the increasing drive for cargo segregation, especially as the R-R and G-G affinities exceed the R-G 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 C and Fig. 3D.

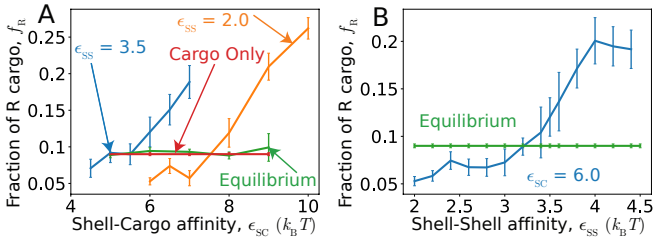

FIG. 6. **Effects of shell-cargo and shell-shell affinity on cargo composition reveal thermodynamic and kinetic influences on encapsulation.** **(A)** The fraction of encapsulated R particles f_R as a function of shell-cargo affinity ε_{SC}, for shell-shell affinities $\varepsilon_{SS} = 2.0$, and $\varepsilon_{SS} = 3.5$. 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 ε_{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 ε_{SS}.

Due to interfacial tension, the boundary between the two domains provides a 'weak point' for insertion of shell subunits, and thus shell closure tends to occur in the vicinity of the interface. The closing shell becomes enriched in the cargo species that dominates the domain on the same side of the interface as the assembling shell; 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 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 affinities, the encapsulated fraction of R particles is significantly smaller than the equilibrium value, but that it monotonically increases with ε_{SC}. The trend is the same for both shell-shell affinities, except that higher shell-cargo affinities are required to observe assembly for lower shell-shell affinities. Similarly, Fig. 6B shows that f_R increases monotonically with shell-shell affinity. These trends can be understood as follows. At low ε_{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 ε_{SC} or ε_{SS} increases, the strong shell-cargo affinities 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. To elucidate mechanisms that control the amount and composition of cargo encapsulation, we compared these results against simulations of cargo coalescence in the absence of microcompartment assembly.

Our simulation results show that the binding affinities among the different cargo species are the strongest determinant of the assembly pathway, as well as the amount, composition, and spatial organization of cargo within assembled shells. Analogous to the case of encapsulation of a single component cargo[57,60,61,120,122,123,161], the simulations exhibit two broad classes of assembly pathways. Relatively strong cargo-cargo affinities and low shell-shell affinities lead to two-step assembly, in which the cargo first forms a globule, followed by adsorption and assembly of the shell around the globule. Weaker cargo-cargo affinities and higher shell-shell affinities drive one-step assembly, in which cargo coalescence and shell assembly occur simultaneously. However, the bicomponent cargo allows for more diverse assembly pathways and morphologies, depending on when and if the two components mix within the encapsulated globule. In particular, we observe assembly pathways ranging from coalescence and encapsulation of a nearly pure globule of a single cargo species, to coalescence of a phase-separated globule with each phase becoming encapsulated in different shells, to encapsulation of a globule consisting of a uniform mixture of both species. Correspondingly, the assembled shells can contain a variety of cargo compositions and morphologies, ranging from highly enriched in a single species to nearly uniform mixing of the two species. In regimes where both cargo species are encapsulated, the two species may be uniformly mixed within each shell, or separated, with each individual shell containing an encapsulated cargo that is nearly pure in one or the other species.

While the relative strengths of the different cargo-cargo binding affinities (ε_{RR}, ε_{GG}, and ε_{RG}) determine cargo composition and mixing, the net cargo-cargo affinity most strongly influences the amount of cargo that is encapsulated. In particular, for low cargo-cargo affinities we observe poor cargo loading (*e.g.* shells that are nearly empty except for a layer of cargo at the surface). Further, depending on the strength of the subunit-subunit affinities, there is a threshold value of cargo-cargo affinities below which assembled microcompartments are either unstable or fail to nucleate on relevant timescales (Figs. 4D and 8).

While cargo composition and mixing within assembled shells is qualitatively similar to the behavior of cargo coalescence in bulk (in the absence of microcompartments) in many parameter regimes, our simulations also identify parameter regimes in which the shell-cargo and shell-shell binding affinities can be manipulated to control cargo encapsulation (*e.g.* Fig. 5 and Fig. 6). 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 affinities 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 the process of 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 (Fig. 2).

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[162–165]. Similarly, cargo-cargo affinities may be tuned through mutagenesis, but they 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[31,36,39,121,122,126,166,167].

Outlook. It will be interesting to extend our model to explicitly incorporate scaffold-mediated cargo-cargo and shell-cargo interactions, for example as Mohajerani *et al.*[62] did for a single-component cargo. More broadly, the predictions from our models can be extended beyond microcompartments, for example to re-engineered viral capsids[46–48,51,147–155] or synthetic capsids constructed via protein engineering (*e.g.*[168–170]) or DNA origami[171] to assemble around specific cargoes. In some of these cases, the shell-cargo interactions are driven by 'scaffolds' attached to the inner surface of the capsid. For example, many virus capsid proteins have flexible peptide tails that drive RNA encapsulation[172–174], while Sigl et al.[171] designed DNA origami capsid subunits with single-stranded DNA oligomers that bound to complementary strands on cargo particles. Previous simulations of microcompartments[62] and virus assembly[106,175–178] suggest that the properties of such scaffolds can be designed to manipulate the spatial organization of the cargo. Finally, it will be important to allow for shell assembly into different geometries, to study coupling between shell size and cargo composition.

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.[18,19,179]). Experiments suggest that these oligomers are the basic assembly units, meaning that smaller protein complexes do not contribute significantly to the assembly process[19,180]. 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[76,79,80,83,102,175–177,181–183].

Shell-shell interactions. In our model, shell subunits interact through both repulsive and attractive forces. The repulsions consist of 'Top' (T) pseudoatoms that exist above the plane of the hexamer and 'Bottom' (B) pseudoatoms that exist below the plane. More specifically, 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. B3). The Top-Top diameters and repulsion strengths set the shell spontaneous curvature and bending modulus (see section B).

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. (B4) in section B). 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, $\varepsilon_{\mathrm{SS}}$.

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 subunit 'B' 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[31,39,121,122,166]. 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. Mapping simulation parameters to biological microcompartments

Relevant ranges of affinities. Although affinities among microcompartment components have not been measured, we estimate relevant ranges of our control parameters based on experimental observations of carboxysomes as follows.

Shell-shell affinities: Because empty carboxysome shells are not typically observed in cells (for wild-type carboxysomes) we have focused on shell-shell affinities that are below the threshold for empty shell assembly ($\varepsilon_{\mathrm{SS}} \approx 4k_{\mathrm{B}}T$). Further, experimental observations suggest that α-carboxysomes assemble by one-step assembly pathways[57,120,161], whereas β-carboxysomes[122,123] and many other microcompartments[23,179] assemble by two-step pathways. Therefore, except where mentioned otherwise, we consider two cases ($\varepsilon_{\mathrm{SS}}/k_{\mathrm{B}}T = 2, 3.5$), which are respectively below and above the transition between one-step and two-step assembly pathways (discussed in section II C).

Shell-cargo affinities. Since shell subunits robustly assemble around RuBisCO in both α and β carboxysomes, we have focused on shell-cargo affinities that are strong enough to observe this ($\varepsilon_{\mathrm{SC}} = 6k_{\mathrm{B}}T$, free energy $g_{\mathrm{SC}} \approx -14k_{\mathrm{B}}T$) except where mentioned otherwise.

