Free energy profile and kinetics of the formation of membrane invaginations by membrane-bending peripheral proteins

Mohsen Sadeghi[1], *

[1] *Department of Mathematics and Computer Science,*
Freie Universität Berlin, Arnimallee 12, 14195 Berlin, Germany

Curvature-inducing peripheral proteins have been observed to spontaneously remodel bilayer membranes, resulting in membrane invaginations and formation of membrane tubules. In case of proteins such as cholera and Shiga toxins that bend the membrane with locally isotropic curvatures, the resulting membrane-mediated interactions are rather small. Thus, the process in which these proteins form dense clusters on the membrane and collectively induce invaginations is extremely slow, progressing over several minutes. This makes it virtually impossible to directly simulate the pathway leading to membrane tubulation even with highly coarse-grained models. Here, we present a steered molecular dynamics protocol through which the peripheral proteins are forced to gather on a membrane patch and form a tubular invagination. Using thermodynamic integration, we obtain the free energy profile of this process and discuss its different stages. We show how protein stiffness, which also determines local membrane curvatures, affects the free energy landscape and the organization of proteins in the invaginated region. Furthermore, we estimate the kinetics of the described pathway modeled as a Markovian stochastic process, and compare the implied timescales with their experimental counterparts.

I. INTRODUCTION

Transport of material between living cells and their environment involves their passage through the cell membrane. While small molecules and ions can either directly permeate the membrane or be transported through designated channels, larger molecules, biomolecular assemblies, and particulate substances usually rely on endo/exocytosis, for which the membrane undergoes significant remodeling [1]. This can be mediated by scaffolding proteins such as clathrin and caveolin [2–5], or be accomplished by the active cellular machinery capable of directly exerting force on the membrane, e.g. actin and dynamin assemblies [6]. On the other hand, toxins such as Shiga, cholera and ricin can internalize in the absence of both mechanisms [7–10]. These toxins usually comprise of distinct pathogenic and membrane-bending subunits. The latter, through a collaborative effort by several protein copies gathered in a dense cluster, can bend the membrane into long tubules necessary for the internalization of the former (the glycolipid-lectin (GL-Lect) hypothesis) [11, 12]. Apart from its toxicological importance, this pathway can potentially be manipulated to deliver drugs or antigens to the living cells [12].

Spontaneous formation of tubular membrane invaginations can in general occur due to different physical mechanisms. For example, adsorption of molecules that change the spontaneous curvature of the membrane due to asymmetric tension can result in spontaneous tubulation [13]. Protein-protein crowding is also a significant driver of membrane tubulation [14]. Osmotic pressure resulting in volume reduction, or a growth in membrane area can also lead to the formation of stable membrane tubules [15]. We direct the interested reader for more examples to see [16] and references within. To understand how membrane budding and tubulation can progress under the influence of membrane-bending peripheral proteins, the most reliable approach is to estimate the free energy landscape of this process. Several studies have previously addressed this problem. Ewers et al. offered a theoretical model that included the local curvature induced (and preferred) by each protein and its corresponding energy in an interplay with the large energy barrier due to neck formation; to investigate the tubulation free energy profile and the possibility of spontaneous tubule nucleation [17]. Tourdot et al. used a continuum membrane model with curvature fields and monitored the change in excess chemical potential of membrane-bending proteins to mark the tubulation threshold [18]. Mathijs et al. considered a Ginzburg-Landau-type free energy for the protein coat and discussed the stability of the cylindrical tubes covered with proteins [19]. Mahapatra and Rangamani developed a theoretical model to estimate the free energy of membrane tubulation due to bound BAR-domain proteins with anisotropic curvature, and showed the snap-through characteristics in the dome-to-cylinder transitions [20]

Here we offer a fresh take on this problem, using a mesoscopic dynamical membrane model [21]. We had previously shown this model to reproduce realistic kinetics attributed to membrane fluctuations [22, 23] as well as lateral diffusion of membrane-bound peripheral proteins [24]. While the dynamical nature of this model allows for the kinetics of the proposed pathway to be studied, it also limits the timescales available to unbiased simulations. To remedy the latter, while benefiting from the former, we lay out a scheme for steered simulations, in which peripheral proteins are artificially

---

* mohsen.sadeghi@fu-berlin.de

directed along the tubulation pathway. We show how the free energy profile of all stages of this process can be recovered, and kinetics of its different steps approximated. We discover that the difference in protein stiffness (which affects how it imprints its curvature on the membrane) plays a deciding role in the spontaneous formation of membrane invaginations.

II. THEORY

Free energy estimation from steered simulations: In the most general sense, the free energy difference between two thermodynamic states A and B is defined as the external work done on the system as it evolves between the two states in a *slow transition*. The transition should be slow enough that thermodynamic equilibrium can be reasonably assumed for the intermediate states. Considering $W_{A\rightarrow B}$ to represent this external work,

$$F_B - F_A = W_{A\rightarrow B} = \int_{\lambda_A}^{\lambda_B} f\left(\lambda\right) d\lambda \tag{1}$$

where λ denotes a specific (macroscopic) parametrization of the path taken from A to B, and $f\left(\lambda\right)$ represents the external force exerted on the system at each step along the transition path. We assume that the microscopic state of the system can be fully described by the phase vector $\mathbf{z} = \left(\mathbf{q}, \mathbf{p}\right)$, where components of $\mathbf{q}$ are positions of particles, plus any other degrees of freedom necessary to the system's configuration (e.g. box size), and components of $\mathbf{p}$ are the conjugate momenta. We can thus connect the macroscopic and microscopic descriptions of the transition by substituting $f\left(\lambda\right) = \langle f\left(\mathbf{z}\right)\rangle_\lambda$, where $\langle\cdots\rangle_\lambda$ is an ensemble average over the generally time-dependent distribution of microscopic states constrained with an instantaneous value of λ.

