Symmetry of membrane protein polyhedra with heterogeneous protein size

Mingyuan Ma, Di Li, Osman Kahraman, and Christoph A. Haselwandter

Phys. Rev. E **101**, 022417 — Published 24 February 2020

DOI: 10.1103/PhysRevE.101.022417

Symmetry of membrane protein polyhedra with heterogeneous protein size

Mingyuan Ma, Di Li, Osman Kahraman, and Christoph A. Haselwandter
Department of Physics & Astronomy and Department of Biological Sciences,
University of Southern California, Los Angeles, CA 90089, USA
(Dated: January 6, 2020)

In experiments on Membrane Protein Polyhedral Nanoparticles (MPPNs) [T. Basta *et al.*, Proc. Natl. Acad. Sci. U.S.A. **111**, 670 (2014)] it has been observed that membrane proteins and lipids can self-assemble into closed lipid bilayer vesicles with a polyhedral arrangement of membrane proteins. In particular, MPPNs formed from the mechanosensitive channel of small conductance (MscS) were found to have the symmetry of the snub cube—a chiral, Archimedean solid—with one MscS protein located at each one of the 24 vertices of the snub cube. It is currently unknown whether MPPNs with heterogeneous protein composition maintain a high degree of symmetry. Inspired by previous work on viral capsid symmetry, we employ here computational modeling to study the symmetry of MPPNs with heterogeneous protein size. We focus on MPPNs formed from MscS proteins, which can exist in closed or open conformational states with distinct sizes. We find that, as an increasing number of closed-state MscS proteins transitions to the open conformational state of MscS, the minimum-energy MscS arrangement in MPPNs follows a strikingly regular pattern, with the dominant MPPN symmetry always being provided by the snub cube. Our results suggest that MPPNs with heterogeneous protein size can be highly symmetric, with a well-defined polyhedral ordering of membrane proteins of different sizes.

PACS numbers: 87.16.D-, 87.17.-d

I. INTRODUCTION

Self-assembly of proteins into ordered two-dimensional (2D) structures provides a general design principle for biological systems [1–5]. Prominent examples of regular 2D protein assemblies are viral capsid shells [4–8] and membrane protein lattices in cell membranes [9–13]. Such 2D protein assemblies can, for instance, serve as protective barriers [7] and allow cooperative signaling [14]. The symmetry of viral capsids is affected crucially by topological constraints arising from the spherical surface shape [1, 4–8]. In contrast, the basic architecture of membrane protein lattices can often be understood from local interactions between proteins [15–21]. Membrane protein polyhedral nanoparticles (MPPNs) [22] provide an interesting example of a system that combines some of the key physical principles governing viral capsid and membrane protein lattice symmetry. MPPNs are closed lipid bilayer vesicles composed of lipids and membrane proteins. Experiments on MPPNs [22, 23] formed from the mechanosensitive channel of small conductance (MscS) [24] found a dominant MPPN diameter ≈ 20 nm at the bilayer midplane, with each MPPN containing 24 MscS proteins arranged at the vertices of a snub cube. MPPNs are thus—akin to membrane protein lattices—composed of membrane proteins embedded in a lipid bilayer but show—akin to viral capsids—polyhedral symmetry.

Experiments on MPPNs have so far [22, 23] focused on MPPNs with homogeneous protein composition. A mean-field model of MPPN self-assembly and shape [25, 26] successfully predicts how the observed symmetry and size of MPPNs with homogeneous protein composition emerge from the interplay of protein-induced lipid bilayer deformations, topological defects in protein packing, and thermal effects. Intriguingly, the closed surface of MPPNs permits chemical or voltage gradients across the MPPN membrane, which could allow trapping of membrane proteins in distinct, physiologically relevant conformational states [22]. Such transitions in protein conformational state are generally accompanied by changes in protein size: for instance, mechanosensitive ion channels often have a larger size, when viewed perpendicularly to the cell membrane, in the open than the closed state [27]. Provided MPPNs with heterogeneous protein size show an ordered arrangement of proteins, MPPNs could be employed to elucidate membrane protein structures stabilized by transmembrane gradients [22]. It is, however, unknown how heterogeneity in protein size affects MPPN symmetry. The purpose of this article is to explore the symmetry of MPPNs with heterogeneous protein size. We thereby use as our benchmark previous experiments on MPPNs containing 24 MscS proteins [22, 23], but allow for open-state as well as closed-state MscS [28–30] with a fixed total number of proteins in MPPNs.

In MPPNs with heterogeneous protein size not all proteins are equivalent. We therefore do not use a mean-field approach to study MPPNs with heterogeneous protein size. Previous work in physical virology has shown that the symmetry of viral capsids can be captured through a minimal molecular model in which individual capsid subunits are represented by Lennard-Jones particles [4, 7, 31, 32]. For a given number of proteins per MPPN, a similar approach can also be used to successfully predict the symmetry of MPPNs with homogeneous protein size [25]. Here we generalize this approach to model the symmetry of MPPNs with heterogeneous protein size. In Sec. II we provide a detailed description of our modeling approach, the simulated annealing Monte Carlo (MC)

simulations we employ to obtain energetically favorable MPPN configurations, and the methods used here for quantifying MPPN symmetry. In Sec. III we survey the minimum-energy protein configurations in MPPNs with heterogeneous protein size. Based on previous experiments on MPPNs composed of MscS proteins [22, 23], we thereby focus on MPPNs with 24 proteins corresponding to closed-state or open-state MscS proteins [28–30]. We find that, as an increasing number of closed-state MscS transitions to the open state, the minimum-energy MscS arrangement in MPPNs follows a strikingly regular pattern, with the dominant MPPN symmetry always being provided by the snub cube. Finally, in Sec. IV we provide a summary and conclusions of the work described here.

II. MODELING MPPN SYMMETRY

Proceeding in analogy to previous work on viral capsid symmetry [4, 31], MPPN symmetry can be described through a simple particle-based model [25] in which lipids and MscS proteins are represented by differently-sized disks on the surface of a sphere. In this section, we first review this previous model of MPPN symmetry [25], which we term the lipid-protein (LP) model (see Sec. II A). In Sec. II B we formulate the composite particle (CP) model of MPPN symmetry, which is the primary focus of this article and which provides a simplified, coarse-grained representation of the LP model. In the CP model, each membrane protein and its surrounding lipid environment are represented by a single disk on the surface of a sphere. We use simulated annealing MC simulations [33, 34] to find the minimum-energy states implied by the LP and CP models. We summarize the pertinent computational methods in Sec. II C. Finally, in Sec. II D we describe the mathematical approaches used here to quantify MPPN symmetry. Throughout this article, we denote the total number of MscS proteins per MPPN by N and the number of open-state MscS proteins per MPPN by n_o, such that the number of closed-state MscS proteins per MPPN is given by $N - n_o$.

A. Lipid-protein (LP) model

As described previously [25], MPPN symmetry can be captured by a minimal molecular model—the LP model—in which proteins and lipids are represented by distinct particles moving on the surface of a sphere. We take this spherical surface to correspond to the outer membrane leaflet of MPPNs [see Figs. 1(a) and 1(b)] and denote its radius by R. In the LP model, lipids interact with other lipids as well as proteins through Lennard-Jones potentials,

$$V_{i,j}^{(\mathrm{LP})}(r) = \epsilon_k \left[\left(\frac{\bar{r}_{i,j}}{r} \right)^{12} - 2 \left(\frac{\bar{r}_{i,j}}{r} \right)^{6} \right] , \qquad (1)$$

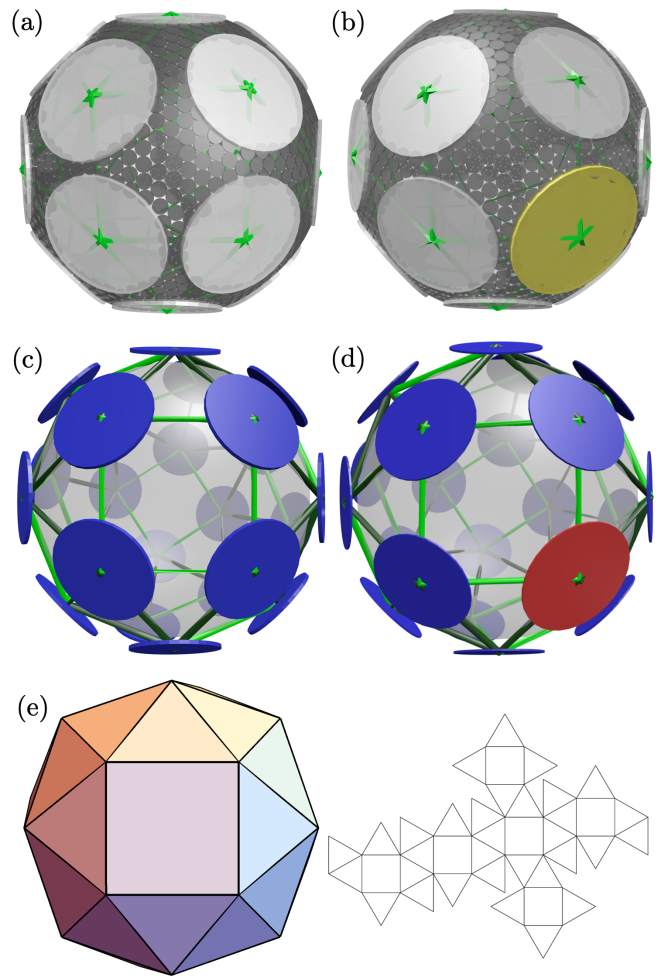

Figure 1. Minimum-energy MPPN configurations obtained from simulated annealing MC simulations of the LP model in Sec. II A with (a) 24 closed-state MscS proteins and (b) one MscS in the open state and 23 closed-state MscS. The small and large disks represent the lipids in a diC14:0 lipid bilayer [35] and MscS proteins [29, 30], respectively, with closed-state MscS corresponding to the large gray disks and open-state MscS to the large yellow disk. The closed-state MscS, open-state MscS, and lipid disk sizes are given by r_c, r_o, and r_l, respectively (see Sec. II A). (c,d) Minimum-energy MPPN configurations obtained as in panels (a) and (b), respectively, but using the coarse-grained CP model in Sec. II B. Particles corresponding to closed-state and open-state MscS are illustrated by blue and red disks, respectively. For ease of visualization, the radii of these disks were decreased by some fixed scale factor relative to the disk radii implied by r_o' and r_c' (see Sec. II B). The green lines in panels (a–d) are obtained by connecting the centers of neighboring proteins. (e) 3D (left panel) and net (right panel) representations of a snub cube.

where we use the notation $(i, j) = (l, l)$, (l, c), and (l, o) to denote interactions between lipids, lipids and closed-state MscS, and lipids and open-state MscS, respectively, the index $k = 1, 2$ denotes lipid-lipid and lipid-protein interactions, r is the Euclidean particle separation in three-dimensional (3D) space, and the $\bar{r}_{i,j}$ denote the energet-

ically most favorable particle separations implied by the Lennard-Jones potentials in Eq. (1). MscS proteins are not expected to aggregate in the absence of lipids, and are therefore assumed here [25] to interact with other MscS proteins only through hardcore steric constraints so that all MscS-MscS separations $r > r_{m,n}$, where the indices $(m, n) = (c, c)$, (c, o), and (o, o) denote pairs of closed-state, closed-state and open-state, and open-state MscS, respectively.