Cargo-cargo affinities: For the case of α-carboxysomes, the RuBisCO does not undergo phase separation in the absence of shell proteins at physiological conditions[57,120,161], but can phase separate without shell proteins under non-physiological conditions[126,128]. Thus, we infer that the effective cargo-cargo affinity for α-carboxysomes is near but below the threshold for phase separation, which corresponds to a cargo-cargo affinity of $\varepsilon^{\mathrm{vl}} \approx 1.3k_{\mathrm{B}}T$ at our simulated cargo concentration. In contrast, for β-carboxysomes the RuBisCO undergoes phase separation before shell assembly, suggesting a cargo-cargo affinity above $\varepsilon^{\mathrm{vl}}$, but below crystallization ($\varepsilon^{\mathrm{fs}}$). Our simulations span this range.

Length scales. Although the model is designed to be generic, we can *approximately* map it to carboxysomes by setting the cargo diameter (the unit length scale in the model) to that of the RuBisCO holoenzyme, implying $r_* \approx 13$ nm. However, to enable tractable simulation of long assembly timescales, we have set the subunit/cargo size ratio to be larger than in carboxysomes — model subunits have a side length of r_* and are thus about three times larger than carboxysome hexamers (side-length ≈ 4 nm). See Ref. [62] for further dis-

cussion.

C. Simulations and systems

We performed simulations using the Langevin dynamics algorithm in HOOMD (which uses GPUs to accelerate molecular dynamics simulations[184]), and periodic boundary conditions in a cubic box to represent a bulk system. The subunits were modeled as rigid bodies[185]. Each simulation was performed in the NVT ensemble, using the HOOMD fundamental units[186], 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. With $r_* = 13\,\mathrm{nm}$, this corresponds to concentrations of $c_\mathrm{p} = 0.6$, $c_\mathrm{h} = 1$, and $c_\mathrm{c} = 10\,\mu\mathrm{M}$ for pentamers, hexamers, and cargo.

We performed most simulations for 2×10^6 timesteps; simulations in the two-step assembly regime were performed for longer, 2×10^7 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 affinities. 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 2×10^7 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 ε_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 2×10^6 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 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.

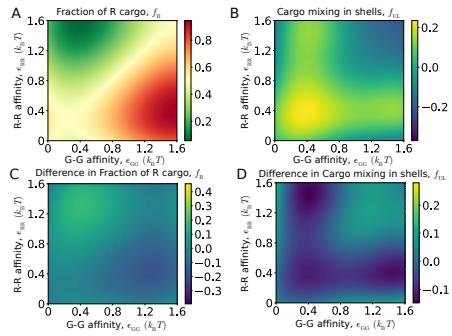

FIG. 7. **Comparison of encapsulated cargo between equilibrium and dynamical simulations. (A,B)** The fraction of R particles **(A)** and mixing **(B)** are shown as a function of R-R and G-G affinities. Other parameters are $\varepsilon_\mathrm{RG} = 1.3$ and $\varepsilon_\mathrm{SC} = 6.0$, as in Fig. 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 $\varepsilon_\mathrm{SS} = 3.5$) and equilibrium simulations (panels (A,B) of this figure).

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.

ACKNOWLEDGMENTS

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.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Open Science Framework at http://doi.org/10.17605/OSF.IO/SKZ7F

Appendix A: Shell-shell and shell-cargo binding free energy values

We estimated the subunit-subunit binding free energy values as functions of the well-depth parameter ε_SS by measuring

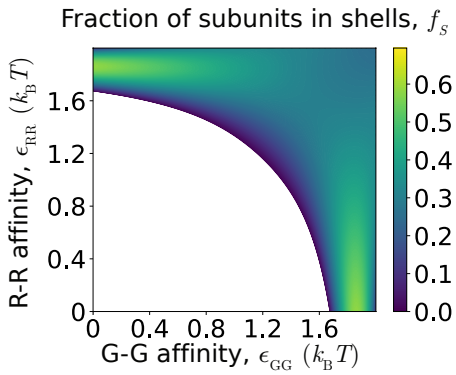

FIG. 8. Fraction of shell subunits in complete shells in the two-step assembly pathway regime, $\varepsilon_{SS} = 2.0$ and $\varepsilon_{RG} = 1.3$.

the dimerization equilibrium constant in simulations of subunits only capable of forming dimers. For both pentamer-hexamer and hexamer-hexamer dimers, we obtain binding free energies which are linear functions of the well-depth ε_{SS}[60]. We interpret the y-intercept as the binding entropy. Setting $r_* = 13\,nm$ and using the conventional standard state concentration of 1 M (rather than r_*^3 as in Ref. [60]) gives

$$\begin{aligned} g_{ph} =& \varepsilon_{ph}\varepsilon_{SS} - T s_{pent} \\ \varepsilon_{ph} =& -2.95; \quad s_{pent} = -10.0 k_B \\ g_{hh} =& \varepsilon_{hh}\varepsilon_{SS} - T s_{hex} \\ \varepsilon_{hh} =& -3.15; \quad s_{hex} = -10.5 k_B \end{aligned} \tag{A1}$$

where g_{ph} and g_{hh} are the binding affinities for pentamer-hexamer and hexamer-hexamer pairs. The dissociation constants are then given by $K_d^{ph} = \exp(g_{ph}/k_B T)$ and $K_d^{ph} = \exp(g_{ph}/k_B T)$. Table I shows the results for hexamer-hexamer dimerization, the values for pentamer-hexamer association are approximately the same. Similarly, we estimated the shell subunit adsorption free energy by performing simulations of subunits which cannot assemble (ε_{SS}=0) in the presence of a cargo globule as[60]

$$g_{SC} = -1.17\varepsilon_{SC} + 0.093 k_B T. \tag{A2}$$

For the free energy of shell assembly, we consider a shell comprised of n_{pent}=12 pentamers and n_{hex} hexamers, which have n_{ph} pentamer-hexamer contacts with binding energy ε_{ph} and n_{hh} hexamer-hexamer contacts with energy ε_{hh}. For our T=3 model, n_{hex}=20, n_{ph}=60, and n_{hh}=30. The assembly free energy is then given by

$$\begin{aligned} \Delta G_{shell} =& n_{ph}\varepsilon_{ph} + n_{hh}\varepsilon_{hh} \\ & - T\left(n_{pent}s_{pent} + n_{hex}s_{hex} + s_{config}\right) \\ & + g_{SC}(n_{pent} + n_{hex}), \end{aligned} \tag{A3}$$

where s_{config} accounts for the configurational entropy associated with subunit and shell symmetries: $s_{config} = k_B \ln(5^{n_{pent}} 3^{n_{hex}}/60)$, and γ and A_C are the surface tension and area of the cargo globule. Finally, we define an effective dissociation constant for the complex as $K_d^{shell} \equiv k_B T \ln \Delta g_{shell}$ with $\Delta g_{shell} = \Delta G_{shell}/(n_{pent} + n_{hex})$ the free energy per subunit averaged over pentamers and hexamers (see Table I). However, this calculation is only approximate since it neglects changes in binding entropy between dimerization and the full complex. Further, to simplify the calculation and focus on the shell subunits, we have neglected cargo-cargo contributions to the complex stability. This provides a reasonable estimate for cargo-cargo affinities near liquid-vapor coexistence, but over/under estimates the shell stability for weaker/stronger cargo-cargo affinities. For a full analysis including cargo-cargo contributions, see Perlmutter et al.[60].

