

## Supplemental Information for: Metastable prepores in tension-free lipid bilayers

Christina L. Ting,<sup>1</sup> Neha Awasthi,<sup>2,3</sup> Marcus Müller,<sup>4\*</sup> and Jochen S. Hub<sup>2,3\*</sup>

<sup>1</sup>Sandia National Laboratories, Albuquerque, NM 87185, USA

<sup>2</sup>University of Goettingen, Institute for Microbiology and Genetics, 37077 Göttingen, Germany

<sup>3</sup>Göttingen Center for Molecular Biosciences, 37077 Göttingen, Germany

<sup>4</sup>University of Goettingen, Institute for Theoretical Physics, 37077 Göttingen, Germany

### I. ADDITIONAL FIGURES DISCUSSED IN THE MAIN TEXT

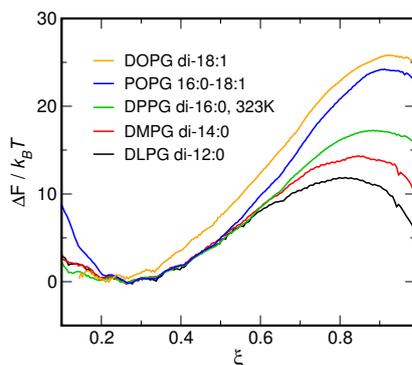


Figure S1: PMFs of prepore formation from MD simulations in tensionless phosphatidylglycerol (PG) membranes with increasing tail length and tail unsaturation. The numbers in the legend indicate the structure of the two tails in format number of carbon atoms:number of double bonds. Rationalized by the larger PG head groups as compared to PC head groups, the PMFs suggest increased tendencies of forming a metastable prepore, as evident from the pronounced nucleation barriers for DLPG and DMPG, and from the shallow barrier for DPPG, POPG, and DOPG (compare Fig. 1A).

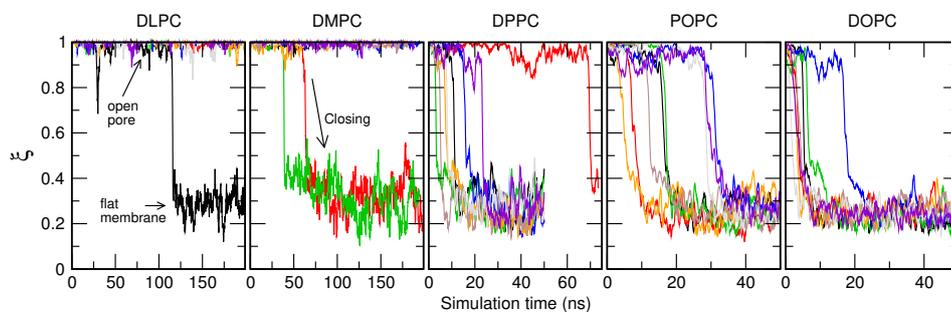


Figure S2: Trajectories of independent free MD simulations starting from randomly picked frames with an open pore (restrained to  $\xi = 1$ ). No bias was applied to the simulations, and all parameters were chosen as described in Sec. III A. In such free atomistic simulations, the simulation time plotted on the abscissa corresponds to real physical time. To visualize the spontaneous closing of the pores during the free simulations, the trajectories are projected onto the reaction coordinate  $\xi$ .  $\xi \approx 1$  and  $\xi \approx 0.3$  correspond to open pores and flat unperturbed membranes, respectively (arrows in left panel). Eight independent simulations were conducted for each lipid type: (from left to right) DLPC, DMPC, DPPC, POPC, and DOPC. Curves with different colors indicate independent trajectories for improved statistics. All pores in bilayers of DPPC, POPC, and DOPC close within 75 ns, compatible with the absence of a nucleation barrier in these membranes (Fig. 1). For bilayers of DLPC and DMPC, in contrast, the pore closes in only 3 out of 16 200-nanosecond simulations, confirming the metastability of the prepores, and compatible with the nucleation barriers revealed by the PMFs (Fig. 1).