The parameters ϵ_k in Eq. (1) set the energy scale of lipid-lipid and lipid-protein interactions, and can be viewed as the energy penalty for exposing lipids or membrane proteins to an aqueous environment. Experiments and previous calculations [36–38] suggest [25] $\epsilon_1 \approx 10\ k_BT$ and $\epsilon_2 \approx 20\ k_BT$ for the diC14:0 lipids and MscS proteins used in experiments on MPPNs [22, 23], where k_B is Boltzmann's constant and T is the room temperature. We employ these values of ϵ_k throughout this article. It was found previously [25, 39] that the minimum-energy MPPN configurations implied by the LP model are robust with respect to the values of ϵ_k. Approximating the shapes of lipids and proteins by disks that are tangent to the spherical MPPN surface, the values of $\bar{r}_{i,j}$ in Eq. (1) can be estimated from the lipid radius r_l, the closed-state MscS radius r_c, and the open-state MscS radius r_o. Assuming that, in their energetically most favorable configuration, lipid and protein disks touch each other but do not overlap, we have

$$\bar{r}_{i,j} = 2R \sin\left[\frac{1}{2}\left(\arctan\frac{r_i}{R} + \arctan\frac{r_j}{R}\right)\right] . \qquad (2)$$

Experiments on the diC14:0 lipids used for MPPNs formed from MscS [22, 23] give $r_l \approx 0.45$ nm [35] while structural studies of MscS yield $r_c \approx 4.0$ nm and $r_o \approx 4.5$ nm [29, 30] for the outer membrane leaflet of MPPNs, respectively.

In experiments on MPPNs formed from 24 MscS proteins, each MPPN was found to be composed of approximately 1700 lipids [22], which corresponds to approximately 1200 lipids in the outer membrane leaflet of MPPNs [39, 40]. To be consistent with these previous experiments on MPPNs [22, 23], we therefore use here a fixed lipid-protein ratio 1200 : 24. Note that lipids are much more abundant in MPPNs than proteins. As a result, while the protein configuration in MPPNs is of primary interest, MC simulations of the LP model devote considerable computational resources to updating the lipid configuration, which can make it computationally challenging to escape from local energy minima in the particle configuration. This issue becomes particularly significant for MPPNs with heterogeneous protein composition. In Sec. II B we thus develop a simplified model of MPPN symmetry.

B. Composite particle (CP) model

The LP model successfully predicts the dominant symmetry of MPPNs composed solely of closed-state MscS proteins [25]. However, as pointed out in Sec. II A, the explicit representation of lipids in the LP model leads to difficulties when simulating MPPNs with heterogeneous protein size. Indeed, most of the degrees of freedom in the LP model correspond to lipid positions, which are not of primary interest. In both the LP model [25] and experiments [22], MscS proteins tend to be surrounded by an annulus of lipids. This motivates us to formulate the CP model, which provides a simplified, coarse-grained description of MPPN symmetry. In the CP model, we take each particle on the MPPN surface to be composed of one protein surrounded by an annulus of lipids [see Figs. 1(c) and 1(d)]. Similarly as in Sec. II A, we let these particles interact with each other via Lennard-Jones potentials,

$$V_{i,j}^{(\mathrm{CP})}(r) = \epsilon\left[\left(\frac{\bar{r}'_{i,j}}{r}\right)^{12} - 2\left(\frac{\bar{r}'_{i,j}}{r}\right)^{6}\right] , \qquad (3)$$

where the parameter ϵ sets the energy scale of particle interactions and, in analogy to Sec. II A, we use the notation $(i, j) = (c, c)$, (c, o), and (o, o) to denote interactions between particles corresponding to two closed-state MscS, one closed- and one open-state MscS, and two open-state MscS, respectively. We take here each MscS protein to be surrounded by one layer of lipids. The values of $\bar{r}'_{i,j}$ in Eq. (3) are then fixed via an expression analogous to Eq. (2), but using the effective particle radii $r'_c = r_c + 2r_l \approx 4.9$ nm and $r'_o = r_o + 2r_l \approx 5.4$ nm instead of r_l, r_c, or r_o in Eq. (2). We use here $\epsilon = 20\ k_BT$ for the energy scale in Eq. (3), but the minimum-energy MPPN configurations implied by the CP model are independent of the value of ϵ. From a practical standpoint, the central advantage of the CP model over the LP model is that the CP model focuses on the protein configuration, which is what defines the MPPN symmetry, and does not allow for any additional degrees of freedom.

C. Simulated annealing MC simulations

For a given set of values of (N, n_o), we employ simulated annealing MC simulations [33, 41] with linear cooling to numerically determine the minimum-energy configurations associated with the LP and CP models described in Secs. II A and II B [25]. For the LP model with $n_o = 0$ we generate the initial conditions for our MC simulations from a random, uniform distribution of lipids and proteins on the MPPN surface with no overlap of proteins. For the CP model with $n_o = 0$ we employ a uniform, random distribution of CPs with no constraints on the relative particle positions.

In our simulated annealing MC simulations [33, 41] of the LP and CP models we randomly pick, in each MC step, one of the particles on the MPPN surface as the

target particle for this MC step. In particular, in the LP model we first randomly decide whether to update the position of a lipid or protein particle (with probabilities 0.8 and 0.2, respectively) and then pick with equal probability a lipid or protein particle among all lipid or protein particles on the MPPN surface. In the CP model we choose the target particle with equal probability among all particles. Next, we generate a unit vector with its initial point at the MPPN center and a target point that is chosen randomly from a uniform angular distribution. We update the MPPN configuration by rotating the target particle about the axis defined by this unit vector through an angular step size $\delta\theta = 0.005$ rad. The MC move is accepted with probability

$$p = \min\left(1, e^{-\Delta G/k_B T_{\rm sys}}\right), \tag{4}$$

where $T_{\rm sys}$ is the system temperature and ΔG is the difference in MPPN energy between the MPPN configurations after and before the attempted MC move. The maximum strength of the interaction potentials in Eqs. (1) and (3) is set by ϵ_2 and ϵ. As discussed in Secs. II A and II B, we use here $\epsilon_2 = \epsilon = 20\ k_B T$. In our simulated annealing MC simulations of the LP and CP models we therefore evolve, for the first 10^5 MC steps, the MPPN configurations using a system temperature that is increased twenty-fold compared to the room temperature, $T_{\rm sys} = 20T$, so as to allow thermal fluctuations to compete with the Lennard-Jones interactions considered here. To access minimum-energy MPPN configurations we then linearly decrease $T_{\rm sys}$ from $T_{\rm sys} = 20T$ to $T_{\rm sys} = 0$ over 10^5 additional MC steps. Finally, we let the system evolve for 10^4 further MC steps with $T_{\rm sys} = 0$, accepting only MC steps that result in a lower-energy MPPN configuration.

Figure 2 illustrates the MC cooling procedure outlined above for the CP model. In particular, Fig. 2 shows the total MPPN energy as a function of MC steps for five representative MC simulations of the CP model at $(N, n_o) = (24, 2)$. As illustrated in Fig. 2, our MC trajectories show a rapid transition from the random initial conditions to energetically more favorable protein configurations. Over the first 10^5 MC steps, we find large fluctuations in the MPPN energy, with substantial overlap in the energy fluctuations for the five representative MC trajectories shown in Fig. 2. This illustrates that $T_{\rm sys} = 20T$ is large enough for the system to explore different energetically favorable protein configurations irrespective of the initial conditions used. Once the linear cooling process is started (after 10^5 MC steps), the MC trajectories converge to the same minimum-energy MPPN configuration with, within numerical accuracy, the same MPPN energy. This convergence of different MC trajectories illustrates that our cooling process is slow enough to allow distinct MC trajectories to "find" the same energy minimum.

For both the LP and CP models, we repeat the above simulated annealing MC procedure for a range of MPPN

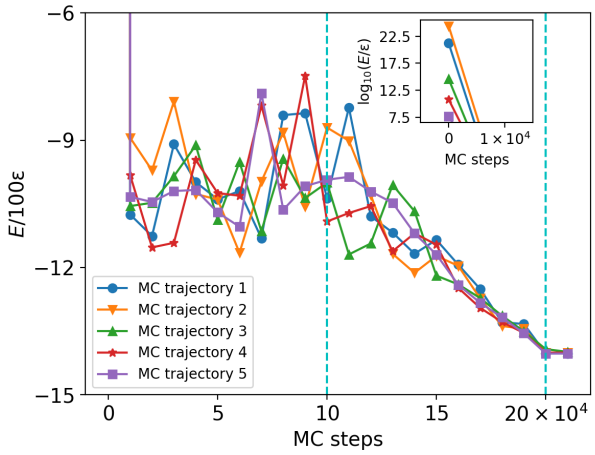

Figure 2. Illustrative MC trajectories of the CP model for five independent MC simulations showing the total MPPN energy calculated from Eq. (3) versus number of MC steps. We set $(N, n_o) = (24, 2)$, start from random initial conditions, and use our standard MC simulation procedure for the CP model including swapping moves (see Sec. II C). The inset shows the MPPN energy associated with the random initial conditions used for the five independent MC simulations. In the main panel, the curves leading to the first data point shown overlap for the five independent MC trajectories.