Appendix B: 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:

$$\begin{aligned} U = & \sum_{\text{shell } i} \sum_{\text{shell } j<i} U_{SS} + \sum_{\text{cargo } i} \sum_{\text{cargo } j<i} U_{CC} \\ & + \sum_{\text{shell } i} \sum_{\text{cargo } j} U_{SC} \end{aligned} \tag{B1}$$

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 α, 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:

$$\begin{aligned} U_{SS}(\{\mathbf{a}_{i\alpha}\}, \mathbf{t}_i, \mathbf{a}_j, \mathbf{t}_j) = & \varepsilon_{SS} Ł\left(\left|\mathbf{t}_i - \mathbf{t}_j\right|, \sigma_{t,ij}\right) \\ & + \varepsilon_{SS} Ł\left(\left|\mathbf{b}_i - \mathbf{b}_j\right|, \sigma_b\right) \\ & + \varepsilon_{SS} Ł\left(\left|\mathbf{b}_i - \mathbf{t}_j\right|, \sigma_{tb}\right) \\ & + \sum_{\alpha,\beta}^{N_{ai}, N_{aj}} \varepsilon_{SS} \mathscr{M}\left(\left|\mathbf{a}_{i\alpha} - \mathbf{a}_{j\beta}\right|, r_0, \rho, r_{cut}^{att}\right) \end{aligned} \tag{B2}$$

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

diameter:

$$Ł(x,\sigma) \equiv \theta(\sigma - x)\left[\left(\frac{\sigma}{x}\right)^{12} - 1\right] \tag{B3}$$

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

$$\mathscr{M}(x, r_0, \rho, r_{\text{cut}}) = \theta(r_{\text{cut}} - x)\times \left[\left(e^{\rho\left(1-\frac{x}{r_0}\right)} - 2\right)e^{\rho\left(1-\frac{x}{r_0}\right)} - V_{\text{shift}}(r_{\text{cut}})\right] \tag{B4}$$

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

The parameter ε_{SS} sets the strength of the shell-shell attraction at each Attractor site, $N_{\text{a}i}$ 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 ε_{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 ε_{SS}. Because vertex Attractors (A1, A4) have multiple partners in an assembled structure, whereas edge Attractors have only one, the well-depth for the vertex pairs (A1-A4 and A4-A4) is set to $0.5\varepsilon_{\text{SS}}$. Similarly, for pentamer-hexamer interactions, the well-depth for edge Attractor pairs (A2-A5, A3-A6) is ε_{SS}, while the vertex interaction pairs (A1-A4 and A4-A4) have $0.5\varepsilon_{\text{SS}}$.

Repulsive interactions: The 'Top' and 'Bottom' heights, or distance out of the Attractor plane, are set to $h = 1/2r_{\text{b}}$, with $r_{\text{b}} = 1$ the distance between a vertex Attractor and the center of the pentagon. For all simulations shells have T=3 preferred curvature and,$\sigma_{tb} = 1.8r_{\text{b}}$ is the diameter of the 'Top' - 'Bottom' interaction (this prevents subunits from binding in inverted configurations[76]), and $\sigma_b = 1.5r_{\text{b}}$ 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, $\sigma_{\text{t,pp}} = 2.1r_{\text{b}}$, $\sigma_{\text{t,hh}} = 2.4r_{\text{b}}$, and $\sigma_{\text{t,ph}} = 2.2r_{\text{b}}$. The parameter r_0 is the minimum energy Attractor distance, set to $0.2r_{\text{b}}$, $\rho = 4r_{\text{b}}$ determines the width of the attractive interaction, and $r_{\text{cut}}^{\text{att}} = 2.0r_{\text{b}}$ 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:

$$U_{\text{CC}}(\{\mathbf{l}_i\},\{\mathbf{l}_j\}) = \sum_{z=\text{RR,GG,RG}} \sum_{i<j}^{N_z} \varepsilon_z \mathscr{L}\left(\left|\mathbf{l}_i - \mathbf{l}_j\right|,\ \sigma_{\text{C}}, r_{\text{cut}}^{\text{C}}\right) \tag{B5}$$

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

$$\mathscr{L}(x,\sigma,r_{\text{cut}}) = \theta(x - r_{\text{cut}})\times \left\{4\left[\left(\frac{x}{\sigma}\right)^{12} - \left(\frac{x}{\sigma}\right)^{6}\right] - V_{\text{shift}}(r_{\text{cut}})\right\} \tag{B6}$$

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 $\sigma_{\text{C}} = r_{\text{b}}$ and cutoff $r_{\text{cut}}^{\text{C}} = 3\sigma_{\text{C}}$ 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 shell subunit 'Bottom' pseudoatoms. 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:

$$U_{\text{SC}}(\{\mathbf{x}_{i\alpha}\},\mathbf{R}_j) = \sum_{\alpha}^{N_x} Ł\left(|\mathbf{x}_{i\alpha} - \mathbf{R}_j|, \sigma_{ex}\right) \tag{B7}$$

$$+ \sum_{\alpha}^{N_t} Ł\left(|\mathbf{t}_{i\alpha} - \mathbf{R}_j|, \sigma_t\right) \tag{B8}$$

$$+ \sum_{\alpha}^{N\text{b}} \varepsilon_{\text{SC}} \mathscr{M}\left(\left|\mathbf{c}_{i\alpha} - \mathbf{R}_j\right|,\ r_0, \rho^{\text{SC}}, r_{\text{cut}}^{\text{SC}}\right) \tag{B9}$$

where ε_{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, $\sigma_{\text{ex}} = 0.5r_{\text{b}}$ and $\sigma_{\text{t}} = 0.5r_{\text{b}}$ are the effective diameters of the Excluder - cargo and 'Top' - cargo repulsions, $r_0^{\text{SC}} = 0.5r_{\text{b}}$ is the minimum energy Attractor distance, the width parameter is $\rho^{\text{SC}} = 2.5r_{\text{b}}$, and the cutoff is set to $r_{\text{cut}}^{\text{SC}} = 3.0r_{\text{b}}$.

[1] C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher, and A. A. Hyman, "Germline p granules are liquid droplets that localize by controlled dissolution/condensation," Science **324**, 1729–1732 (2009).

[2] C. Brangwynne, P. Tompa, and R. Pappu, "Polymer physics of intracellular phase transitions," Nat. Phys. **11**, 899–904 (2015).

[3] C. F. Lee, C. P. Brangwynne, J. Gharakhani, A. A. Hyman, and F. Julicher, "Spatial organization of the cell cytoplasm by position-dependent phase separation," Phys Rev Lett **111**, 088101 (2013).

[4] C. P. Brangwynne, "Phase transitions and size scaling of membrane-less organelles," J. Cell Biol. **203**, 875–881 (2013).

[5] S. C. Weber and C. P. Brangwynne, "Getting rna and protein in phase," Cell **149**, 1188–1191 (2012).

[6] A. A. Hyman and C. P. Brangwynne, "Beyond stereospecificity: liquids and mesoscale organization of cytoplasm," Dev. Cell **21**, 14–16 (2011).

[7] A. A. Hyman and K. Simons, "Beyond oil and water,Äîphase transitions in cells," Science **337**, 1047–1049 (2012).

[8] L. P. Bergeron-Sandoval, N. Safaee, and S. W. Michnick, "Mechanisms and consequences of macromolecular phase separation," Cell **165**, 1067–79 (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, "Coexisting liquid phases underlie nucleolar subcompartments," Cell **165**, 1686–97 (2016).