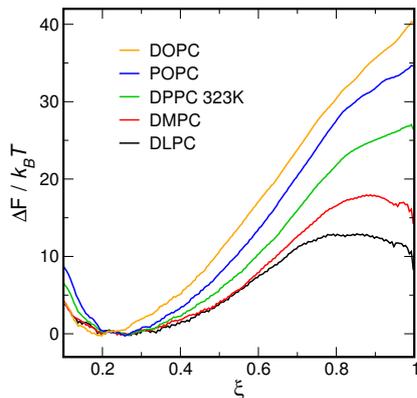


Figure S3: PMFs of prepre formation from MD simulations in tensionless phosphatidylcholine membranes, computed with the Slipids [S1] instead of the Charmm36 lipid force field [S2]. The PMFs reasonably agree between Slipids and Charmm36 force fields.

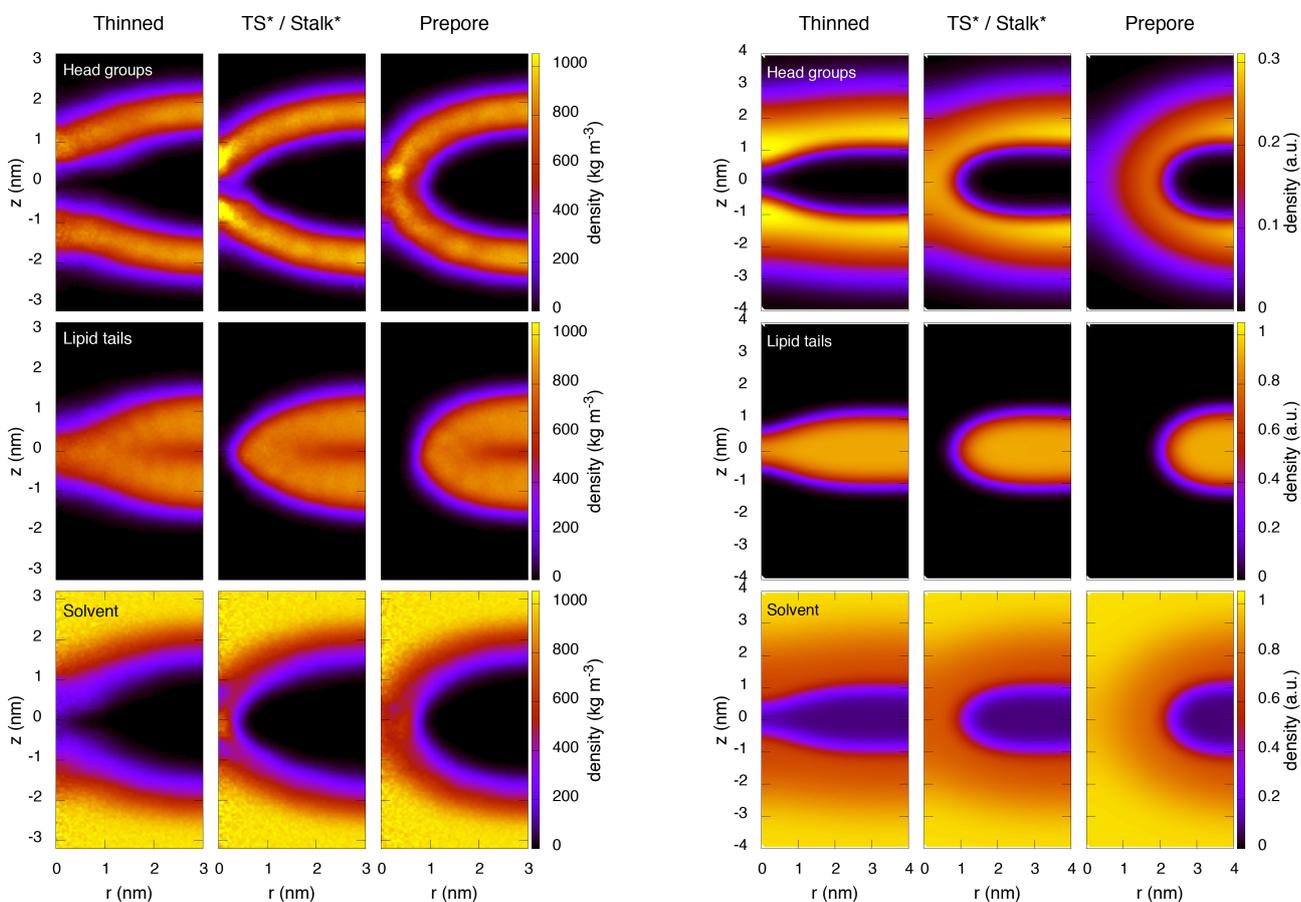


Figure S4: Density of head groups (top row), tails (middle row), and water (bottom row) from atomistic MD simulations. Left column: thinned membrane; middle column: transition state; right column: prepore. The densities are computed from the umbrella windows indicated by red dots in Fig. 1A.

Figure S5: Density of head groups (top row), tails (middle row), and solvent (bottom row) from SCFT of the coarse-grained (CG) model. Left column: thinned membrane; middle column: transition state; right column: prepore. The three states are highlighted along the MFEP as grey dots in Fig. 1B.

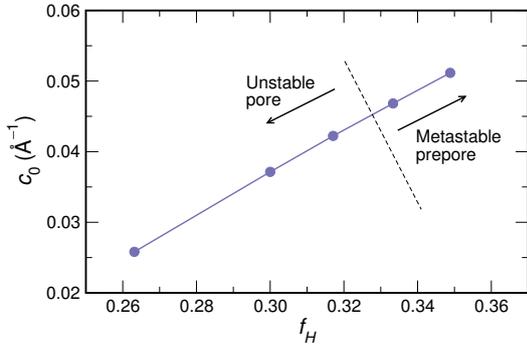