radii R to find the optimal MPPN radius R^* minimizing the MPPN energy. The value of R^* depends on (N, n_o) as well as the model under consideration. For the LP model with $N = 24$ we find, in agreement with previous work [25, 39], $R^* \approx 12.3$ nm for $n_o = 0$ and $R^* \approx 12.4$ nm for $n_o = 1$ with a snub cube symmetry of protein centers [see Figs. 1(a) and 1(b)]. To obtain R^* in the CP model for $N = 24$ and $n_o \geq 1$ we carry out simulated annealing MC simulations first with the value of $R = R^*$ found for $n_o - 1$, and then increase R to determine R^* with a resolution of 0.01 nm. For $n_o = 0$ we start our search for R^* at $R = 11$ nm. We thus find the optimal MPPN radii 12.06 nm $\lessapprox R^* \lessapprox$ 13.29 nm for $0 \leq n_o \leq 24$. A simple way to rationalize this increase in R^* with n_o is to regard the ratio of MPPN surface area to the area occupied by the disks representing closed-state and open-state MscS proteins as being approximately constant, which implies

$$\frac{(N - n_o + 1)\,\pi r_c'^2 + (n_o - 1)\,\pi r_o'^2}{4\pi\left[R^*(n_o - 1)\right]^2} \approx \frac{(N - n_o)\,\pi r_c'^2 + n_o \pi r_o'^2}{4\pi\left[R^*(n_o)\right]^2} \tag{5}$$

for $R^* = R^*(n_o - 1)$ and $R^* = R^*(n_o)$. Using the value $R^*(0) = 12.06$ nm found in our simulated annealing MC simulations, Eq. (5) allows us to recursively estimate $R^*(n_o)$ for $1 \leq n_o \leq 24$. Note that, as shown in Fig. 3, Eq. (5) implies that $(R^*)^2$, which is proportional to the MPPN surface area, depends approximately linearly on n_o. Equation (5) yields the optimal MPPN radii

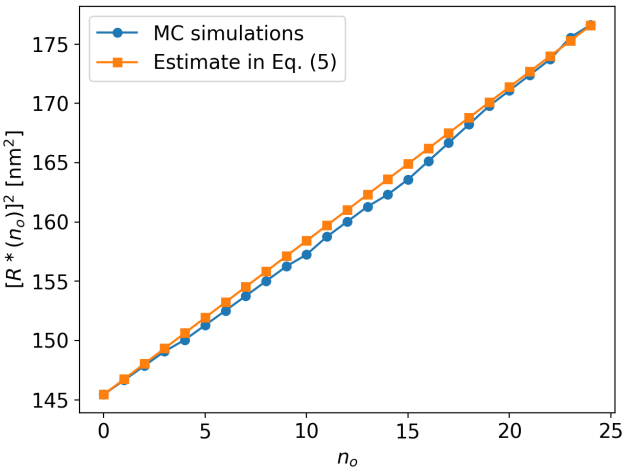

Figure 3. Square of the optimal MPPN radius, $(R^*)^2$, versus number of open-state MscS proteins, n_0, for $N = 24$ obtained from simulated annealing MC simulations of the CP model and the estimate in Eq. (5). For the MC simulations we used our standard MC simulation procedure including swapping moves (see Sec. II C).

12.11 nm $\lessapprox R^* \lessapprox$ 13.29 nm for $1 \leq n_o \leq 24$, which is in approximate agreement with the range of $R^*(n_o)$ obtained through our simulated annealing MC simulations (Fig. 3). Proceeding as for $N = 24$, we find $R^* \approx 7.84$ nm and $R^* \approx 8.6$ nm for the CP and LP models with $(N, n_o) = (12, 0)$, and $R^* \approx 14.07$ nm and $R^* \approx 15.0$ nm for the CP and LP models with $(N, n_o) = (32, 0)$, respectively (see Sec. III). For both the LP and CP models, we estimated R^* on the basis of 50 independent simulated annealing MC simulations for each MPPN radius considered, for which we used different random seeds. To determine the lowest-energy MPPN configuration implied by the LP model, we selected from the resulting set of 50 MPPN configurations at $R = R^*$ the MPPN configuration with the lowest energy. To determine the lowest-energy MPPN configuration implied by the CP model, we carried out 200 further simulated annealing MC simulations at $R = R^*$, and selected from the resulting set of 250 MPPN configurations at $R = R^*$ the MPPN configuration with the lowest energy.

For simulations of the LP and CP models with $n_o = 1$ we use as initial conditions the lowest-energy MPPN configurations obtained at $n_o = 0$ and randomly replace one closed-state MscS protein by an open-state MscS protein. When using the CP model to simulate MPPNs with $n_o > 1$ we take MscS proteins to gate sequentially, i.e., we use as the initial conditions for MPPNs with (N, n_o) and $n_o > 1$ the minimum-energy configurations found for $(N, n_o - 1)$ and randomly replace a closed-state MscS protein with an open-state MscS protein. Our simulations suggest that similar results for the minimum-energy MPPN configurations are obtained if one does not make this assumption. Here a technical difficulty arises in that the simulated annealing MC procedure described above fails to robustly identify the minimum-energy arrangements of open-state MscS in MPPNs with $n_o > 1$. To illustrate this issue we consider $(N, n_o) = (24, 2)$, in which case the arrangement of open-state MscS can be characterized by the angle α between the vectors pointing from the MPPN center to the particles representing open-state MscS [see inset in Fig. 4(a)]. The minimum-energy MPPN configuration corresponds to a snub cube arrangement of protein centers with the two open-state MscS being located across the diagonal of one of the square faces of the snub cube. This configuration corresponds to $\alpha \approx 1.12$ rad [see Figs. 1(e) and 4(a)]. Plotting the values of α associated with the final MPPN configurations obtained in our simulated annealing MC simulations, α_{final}, versus the values of α associated with the initial MPPN configurations, α_{init}, we find that our simulated annealing MC simulations are, in general, unable to overcome the energy barriers for interchanging closed-state and open-state MscS and, hence, fail to yield the minimum-energy MPPN configuration [see Figs. 4(a) and 4(b)].

To address the computational issue described above and illustrated in Figs. 4(a) and 4(b), we augment our simulated annealing MC simulations to allow for "swapping moves," in which the positions of closed-state and open-state MscS are randomly interchanged. In particular, we allow the "newly gated" MscS protein to randomly swap its position with a randomly selected closed-state MscS protein every 2×10^3 MC steps. The swapping move is accepted with a probability of the same form as in Eq. (4), but with ΔG being given by the difference in MPPN energy between the MPPN configurations after and before the attempted swapping move. If this MC move is not accepted, we let the system evolve as if it was accepted and, after 100 further MC steps, again apply Eq. (4), where ΔG is now calculated with respect to this evolved MPPN configuration. If, after this second attempt, the swapping move is still rejected, we revert the MPPN configuration to its state immediately prior to the attempted swapping move, and attempt the next swapping move after 2×10^3 further MC steps.

The motivation behind the "staggered" swapping procedure described above is that, since open-state MscS proteins have a larger size than closed-state MscS proteins, swapping of closed-state and open-state MscS proteins tends to strongly increase the MPPN energy unless the MPPN configuration is allowed to relax following the swapping move. Spurious rejection of swapping moves is thus avoided, speeding up our MC procedure. Our choice to evolve the system over 100 MC steps following an attempted swapping move does not have any deeper significance and, indeed, no such relaxation steps are needed if the number of independent MC simulations is large enough. We find that our results concerning energetically optimal MPPN configurations in Sec. III are robust with respect to the number of MC relaxation steps used. Figure 4(c) shows that, upon implementation of the swapping procedure, our simulated annealing

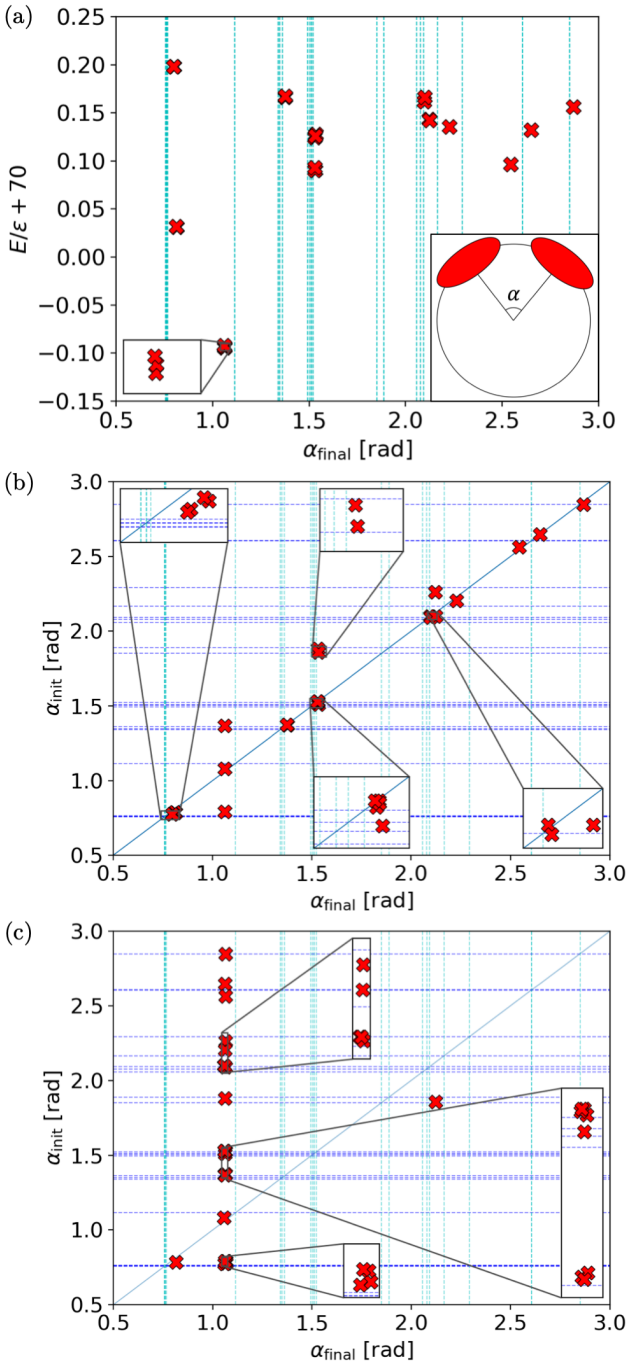

Figure 4. CP model with $(N, n_o) = (24, 2)$. (a) Total MPPN energy versus angular separation of open-state MscS in the final MPPN configuration, $\alpha = \alpha_{\text{final}}$ (see inset), and (b) α_{final} versus corresponding initial angular separation of open-state MscS, $\alpha = \alpha_{\text{init}}$. The results in panels (a) and (b) were obtained through simulated annealing MC simulations with no swapping moves (see Sec. II C). (c) α_{final} versus α_{init} as in panel (b) and using the same initial conditions as in panel (b) but allowing for swapping moves. The dashed vertical and horizontal lines show the values of α associated with perfect snub cube symmetry. The solid lines in panels (b) and (c) indicate $\alpha_{\text{final}} = \alpha_{\text{init}}$. Each cross symbol represents the result of one simulated annealing MC simulation. The insets show enlarged versions of the indicated regions in the plots.