[10] D. Hnisz, K. Shrinivas, R. A. Young, A. K. Chakraborty, and P. A. Sharp, "A phase separation model for transcriptional control," Cell **169**, 13–23 (2017).

[11] M. B. Stone, S. A. Shelby, M. F. Nunez, K. Wisser, and S. L. Veatch, "Protein sorting by lipid phase-like domains supports emergent signaling function in b lymphocyte plasma membranes," Elife **6** (2017), 10.7554/eLife.19891.

[12] L. Zhu and C. P. Brangwynne, "Nuclear bodies: the emerging biophysics of nucleoplasmic phases," Curr Opin Cell Biol **34**, 23–30 (2015).

[13] S. F. Banani, H. O. Lee, A. A. Hyman, and M. K. Rosen, "Biomolecular condensates: organizers of cellular biochemistry," Nat Rev Mol Cell Biol **18**, 285–298 (2017).

[14] J. A. Joseph, J. R. Espinosa, I. Sanchez-Burgos, A. Garaizar, D. Frenkel, and R. Collepardo-Guevara, "Thermodynamics and kinetics of phase separation of protein-rna mixtures by a minimal model," Biophys J **120**, 1219–1230 (2021).

[15] T. O. Yeates, C. A. Kerfeld, S. Heinhorst, G. C. Cannon, and J. M. Shively, "Protein-based organelles in bacteria: carboxysomes and related microcompartments," Nat. Rev. Microbiol. **6**, 681–691 (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.*, "Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles," Mol. Cell **57**, 936–947 (2015).

[17] B. Y. Zaslavsky, L. A. Ferreira, A. L. Darling, and V. N. Uversky, "The solvent side of proteinaceous membrane-less organelles in light of aqueous two-phase systems," Int. J. Biol. Macromol. **117**, 1224–1251 (2018).

[18] C. A. Kerfeld, S. Heinhorst, and G. C. Cannon, "Bacterial microcompartments," Microbiology+ **64**, 391 (2010).

[19] S. Tanaka, C. A. Kerfeld, M. R. Sawaya, F. Cai, S. Heinhorst, G. C. Cannon, and T. O. Yeates, "Atomic-level models of the bacterial carboxysome shell," Science **319**, 1083–1086 (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, "Functions, compositions, and evolution of the two types of carboxysomes: Polyhedral microcompartments that facilitate co2 fixation in cyanobacteria and some proteobacteria," Microbiol. Mol. Biol. Rev. **77**, 357–379 (2013), http://mmbr.asm.org/content/77/3/357.full.pdf+html.

[21] T. A. Bobik, B. P. Lehman, and T. O. Yeates, "Bacterial microcompartments: widespread prokaryotic organelles for isolation and optimization of metabolic pathways," Mol. Microbiol. **98**, 193–207 (2015).

[22] C. Chowdhury, S. Sinha, S. Chun, T. O. Yeates, and T. A. Bobik, "Diverse bacterial microcompartment organelles," Microbiol. Mol. Biol. Rev. **78**, 438–68 (2014).

[23] C. A. Kerfeld and M. R. Melnicki, "Assembly, function and evolution of cyanobacterial carboxysomes," Curr. Opin. Plant Biol. **31**, 66–75 (2016).

[24] J. K. Polka, S. G. Hays, and P. A. Silver, "Building spatial synthetic biology with compartments, scaffolds, and communities," Cold Spring Harb Perspect Biol **8** (2016), 10.1101/cshperspect.a024018.

[25] C. A. Kerfeld, C. Aussignargues, J. Zarzycki, F. Cai, and M. Sutter, "Bacterial microcompartments," 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, "Structural basis of enzyme encapsulation into a bacterial nanocompartment," Nat. Struct. Mol. Biol. **15**, 939–947 (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, "Discovery and characterization of a novel family of prokaryotic nanocompartments involved in sulfur metabolism," bioRxiv (2020).

[28] F. Pfeifer, "Distribution, formation and regulation of gas vesicles," Nat Rev Micro **10**, 705–715 (2012).

[29] V. A. Kickhoefer, K. S. Rajavel, G. L. Scheffer, W. S. Dalton, R. J. Scheper, and L. H. Rome, "Vaults are up-regulated in multidrug-resistant cancer cell lines," J. Biol. Chem. **273**, 8971–8974 (1998), http://www.jbc.org/content/273/15/8971.full.pdf+html.

[30] A. Chessher, R. Breitling, and E. Takano, "Bacterial microcompartments: Biomaterials for synthetic biology-based compartmentalization strategies," ACS Biomaterials Science & Engineering **1**, 345–351 (2015).

[31] B. P. Lehman, C. Chowdhury, and T. A. Bobik, "The n terminus of the pdub protein binds the protein shell of the pdu microcompartment to its enzymatic core," J. Bacteriol. **199** (2017), 10.1128/JB.00785-16.

[32] B. M. Long, M. R. Badger, S. M. Whitney, and G. D. Price, "Analysis of carboxysomes from synechococcus pcc7942 reveals multiple rubisco complexes with carboxysomal proteins ccmm and ccaa," J. Biol. Chem. **282**, 29323–35 (2007).

[33] B. M. Long, L. Tucker, M. R. Badger, and G. D. Price, "Functional cyanobacterial beta-carboxysomes have an absolute requirement for both long and short forms of the ccmm protein," Plant Physiol. **153**, 285–93 (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, "Effect of bio-engineering on size, shape, composition and rigidity of bacterial microcompartments," Sci. Rep. **6**, 36899 (2016).

[35] M. B. Quin, S. A. Perdue, S. Y. Hsu, and C. Schmidt-Dannert, "Encapsulation of multiple cargo proteins within recombinant eut nanocompartments," Appl. Microbiol. Biotechnol. **100**, 9187–9200 (2016).

[36] M. Slininger Lee and D. Tullman-Ercek, "Practical considerations for the encapsulation of multi-enzyme cargos within the bacterial microcompartment for metabolic engineering," 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, "Evidence for improved encapsulated pathway behavior in a bacterial microcompartment through shell protein engineering," ACS Synth. Biol. (2017), 10.1021/acssynbio.7b00042.

[38] A. Hagen, M. Sutter, N. Sloan, and C. A. Kerfeld, "Programmed loading and rapid purification of engineered bacterial microcompartment shells," Nat. Comm. **9**, 1–10 (2018).

[39] J. N. Kinney, A. Salmeen, F. Cai, and C. a. Kerfeld, "Elucidating essential role of conserved carboxysomal protein CcmN reveals common feature of bacterial microcompartment assembly," J. Biol. Chem. **287**, 17729–17736 (2012).

[40] B. Ferlez, M. Sutter, and C. A. Kerfeld, "A designed bacterial microcompartment shell with tunable composition and precision cargo loading," Metab. Eng. **54**, 286–291 (2019).

[41] G. Kalnins, E.-E. Cesle, J. Jansons, J. Liepins, A. Filimonenko, and K. Tars, "Encapsulation mechanisms and structural studies of grm2 bacterial microcompartment particles," Nat. Comm. **11**, 1–13 (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, "Effect of metabolosome encapsulation peptides on enzyme activity, coaggregation, incorporation, and bacterial microcompartment formation," MicrobiologyOpen **9**, e1010 (2020).

[43] M. Künzle, J. Mangler, M. Lach, and T. Beck, "Peptide-directed encapsulation of inorganic nanoparticles into protein containers," Nanoscale **10**, 22917–22926 (2018).

[44] B. Alberts, *Essential cell biology* (Garland Science, New York; London, 2010).