On the other hand, with the assumption of slow transition, we can consider any macrostate tagged by a value of λ to follow the Boltzmann distribution with the partition function $Q \propto \int \exp\left(-\beta u_\lambda\left(\mathbf{z}\right)\right) d\mathbf{z}$ and $\beta = 1/kT$. To simplify the formulation, we have used the notion of generalized potential u_λ, which, for example, in the isothermal-isobaric ensemble would be $u_\lambda = \mathcal{H} + pV$, with $\mathcal{H}$ being the system's Hamiltonian, p the surrounding pressure, and V the volume [25, 26]. In general, in the presence of an external force $\mathbf{J}$, the generalized potential has the form $u_\lambda = \mathcal{H} - \mathbf{J}\cdot\mathbf{x}$, where $\mathbf{x}$ are configurational variables conjugates to the external force $\mathbf{J}$ [27]. Equilibrium free energy of each state is simply $\beta F = -\ln Q$, resulting in,

$$F_B - F_A = \int_{\lambda_A}^{\lambda_B} \frac{\partial F}{\partial \lambda} d\lambda = \frac{-1}{\beta}\int_{\lambda_A}^{\lambda_B} \frac{1}{Q}\frac{\partial Q}{\partial \lambda} d\lambda = \int_{\lambda_A}^{\lambda_B} \left\langle \frac{\partial u_\lambda}{\partial \lambda}\right\rangle_\lambda d\lambda \tag{2}$$

hence reiterating the relationship $f\left(\mathbf{z}\right) = \partial u_\lambda\left(\mathbf{z}\right)/\partial\lambda$, which can be easily understood as the microscopic balancing between (external) force and the gradient of the generalized potential energy of the system.

Reaction coordinate and the external force: In this study, we have performed steered simulations in which peripheral proteins on the surface of the membrane are drawn towards the center of the membrane patch (figure 1A). This happens under the influence of two-dimensional centrifugal forces of constant magnitude $\tilde{f}$. For the macroscopic parameter λ that describes the transition of the system (i.e. the reaction coordinate), we have chosen the mean in-plane radial distance of protein particles from the center of simulation box,

$$\lambda = \bar{r}_c = \frac{1}{N_\mathrm{p}} \sum_{i=1}^{N_\mathrm{p}} r_{c,i} \tag{3}$$

where $r_{c,i}$ is the two-dimensional radial distance of particle i from the center and N_p is the total number of proteins on the membrane surface. Simulations are done in a box with dimensions L_x, L_y, and L_z and the membrane initially lies in the xy plane, coupled to a barostat controlling xy in-plane pressure.

For the generalized potential, we use the expression $u_\lambda = \mathcal{H} - \sum_j \tilde{f} r_{c,j} + \frac{1}{2}\left(p_x + p_y\right) A\, L_z$, with the $A = L_x \times L_y$ being the projected area of the membrane patch that may fluctuate due to the action of the barostat. Using chain-rule of differentiation, we have,

$$\frac{\partial u_\lambda}{\partial \lambda} = \frac{\partial}{\partial \lambda}\left(N_\mathrm{p}\lambda\tilde{f}\right)_A + \frac{\partial u_\lambda}{\partial A}\frac{\partial A}{\partial \lambda} = N_\mathrm{p}\tilde{f} + \frac{1}{\lambda}\left(p_x + p_y\right) A\, L_z \tag{4}$$

Thus, by keeping the radial external forces constant, and monitoring the pressure tensor during the simulation, we can use eq. 2 to obtain free energies.

Kinetics: Given that the free energy profile as a function of the reaction coordinate is available via the described approach, one simple approach to estimating the kinetics of spontaneous processes is to consider a diffusion process on this one-dimensional energy landscape [28]. To estimate the timescales of slow processes, we took two different approaches. The first approach, which is less accurate, but more interpretable, is based on the approximate model of a double-well potential. Partial transition times between a pair of connected harmonic potential wells, with energy minima $E_A = E\left(x_A\right)$ and $E_C = E\left(x_C\right)$, separated by a barrier of maximum energy $E_B = E\left(x_B\right)$, can be estimated by the Kramers equation [29–31],

$$\langle\tau_{A\rightarrow C}\rangle = \frac{2\pi kT}{D\sqrt{\left|E''\left(x_A\right)E''\left(x_B\right)\right|}} \exp\left(\frac{E_B - E_A}{kT}\right) \tag{5}$$

where $E''\left(x\right)$ is the second spatial derivatives of the energy function, and D is the appropriate diffusion coefficient for the random walk on the energy landscape. The assumption in Kramers model is that there exist a separation of timescales between the fast local dynamics in each potential well, and the slow dynamics of barrier-crossing. In our case, we case identify several double-well-like regions in the free energy landscape and apply eq. 5 sequentially between them. This would be an approximation, as for each double-well, we would ignore all transitions other than the intermediate barrier crossing. While there exist models in which transitions between multiple energy wells have been considered [32], they also come with other sets of simplifying assumptions that do not hold in our case.

The second approach is to use the transfer operator that describes the evolution of probability densities resulting from the diffusion on the whole energy landscape. Considering the time-dependent probability density $p_t(x)$, and its weighted counterpart $u_t(x) = p_t(x)/\mu(x)$ with $\mu(x)$ being the equilibrium (stationary) probability density, the transfer operator $\mathcal{T}_\tau$ with the lag-time τ is defined as [33, 34],

$$u_{t+\tau}(y) = \mathcal{T}_\tau \circ u_t(y) = \frac{1}{\mu(y)} \int p(x,y;\tau)\,\mu(x)\,dx \tag{6}$$

where $p(x,y;\tau)$ is the transition probability density between states x and y (i.e. $p(x,y;\tau)\,dxdy$ is the probability of transitioning between states in $[x, x+dx]$ to states in $[y, y+dy]$). Eigenvalues of the transfer operator (ω_i's) are related to the timescales of transitions as $t_i = -\tau/\ln\omega_i$ [34, 35]. Making the assumption of over-damped diffusion in the reaction-coordinate space, we can estimate the transition probabilities from Brownian dynamics,

$$p(x,y;\tau) \propto \exp\left[\frac{\left((y-x)+\frac{D_{\overline{r}_c}\tau}{kT}E'(x)\right)^2}{4D_{\overline{r}_c}\tau}\right] \tag{7}$$

where $D_{\overline{r}_c}$ is the diffusion coefficient associated with random walks in the $\overline{r}_c$ space and $-E'(x)$ is the force at position x.