Figure S6: Spontaneous curvature,  $c_0$ , of coarse-grained membranes versus head-to-tail volume ratio,  $f_H$ .  $c_0$  increases linearly with  $f_H$ , which supports the intuition that as the head-group volume fraction increases, the effective shape of the lipid becomes more conical and the lipid prefers structures with higher curvature. For atomistic PC lipids, no such linear relation holds.

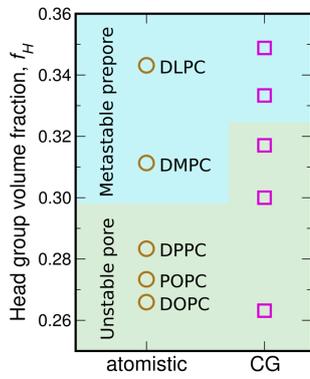


Figure S7: Head-group volume fraction,  $f_H$ , of atomistic (tan circles) and coarse-grained (CG, magenta squares) lipids.  $f_H$  of atomistic models were computed using Voronoi tessellation. The blue and green backgrounds highlight lipids that form metastable and unstable pores, respectively.

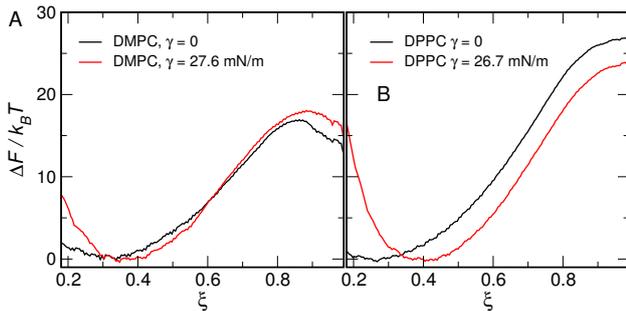


Figure S8: PMFs of prepore formation from atomistic MD simulations for (A) DMPC and (B) DPPC with and without tension,  $\gamma$ , using the Charmm36 lipid force field [S2].  $\gamma$  may influence the PMF by thinning of the membrane, as indicated by a shift of the PMF minimum (flat membrane) to larger  $\xi$ . However,  $\gamma$  does not have a strong influence on the stability of the prepore, i.e., on a (possible) PMF minimum at  $\xi \approx 1$ .