[45] H. M. Huttanus and X. Feng, "Compartmentalized metabolic engineering for biochemical and biofuel production," Biotechnol. J. (2017), 10.1002/biot.201700052.

[46] D. P. Patterson, P. E. Prevelige, and T. Douglas, "Nanoreactors by programmed enzyme encapsulation inside the capsid of the bacteriophage p22," ACS Nano **6**, 5000–5009 (2012), http://dx.doi.org/10.1021/nn300545z.

[47] D. P. Patterson, B. Schwarz, R. S. Waters, T. Gedeon, and T. Douglas, "Encapsulation of an enzyme cascade within the bacteriophage p22 virus-like particle," ACS Chem. Biol. **9**, 359–365 (2014), http://dx.doi.org/10.1021/cb4006529.

[48] B. Worsdorfer, Z. Pianowski, and D. Hilvert, "Efficient in vitro encapsulation of protein cargo by an engineered protein container," J. Am. Chem. Soc. **134**, 909–911 (2012), http://dx.doi.org/10.1021/ja211011k.

[49] F. Cai, S. L. Bernstein, S. C. Wilson, and C. A. Kerfeld, "Production and characterization of synthetic carboxysome shells with incorporated luminal proteins," Plant Physiol. **170**, 1868–77 (2016).

[50] S. M. Lee, N. Bond, C. Callaway, B. Clark, E. Farmer, M. Mallard, and S. S. Jang, "Dissipative particle dynamics simulation of multicompartment micelle nanoreactor with channel for reactants," RSC Adv. **8**, 37866–37871 (2018).

[51] S. Chakraborti, A. Korpi, M. Kumar, P. Stepien, M. A. Kostiainen, and J. G. Heddle, "Three-dimensional protein cage array capable of active enzyme capture and artificial chaperone activity," Nano Lett **19**, 3918–3924 (2019).

[52] T. G. W. Edwardson and D. Hilvert, "Virus-inspired function in engineered protein cages," J Am Chem Soc **141**, 9432–9443 (2019).

[53] G. D. Price and M. R. Badger, "Evidence for the role of carboxysomes in the cyanobacterial co2-concentrating mechanism," Can. J. Bot. **69**, 963–973 (1991).

[54] J. M. Shively, F. Ball, D. H. Brown, and R. E. Saunders, "Functional organelles in prokaryotes: Polyhedral inclusions (carboxysomes) of thiobacillus neapolitanus," Science **182**, 584–586 (1973).

[55] J. M. Shively, F. L. Ball, and B. W. Kline, "Electron microscopy of the carboxysomes (polyhedral bodies) of thiobacillus neapolitanus," J. Bacteriol. **116**, 1405–1411 (1973).

[56]C. V. Iancu, H. J. Ding, D. M. Morris, D. P. Dias, A. D. Gonzales, A. Martino, and G. J. Jensen, "The Structure of Isolated Synechococcus Strain WH8102 Carboxysomes as Revealed by Electron Cryotomography," J. Mol. Biol. **372**, 764–773 (2007).

[57]C. V. Iancu, D. M. Morris, Z. Dou, S. Heinhorst, G. C. Cannon, and G. J. Jensen, "Organization, structure, and assembly of α-carboxysomes determined by electron cryotomography of intact cells," J. Mol. Biol. **396**, 105 – 117 (2010).

[58]M. Sutter, B. Greber, C. Aussignargues, and C. A. Kerfeld, "Assembly principles and structure of a 6.5-mda bacterial microcompartment shell," Science **356**, 1293 (2017).

[59]M. Sutter, T. G. Laughlin, N. B. Sloan, D. Serwas, K. M. Davies, and C. A. Kerfeld, "Structure of a synthetic β-carboxysome shell," Plant Physiol. **181**, 1050–1058 (2019).

[60]J. D. Perlmutter, F. Mohajerani, and M. F. Hagan, "Many-molecule encapsulation by an icosahedral shell," eLife **5**, e14078 (2016).

[61]F. Mohajerani and M. F. Hagan, "The role of the encapsulated cargo in microcompartment assembly," PLoS Comput. Biol. **14**, e1006351 (2018).

[62]F. Mohajerani, E. Sayer, C. Neil, K. Inlow, and M. F. Hagan, "Mechanisms of scaffold-mediated microcompartment assembly and size control," ACS Nano **15**, 4197–4212 (2021), pMID: 33683101, https://doi.org/10.1021/acsnano.0c05715.

[63]J. P. Mahalik, K. A. Brown, X. Cheng, and M. Fuentes-Cabrera, "Theoretical study of the initial stages of self-assembly of a carboxysome's facet," ACS Nano **10**, 5751–8 (2016).

[64]E. J. Young, R. Burton, J. P. Mahalik, B. G. Sumpter, M. Fuentes-Cabrera, C. A. Kerfeld, and D. C. Ducat, "Engineering the bacterial microcompartment domain for molecular scaffolding applications," Front. Microbiol. **8** (2017), 10.3389/fmicb.2017.01441.

[65]G. M. Rotskoff and P. L. Geissler, "Robust nonequilibrium pathways to microcompartment assembly," Proc. Natl. Acad. Sci. **115**, 6341–6346 (2018).

[66]Y. Li, N. W. Kennedy, S. Li, C. E. Mills, D. Tullman-Ercek, and M. Olvera de la Cruz, "Computational and experimental approaches to controlling bacterial microcompartment assembly," ACS Cent Sci **7**, 658–670 (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, "Spontaneous non-canonical assembly of ccmk hexameric components from β-carboxysome shells of cyanobacteria," (2017).

[68]T. Chen and S. C. Glotzer, "Simulation studies of a phenomenological model for elongated virus capsid formation," Physical Review E. **75**, 051504 (2007).

[69]R. F. Bruinsma, W. M. Gelbart, D. Reguera, J. Rudnick, and R. Zandi, "Viral self-assembly as a thermodynamic process," Phys. Rev. Lett. **90**, 248101 (2003).

[70]R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, and J. Rudnick, "Origin of icosahedral symmetry in viruses," Proc. Natl. Acad. Sci. U. S. A. **101**, 15556–15560 (2004).

[71]B. Berger, P. W. Shor, L. Tuckerkellogg, and J. King, "Local Rule-Based Theory of Virus Shell Assembly," Proc. Natl. Acad. Sci. U. S. A. **91**, 7732–7736 (1994).

[72]R. Schwartz, R. L. Garcea, and B. Berger, ""Local Rules" Theory Applied to Polyomavirus Polymorphic Capsid Assemblies," Virology **268**, 461–470 (2000).

[73]D. Rapaport, "Self-assembly of polyhedral shells: a molecular dynamics study," Phys. Rev. E. **70**, 051905 (2004).

[74]H. D. Nguyen, V. S. Reddy, and C. L. Brooks, "Invariant Polymorphism in Virus Capsid Assembly," J. Am. Chem. Soc. **131**, 2606–14 (2009).

[75]O. M. Elrad and M. F. Hagan, "Mechanisms of Size Control and Polymorphism in Viral Capsid Assembly," Nano Lett. **8**, 3850–3857 (2008).

[76]I. G. Johnston, A. A. Louis, and J. P. K. Doye, "Modelling the self-assembly of virus capsids," J. Phys.: Condens. Matter **22**, 104101 (2010).

[77]D. C. Rapaport, "Studies of reversible capsid shell growth," J. Phys.: Condens. Matter **22**, 104115 (2010).