MC simulations robustly identify the minimum-energy arrangement of open-state MscS proteins irrespective of the initial conditions used, and only in rare instances fail to produce the energetically preferred value of α_{final}.

Our simulations suggest that if one allows any open-state MscS protein—rather than just the newly opened MscS protein—to engage in swapping moves one obtains similar results for the minimum-energy MPPN configurations as with the sequential gating procedure used here. We find, however, that if any open-state MscS protein is allowed to engage in swapping moves a given simulated annealing MC simulation is less likely to successfully identify the energetically preferred MPPN configuration, making the sequential gating procedure employed here more efficient from a computational perspective. Furthermore, our simulations indicate that if one uses as the "initial" MPPN state $(N, n_o) = (24, 24)$—rather than $(N, n_o) = (24, 0)$—and sequentially closes open-state MscS proteins—rather than opens closed-state MscS proteins—one obtains similar minimum-energy MPPN configurations as those described here (see Sec. III).

D. Quantifying MPPN symmetry

The most straightforward approach for quantifying the symmetry of the MPPN configurations obtained through our simulated annealing MC simulations of the LP and CP models is to fit the protein centers to polyhedral vertices. In particular, we proceed as in experiments on MPPNs [22] and previous simulations of the LP model [25] and compare the protein arrangements in MPPNs with the symmetries implied by the 132 convex polyhedra with regular faces [42]: the Platonic (P), Archimedean (A), Catalan (C), and Johnson (J) solids. We denote the Platonic, Archimedean, and Catalan solids using the Conway polyhedron notation [43], and the Johnson solids using the indexing scheme developed in Ref. [44]. For each simulated MPPN configuration we quantify the fits to the aforementioned 132 polyhedral symmetries by calculating the fit error

$$\mathscr{E} = \sum_{i=1}^{N} (\vec{v_i} - \vec{v_{0i}})^2 \,, \tag{6}$$

where the vectors $\vec{v_i}$ point from the MPPN center to the positions of the protein centers on the MPPN surface obtained in our simulated annealing MC simulations, and the vectors $\vec{v_{0i}}$ denote the corresponding positions of the closest fitted polyhedron vertices. The latter are obtained [22, 45] by freely moving, rotating, and rescaling the polyhedron models until $\mathscr{E}$ is minimized for each polyhedral symmetry considered. Note that this fitting procedure yields for each MscS particle a closest polyhedron vertex. If the number of polyhedron vertices is smaller than N, each polyhedron vertex is associated with multiple MscS particles. In contrast, if the number of polyhedron vertices is greater than N, the fitting

procedure used here determines the subset of polyhedron vertices yielding the best fit to the MscS positions. We also note that in our simulated annealing MC simulations the preference of the minimum-energy MPPN configuration for one chiral polyhedral symmetry over its mirror-symmetric configuration results from the random numbers and initial conditions used [25], and is therefore not a model prediction.

In addition to Eq. (6), we characterize MPPN symmetry through bond-orientational order (BOO) parameters. BOO parameters have been employed to quantify local ordering in liquids and glasses [46–48] and, more recently, have been used to characterize the symmetry of protein shells [4]. BOO parameters are rotational invariants and, for MPPNs, can be constructed from the spherical harmonics of the protein positions on the MPPN surface. In particular, we employ here the BOO parameters Q_l [46],

$$Q_l = \left(\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |Q_{lm}|^2 \right)^{1/2}, \tag{7}$$

where

$$Q_{lm} = \frac{1}{N} \sum_{i=1}^{N} Y_{lm}(\vec{v_i}), \tag{8}$$

in which, as in Eq. (6), the vectors $\vec{v_i}$ point from the MPPN center to the centers of the proteins on the MPPN surface obtained in our simulated annealing MC simulations, and the $Y_{lm}(\vec{v_i})$ denote the corresponding spherical harmonics. We take here l in Eqs. (7) and (8) to be even so that Q_l is independent of the direction of a particular bond [46].

In Sec. III C we employ the BOO parameters Q_l in Eq. (7) to quantify how closely the MPPN configurations implied by the CP model resemble a snub cube [Fig. 1(e)]. To this end, we first use Eq. (7) to calculate Q_l for a perfect snub cube for even l starting from $l = 0$. We denote the values of Q_l associated with a perfect snub cube by $Q_l^{(\mathrm{sc})}$. We note that these values of $Q_l^{(\mathrm{sc})}$ are independent of the chirality of the snub cube. The relative difference between the Q_l associated with a simulated MPPN configuration and $Q_l^{(\mathrm{sc})}$ can then be expressed in the form

$$\widehat{Q_l} = \left| 1 - \frac{Q_l}{Q_l^{(\mathrm{sc})}} \right| . \tag{9}$$

We find that for even $l > 0$ the first non-zero $Q_l^{(\mathrm{sc})}$ occurs at $l = 4$, and $Q_l^{(\mathrm{sc})} = 0.0525$ and 0.0412 for $l = 4$ and $l = 6$. In Sec. III C we use $\widehat{Q_4}$ and $\widehat{Q_6}$ to characterize the symmetry of MPPNs with $N = 24$.

III. MINIMUM-ENERGY MPPN CONFIGURATIONS

In this section we discuss the results of our simulated annealing MC simulations of the LP and CP models of MPPN symmetry. We first compare the MPPN symmetries predicted by the LP and CP models (see Sec. III A). We then use the CP model to survey the minimum-energy MscS configurations in MPPNs with $N = 24$ and $0 \leqslant n_o \leqslant 24$. We thereby first consider qualitative features of the protein arrangement in MPPNs with heterogeneous protein size (see Sec. III B) and then quantify the symmetry of MPPNs with heterogeneous protein size (see Sec. III C).

A. Comparison of LP and CP models

As discussed in Sec. II, we consider here two models of MPPN symmetry: the LP model (see Sec. II A) and the CP model (see Sec. II B). The LP model separately accounts for the lipids and proteins in MPPNs. In contrast, the CP model focuses on the protein configurations in MPPNs and does not explicitly consider the arrangement of lipids in MPPNs. Table I shows the polyhedral symmetries obtained through simulated annealing MC simulations of the LP and CP models (see Sec. II C) for $N = 12$, 24, and 32 with $n_o = 0$. We fitted the simulated MPPN configurations to polyhedral symmetries as described in Sec. II D. For both the LP and CP models we find the minimum-energy MPPNs to have the symmetry of the icosahedron for $N = 12$ [see Fig. 5(a)], the snub cube for $N = 24$ [see Fig. 1(e)], and the pentakis dodecahedron for $N = 32$ [see Fig. 5(b)]. For $N = 12$ and $N = 32$, the LP and CP models give different results for the 2nd-best polyhedral fits to the minimum-energy MPPNs, but identical results for $N = 24$.

The icosahedron is a Platonic solid with 12 vertices,

		Best polyhedral fit		2nd-best polyhedral fit	
N	Model	Symmetry	$\mathscr{E}$ [nm^2]	Symmetry	$\mathscr{E}$ [nm^2]
12	CP	I	1.50×10^{-2}	J60	6.07×10
12	LP	I	4.18	tO	1.21×10
24	CP	sC (dextro)	4.70×10^{-2}	gD (laevo)	3.62×10
24	LP	sC (dextro)	2.96×10	gD (laevo)	6.60×10
32	CP	kD	1.06×10	jD	1.95×10
32	LP	kD	7.27×10	mD	7.78×10

Table I. Symmetries and associated fit errors $\mathscr{E}$ in Eq. (6) of the best two polyhedral fits to the minimum-energy MPPN configurations implied by the LP and CP models of MPPN symmetry for $N = 12$, 24, and 32 with $n_o = 0$. All results were obtained through simulated annealing MC simulations (see Sec. II C). We denote [43, 44] the icosahedron by I, the snub cube by sC, the pentakis dodecahedron by kD, the metabiaugmented dodecahedron by J60, the truncated octahedron by tO, the pentagonal hexecontahedron by gD, the rhombic triacontahedron by jD, and the disdyakis triacontahedron by mD. We proceeded as described in Sec. II D when searching for optimal polyhedral fits. The polyhedral chiralities result from the random numbers and initial conditions used and are not model predictions.

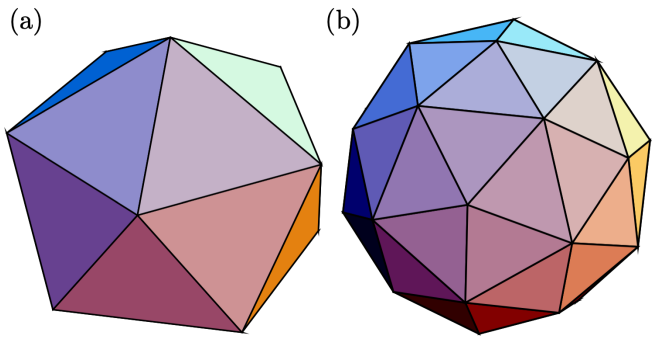

Figure 5. 3D representations of (a) the icosahedron (I) and (b) the pentakis dodecahedron (kD) in Table I.

while the pentakis dodecahedron is a lower-symmetric Catalan solid with 32 vertices. Experiments on MPPNs formed from MscS [22, 23] yielded MPPNs with $N = 24$ and snub cube symmetry as the dominant MPPN symmetry, which we also find through our simulated annealing MC simulations of the LP and CP models. We note that the snub cube is a chiral polyhedron. The right-handed (dextro) and left-handed (laevo) chiralities of the snub cube can be constructed by translating outward the faces of a cube, and rotating them clockwise (dextro) or counter-clockwise (laevo) as viewed from the polyhedron center until the polyhedral shell can be closed up with equilateral triangles. As noted in Sec. II D, the preference of the minimum-energy MPPN configuration for one chiral polyhedral symmetry over its mirror-symmetric configuration results from the random numbers and initial conditions used [25] for our simulated annealing MC simulations, and is therefore not a model prediction.

The polyhedra in Table I that provide the best fits to the simulated MPPN configurations all have N vertices. Interestingly, the 2nd-best polyhedral fits in Table I do not necessarily correspond to polyhedra with the same number of vertices as the number of proteins in MPPNs. For instance, the 2nd-best polyhedral fits in Table I for $(N, n_o) = (12, 0)$ are provided by the metabiaugmented dodecahedron (J60) and the truncated octahedron (tO), which have 22 and 24 vertices, respectively, and not by the truncated tetrahedron or the cuboctahedron, which both have 12 vertices. While we considered in Table I MPPNs with $N = 12$, 24, and 32, we find similar agreement of the dominant symmetries predicted by the LP and CP models for other values of N. Furthermore, we checked whether the LP and CP models yield identical results for the dominant MPPN symmetry for MPPNs with heterogeneous protein size. In particular, we find that the LP and CP models both yield the snub cube as the dominant symmetry of MPPNs with $(N, n_o) = (24, 1)$ [Figs. 1(b) and 1(d)]. As discussed in Secs. II A and II B, the CP model is conceptually simpler than the LP model and avoids some of the computational difficulties associated with finding the minimum-energy MPPN configurations in the LP model. In the remainder of this article we therefore focus on the CP model.