## II. SELF-CONSISTENT FIELD THEORY (SCFT) OF COARSE-GRAINED (CG) LIPID MEMBRANES

### A. Minimal coarse-grained (CG) model

Our coarse-grained (CG) model consists of a membrane bilayer assembled from double-tailed lipids,  $L$ , in an explicit solvent,  $S$ . The solvent molecules are described as monomers, and the lipids are represented as graft copolymers, which we model as discrete Gaussian chains with  $N_H$  head monomers and  $2N_T$  tail monomers (cf. Fig. S9). In using the discrete Gaussian chain as our model, we capture the finite-size bond length, but ignore the bending rigidity of real lipids. However, bond-order parameters calculated from lattice models [S3–S5] in the same spirit as our discrete Gaussian chain model have shown good qualitative agreement with experimental findings [S6, S7] and even excellent agreement with molecular dynamics simulations [S8–S10]. Similar models have been used to study other topology-altering membrane processes [S11, S12].

The particle-based Hamiltonian for this model, accounting for bonded and non-bonded interactions, is defined by

$$\frac{\mathcal{H}'}{k_B T} = \sum_{i=1}^{n_L} h_i(\mathbf{r}^{N_L}) + \frac{\rho_0}{2} \int d\mathbf{r} d\mathbf{r}' u_{\alpha\beta}(\mathbf{r}, \mathbf{r}') \hat{\phi}_\alpha(\mathbf{r}) \hat{\phi}_\beta(\mathbf{r}'). \quad (\text{S1})$$

In this expression, the first term represents the bonded interactions of the  $n_L$  lipids, and takes the form

$$h_i(\mathbf{r}^{N_L}) = \frac{3k_B T}{2b_0^2} \sum_{j=1}^{N_L-1} (\mathbf{r}_{j+1} - \mathbf{r}_j)^2, \quad (\text{S2})$$

where  $N_L = N_H + 2N_T$  is the number of monomers,  $k_B T$  is the thermal energy,  $b_0$  is the statistical bond length, and  $\mathbf{r}_j$  is the position of monomer  $j$ . The second term in the Hamiltonian represents the pair-wise interaction potentials between monomer species, where  $\rho_0$  is the reference density and the summation is over pairs of Greek indices:  $\alpha, \beta = H, T, S$ . The microscopic volume fraction of monomer species is defined by

$$\hat{\phi}_S(\mathbf{r}) = \frac{1}{\rho_0} \sum_{i=1}^{n_S} \delta(\mathbf{r} - \mathbf{r}_i), \quad (\text{S3})$$

$$\hat{\phi}_H(\mathbf{r}) = \frac{1}{\rho_0} \sum_{i=1}^{n_L} \sum_{j=1}^{N_H} \delta(\mathbf{r} - \mathbf{r}_{ij}), \quad (\text{S4})$$

$$\hat{\phi}_T(\mathbf{r}) = \frac{2}{\rho_0} \sum_{i=1}^{n_L} \sum_{j=1}^{N_T} \delta(\mathbf{r} - \mathbf{r}_{ij}), \quad (\text{S5})$$

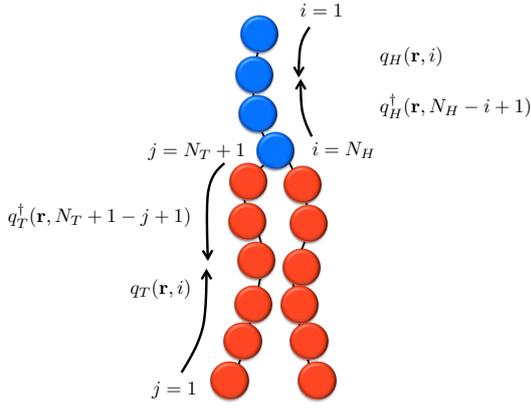


Figure S9: Cartoon of the double-tailed lipid model consisting of head (blue) and tail (red) monomers. The chain propagators  $q_H$  and  $q_T$  are used to obtain the single-molecule partition function in Eq. (S16), whereas the complementary chain propagators  $q_H^\dagger$  and  $q_T^\dagger$  are additionally required for computing the volume fractions in Eqs. (S17) and (S18).