[78]H. Nguyen and C. Brooks, "Generalized structural polymorphism in self-assembled viral particles," Nano Lett. **8**, 4574 (2008).

[79]M. G. Mateu, "Assembly, stability and dynamics of virus capsids," Arch. Biochem. Biophys. **531**, 65–79 (2013).

[80]M. F. Hagan, "Modeling Viral Capsid Assembly," Adv. Chem. Phys. **155**, 1–68 (2014).

[81]J. D. Perlmutter and M. F. Hagan, "Mechanisms of Virus Assembly," Annu. Rev. Phys. Chem. **66**, 217–239 (2015), http://dx.doi.org/10.1146/annurev-physchem-040214-121637.

[82]R. F. Bruinsma and W. S. Klug, "Physics of viral shells," Annu. Rev. Condens. Matter Phys. **6**, 245–268 (2015), https://doi.org/10.1146/annurev-conmatphys-031214-014325.

[83]M. F. Hagan and R. Zandi, "Recent advances in coarse-grained modeling of virus assembly," Curr. Opin. Vir. **18**, 36 (2016).

[84]D. C. Rapaport, "Molecular dynamics simulation of reversibly self-assembling shells in solution using trapezoidal particles," Phys. Rev. E **86**, 051917 (2012).

[85]D. C. Rapaport, J. E. Johnson, and J. Skolnick, "Supramolecular self-assembly: molecular dynamics modeling of polyhedral shell formation," Comput. Phys. Commun. **122**, 231–235 (1999).

[86]C. I. Mendoza and D. Reguera, "Shape selection and mis-assembly in viral capsid formation by elastic frustration," eLife **9**, e52525 (2020).

[87]J. Wagner and R. Zandi, "The robust assembly of small symmetric nanoshells," Biophys. J. **109**, 956 (2015).

[88]R. Schwartz, P. W. Shor, P. E. Prevelige, and B. Berger, "Local Rules Simulation of the Kinetics of Virus Capsid Self-Assembly," Biophys. J. **75**, 2626–2636 (1998).

[89]M. F. Hagan and D. Chandler, "Dynamic Pathways for Viral Capsid Assembly," Biophys. J. **91**, 42–54 (2006).

[90]A. W. Wilber, J. P. K. Doye, A. A. Louis, E. G. Noya, M. A. Miller, and P. Wong, "Reversible Self-Assembly of Patchy Particles into Monodisperse Icosahedral Clusters," J. Chem. Phys. **127**, 085106 (2007).

[91]J. E. Baschek, H. C. R. Klein, and U. S. Schwarz, "Stochastic dynamics of virus capsid formation: direct versus hierarchical self-assembly," Bmc Biophysics **5** (2012), 10.1186/2046-1682-5-22.

[92]M. Castelnovo, T. Verdier, and L. Foret, "Comparing open and closed molecular self-assembly," Epl-europhys Lett **105** (2014), 10.1209/0295-5075/105/28006.

[93]M. Castelnovo, D. Muriaux, and C. Faivre-Moskalenko, "Entropic control of particle sizes during viral self-assembly," 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, "Role of dynamic capsomere supply for viral capsid self-assembly," Phys. Biol. **12** (2015), 10.1088/1478-3975/12/1/016014.

[95]S. N. Fejer, D. Chakrabarti, and D. J. Wales, "Emergent Complexity from Simple Anisotropic Building Blocks: Shells, Tubes, and Spirals," ACS Nano **4**, 219–228 (2010).

[96]R. Kusters, H.-K. Lin, R. Zandi, I. Tsvetkova, B. Dragnea, and P. van der Schoot, "Role of Charge Regulation and Size Polydispersity in Nanoparticle Encapsulation by Viral Coat Proteins," J. Phys. Chem. B **119**, 1869–1880 (2015).

[97]R. Zandi and P. van der Schoot, "Size Regulation of ss-RNA Viruses," Biophys. J. **96**, 9–20 (2009).

[98]M. F. Hagan, "Controlling Viral Capsid Assembly with Templating," Phys. Rev. E **77**, 051904 (2008).

[99]M. F. Hagan, "A Theory for Viral Capsid Assembly around Electrostatic Cores," J. Chem. Phys. **130**, 114902 (2009).

[100]V. A. Belyi and M. Muthukumar, "Electrostatic origin of the genome packing in viruses," Proc. Natl. Acad. Sci. U. S. A. **103**, 17174–17178 (2006).

[101]R. Twarock, R. J. Bingham, E. C. Dykeman, and P. G. Stockley, "A modelling paradigm for rna virus assembly," Curr. Opin. Vir. **31**, 74–81 (2018).

[102]R. Zandi, B. Dragnea, A. Travesset, and R. Podgornik, "On virus growth and form," Phys. Rep. (2020).

[103]S. Li, P. Roy, A. Travesset, and R. Zandi, "Why large icosahedral viruses need scaffolding proteins," Proc. Natl. Acad. Sci. U. S. A. **115**, 10971–10976 (2018).

[104]S. Li, R. Zandi, A. Travesset, and G. M. Grason, "Ground States of Crystalline Caps: Generalized Jellium on Curved Space," Phys. Rev. Lett. (2019), 10.1103/PhysRevLett.123.145501, arXiv:1906.03301.

[105]P. van der Schoot and R. Zandi, "Impact of the topology of viral rnas on their encapsulation by virus coat proteins," J. Biol. Phys. **39**, 289–299 (2013).

[106]G. Erdemci-Tandogan, J. Wagner, P. van der Schoot, R. Podgornik, and R. Zandi, "RNA topology remolds electrostatic stabilization of viruses," Phys. Rev. E **89**, 032707 (2014).