III. MODELING AND SIMULATION

All simulations have been performed using the membrane model developed by the author [21–24, 36]. In this model, the membrane comprises of a series of particle dimers, with each particle representing a patch of lipids on one leaflet. Particle dimers interact with their nearest neighbors via Morse-type bond-stretching and harmonic angle-bending potentials, and in-plane bond-flipping moves are implemented to mimic the fluidity of the bilayer [21]. More recently, we showed how peripheral proteins can be incorporated in the model via force field masking [24, 36]. Here we have used the same set of parameters as presented in [24, 36].

We have modeled two types of proteins with the stiffness values of $Y_\mathrm{p} = 100\,\mathrm{MPa}$ and $Y_\mathrm{p} = 200\,\mathrm{MPa}$. Both proteins are considered to have a geometry of the membrane-binding subunit similar to Shiga toxin subunit B (STxB). We have used this geometric representation to connect the stiffness values of a solid model of the protein to their elastic contribution to the membrane-protein energy landscape (see [24] and its supplementary information). Protein stiffness has a distinct effect on the distributions of local membrane curvature. This is reflected in the data we obtained from unbiased simulations with the two protein types (figure 1A). The range of curvatures obtained here can be compared with previously reported values, (i) membrane curvature of $0.055 \pm 0.012\,\mathrm{nm}^{-1}$ measured in curvature sorting study of Cholera toxin subunit B (CTxB) [37], (ii) curvature of $0.02\,\mathrm{nm}^{-1}$ extrapolated from the change in area per molecule when STxB binds to gel-phase monolayers [38], and (iii) molecular dynamics simulation of STxB binding to membranes containing the receptor Gb3, which predicts curvatures of $0.034 \pm 0.004\,\mathrm{nm}^{-1}$ or $0.035 \pm 0.003\,\mathrm{nm}^{-1}$, depending on acyl chain saturation [39].

Flat membrane patches of $300\,\mathrm{nm}$ lateral size with a number of bound peripheral proteins randomly scattered are used as initial configuration of each simulation. For each two protein type, 10 different realizations of the steered simulations that start from different initial configuration are performed (figure 1B). External forces with the magnitude of $0.5\,\mathrm{pN}$, and the direction continually passing through the central axis of the simulation box, are applied on each protein (figure 1A). The centrifugal forces cause a drift in the stochastic motion of the proteins, guiding them toward clustering in the center of the box within the time limit of each realization. Simulations are allowed to run until no further development in the membrane remodeling is observed. In some cases, the simulations were terminated prematurely due to instabilities caused in the coarse-grained model under the influence of the external force. In the end, all the stable parts of trajectories were cropped and used together.

During each simulation, particle positions have been propagated using anisotropic over-damped Langevin dynamics that includes hydrodynamic interactions between nearest neighbors [22, 23]. Water, with the dynamic viscosity of $0.7\,\mathrm{mPa\,s}$, has been used as the surrounding fluid environment. Simulation box is laterally coupled to a zero-pressure bath via a stochastic Langevin-piston barostat [40, 41]. All simulations have been performed at constant $310\,\mathrm{K}$ temperature.

IV. RESULTS AND DISCUSSION

Under the influence of the centrifugal external forces, proteins cluster in the center of the membrane patch and eventually induce a pit that further develops into an invagination (figures 1B and 1C and supplementary movies). Before moving forward with the free energy estimation procedure, we need to verify that the steered simulations are indeed performed slow enough for the membrane system to be considered at quasi-equilibrium throughout. As a first step, we can check simulation times against the physical timescales of the system. Fluctuations of the membrane, with the wave vector $\mathbf{q}$, have a hydrodynamic dissipation timescale of $4\eta/\kappa q^3$, where η is the viscosity of the solvent and $\kappa \approx 20\,kT$ is the bending rigidity of the membrane. This gives the hydrodynamic dissipation timescale to be in the $0.01\,\mathrm{ms}$ range. The proteins in the present model have a lateral diffusion constant of $D_\mathrm{p} \approx 0.3\,\mathrm{\mu m^2\,s^{-1}}$ [24]. The average time for proteins to diffuse the mean distance Δr between nearest neighbors can be estimated as $\Delta r^2/4D_\mathrm{p}$. With the surface density of proteins used in these simulations, we arrive at an upper limit of $0.1\,\mathrm{ms}$ for the protein diffusion. As the next step, and to get a more reliable estimate of the timescale of slow diffusive processes, we consider unbiased simulations of proteins diffusing on the membrane in the absence of external forces. We used a similar collective variable $\overline{r}_c$, measured with respect to the center of mass of proteins scattered on the membrane patch. By looking at the initial decay in the time-correlation function of $\overline{r}_c$, we estimated timescales between $0.05\,\mathrm{ms}$ and $0.12\,\mathrm{ms}$, depending on the protein stiffness and surface concentration of proteins (figure 1B). Thus, in comparison, our steered simulations progress on timescales at least one order of magnitude slower than the slowest process involved in our system. This justifies the quasi-equilibrium picture used in the thermodynamic integration scheme. Finally, from a different perspective, we can compare the magnitude of the external forces on the proteins in our simulations with more relatable biological examples. Using optical traps, the unidirectional force that the molecular motor kinesin produces to move cargo along microtubules has been measured to be as large as $6.5\,\mathrm{pN}$ [42]. Another example is the forces generated by the growth of a single actin filament, which has been estimated to be $0.8\,\mathrm{pN}$ [43]. Thus, the forces used in our steered simulation are one order of magnitude smaller than the force exerted by the motor protein, and comparable with the action of the cytoskeleton (figure 1B).

Being assured of the soundness of the protocol developed for steered simulation, we have used eqs. 2 and 4 to estimate the free energy profile of the formation of membrane invaginations. Figure 2A shows the estimated ensemble averages $\langle \partial u_\lambda/\partial\lambda \rangle_\lambda$ for the two proteins, obtained by binning trajectories from all realizations along $\overline{r}_c$. The free energy values are obtained form eq. 2 via numerical integration of $\langle \partial u_\lambda/\partial\lambda \rangle_\lambda$ using the trapezoid rule. The uncertainty in $\langle \partial u_\lambda/\partial\lambda \rangle_\lambda$ ensemble averages have been estimated via bootstrapping, and they have been propagated through the thermodynamic integration. For that, we have resampled each discrete value along the integration path from its corresponding distribution. For a better estimate, numerical integrations are repeated from both directions (λ_A to λ_B and vice versa) and a mean of two values with the accompanying error have been used (figure 2B).