B. Protein arrangement in MPPNs with heterogeneous protein size

To determine the minimum-energy arrangement of MscS proteins in MPPNs [22, 23] with heterogeneous MscS size, we carried out simulated annealing MC simulations of the CP model for $N = 24$ and $0 \leqslant n_o \leqslant 24$. Independent of the value of n_o considered, we find that the minimum-energy MPPNs have the symmetry of a snub cube with one (closed-state or open-state) MscS protein being located at each one of the 24 vertices of the snub cube (see also Sec. III C). We find that, as n_o is increased from $n_o = 0$, the minimum-energy MscS arrangement in MPPNs follows a strikingly regular pattern (see Figs. 6 and 7). To specify, at each n_o, the minimum-energy MscS arrangement obtained in our simulated annealing MC simulations, we label in Figs. 6 and 7 the vertices of the snub cube by n_o to denote the position of the n_o^{th} open-state MscS protein. If there is more than one equivalent choice for the position of the n_o^{th} open-state MscS protein we introduce a subscript specifying the degree of degeneracy. In Fig. 6 we show the pattern of open-state MscS proteins found in our simulated annealing MC simulations. In Fig. 7 we provide 3D illustrations of selected MPPN configurations in Fig. 6, and in Fig. 8 we show some of the MPPN configurations corresponding to Fig. 6.

To understand the protein arrangement in MPPNs with $N = 24$ and heterogeneous protein size it is instructive to briefly recall the symmetry properties of the (undeformed) snub cube [49]. The snub cube is an Archimedean solid with 24 vertices and 38 faces corresponding to six squares, none of which share a vertex, and 32 equilateral triangles [Fig 1(e)]. All vertices in the snub cube are equivalent. The six square faces of the snub cube are associated with six four-fold rotational symmetry axes. Eight of the 32 triangular faces of the snub cube are associated with three-fold rotational symmetry axes, while 24 of the 32 triangular faces of the snub cube are associated with two-fold rotational symmetry axes. Figure 7(a) provides an illustration of the symmetry properties of the snub cube.

For $n_o = 1$, the single open-state MscS protein can equivalently occupy each one of the 24 vertices of the snub cube. We therefore have a 24-fold degeneracy in the position of the first open-state MscS protein (Fig. 6). In contrast, once the position of the first open-state MscS protein has been set, the energetically most favorable position of the 2nd open-state MscS protein is fixed (Fig. 6): The 2nd open-state MscS protein is arranged so that it is located diagonally across the square-shaped face of the snub cube from the first open-state MscS protein [see Fig. 8(a)]. In other words, the two open-state MscS proteins form a next-nearest-neighbor pair across a square

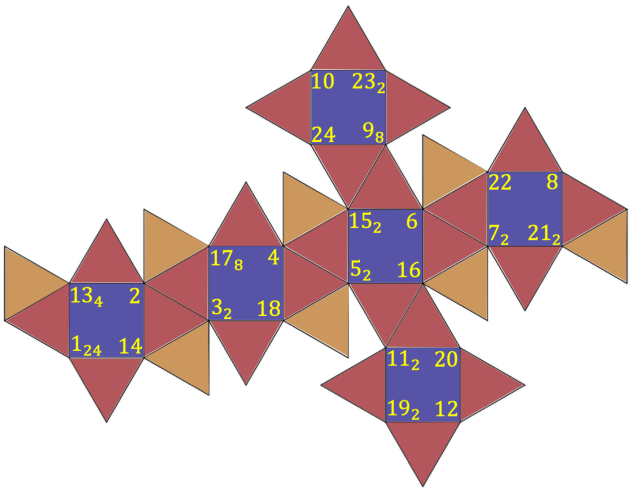

Figure 6. Minimum-energy protein arrangements in MPPNs implied by the CP model of MPPN symmetry for $N = 24$. The numbers labeling the vertices denote the positions of the n_o^{th} open-state MscS protein, with subscripts denoting the degree of degeneracy in placing the n_o^{th} open-state MscS protein. We omit this subscript if the position of the n_o^{th} open-state MscS protein is uniquely determined by the protein arrangement in MPPNs with $(n_o - 1)$ open-state MscS proteins. The degeneracy in placing open-state MscS proteins follows from the symmetry of the snub cube. The faces are colored according to their symmetry properties with red (gray), yellow (light gray), and blue (dark gray) colors indicating two-fold, three-fold, and four-fold symmetry axes, respectively (see also Fig. 7).

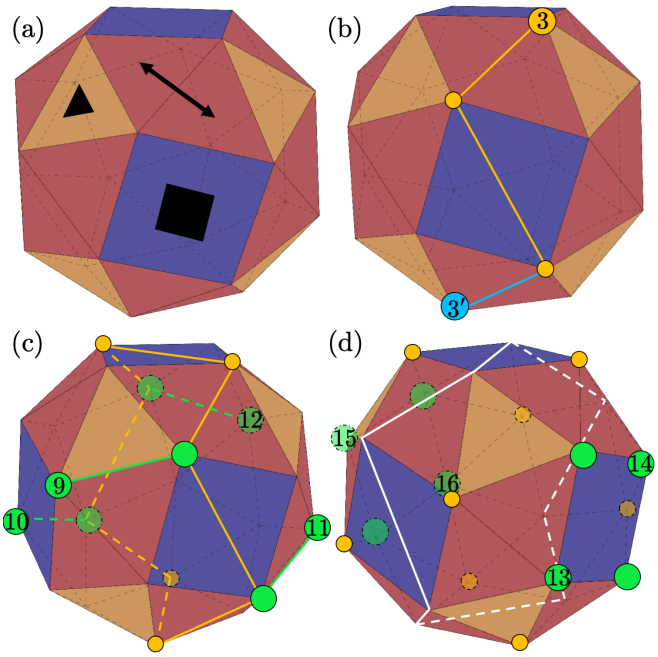

Figure 7. 3D illustrations of selected minimum-energy protein arrangements in Fig. 6 for (a) $n_o = 0$ or $n_o = 24$ (undeformed snub cube), (b) $n_o = 3$, (c) $n_o = 12$, and (d) $n_o = 16$. Following the labeling scheme in Fig. 6, the faces of the snub cube are colored according to their symmetry properties. In panel (a), the locations of some of the two-, three-, and four-fold symmetry axes of the snub cube are indicated by arrow, triangle, and square symbols, respectively, with the symmetry axes perpendicularly intersecting these symbols at their geometric centers. In panels (b), (c), and (d), vertices of the snub cube occupied by open-state MscS proteins are indicated by disks. In panel (b), the two geometrically equivalent choices for placing the 3rd open-state MscS protein are labeled as 3 and $3'$, respectively. In panels (c) and (d) we highlight the positions of the open-state MscS proteins occupying the "front" and "back" square faces of the snub cube through increased disk sizes with no labels (see main text), and label the positions of selected open-state MscS proteins with $n_o > 8$ by n_o. In panel (c), the closed zig-zag loop formed by the first eight open-state MscS proteins is indicated by small (orange) disks, while large (green) disks indicate the nearest-neighbor pairs formed by the 9th to 12th open-state MscS proteins with the open-state MscS proteins occupying the front and back square faces of the snub cube. In panel (d), the white lines show the closed zig-zag loop of closed-state MscS proteins formed at $n_o = 16$. Portions of the loops in panels (c) and (d) located at the back of the snub cube are indicated by dashed curves.

face of the snub cube. As a result, the symmetry axis associated with this square face is changed from a four-fold symmetry axis (for a perfect snub cube) to an axis with approximate two-fold rotational symmetry. The 3rd open-state MscS protein forms a nearest-neighbor pair across a triangular face of the snub cube with either one of the two other open-state MscS proteins so as to trace out a "zig-zag" pattern (Fig. 6). The two equivalent nearest-neighbor "bonds" associated with the 3rd open-state MscS protein intersect two-fold symmetry axes of the snub cube [see Fig. 7(b)]. The 4th open-state MscS protein is again located diagonally across a square face of the snub cube from an open-state MscS protein [see Fig. 8(b)]. Hence, its position is uniquely determined by the position of the 3rd open-state MscS protein (Fig. 6).

The aforementioned zig-zag pattern of alternating next-nearest-neighbor bonds of open-state MscS proteins across the square faces of the snub cube and nearest-neighbor bonds intersecting two-fold symmetry axes of the snub cube continues for $0 \leq n_o \leqslant 8$ (Fig. 6). At $n_o = 8$, the eight open-state MscS proteins thus connect up to form a closed loop [see Fig. 7(c)]. As a result, two square faces of the snub cube, located at opposite sides of the snub cube, are left devoid of any open-state MscS proteins [see the lower panel in Fig. 8(c)]. At $n_o = 9$, one of the eight vertices associated with these two square faces is occupied by the 9th open-state MscS protein, resulting in a degeneracy of eight for placing the 9th open-state MscS protein (Fig. 6). Note that the 9th open-state MscS protein forms a nearest-neighbor pair with some other open-state MscS protein located across a polyhedral ridge separating triangular faces associated with two- and three-fold symmetry axes [Fig. 7(c)].