for the solvent, head, and tail monomers, respectively. Importantly, a strict incompressibility condition at all positions within the system volume is also accounted for by requiring the solvent volume fraction obeys the constraint,  $\hat{\phi}_S(\mathbf{r}) = 1 - \hat{\phi}_H(\mathbf{r}) - \hat{\phi}_T(\mathbf{r})$ .

## B. Self-consistent field theory (SCFT)

We work in the grand-canonical ensemble, where the numbers of lipids and solvent molecules are controlled by their chemical potentials  $\mu_L$  and  $\mu_S$ . Due to the incompressibility constraint, only the difference  $\mu_L - \mu_S$  is relevant. Note that in the grand-canonical ensemble, the excess free energy (relative to the homogeneous solution) per unit area directly gives the membrane tension,  $\gamma$ , where the chemical potential difference,  $\mu_L - \mu_S$ , is used to control the excess free energy.

In self-consistent field theory (SCFT), the interactions among molecules are replaced with the interactions between a single molecule and an effective external field [S13]. As a result, the field-theoretic partition function may be written as a functional integral over fluctuating fields

$$\Omega[W_\alpha, \Phi_\alpha] = \int \mathcal{D}W_\alpha \mathcal{D}\Phi_\alpha \exp\left(-\frac{\mathcal{G}[W_\alpha, \Phi_\alpha]}{k_B T}\right), \quad (\text{S6})$$

where  $\Phi_\alpha$  and  $W_\alpha$  now denote the spatially varying volume fraction and conjugate potential field variables, respectively. The term in the exponent of Eq. (S6) is the

field-theoretic free energy, defined by

$$\begin{aligned} \frac{\mathcal{G}}{k_B T} &= e^{\frac{\mu_L}{k_B T}} \mathcal{Q}_L[W_H, W_T] - e^{\frac{\mu_S}{k_B T}} \mathcal{Q}_S[W_S] \\ &+ \rho_0 \int d\mathbf{r} \left[ \frac{\chi_{\alpha\beta}}{2} \Phi_\alpha \Phi_\beta + \frac{\kappa_\alpha}{2} (\nabla \Phi_\alpha)^2 - W_\alpha \Phi_\alpha \right], \end{aligned} \quad (\text{S7})$$

where  $\mathcal{Q}_L[W_H, W_T]$  and  $\mathcal{Q}_S[W_S]$  are the single-molecule partition functions of the lipids and solvent molecules, respectively. The second line in Eq. (S7) contains the Flory-Huggins  $\chi_{\alpha\beta}$  and square-gradient  $\kappa_\alpha$  terms, which capture the local and nonlocal interactions [S14], as well as a term that couples the volume fraction  $\Phi_\alpha$  to its conjugate potential field  $W_\alpha$ . Again, the incompressibility condition is accounted for by constraining  $\Phi_S = 1 - \Phi_H - \Phi_T$ .

Extremizing Eq. (S7) we obtain a set of SCFT equations that determine the saddle-point values for the field variables:

$$\left. \frac{\delta \mathcal{G}}{\delta \Phi_\alpha} \right|_{\Phi_\alpha = \phi_\alpha} = 0; \quad \left. \frac{\delta \mathcal{G}}{\delta W_\alpha} \right|_{W_\alpha = w_\alpha} = 0. \quad (\text{S8})$$