Overall, free energy profiles reveal that the formation of a membrane invagination is energetically favorable for both proteins (figure 2B). To put the range of free energies into perspective, we can use the Helfrich Hamiltonian $\mathcal{H} = \int [2\kappa\,(H-H_0)^2 + \bar{\kappa}K]dA$, where H, H_0 and K are the mean, spontaneous and Gaussian curvatures and κ and $\bar{\kappa}$ are the corresponding elastic constants [44, 45], to estimate the free energies of idealized membrane geometries relevant to the invagination process. For a hemispherical cap, we obtain an energy of $2\pi\,(2\kappa+\bar{\kappa})$, which for $\kappa = 20\,kT$ and $\bar{\kappa} \approx -0.8\kappa$ [46, 47], gives a value of $150\,kT$. Also, a cylindrical membrane tubule of length L and radius R would have an energy of $4\pi\kappa L/R$, which for the approximate maximum length of $150\,\mathrm{nm}$ and mean tube radius of $50\,\mathrm{nm}$ observed in our simulations, results in a value as high as $750\,kT$. This demonstrates the significant contribution of the elastic energy of proteins. The dip in free energy occurs when the proteins successfully coat the pits, and thus compensate the rather large energies needed to deform the membrane.

We observe significant free energy barriers at the onset of pit formation ($3 \rightarrow 4 \rightarrow 5$ transitions for the $100\,\mathrm{MPa}$ and $4 \rightarrow 5 \rightarrow 6$ transitions for the $200\,\mathrm{MPa}$ proteins in figure 2B). Heights of these barriers are comparable to the $30-90\,kT$ range of excess chemical potentials estimated by Tourdot et al. for the onset of membrane tubule formation [18]. While the curved region of the pit complies with the preferred curvature of peripheral proteins, and reduces the overall elastic energy, the connecting neck region is energetically unfavorable and poses the barrier [9]. The organization of proteins in dense clusters on the membrane would also affect the barriers. We observe the $200\,\mathrm{MPa}$ proteins to experience a significantly smaller barrier at the onset of pit formation compared to their less stiff counterpart (figure 2B). Comparing snapshots along the reaction coordinate (figure 2B), as well as observing the trajectories of steered simulations (supplementary movies) reveals that stiffer $200\,\mathrm{MPa}$ proteins have a higher tendency for forming clusters as they are pushed together. We had previously observed the same trend in unbiased simulations [24]. We also measured larger pairwise attractions between these proteins [48]. We can therefore argue that the stiffer proteins can be organized in a pit with less energy expenditure, and experience an overall smaller barrier at the onset of pit formation.

While simulations with $100\,\mathrm{MPa}$ proteins progress to smaller $\overline{r}_c$ values, this progression corresponds to the formation of a pit densely covered by proteins (figure 2B and supplementary movie 1). With the $200\,\mathrm{MPa}$ proteins, on the other hand, a deep invagination at larger $\overline{r}_c$ values is fully formed (figure 2B and supplementary movie 2. Notice the $1/\overline{r}_c$ scaling on the x-axis of figure 2B). Then $200\,\mathrm{MPa}$ proteins also disperse over the invaginated area less densely. This may be attributed to the repulsive curvature-mediated interactions between these proteins at close range, due to large local curvatures induced by each (figure 1A) [49]. Considering the free energy differences between the invaginated and flat states, it can be stated that while formation of tubular invaginations is a slow process governed by large energy barriers, it is essentially irreversible in the current setup, unless other parameters such as the membrane tension are changed.

To estimate the kinetics of the stages of the invagination process, as highlighted with the free energy landscape, we used the two approaches described in section II. For both approaches, we needed an estimate of $D_{\overline{r}_c}$. This value is generally not constant and depends on $\overline{r}_c$. Also, as was shown in figure 1B, the assumption of Markovian dynamics in $\overline{r}_c$ space (equivalent to the overdamped dynamics described in eq. 7) only holds over a large lag-time of $\approx 0.1\,\mathrm{ms}$. Thus, we approached this by binning the trajectories along $\overline{r}_c$ and using asymptotic fits to the mean squared displacements in each bin to obtain long-time estimates to the diffusion coefficient. The result is an estimate of $D_{\overline{r}_c}$ accompanied by the uncertainties resulting from limited sample sizes in each bin (figure 3A).

The local double-well approximation of the free energy profile yields transition times that depict a path severely hindered by the intermediate high barriers (figure 3B). While the maximum estimated barrier height for the $200\,\mathrm{MPa}$ proteins in transition from flat membrane and dispersed protein to the pit formation is $\approx 25\,kT$, for $100\,\mathrm{MPa}$ proteins, this value can reach $\approx 45\,kT$. Such a high energy barrier in effect renders it impossible for these proteins to spontaneously form an invagination.

The more accurate analysis is possible using the transfer operator framework (eq. 6). We have used the same lag-time of $0.1\,\mathrm{ms}$ to build discretized transition matrices (figure 3C). Comparing the spectrum of eigenvalues of the transfer operator for the two cases shows how for the $100\,\mathrm{MPa}$ proteins, the eigenvalues persist in the vicinity of 1 for several transition modes, while for $200\,\mathrm{MPa}$ proteins they soon start to drop to lower values (figure 3C). Noteworthy is that the timescales implied by these transition modes also include the reverse process of going from invaginated states to the flat configuration, and hence the number of eigenvalues close to one. We can distinguish which transition is represented by an eigenvalue by looking at the corresponding eigenvector [34]. For the two protein types studied here, we have selected the two eigenvectors that best match the dynamics of final barrier crossing before a pit is successfully formed (figure 3C). The implied timescales for these transitions for the $100\,\mathrm{MPa}$ proteins are prohibitively large, while for $200\,\mathrm{MPa}$ proteins we arrive at an upper limit of $50\,\mathrm{min}$ (figure 3C). We may compare this timescale with the time it takes STxB proteins to induce tubules when incubated with giant unilamellar vesicles (GUV's). In these experiments, tubulation had been observed after $5\,\mathrm{min}$ [8]. While our estimated timescale might seem large in comparison, considering the uncertainties in the free energy estimate, and the fact that the exponential dependence causes each kT of barrier energy to almost triple the timescale, we have indeed achieved remarkably comparable estimates.