[107]A. Siber, R. Zandi, and R. Podgornik, "Thermodynamics of nanospheres

encapsulated in virus capsids," Phys. Rev. E **81**, 051919 (2010).
[108]E. C. Dykeman, P. G. Stockley, and R. Twarock, "Packaging signals in two single-stranded RNA viruses imply a conserved assembly mechanism and geometry of the packaged genome." J. Mol. Biol. **425**, 3235–49 (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, "Revealing the density of encoded functions in a viral rna." Proc. Natl. Acad. Sci. U.S.A. (2015), 10.1073/pnas.1420812112.
[110]E. C. Dykeman, P. G. Stockley, and R. Twarock, "Solving a Levinthal's paradox for virus assembly identifies a unique antiviral strategy," Proc. Natl. Acad. Sci. U. S. A. **111**, 5361–5366 (2014).
[111]E. C. Dykeman, P. G. Stockley, and R. Twarock, "Building a Viral Capsid in the Presence of Genomic RNA," 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, "The Impact of Viral RNA on Assembly Pathway Selection," J. Mol. Biol. **401**, 298–308 (2010).
[113]R. Zhang, E. Wernersson, and P. Linse, "Icosahedral capsid formation by capsomer subunits and a semiflexible polyion," RSC Adv. **3**, 25258–25267 (2013).
[114]R. Zhang and P. Linse, "Topological effects on capsomer-polyion coassembly," J. Chem. Phys. **140** (2014), 10.1063/1.4883056, zhang, Ran Linse, Per 1089-7690.
[115]R. Zhang and P. Linse, "Icosahedral capsid formation by capsomers and short polyions," J. Chem. Phys. **138** (2013), 10.1063/1.4799243, zhang, Ran Linse, Per 1089-7690.
[116]D. G. Angelescu, J. Stenhammar, and P. Linse, "Packaging of a flexible polyelectrolyte inside a viral capsid: Effect of salt concentration and salt valence," J. Phys. Chem. B **111**, 8477–8485 (2007).
[117]D. G. Angelescu and P. Linse, "Monte Carlo simulations of flexible polyelectrolytes inside viral capsids with dodecahedral charge distribution," Phys. Rev. E **75**, 051905 (2007).
[118]S. Panahandeh, S. Li, and R. Zandi, "The equilibrium structure of self-assembled protein nano-cages," Nanoscale **10**, 22802–22809 (2018).
[119]S. Panahandeh, S. Li, L. Marichal, R. Leite Rubim, G. Tresset, and R. Zandi, "How a virus circumvents energy barriers to form symmetric shells," ACS Nano **14**, 3170–3180 (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, "Advances in understanding carboxysome assembly in prochlorococcus and synechococcus implicate CsoS2 as a critical component," 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, "Short N-terminal sequences package proteins into bacterial microcompartments." Proc. Natl. Acad. Sci. U. S. A. **107**, 7509–7514 (2010).
[122]J. C. Cameron, S. C. Wilson, S. L. Bernstein, and C. A. Kerfeld, "Biogenesis of a bacterial organelle: the carboxysome assembly pathway." Cell **155**, 1131–40 (2013).
[123]A. H. Chen, A. Robinson-Mosher, D. F. Savage, P. A. Silver, and J. K. Polka, "The bacterial carbon-fixing organelle is formed by shell envelopment of preassembled cargo," PLoS One **8**, e76127 (2013).
[124]T. Wunder, S. L. H. Cheng, S.-K. Lai, H.-Y. Li, and O. Mueller-Cajar, "The phase separation underlying the pyrenoid-based microalgal rubisco supercharger," Nat. Comm. **9**, 1–10 (2018).
[125]T. Chaijarasphong, R. J. Nichols, K. E. Kortright, C. F. Nixon, P. K. Teng, L. M. Oltrogge, and D. F. Savage, "Programmed ribosomal frameshifting mediates expression of the α-carboxysome," J. Mol. Biol. **428**, 153–164 (2016).
[126]L. M. Oltrogge, T. Chaijarasphong, A. W. Chen, E. R. Bolin, S. Marqusee, and D. F. Savage, "Multivalent interactions between csos2 and rubisco mediate α-carboxysome formation," Nature Structural & Molecular Biology , 1–7 (2020).
[127]J.-D. Rochaix, "The pyrenoid: an overlooked organelle comes out of age," Cell **171**, 28–29 (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, "Rubisco condensate formation by ccmm in β-carboxysome biogenesis," Nature **566**, 131–135 (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, "The structural basis of rubisco phase separation in the pyrenoid," Nat. Plants , 1–11 (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, "Assembly of the algal co2-fixing organelle, the pyrenoid, is guided by a rubisco-binding motif," Science advances **6**, eabd2408 (2020).
[131]C. P. Brangwynne, "Soft active aggregates: mechanics, dynamics and self-assembly of liquid-like intracellular protein bodies," 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, "Liquid-liquid phase separation in an elastic network," Phys. Rev. X **8** (2018), 10.1103/PhysRevX.8.011028.
[133]W. M. Jacobs and D. Frenkel, "Phase transitions in biological systems with many components," Biophys. J. **112**, 683–691 (2017).
[134]W. M. Jacobs, "Self-assembly of biomolecular condensates with shared components," Phys. Rev. Lett. **126**, 258101 (2021).
[135]S. L. Perry, "Phase separation: Bridging polymer physics and biology," Current Opinion in Colloid & Interface Science **39**, 86–97 (2019).
[136]I. Sanchez-Burgos, J. R. Espinosa, J. A. Joseph, and R. Collepardo-Guevara, "Valency and binding affinity variations can regulate the multilayered organization of protein condensates with many components," 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, "Characterization of virus capsids and their assembly intermediates by multicycle resistive-pulse sensing with four pores in series," Anal. Chem. **90**, 7267–7274 (2018).
[138]R. B. Best, "Computational and theoretical advances in studies of intrinsically disordered proteins," Curr Opin Struct Biol **42**, 147–154 (2017).
[139]G. L. Dignon, W. Zheng, R. B. Best, Y. C. Kim, and J. Mittal, "Relation between single-molecule properties and phase behavior of intrinsically disordered proteins," Proc Natl Acad Sci U S A **115**, 9929–9934 (2018).
[140]G. L. Dignon, W. Zheng, Y. C. Kim, R. B. Best, and J. Mittal, "Sequence determinants of protein phase behavior from a coarse-grained model," PLoS Comput Biol **14**, e1005941 (2018).
[141]G. L. Dignon, W. Zheng, and J. Mittal, "Simulation methods for liquid-liquid phase separation of disordered proteins," Curr Opin Chem Eng **23**, 92–98 (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, "Liquid-liquid phase separation during amphiphilic self-assembly," Nat Chem **11**, 320–328 (2019).
[143]A. E. Posey, A. S. Holehouse, and R. V. Pappu, "Phase separation of intrinsically disordered proteins," Methods Enzymol **611**, 1–30 (2018).
[144]K. M. Ruff, R. V. Pappu, and A. S. Holehouse, "Conformational preferences and phase behavior of intrinsically disordered low complexity sequences: insights from multiscale simulations," Curr Opin Struct Biol **56**, 1–10 (2018).
[145]S. Alberti, A. Gladfelter, and T. Mittag, "Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates," Cell **176**, 419–434 (2019).
[146]D. L. D. Caspar and A. Klug, "Physical Principles in Construction of Regular Viruses," Cold Spring Harbor Symp. Quant. Biol. **27**, 1–24 (1962).
[147]T. Douglas and M. Young, "Host-guest encapsulation of materials by assembled virus protein cages," Nature **393**, 152–155 (1998).
[148]T. Douglas and M. Young, "Viruses: making friends with old foes.(PERSPECTIVE)," Science **312**, 873(3) (2006).
[149]C. P. Goodall, B. Schwarz, E. Selivanovitch, J. Avera, J. Wang, H. Miettinen, and T. Douglas, "Controlled modular multivalent presentation of the cd40 ligand on p22 virus-like particles leads to tunable amplification of cd40 signaling," ACS Applied Bio Materials **4**, 8205–8214 (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, "Engineering a virus-like particle to display peptide insertions using an apparent fitness landscape," Biomacromolecules **21**, 4194–4204 (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, "Controlled encapsulation of multiple proteins in virus capsids," J. Am. Chem. Soc. **131**, 17771–17773 (2009), http://dx.doi.org/10.1021/ja907843s.
[152]J.-K. Rhee, M. Hovlid, J. D. Fiedler, S. D. Brown, F. Manzen-

rieder, H. Kitagishi, C. Nycholat, J. C. Paulson, and M. G. Finn, "Colorful virus-like particles: Fluorescent protein packaging by the q?? capsid," Biomacromolecules **12**, 3977–3981 (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, "Self-sorting of foreign proteins in a bacterial nanocompartment," J. Am. Chem. Soc. **136**, 3828–3832 (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, "Predicting the loading of virus-like particles with fluorescent proteins," Biomacromolecules **15**, 558–563 (2014), http://dx.doi.org/10.1021/bm4015792.

[155]N. H. Dashti, R. S. Abidin, and F. Sainsbury, "Programmable in vitro coencapsidation of guest proteins for intracellular delivery by virus-like particles," ACS Nano **12**, 4615–4623 (2018).