Variation with respect to  $\Phi_H$  and  $\Phi_T$  yield the conjugate potential fields

$$\begin{aligned} w_H(\mathbf{r}) &= w_S(\mathbf{r}) + \chi_{SH} [\phi_S(\mathbf{r}) - \phi_H(\mathbf{r})] \\ &+ [\chi_{HT} - \chi_{TS}] \phi_T(\mathbf{r}) - \kappa_H \Delta \phi_H(\mathbf{r}), \quad (\text{S9}) \\ w_T(\mathbf{r}) &= w_S(\mathbf{r}) + \chi_{TS} [\phi_S(\mathbf{r}) - \phi_T(\mathbf{r})] \\ &+ [\chi_{HT} - \chi_{SH}] \phi_H(\mathbf{r}) - \kappa_T \Delta \phi_T(\mathbf{r}); \quad (\text{S10}) \end{aligned}$$

and variation with respect to  $W_\alpha$  yields the volume fraction fields

$$\phi_H(\mathbf{r}) = -\frac{e^{\frac{\mu_L}{k_B T}}}{\rho_0} \frac{\delta \mathcal{Q}_L}{\delta w_H(\mathbf{r})}, \quad (\text{S11})$$

$$\phi_T(\mathbf{r}) = -\frac{e^{\frac{\mu_L}{k_B T}}}{\rho_0} \frac{\delta \mathcal{Q}_L}{\delta w_T(\mathbf{r})}, \quad (\text{S12})$$

$$\phi_S(\mathbf{r}) = -\frac{e^{\frac{\mu_S}{k_B T}}}{\rho_0} \frac{\delta \mathcal{Q}_S}{\delta w_S(\mathbf{r})} = \frac{e^{\frac{\mu_S}{k_B T}}}{\rho_0} \mathcal{Z}_0 e^{-w_S(\mathbf{r})}, \quad (\text{S13})$$

where the spatially-varying volume fraction,  $\phi_\alpha(\mathbf{r})$ , and conjugate field,  $w_\alpha(\mathbf{r})$ , variables now correspond to their saddle-point values and are denoted by lower-case letters.

The solvent volume fraction,  $\phi_S(\mathbf{r})$ , in Eq. (S13) is obtained by noting that the single-molecule partition function for the solvents is defined by  $\mathcal{Q}_S[w_S] = \mathcal{Z}_0 \int d\mathbf{r} e^{-w_S(\mathbf{r})}$ , where  $\mathcal{Z}_0 V$  is the partition function in the absence of an external field. To compute the volume fraction of the head and tail monomers,  $\phi_H(\mathbf{r})$  and  $\phi_T(\mathbf{r})$ , in Eqs. (S11) and (S12), we note that the single-molecule partition function for the lipids,  $\mathcal{Q}_L[w_H, w_T]$ , must additionally take into account the chain connectivity. Here, we introduce the chain propagator  $q(\mathbf{r}, i)$ , which gives the statistical weight for a chain to have monomer  $i$  at position  $\mathbf{r}$ . The chain propagator for the discrete Gaussian chain begins with the initial condition

$q(\mathbf{r}, 1) = e^{-w(\mathbf{r})/k_B T}$  and is built up recursively from the ends of the chain according to

$$q(\mathbf{r}, i) = e^{-\frac{w(\mathbf{r})}{k_B T}} \int d\mathbf{r}' p(\mathbf{r} - \mathbf{r}') q(\mathbf{r}', i - 1), \quad (\text{S14})$$

where  $p(\mathbf{b})$  is the normalized bond transition probability, and is defined by

$$p(\mathbf{b}) = \left( \frac{3}{2\pi b_0^2} \right)^{d/2} \exp\left(-\frac{3\mathbf{b}^2}{2b_0^2}\right). \quad (\text{S15})$$

In this expression,  $d$  is the dimensionality and the statistical segment length is given by  $b_0^2 \equiv R_{e0}^2/(N - 1)$ . Standard methods for numerically solving for the chain propagators are described in detail in Ref. [S13]. We use the Scheutjens and Fler lattice SCFT method [S15], where additional substeps between physical monomers are necessary for numerical stability and accuracy of the Scheutjens-Fler propagator approximation to the Gaussian propagator. These substeps also satisfy Eq. (S14), but without the Boltzmann weight due to the external field.