V. CONCLUSION

We have used a particle-based membrane model including particles representing curvature-inducing proteins to study the formation of membrane tubular invaginations using steered simulations. We obtained the free energy profile of this process from thermodynamic integration (figure 2). Formation of similar invaginations have been observed with dynamically triangulated membrane models in Monte Carlo simulations [18, 39]. Compared to that, our work has two distinctive features: (i) our model reproduces the long-range effects of the curvatures induced by peripheral proteins and the resulting static and dynamic membrane-mediated interactions intrinsically and without the need for explicit enforcement [24, 48], and (ii) using steered molecular dynamics, we get a glimpse of the actual process of proteins clustering on the membrane, followed by a pit being formed and developed into a membrane invagination. Our steered simulations not only highlighted the rather rough free energy landscape stretching between the flat to the invaginated configurations, but also made it possible to consider the kinetics of this process in details. We went to great lengths to validate the kinetics of this model [22–24]. We also took great care in establishing the protocol for the steered simulations in this work. Because of that we are able to trust the kinetic predictions of the model and compare the resulting timescales with the experimental observations.

We have modeled the invagination process as occurring under the influence of proteins already bound to the membrane and not being adsorbed on it during the deformation. This follows the experimental observation of STxB's binding to GUV's and forming dense clusters before any invaginations are observed [8]. The recruitment and binding of toxins such as Shiga and Cholera with pentameric membrane-binding subunits is in itself a process worthy of detailed study [50], and can be addressed in mesoscopic simulations with other free energy methods [51]. It is imaginable that including a reservoir of proteins ready to be recruited [52] or performing grand canonical simulations can lead to smoother energy landscapes and faster kinetics. This might also explain the generally larger timescales that we obtained with our simulations compared to experimental values.

We can imagine the general approach laid out in this work to be applicable to a variety of similar systems, especially including BAR-domain proteins with anisotropic curvatures. Our approach to estimating the free energy of these complex membrane and protein systems results in detailed landscapes across a very wide range of protein aggregation and membrane deformation. Thus, it can help provide mechanistic understanding of processes such as clathrin-independent endocytosis [53], or endoplasmic reticulum reorganization [54] at their native scale. Besides using more realizations of steered simulations and improving estimation accuracies, our approach is amenable to further improvement in the future via: (i) using more realistic models of peripheral proteins interacting with the membrane, (ii) including membrane tension as an important factor in the invagination process, and (iii) including effects due to the presence of cytoskeletal structures. We believe the tools developed in our modeling framework and the introduced methods for free energy estimation and kinetics analysis pave the way toward investigating these scenarios and offer powerful means of quantitatively understanding membrane-protein interplay at biological spatiotemporal scales.

CONFLICT OF INTEREST STATEMENT

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

The author developed the model and the theory, performed the simulations and analyzes, and prepared the manuscript.

FUNDING

This research has been funded by Deutsche Forschungsgemeinschaft (DFG) through grants SFB 958/A04 and SFB 1114/C03.

ACKNOWLEDGMENTS

The author is especially grateful to Frank Noé (FU Berlin) for constant support and valuable comments and discussions.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the author upon reasonable request.

[1] Alberts, B.; Johnson, A.; Lewis, J.; Morgan, D.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*, 6th ed.; Garland Science, Taylor & Francis Group, LLC: New York, 2015.

[2] Kaksonen, M.; Roux, A. Mechanisms of clathrin-mediated endocytosis. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 313–326.

[3] Jin, M.; Shirazinejad, C.; Wang, B.; Yan, A.; Schöneberg, J.; Upadhyayula, S.; Xu, K.; Drubin, D. G. Branched actin networks are organized for asymmetric force production during clathrin-mediated endocytosis in mammalian cells. *Nat. Commun.* **2022**, *13*, 1–12.

[4] Nawara, T. J.; Williams, Y. D.; Rao, T. C.; Hu, Y.; Sztul, E.; Salaita, K.; Mattheyses, A. L. Imaging vesicle formation dynamics supports the flexible model of clathrin-mediated endocytosis. *Nat. Commun. 2022 131* **2022**, *13*, 1–14.

[5] Alaoui, F. E.; Casuso, I.; Sanchez-Fuentes, D.; Arpin-Andre, C.; Rathar, R.; Baecker, V.; Castro, A.; Lorca, T.; Viaud, J.; Vassilopoulos, S.; Carretero-Genevrier, A.; Picas, L. Structural organization and dynamics of FCHo2 docking on membranes. *Elife* **2022**, *11*, 1–19.

[6] Shin, W.; Zucker, B.; Kundu, N.; Lee, S. H.; Shi, B.; Chan, C. Y.; Guo, X.; Harrison, J. T.; Turechek, J. M.; Hinshaw, J. E.; Kozlov, M. M.; Wu, L. G. Molecular mechanics underlying flat-to-round membrane budding in live secretory cells. *Nat. Commun.* **2022**, *13*, 2022.02.02.478826.

[7] Torgersen, M. L.; Skretting, G.; Van Deurs, B.; Sandvig, K. Internalization of cholera toxin by different endocytic mechanisms. *J. Cell Sci.* **2001**, *114*, 3737–3747.

[8] Römer, W. et al. Shiga toxin induces tubular membrane invaginations for its uptake into cells. *Nature* **2007**, *450*, 670–675.

[9] Ewers, H.; Helenius, A. Lipid-mediated endocytosis. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, 1–14.

[10] Sandvig, K.; Kavaliauskiene, S.; Skotland, T. The protein toxins ricin and shiga toxin as tools to explore cellular mechanisms of internalization and intracellular transport. *Toxins (Basel).* **2021**, *13*, 377.

[11] Johannes, L. Shiga toxin – a model for glycolipid-dependent and lectin-driven endocytosis. *Toxins (Basel).* **2017**, *9*, 340.

[12] Tartour, E.; Johannes, L. STxB as an Antigen Delivery Tool for Mucosal Vaccination. *Toxins (Basel).* **2022**, *14*, 202.

[13] Lipowsky, R. Spontaneous tubulation of membranes and vesicles reveals membrane tension generated by spontaneous curvature. *Faraday Discuss.* **2012**, *161*, 305–331.

[14] Ruhoff, V. T.; Moreno-Pescador, G.; Pezeshkian, W.; Bendix, P. M. Strength in numbers: effect of protein crowding on the shape of cell membranes. *Biochem. Soc. Trans.* **2022**, BST20210883.