The 10th open-state MscS protein forms a next-nearest-neighbor bond across a square face of the snub

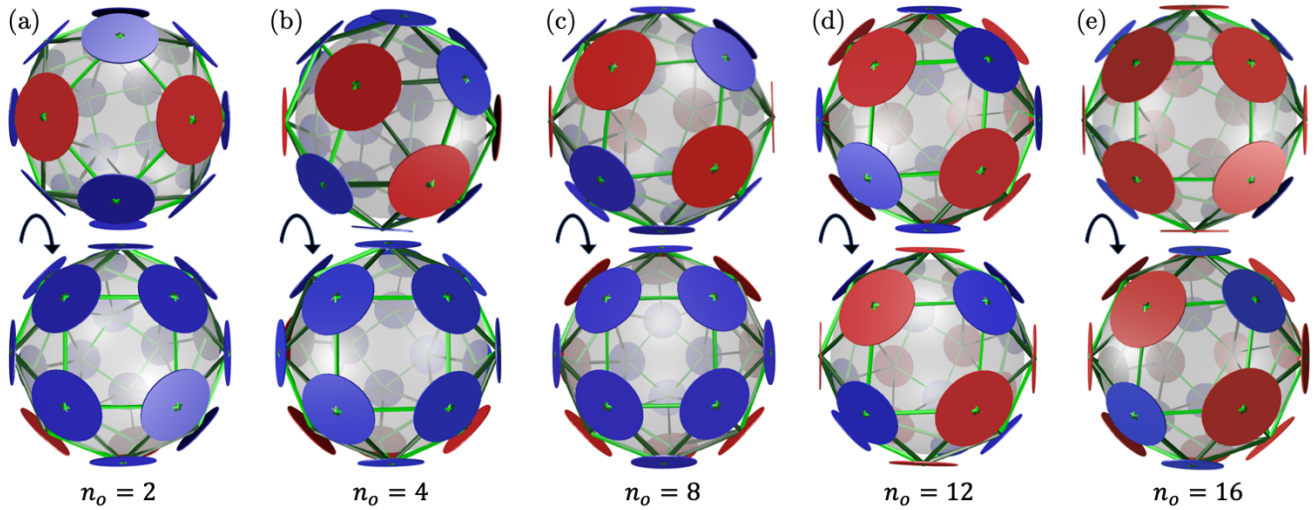

Figure 8. 3D representations of the minimum-energy MPPN configurations implied by the CP model of MPPN symmetry for $N = 24$ and (a) $n_o = 2$, (b) $n_o = 4$, (c) $n_o = 8$, (d) $n_o = 12$, and (e) $n_o = 16$ (see also Fig. 6). As in Fig. 1(d), closed-state and open-state MscS proteins are represented by small (blue) and large (red) disks, respectively, with the radii of these disks being decreased by some fixed scale factor relative to the disk radii implied by r'_o and r'_c. For clarity, polyhedral ridges enclosing (distorted) square faces of the snub cube are shown in bright green (light gray), while polyhedral ridges associated only with triangular faces of the snub cube are shown in dark green (dark gray).

cube with the 9th open-state MscS protein and, hence, has a position that is uniquely determined by the position of the 9th open-state MscS protein (Fig. 6). The polyhedral geometry of the snub cube mandates that, similarly as the 9th open-state MscS protein, the 10th open-state MscS protein forms a nearest-neighbor bond with some other open-state MscS protein located across a polyhedral ridge separating triangular faces with two- and three-fold symmetry axes [Fig. 7(c)]. As a result, the square face of the snub cube containing the 9th and 10th open-state MscS proteins connects two square faces of the snub cube, each containing two open-state MscS proteins, that are part of the zig-zag loop formed at $n_o = 8$. Following Fig. 7(c), we denote the latter two faces as the "front" and "back" faces of the zig-zag loop formed at $n_o = 8$, respectively.

At $n_o = 11$, the last remaining square face of the snub cube containing only closed-state MscS proteins starts to be occupied by open-state MscS proteins [Figs. 6 and 7(c)]. Two vertices on that square face are—via open-state MscS proteins that are part of the front and back faces of the zig-zag loop of open-state MscS proteins formed at $n_o = 8$ in Fig. 7(c)—next-next-nearest neighbors of the 9th and 10th open-state MscS proteins. We find that either one of these two vertices is occupied by the 11th open-state MscS protein. Hence, there is a two-fold degeneracy in the position of the 11th open-state MscS protein. The 12th open-state MscS protein is again located diagonally across a square face of the snub cube from an open-state MscS protein, and its position is therefore uniquely determined by the position of the 11th open-state MscS protein [Figs. 6 and 7(c)]. Thus, the square face occupied by open-state MscS proteins at $n_o = 11$ and $n_o = 12$ connects, similarly as the square face containing the 9th and 10th open-state MscS proteins, the front and back faces of the zig-zag loop of open-state MscS proteins formed at $n_o = 8$ in Fig. 7(c). At $n_o = 12$, all square faces of the snub cube contain two open-state MscS proteins, which form next-nearest-neighbor pairs across the diagonals of the square faces [see Figs. 7(c) and 8(d)].

As n_o is increased beyond $n_o = 12$, open-state MscS proteins start to form nearest-neighbor bonds on the square faces of the snub cube. We find that first the front and back faces of the zig-zag loop of open-state MscS proteins formed at $n_o = 8$ in Fig. 7(c) are fully occupied by open-state MscS proteins. In particular, at $n_o = 13$, there are four geometrically equivalent choices for the location of the 13th open-state MscS protein, resulting in a four-fold degeneracy for placing the 13th open-state MscS protein (Fig. 6). The 14th open-state MscS protein is located diagonally across a square face of the snub cube from an open-state MscS protein, and its position is therefore uniquely determined by the position of the 13th open-state MscS protein (Fig. 6). Similarly, we have a two-fold degeneracy in the protein position at $n_o = 15$, with the position of the 16th open-state MscS protein being determined uniquely by the position of the 15th open-state MscS protein (Fig. 6).

Note that, at $n_o = 16$, there are four geometrically equivalent square faces of the snub cube containing two closed-state MscS proteins each. Connecting the vertices associated with these closed-state MscS proteins, we obtain a closed zig-zag loop with similar geometric properties as the loop of open-state MscS proteins found at $n_o = 8$ [see Fig. 7(d)]. As n_o is increased beyond $n_o = 16$,

the vertices of this loop are occupied by open-state MscS proteins following a pattern that is analogous to that obtained for $0 \leq n_o \leq 8$ (Fig. 6). We find an eight-fold degeneracy in the protein position at $n_o = 17$ and two-fold degeneracies in the protein position at $n_o = 19$, 21, and 23, respectively, with the positions of open-state MscS proteins at even n_o being determined uniquely by the positions of open-state MscS proteins at odd n_o. At $n_o = 24$, all vertices of the snub cube are occupied by open-state MscS proteins.

C. Quantifying the symmetry of MPPNs with heterogeneous protein size

In this section we employ the mathematical approaches described in Sec. II D to quantify the symmetry of MPPNs with heterogeneous protein size. To determine how closely the protein configurations in MPPNs follow polyhedral symmetry, it is convenient to introduce, based on the fit error $\mathscr{E}$ in Eq. (6), the dimensionless root-mean-square fit error

$$\widehat{\sigma} = \frac{1}{\lambda}\sqrt{\frac{\mathscr{E}}{N}} \tag{10}$$

with $\lambda = 10^{-2} r_c'$ as the characteristic length scale. In Table II we list $\widehat{\sigma}$ in Eq. (10) for the best two polyhedral fits for MPPNs with $N = 24$ and $0 \leqslant n_o \leqslant 24$. We calculated these values of $\widehat{\sigma}$ from the minimum-energy MPPN configurations obtained in our simulated annealing MC simulations of the CP model. As already noted in Sec. III B we find that, independent of the value of n_o considered, the best polyhedral fits in Table II always correspond to snub cube (sC) symmetry.

Figure 9 shows $\widehat{\sigma}$ in Table II versus n_o for the best polyhedral fits. We also provide in Fig. 9 the range in $\widehat{\sigma}$ associated with the ten lowest-energy MPPN configurations obtained, at each n_o, in our simulated annealing MC simulations, as well as the average $\widehat{\sigma}$ associated with these ten lowest-energy MPPN configurations. Each one of these ten lowest-energy MPPN configurations corresponds to one independent MC trajectory. At each n_o, the best polyhedral fits to the ten lowest-energy MPPN configurations in Fig. 9 all correspond to snub cube symmetry. Note from Fig. 9 that the root-mean-square fit error in Eq. (10) is smallest for MPPNs with homogeneous protein composition ($n_o = 0$ and $n_o = 24$). Similarly, comparison of fits of the simulated MPPN configurations to the snub cube with fits to competing polyhedral symmetries shows that the snub cube symmetry is most dominant for homogeneous or nearly homogeneous protein compositions (see Table II).

Starting from $n_o = 0$, the increase in $\widehat{\sigma}$ with n_o in Fig. 9 and Table II can be understood by noting that for MPPNs with heterogeneous protein composition the polyhedral symmetry must deform so as to accommodate proteins of different size. For instance, the square faces of the snub cube in Fig. 8 containing a mixture of closed-state and open-state MscS proteins are seen to deviate from a perfect square. We find a maximum in $\widehat{\sigma}$ in Fig. 9 at $n_o = 8$. As already noted in Sec. III B, $n_o = 8$ yields an MPPN configuration with four deformed square faces of the snub cube forming a closed loop, which "flattens" the polyhedron and increases the fit error [Figs. 6 and 8(c)]. An analogous protein configuration is obtained at $n_o = 16$ [Figs. 6 and 8(e)], which may explain the large values of $\widehat{\sigma}$ found in our simulations in the vicinity of $n_o = 16$. A (weak) local minimum occurs in Fig. 9 at $n_o = 12$. As noted in Sec. III B, at $n_o = 12$ all square faces of the snub cube are composed of two closed-state and two

	Best polyhedral fit		2nd-best polyhedral fit	
n_o	Symmetry	$\widehat{\sigma}$	Symmetry	$\widehat{\sigma}/10$
0	sC (dextro)	1.13	gD (laevo)	3.06
1	sC (laevo)	3.16	eC	3.66
2	sC (dextro)	5.52	gD (laevo)	3.31
3	sC (laevo)	7.21	gD (dextro)	3.22
4	sC (dextro)	8.31	gD (dextro)	3.24
5	sC (laevo)	8.93	gD (dextro)	3.34
6	sC (laevo)	1.08×10	gD (dextro)	3.04
7	sC (dextro)	1.21×10	gD (laevo)	3.03
8	sC (dextro)	1.44×10	sC (laevo)	3.15
9	sC (laevo)	1.37×10	sC (dextro)	3.25
10	sC (laevo)	1.29×10	gD (laevo)	3.67
11	sC (laevo)	1.24×10	gD (laevo)	3.35
12	sC (laevo)	1.20×10	sC (dextro)	3.58
13	sC (dextro)	1.23×10	J45 (dextro)	3.56
14	sC (laevo)	1.30×10	gD (dextro)	3.4
15	sC (dextro)	1.30×10	gD (laevo)	3.14
16	sC (laevo)	1.29×10	J45 (laevo)	3.52
17	sC (dextro)	1.18×10	gD (laevo)	3.21
18	sC (laevo)	1.07×10	gD (dextro)	3.15
19	sC (dextro)	9.23	gD (laevo)	3.24
20	sC (laevo)	8.23	gD (laevo)	3.49
21	sC (dextro)	7.04	gD (dextro)	3.48
22	sC (laevo)	6.16	gD (dextro)	3.42
23	sC (laevo)	3.47	gD (dextro)	3.75
24	sC (dextro)	1.08	gD (dextro)	3.54

Table II. Symmetries and associated root-mean-square fit errors $\widehat{\sigma}$ in Eq. (10) of the best two polyhedral fits to the minimum-energy MPPN configurations implied by the CP model of MPPN symmetry for $N = 24$ and the indicated values of n_o. All results were obtained through simulated annealing MC simulations (see Sec. II C). We use the same notation for polyhedral symmetries as in Table I [43, 44] with, in particular, gD corresponding to the pentagonal hexecontahedron, and denote the rhombicuboctahedron by eC and the gyroelongated square bicupola by J45. We proceeded as described in Sec. II D when searching for optimal polyhedral fits. The polyhedral chiralities result from the random numbers and initial conditions used and are not model predictions.