[156]M. Faulkner, L. S. Zhao, S. Barrett, and L. N. Liu, "Self-assembly stability and variability of bacterial microcompartment shell proteins in response to the environmental change," Nanoscale Res Lett **14**, 54 (2019).

[157]S. Whitelam and R. L. Jack, "The Statistical Mechanics of Dynamic Pathways to Self-assembly," Ann Rev Phys Chem **66**, 143–63 (2015).

[158]B. Smit, "Phase diagrams of lennard-jones fluids," J. Chem. Phys. **96**, 8639–8640 (1992), https://doi.org/10.1063/1.462271.

[159]J. N. C. LOPES, "Phase equilibra in binary lennard-jones mixtures: phase diagram simulation," Mol. Phys. **96**, 1649–1658 (1999), https://doi.org/10.1080/00268979909483108.

[160]V. Shen and J. Errington, "Determination of fluid-phase behavior using transition-matrix monte carlo: Binary lennard-jones mixtures," J. Chem. Phys. (2005).

[161]W. Dai, M. Chen, C. Myers, S. J. Ludtke, B. M. Pettitt, J. A. King, M. F. Schmid, and W. Chiu, "Visualizing individual rubisco and its assembly into carboxysomes in marine cyanobacteria by cryo-electron tomography," J. Mol. Biol. **430**, 4156–4167 (2018).

[162]J. B. Parsons, S. Frank, D. Bhella, M. Liang, M. B. Prentice, D. P. Mulvihill, and M. J. Warren, "Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement," Mol. Cell **38**, 305–315 (2010).

[163]S. Choudhary, M. B. Quin, M. A. Sanders, E. T. Johnson, and C. Schmidt-Dannert, "Engineered protein nano-compartments for targeted enzyme localization," PLoS One **7**, e33342 (2012).

[164]J. K. Lassila, S. L. Bernstein, J. N. Kinney, S. D. Axen, and C. A. Kerfeld, "Assembly of robust bacterial microcompartment shells using building blocks from an organelle of unknown function." J. Mol. Biol. **426**, 2217–28 (2014).

[165]G. N. Condezo and C. San Martin, "Localization of adenovirus morphogenesis players, together with visualization of assembly intermediates and failed products, favor a model where assembly and packaging occur concurrently at the periphery of the replication center," PLoS Pathog. **13** (2017).

[166]C. Aussignargues, B. C. Paasch, R. Gonzalez-Esquer, O. Erbilgin, and C. A. Kerfeld, "Bacterial microcompartment assembly: The key role of encapsulation peptides," Communicative & integrative biology **8**, e1039755 (2015).

[167]P. Ryan, T. J. Forrester, C. Wroblewski, T. M. Kenney, E. N. Kitova, J. S. Klassen, and M. S. Kimber, "The small rbcs-like domains of the β-carboxysome structural protein ccmm bind rubisco at a site distinct from that binding the rbcs subunit," J. Biol. Chem. **294**, 2593–2603 (2019).

[168]J. B. Bale, S. Gonen, Y. Liu, W. Sheffler, D. Ellis, C. Thomas, D. Cascio, T. O. Yeates, T. Gonen, N. P. King, and D. Baker, "Accurate Design of Megadalton-Scale Two-Component Icosahedral Protein Complexes," Science **353**, 389–94 (2016), type: Journal Article.

[169]G. L. Butterfield, 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, N. P. King, and D. Baker, "Evolution of a Designed Protein Assembly Encapsulating Its Own RNA Genome," Nature **552**, 415–420 (2017), type: Journal Article.

[170]E. Golub, R. H. Subramanian, J. Esselborn, R. G. Alberstein, J. B. Bailey, J. A. Chiong, X. Yan, T. Booth, T. S. Baker, and F. A. Tezcan, en"Constructing protein polyhedra via orthogonal chemical interactions," Nature **578**, 172–176 (2020), number: 7793 Publisher: Nature Publishing Group.

[171]C. Sigl, E. M. Willner, W. Engelen, J. A. Kretzmann, K. Sachenbacher, A. Liedl, F. Kolbe, F. Wilsch, S. A. Aghvami, U. Protzer, *et al.*, "Programmable icosahedral shell system for virus trapping," Nature Materials , 1–9 (2021).

[172]J. M. Fox, G. J. Wang, J. A. Speir, N. H. Olson, J. E. Johnson, T. S. Baker, and M. J. Young, "Comparison of the native CCMV virion with in vitro assembled CCMW virions by cryoelectron microscopy and image reconstruction," Virology **244**, 212–218 (1998).

[173]A. Schneemann, "The structural and functional role of RNA in icosahedral virus assembly," Annu. Rev. Microbiol. **60**, 51–67 (2006).

[174]J. A. Speir, S. Munshi, G. J. Wang, T. S. Baker, and J. E. Johnson, "Structures of the Native and Swollen Forms of Cowpea Chlorotic Mottle Virus Determined by X-Ray Crystallography and Cryoelectron Microscopy," Structure **3**, 63–78 (1995).

[175]J. D. Perlmutter, C. Qiao, and M. F. Hagan, "Viral genome structures are optimal for capsid assembly," eLife **2**, e00632 (2013).

[176]J. D. Perlmutter, M. R. Perkett, and M. F. Hagan, "Pathways for virus assembly around nucleic acids," J. Mol. Biol. (2014), 10.1016/j.jmb.2014.07.004.

[177]J. D. Perlmutter and M. F. Hagan, "The Role of Packaging Sites in Efficient and Specific Virus Assembly." J. Mol. Biol. (2015), 10.1016/j.jmb.2015.05.008.

[178]S. Li, G. Erdemci-Tandogan, J. Wagner, P. van der Schoot, and R. Zandi, "Impact of a nonuniform charge distribution on virus assembly," Phys. Rev. E **96**, 022401 (2017).

[179]C. a. Kerfeld and O. Erbilgin, "Bacterial microcompartments and the modular construction of microbial metabolism," Trends Microbiol. **23**, 22–34 (2015).

[180]M. Sutter, M. Faulkner, C. Aussignargues, B. C. Paasch, S. Barrett, C. A. Kerfeld, and L.-N. Liu, "Visualization of bacterial microcompartment facet assembly using high-speed atomic force microscopy," Nano Lett. **16**, 1590–1595 (2016), pMID: 26617073, http://dx.doi.org/10.1021/acs.nanolett.5b04259.

[181]D. J. Wales, "The energy landscape as a unifying theme in molecular science," Phil. Trans. R. Soc. A **363**, 357–375 (2005).

[182]S. N. Fejer, T. R. James, J. Hernandez-Rojas, and D. J. Wales, "Energy landscapes for shells assembled from pentagonal and hexagonal pyramids," Phys. Chem. Chem. Phys. **11**, 2098–2104 (2009).

[183]T. Ruiz-Herrero and M. F. Hagan, "Simulations show that virus assembly and budding is facilitated by membrane microdomains," Biophys. J. , 1–13 (2015).

[184]J. A. Anderson, C. D. Lorenz, and A. Travesset, "General purpose molecular dynamics simulations fully implemented on graphics processing units," J. Comput. Phys. **227**, 5342–5359 (2008).

[185]T. D. Nguyen, C. L. Phillips, J. A. Anderson, and S. C. Glotzer, "Rigid body constraints realized in massively-parallel molecular dynamics on graphics processing units," Comput. Phys. Commun. **182**, 2307–2313 (2011).

[186]J. Anderson, (URL), http://codeblue.umich.edu/hoomd-blue/doc/page_units.html.