[15] Bahrami, A. H.; Hummer, G. Formation and Stability of Lipid Membrane Nanotubes. *ACS Nano* **2017**, *11*, 9558–9565.

[16] Mahapatra, A.; Uysalel, C.; Rangamani, P. The Mechanics and Thermodynamics of Tubule Formation in Biological Membranes. *J. Membr. Biol.* **2021**, *254*, 273–291.

[17] Ewers, H. et al. GM1 structure determines SV40-induced membrane invagination and infection. *Nat. Cell Biol.* **2010**, *12*, 11–18; sup pp 1–12.

[18] Tourdot, R. W.; Ramakrishnan, N.; Baumgart, T.; Radhakrishnan, R. Application of a free-energy-landscape approach to study tension-dependent bilayer tubulation me-

diated by curvature-inducing proteins. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **2015**, *92*, 42715.

[19] Janssen, M.; Liese, S.; Carlson, A. Stability of a biomembrane tube covered with proteins. *bioRxiv* **2022**, 2022.09.29.510025.

[20] Mahapatra, A.; Rangamani, P. Formation of protein-mediated tubes is governed by a snapthrough transition. *arXiv* **2022**, 2206.3652.

[21] Sadeghi, M.; Weikl, T. R.; Noé, F. Particle-based membrane model for mesoscopic simulation of cellular dynamics. *J. Chem. Phys.* **2018**, *148*, 044901.

[22] Sadeghi, M.; Noé, F. Large-scale simulation of biomembranes incorporating realistic kinetics into coarse-grained models. *Nat. Commun.* **2020**, *11*, 2951.

[23] Sadeghi, M.; Noé, F. Hydrodynamic coupling for particle-based solvent-free membrane models. *J. Chem. Phys.* **2021**, *155*, 114108.

[24] Sadeghi, M.; Noé, F. Thermodynamics and kinetics of aggregation of flexible peripheral membrane proteins. *J. Phys. Chem. Lett.* **2021**, *12*, 10497–10504.

[25] Frenkel, D.; Smit, B. *Underst. Mol. Simul.*, 2nd ed.; Academic Press, 2002; p 664.

[26] Nilmeier, J. P.; Crooks, G. E.; Minh, D. D.; Chodera, J. D. Nonequilibrium candidate Monte Carlo is an efficient tool for equilibrium simulation. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, E1009–18.

[27] Kardar, M. *Statistical physics of particles*; Cambridge University Press, 2007; p 320.

[28] Palacio-Rodriguez, K.; Pietrucci, F. Free Energy Landscapes, Diffusion Coefficients, and Kinetic Rates from Transition Paths. *J. Chem. Theory Comput.* **2022**, *18*, 4639–4648.

[29] Kramers, H. A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **1940**, *7*, 284–304.

[30] Langer, J. S. Statistical theory of the decay of metastable states. *Ann. Phys. (N. Y).* **1969**, *54*, 258–275.

[31] Berezhkovskii, A. M.; Szabo, A.; Greives, N.; Zhou, H. X. Multidimensional reaction rate theory with anisotropic diffusion. *J. Chem. Phys.* **2014**, *141*, 204106.

[32] Arrayás, M.; Kaufman, I. K.; Luchinsky, D. G.; Mc Clintock, P. V.; Soskin, S. M. Kramers problem for a multiwell potential. *Phys. Rev. Lett.* **2000**, *84*, 2556–2559.

[33] Schütte, C.; Fischer, A.; Huisinga, W.; Deuflhard, P. A Direct Approach to Conformational Dynamics Based on Hybrid Monte Carlo. *J. Comput. Phys.* **1999**, *151*, 146–168.

[34] Prinz, J. H.; Wu, H.; Sarich, M.; Keller, B.; Senne, M.; Held, M.; Chodera, J. D.; Schtte, C.; Noé, F. Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.* **2011**, *134*, 174105.

[35] Roux, B. Transition rate theory, spectral analysis, and reactive paths. *J. Chem. Phys.* **2022**, *156*, 134111.

[36] Sadeghi, M. Investigating the entropic nature of membrane-mediated interactions driving the aggregation of peripheral proteins. *Soft Matter* **2022**, *18*, 3917 – 3927.

[37] Tian, A.; Baumgart, T. Sorting of lipids and proteins in membrane curvature gradients. *Biophys. J.* **2009**, *96*, 2676–2688.

[38] Watkins, E. B.; Majewski, J.; Chi, E. Y.; Gao, H.; Florent, J. C.; Johannes, L. Shiga Toxin Induces Lipid Compression: A Mechanism for Generating Membrane Curvature. *Nano Lett.* **2019**, *19*, 7365–7369.

[39] Pezeshkian, W.; Hansen, A. G.; Johannes, L.; Khandelia, H.; Shillcock, J. C.; Kumar, P. B. S.; Ipsen, J. H. Membrane invagination induced by Shiga toxin B-subunit: from molecular structure to tube formation. *Soft Matter* **2016**, *12*, 5164–5171.

[40] Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **1995**, *103*, 4613–4621.

[41] Bernetti, M.; Bussi, G. Pressure control using stochastic cell rescaling. *J. Chem. Phys.* **2020**, *153*, 114107.

[42] Visscher, K.; Schnltzer, M. J.; Block, S. M. Single kinesin molecules studied with a molecular force clamp. *Nature* **1999**, *400*, 184–189.

[43] Footer, M. J.; Kerssemakers, J. W.; Theriot, J. A.; Dogterom, M. Direct measurement of force generation by actin filament polymerization using an optical trap. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 2181–2186.

[44] Helfrich, W. Elastic Properties of Lipid Bilayers: Theory and Possible Experiments. *Zeitschrift fur Naturforsch. - Sect. C J. Biosci.* **1973**, *28*, 693–703.

[45] Guckenberger, A.; Gekle, S. Theory and algorithms to compute Helfrich bending forces: A review. *J. Phys. Condens. Matter* **2017**, *29*.

[46] Hu, M.; Briguglio, J. J.; Deserno, M. Determining the Gaussian curvature modulus of lipid membranes in simulations. *Biophys. J.* **2012**, *102*, 1403–1410.

[47] Marsh, D. Elastic curvature constants of lipid monolayers and bilayers. *Chem. Phys. Lipids* **2006**, *144*, 146–159.