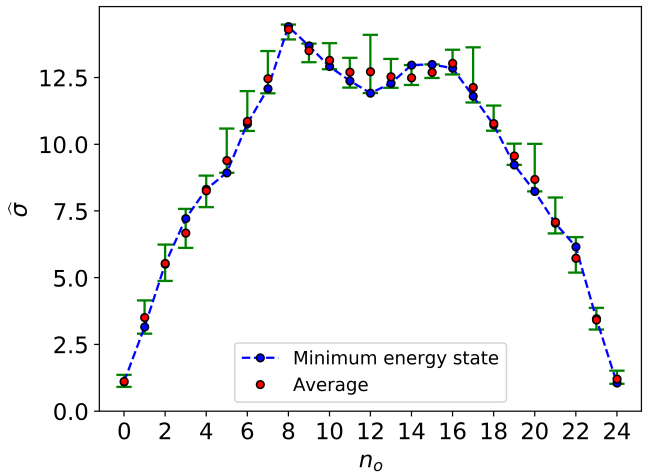

Figure 9. Root-mean-square fit errors $\widehat{\sigma}$ in Eq. (10) of the best polyhedral fit (snub cube symmetry) to the minimum-energy MPPN configurations implied by the CP model of MPPN symmetry for $N = 24$ versus number of open-state MscS proteins, n_o [blue (dark gray) data points]. For each n_o, we also show the range in $\widehat{\sigma}$ associated with the ten lowest-energy MPPN configurations obtained in our simulated annealing MC simulations, which all correspond to snub cube symmetry. This range in $\widehat{\sigma}$ is indicated by bars, with the red (light gray) data points showing the average $\widehat{\sigma}$ for the ten lowest-energy MPPN configurations obtained in our simulated annealing MC simulations. See also Table II.

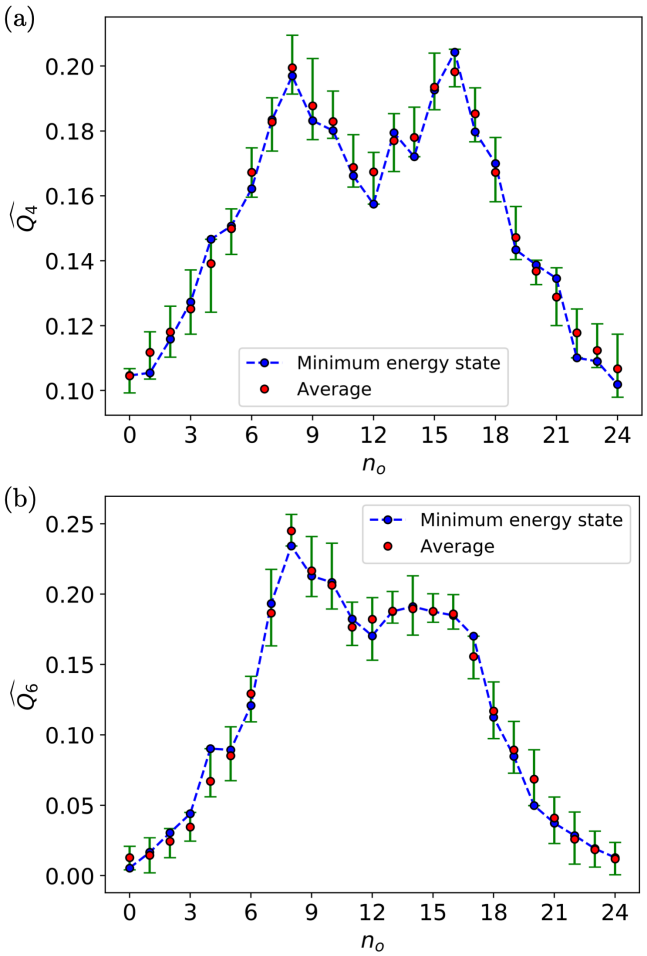

Figure 10. Relative difference in BOO parameters between a snub cube and the MPPN configurations implied by the CP model, $\widehat{Q_l}$, versus number of open-state MscS, n_o, at (a) order $l = 4$ and (b) order $l = 6$ in Eq. (9). For each n_o, we show the range in $\widehat{Q_l}$ associated with the ten lowest-energy MPPN configurations obtained in our simulated annealing MC simulations, which all correspond to snub cube symmetry. This range in $\widehat{Q_l}$ is indicated by bars, with the red (light gray) data points showing the average $\widehat{Q_l}$ for the ten lowest-energy MPPN configurations and the blue (dark gray) data points showing the $\widehat{Q_l}$ associated with the minimum-energy MPPN configurations. The plots in panels (a) and (b) are obtained from the data for low-energy MPPN configurations also used in Fig. 9.

open-state MscS proteins arranged in the same pattern [Figs. 6 and 8(d)], which may explain the relatively small value of $\widehat{\sigma}$ at $n_o = 12$.

In addition to the dimensionless root-mean-square fit error in Eq. (10), the BOO parameters Q_l in Eq. (7) provide a mathematical approach for quantifying MPPN symmetry. In the case of the snub cube, the first two nonzero values of the BOO parameters Q_l occur at $l = 4$ and $l = 6$ (see Sec. II D). In Figs. 10(a) and 10(b) we plot $\widehat{Q_l}$ in Eq. (9), which corresponds to the relative difference between the values of Q_l associated with a (perfect) snub cube and the minimum-energy MPPN configurations implied by the CP model, for $0 \leqslant n_o \leqslant 24$ and $l = 4$ and $l = 6$, respectively. The parameters $\widehat{Q_{4,6}}$ do not depend on the chirality of the snub cube. For completeness, we also show in Fig. 10 the range in $\widehat{Q_{4,6}}$ associated with the ten lowest-energy MPPN configurations obtained, at each n_o, in our simulated annealing MC simulations, as well as the corresponding average values of $\widehat{Q_{4,6}}$. With the definition of $\widehat{Q_l}$ in Eq. (9), smaller values of $\widehat{Q_l}$ indicate a closer resemblance of the protein arrangement in MPPNs to a snub cube. Consistent with the results in Fig. 9, Fig. 10 shows local peaks at $n_o = 8$ and $n_o = 16$, as well as a local minimum at $n_o = 12$. Similarly as for Fig. 9, these features of Fig. 10 can be understood by noting that the protein arrangements found for $n_o = 8$ and $n_o = 16$ correspond to deformed square faces of the snub cube forming closed loops [Figs. 6, 8(c), and 8(e)], while for $n_o = 12$ all square faces of the snub cube are occupied by two closed-state and two open-state MscS proteins [Figs. 6 and 8(d)]. Note that the $\widehat{Q_4}$-versus-n_o curve in Fig. 10(a) shows two peaks at $n_o = 8$ and $n_o = 16$ of approximately equal height. In contrast, the $\widehat{Q_6}$-versus-n_o curve in Fig. 10(b) shows a more pronounced peak at $n_o = 8$ than at $n_o = 16$. At a qualitative level, the $\widehat{Q_6}$-versus-n_o curve in Fig. 10(b) thus resembles the $\widehat{\sigma}$-

versus-n_o curve in Fig. 9.

IV. SUMMARY AND CONCLUSIONS

We have used here computational modeling to explore the symmetry of MPPNs with heterogeneous protein size. Our computational modeling approach is closely related to previous models describing the symmetry of viral capsids [4, 7, 31, 32]. Although, from an experimental perspective, MPPNs and viral capsids are quite distinct—with proteins in MPPNs being embedded in a lipid bilayer environment but viral capsids being composed solely of proteins—our results suggest that MPPN symmetry and viral capsid symmetry are governed by similar physical principles [25]. Analogous modeling approaches may also be applicable to other kinds of systems forming polyhedral shells [1–5]. Motivated by previous experimental studies of MPPNs [22, 23], we have focused here on MPPNs composed of 24 closed-state or open-state MscS proteins [28–30]. However, our modeling approach is easily generalized to other types of MPPNs.

An important distinction between previous studies of protein shells [1, 4, 5, 7, 31, 32] and the model of MPPN symmetry developed here is that, in the former case, a key question concerns the symmetry of protein shells as a function of the number of protein subunits. In contrast, a central question for MPPNs is how the protein arrangement in MPPNs changes after some proteins in MPPNs transition to a different conformational state following, for instance, osmotic shock, while leaving the membrane intact [22, 23]. We have therefore focused here on MPPNs containing a fixed number of proteins. In the spirit of previous work on the symmetry of closed protein shells [4, 7, 31, 32], we have employed highly idealized models of MPPN symmetry. Our modeling approach could be extended in various ways to allow more detailed predictions. For instance, we assumed here similar interactions between closed-state and open-state MscS proteins, with the only difference in their interaction potentials stemming from the distinct sizes of closed-state and open-state MscS proteins [28–30]. In general, different conformational states of a given membrane protein or distinct kinds of membrane proteins may show distinct interactions in MPPNs. Furthermore, we assumed here that the particles representing membrane proteins (and lipids) are confined to the surface of a sphere. While this assumption is justified for the observed MPPNs formed from MscS proteins [22, 23], it may not hold in general. For instance, for large enough protein numbers MPPNs may, in analogy to protein shells [50, 51], buckle into faceted shapes.