[48] Sadeghi, M. Investigating the entropic nature of membrane-mediated interactions driving the aggregation of peripheral proteins. *bioRxiv* **2022**, 2022.01.24.477571.

[49] Weikl, T. R. Membrane-Mediated Cooperativity of Proteins. *Annu. Rev. Phys. Chem.* **2018**, *69*, 521–539.

[50] Groza, R.; Ewers, H. Membrane deformation by the cholera toxin beta subunit requires more than one binding site. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 17467–17469.

[51] Tourdot, R. W.; Ramakrishnan, N.; Radhakrishnan, R. Defining the free-energy landscape of curvature-inducing proteins on membrane bilayers. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **2014**, *90*, 22717.

[52] Bonazzi, F.; Weikl, T. R. Membrane Morphologies Induced by Arc-Shaped Scaffolds Are Determined by Arc Angle and Coverage. *Biophys. J.* **2019**, *116*, 1239–1247.

[53] Sandvig, K.; Kavaliauskiene, S.; Skotland, T. Clathrin-independent endocytosis: an increasing degree of complexity. *Histochem. Cell Biol.* **2018**, *150*, 107–118.

[54] Yedida, G.; Milani, M.; Cohen, G. M.; Varadarajan, S. Apogossypol-mediated reorganisation of the endoplasmic reticulum antagonises mitochondrial fission and apoptosis. *Cell Death Dis.* **2019**, *10*, 1234567890.

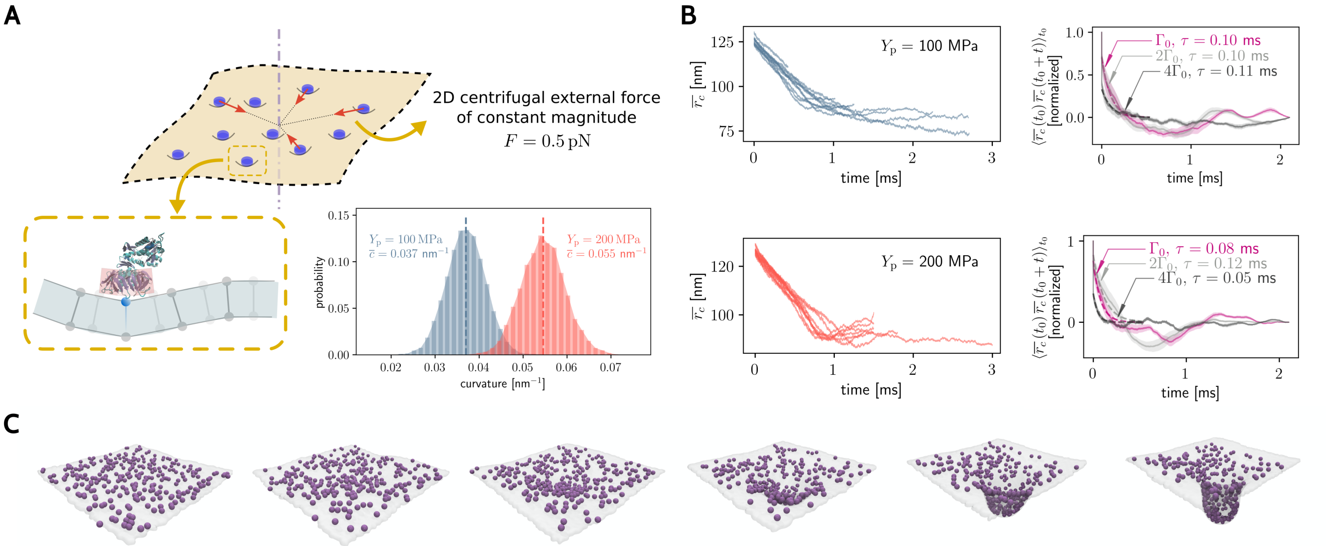

Figure 1: Steered simulations for capturing membrane tubulation induced by membrane-bending peripheral proteins. **(A)** (top) schematic of the simulation setup with peripheral proteins scattered on the membrane and pulled toward the center of the patch under the influence of in-plane centrifugal external forces. (bottom left) schematic of a curvature-inducing peripheral protein on the particle-based membrane model. (bottom right) distribution of the local membrane curvatures induced by proteins with the given stiffness values, Y_p. **(B)** (top and bottom left) mean in-plane radial distance of proteins from the center point of the membrane patch, $\bar{r}_c$, as a function of time in steered simulations. Results are given for different realizations of the simulation. (top and bottom right): time-correlation of the mean in-plane radial distance from unbiased simulations without external forces. Top and bottom plots correspond to the proteins with the stiffness values given in the plots on the left. Data is shown for simulations of the same surface concentration as the steered simulation (Γ_0) in addition to higher surface concentrations of $2\Gamma_0$ and $4\Gamma_0$. Time scales obtained from an exponential decay model are given in each case. **(C)** Consecutive snapshots of one realization of the steered simulations (see supplementary movies).

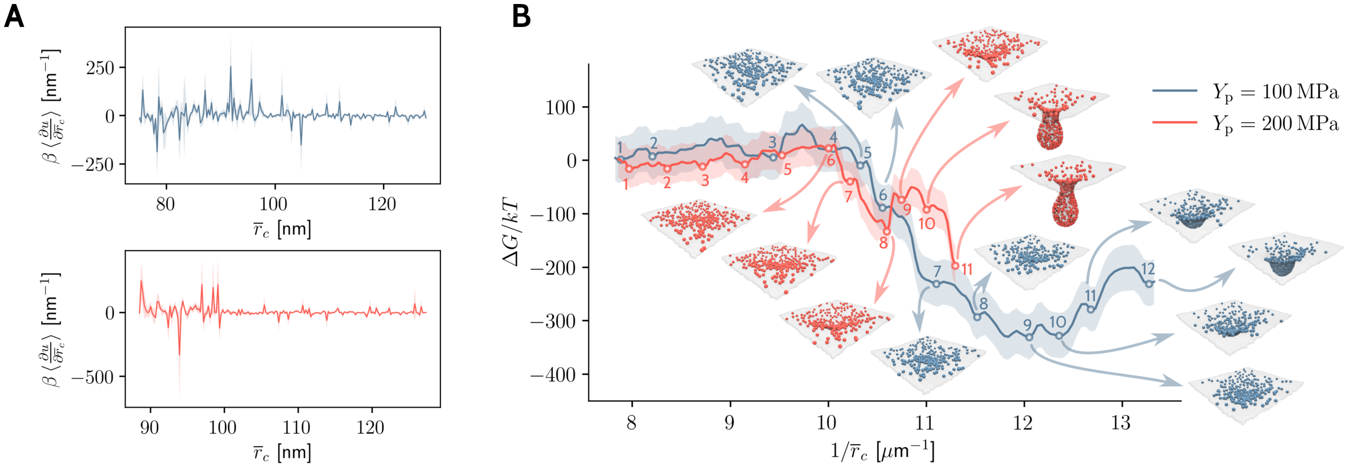

Figure 2: Free energy landscape of the formation of membrane invaginations by peripheral proteins. **(A)** ensemble averages of the derivative of the generalized potential, u, with respect to the reaction coordinate, $\bar{r}_c$, found using eq. 4. Blue and red color plots respectively corresponds to proteins with the stiffness of $Y_\mathrm{p} = 100\,\mathrm{MPa}$ and $200\,\mathrm{MPa}$. **(B)** Free energy landscapes along the reaction coordinate for invaginations caused by proteins with the given stiffness values. Free energy values have been obtained via thermodynamic integration using eq. 2. On each curve, selected states have been numbered for further kinetics analysis. Representative snapshots from one realization of steered simulations are shown for some of the states.

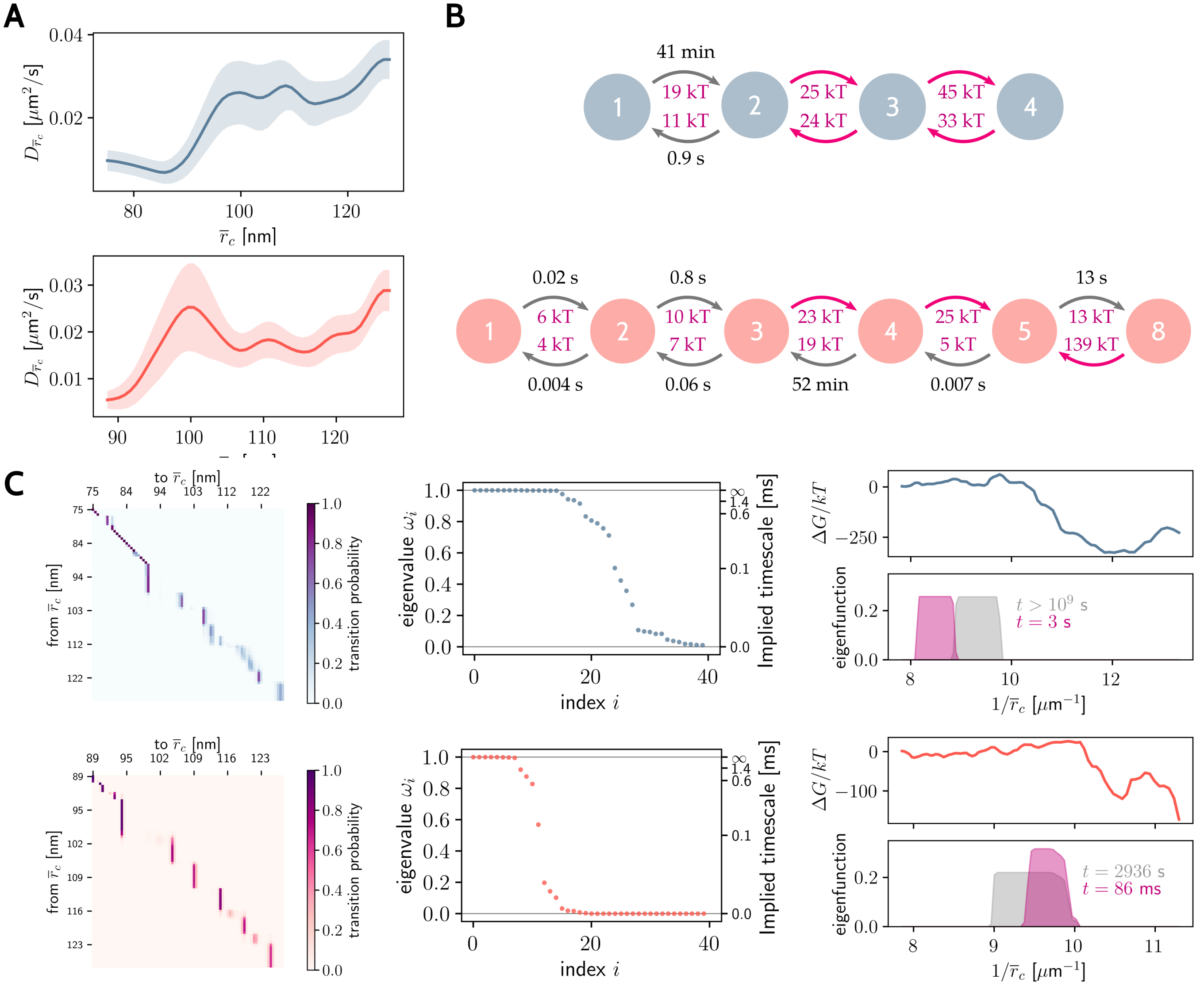

Figure 3: Kinetics of the spontaneous formation of membrane invaginations by the peripheral proteins. **(A)** diffusion coefficient governing the random walk in the space of the reaction coordinate, $\overline{r}_c$. Similar to previous figures, blue and red color plots respectively corresponds to the $Y_\mathrm{p} = 100\,\mathrm{MPa}$ and $200\,\mathrm{MPa}$ proteins. **(B)** Barrier energies and timescales of transition between the designated metastable states, based on eq. 5. Numbers are the same as those given on the free energy plots. **(C)** (left) discretized transfer operator for over-damped diffusion of $\overline{r}_c$ on the free energy landscape. The transfer operator has been discretized by partitioning the range of $\overline{r}_c$ into 100 bins. (middle) sorted eigenvalues of the discretized transfer operator and the resulting implied timescales. (right) selected eigenvectors of the discretized transfer operator and the corresponding timescales. Plots showing free energy profiles have been reproduced for reference.