A key outcome of our study is that MPPNs with heterogeneous protein size can be highly symmetric, with a well-defined polyhedral (snub cube) ordering of membrane proteins of different sizes. MPPNs have been proposed [22] as a means for the structural analysis of membrane proteins in the presence of physiologically relevant transmembrane gradients. Such transmembrane gradients are expected to result in heterogeneous protein size, with different proteins being trapped in different conformational states, while leaving the membrane intact. Our finding that MPPNs with heterogeneous protein size can be highly symmetric suggests that it may be feasible to utilize MPPNs for structural studies [22, 52] even if not all membrane proteins in MPPNs are trapped in the same conformational state. In particular, for MPPNs formed from 24 MscS proteins [22, 23] we predict that the first eight gated (open-state) MscS proteins form a closed zig-zag loop, resulting in two square faces of the snub cube that are devoid of any open-state MscS proteins (Figs. 6–8). For more than eight open-state MscS proteins, these two square faces are gradually filled with open-state MscS proteins until, for twelve open-state MscS proteins, all square faces of the snub are occupied by two open-state MscS proteins located diagonally across the square faces of the snub cube. As the number of open-state MscS proteins is increased further, the square faces of the snub cube that connect—via bonds between open-state MscS proteins—the square faces of the snub cube that were not part of the original closed zig-zag loop of open-state MscS proteins are filled with open-state MscS proteins until, for more than sixteen open-state MscS proteins, the remaining vertices of the snub cube are populated by open-state MscS proteins.

In addition to their potential use for structural studies, MPPNs have also been proposed as a novel vehicle for targeted drug delivery with precisely controlled release mechanisms [22]. To this end, it is desirable to arrive at a quantitative understanding of the physical mechanisms governing MPPN self-assembly, shape, and stability for arbitrary protein compositions. The self-assembly, symmetry, and size of MPPNs with homogeneous protein composition [22, 23] can be understood based on a simple mean-field model combining protein-induced lipid bilayer deformations, topological defects in protein packing, and thermal effects [25, 26]. This mean-field approach cannot be directly applied to MPPNs with heterogeneous protein composition, because in MPPNs with heterogeneous protein composition not all proteins are equivalent. However, the well-defined and regular protein arrangements in MPPNs with heterogeneous protein size found here suggest that it may be practicable to generalize the theory of MPPNs [25, 26] to allow for heterogeneous protein compositions. Such a generalized theory of MPPNs may allow prediction of how MPPN self-assembly must be directed to produce MPPNs with given release mechanisms [22].

ACKNOWLEDGMENTS

This work was supported by NSF award number DMR-1554716 and the USC Center for High-Performance Computing.

[1] G. Vernizzi and M. Olvera de la Cruz. *Proc. Natl. Acad. Sci. U.S.A.*, 104(47):18382–18386, 2007.

[2] T. O. Yeates. *Nature*, 569:340, 2019.

[3] A. D. Malay, N. Miyazaki, A. Biela, S. Chakraborti, K. Majsterkiewicz, I. Stupka, C. S. Kaplan, A. Kowalczyk, B. M. Piette, G. K. Hochberg, and D. Wu. *Nature*, 569(7756):438, 2019.

[4] M. Mosayebi, D. K. Shoemark, J. M. Fletcher, R. B. Sessions, N. Linden, D. N. Woolfson, and T. B. Liverpool. *Proc. Natl. Acad. Sci. U.S.A.*, 114(34):9014–9019, 2017.

[5] S. Panahandeh, S. Li, and R. Zandi. *Nanoscale*, 10(48):22802–22809, 2018.

[6] D. L. D. Caspar and A. Klug. In *Cold Spring Harbor Symp. Quant. Biol.*, volume 27, pages 1–24. Cold Spring Harbor Laboratory Press, 1962.

[7] R. F. Bruinsma and W. S. Klug. *Annu. Rev. Condens. Matter Phys.*, 6(1):245–268, 2015.

[8] S. Li, P. Roy, A. Travesset, and R. Zandi. *Proc. Natl. Acad. Sci. U.S.A.*, 115(43):10971–10976, 2018.

[9] S. Bahatyrova, R. N. Frese, C. A. Siebert, J. D. Olsen, K. O. van der Werf, R. van Grondelle, R. A. Niederman, P. A. Bullough, C. Otto, and C. N. Hunter. *Nature*, 430(7003):1058–1063, 2004.

[10] D. Baddeley, I. D. Jayasinghe, L. Lam, S. Rossberger, M.B. Cannell, and C. Soeller. *Proc. Natl. Acad. Sci. U.S.A.*, 106(52):22275–22280, 2009.

[11] A. Briegel, X. Li, A. M. Bilwes, K. T. Hughes, G. J. Jensen, and B. R. Crane. *Proc. Natl. Acad. Sci. U.S.A.*, 109(10):3766–3771, 2012.

[12] J. Liu, B. Hu, D. R. Morado, S. Jani, M. D. Manson, and W. Margolin. *Proc. Natl. Acad. Sci. U.S.A.*, 109(23):E1481–E1488, 2012.

[13] C. Jiko, K. M. Davies, K. Shinzawa-Itoh, K. Tani, S. Maeda, D. J. Mills, T. Tsukihara, Y. Fujiyoshi, W. Kühlbrandt, and C. Gerle. *eLife*, 4:e06119, 2015.

[14] V. Sourjik and N. S. Wingreen. *Curr. Opin. Cell Biol.*, 24(2):262–268, 2012.

[15] S.-Y. Park, P. P. Borbat, G. Gonzalez-Bonet, J. Bhatnagar, A. M. Pollard, J. H. Freed, A. M. Bilwes, and B. R. Crane. *Nat. Struct. & Mol. Biol.*, 13(5):400, 2006.

[16] J. J. Sieber, K. I. Willig, C. Kutzner, C. Gerding-Reimers, B. Harke, G. Donnert, C. Rammner, B.and Eggeling, S. W. Hell, H. Grubmüller, and T. Lang. *Science*, 317(5841):1072–1076, 2007.

[17] S. Mondal, J. M. Johnston, H. Wang, G. Khelashvili, M. Filizola, and H. Weinstein. *Sci. Rep.*, 3:2909, 2013.

[18] C. A. Haselwandter and N. S. Wingreen. *PLoS Comput. Biol.*, 10(12):e1003932, 2014.

[19] A. Briegel, M. L. Wong, H. L. Hodges, C. M. Oikonomou, K. N. Piasta, M. J. Harris, D. J. Fowler, L. K. Thompson, J. J. Falke, L. L. Kiessling, and G. J. Jensen. *Biochemistry*, 53(10):1575–1585, 2014.

[20] A. M. Pollard and V. Sourjik. *J. Biol. Chem.*, 293(6):2149–2158, 2018.

[21] O. Kahraman and C. A. Haselwandter. *Soft Matter*, 15(21):4301–4310, 2019.

[22] T. Basta, H.-J. Wu, M. K. Morphew, J. Lee, N. Ghosh, J. Lai, J. M. Heumann, K. Wang, Y. C. Lee, D. C. Rees, and M. H. B. Stowell. *Proc. Natl. Acad. Sci. U.S.A.*, 111(2):670–674, 2014.

[23] H.-J. Wu, T. Basta, M. Morphew, D. C. Rees, M. H. B. Stowell, and Y. C. Lee. Microfluidic device for superfast evaluation of membrane protein crystallization. In *The 8th Annual IEEE International Conference on Nano/Micro Engineered and Molecular Systems*, pages 84–87, April 2013.

[24] N. Levina, S. Tötemeyer, N. R. Stokes, P. Louis, M. A. Jones, and I. R. Booth. *The EMBO J.*, 18(7):1730–1737, 1999.

[25] D. Li, O. Kahraman, and C. A. Haselwandter. *Phys. Rev. Lett.*, 117:138103, 2016.

[26] D. Li, O. Kahraman, and C. A. Haselwandter. *EPL*, 117:58001, 2017.

[27] E. S. Haswell, R. Phillips, and D. C. Rees. *Structure*, 19(10):1356–1369, 2011.

[28] R. B. Bass, P. Strop, M. Barclay, and D. C. Rees. *Science*, 298(5598):1582–1587, 2002.

[29] S. Steinbacher, R. B. Bass, P. Strop, and D. C. Rees. *Structures of the Prokaryotic Mechanosensitive Channels MscL and MscS*, volume 58 of *Curr. Top. Membr.* Academic Press, 2007.

[30] W. Wang, S. S. Black, M. D. Edwards, S. Miller, E. L. Morrison, W. Bartlett, C. Dong, J. H. Naismith, and I. R. Booth. *Science*, 321(5893):1179–1183, 2008.

[31] R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, and J. Rudnick. *Proc. Natl. Acad. Sci. U.S.A.*, 101(44):15556–15560, 2004.

[32] S. N. Fejer, D. Chakrabarti, and D. J. Wales. *ACS Nano*, 4(1):219–228, 2010.

[33] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. *Science*, 220(4598):671–680, 1983.

[34] D. Frenkel and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications.* Academic Press, 1996.

[35] K. V. Damodaran and K. M. Merz. *Langmuir*, 9(5):1179–1183, 1993.

[36] D. H. Boal. *Mechanics of the Cell.* Cambridge University Press, 2nd edition, 2012.

[37] N. Ben-Tal, A. Ben-Shaul, A. Nicholls, and B. Honig. *Biophys. J.*, 70(4):1803–1812, 1996.

[38] S. Choe, K. A. Hecht, and M. Grabe. *J. Gen. Physiol.*, 131(6):563–573, 2008.

[39] D. Li. PhD thesis, University of Southern California, 2017.

[40] B. De Kruijff, P. R. Cullis, and G. K. Radda. *Biochim. Biophys. Acta (BBA) - Biomembranes*, 406(1):6 – 20, 1975.

[41] W. Ben-Ameur. *Comput. Optim. Appl.*, 29:369–385, 12 2004.

[42] G. W. Hart. The encyclopedia of polyhedra. `https://www.georgehart.com/virtual-polyhedra/vp.html`, 2000. Accessed: 2019-05-28.

[43] J. H. Conway, H. Burgiel, and C. Goodman-Strauss. *The Symmetries of Things.* Ak Peters Series. Taylor & Francis, 2008.

[44] N. W. Johnson. *Canadian Journal of Mathematics*, 18:169200, 1966.

[45] Polyhedron search. `https://sourceforge.net/projects/polyhedronsrch/`, 2013. Accessed: 2019-02-25.

[46] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti. *Phys. Rev. B*, 28:784–805, 1983.

[47] W. Lechner and C. Dellago. *J. Chem. Phys.*, 129(11):114707, 2008.

[48] W. Mickel, Sebastian C. Kapfer, G. E. Schrder-Turk, and K. Mecke. *J. Chem. Phys.*, 138(4):044501, 2013.

[49] P. R. Cromwell. *Polyhedra*. Cambridge University Press, 1997.

[50] J. Lidmar, L. Mirny, and D. R. Nelson. *Phys. Rev. E*, 68(5):051910, 2003.

[51] E. H. Yong, D. R. Nelson, and L. Mahadevan. *Phys. Rev. Lett.*, 111(17):177801, 2013.

[52] Z. Liu, H. Yan, K. Wang, T. Kuang, J. Zhang, L. Gui, X. An, and W. Chang. *Nature*, 428(6980):287, 2004.