Thus, the total partition function for the lipid is obtained by joining the chain propagators  $q_H(\mathbf{r}, N_H)$  and  $q_T^2(\mathbf{r}, N_T + 1)$  for the head and two tail branches, respectively, at the branch point of the lipid:

$$\mathcal{Q}_L[w_A, w_B] = \mathcal{Z}_0 \int d\mathbf{r} q_H(\mathbf{r}, N_H) e^{2w_H(\mathbf{r})} q_T^2(\mathbf{r}, N_T + 1), \quad (\text{S16})$$

where the extra exponential factor  $e^{2w_H(\mathbf{r})}$  corrects for over-counting the joined monomer, which we define to be a head monomer; see Fig. S9 for a schematic of the chain propagators for the discrete Gaussian chain model of the lipid.

Finally, the spatially-varying head and tail volume fractions in Eqs. (S11) and (S12) can be expressed in terms of the chain propagators:

$$\begin{aligned} \phi_H &= \frac{e^{\frac{\mu_L}{k_B T}}}{\rho_0} \mathcal{Z}_0 \sum_{i=1}^{N_H} q_H(\mathbf{r}, i) e^{\frac{w_H(\mathbf{r})}{k_B T}} q_H^\dagger(\mathbf{r}, N_H - i + 1), \quad (\text{S17}) \\ \phi_T &= 2 \frac{e^{\frac{\mu_L}{k_B T}}}{\rho_0} \mathcal{Z}_0 \sum_{i=1}^{N_T} q_T(\mathbf{r}, i) e^{\frac{w_T(\mathbf{r})}{k_B T}} q_T^\dagger(\mathbf{r}, N_T + 1 - i + 1), \quad (\text{S18}) \end{aligned}$$

where the extra exponential factor again corrects for over-counting the monomer when the two propagators for the discrete Gaussian chain are joined. Note that we have introduced the complementary chain propagators  $q_H^\dagger(\mathbf{r}, i)$  and  $q_T^\dagger(\mathbf{r}, i)$ , which also satisfy Eq. (S14) but start from the branch point with the initial conditions:

$$q_H^\dagger(\mathbf{r}, 1) = e^{w_H(\mathbf{r})} q_T^2(\mathbf{r}, N_T + 1), \quad (\text{S19})$$

$$q_T^\dagger(\mathbf{r}, 1) = e^{w_H(\mathbf{r})} q_T(\mathbf{r}, N_T + 1) q_H(\mathbf{r}, N_H). \quad (\text{S20})$$

Finally, we note that Eq. (S13) may be trivially solved to obtain

$$w_S(\mathbf{r}) = -\log \left[ \frac{\rho_0}{\mathcal{Z}_0} (1 - \phi_H(\mathbf{r}) - \phi_T(\mathbf{r})) \right], \quad (\text{S21})$$

where the incompressibility constraint,  $\phi_H + \phi_T + \phi_S = 1$ , has been used to eliminate  $\phi_S(\mathbf{r})$ , and we have defined  $\mu_S = 0$  because the chemical potentials of the solvents and lipids are not independent for an incompressible system. Eqs. (S9), (S10), and (S21) for the conjugate potential fields, together with Eqs. (S17) and (S18) for the volume fraction fields are solved iteratively, until convergence, subject to the incompressibility constraint.

### C. String method to obtain the minimum free-energy path (MFEP)

Whereas the saddle-point solutions to the SCFT equations correspond to metastable or equilibrium states of a bilayer membrane, our particular interest is in the transition paths connecting an equilibrium, defect-free membrane to one containing a macroscopic pore, through an intermediate prepore state. This transition path is represented by a string, i.e., a sequence of morphologies, on the free-energy landscape. Given an initial string between two states, the string method [S16] locally evolves the path to a minimum free energy path (MFEP), corresponding to the most probable transition path, where the gradient of the free energy perpendicular to the path vanishes. The string method has successfully been combined with field-theoretic models to compute transition pathways in a variety of self-assembling systems [S17–S19].

To initialize our string, we assign our starting morphology to a defect-free planar bilayer, and our ending morphology to a planar bilayer with a pore whose radius is constrained to be larger than the size of the transition state. In general, there may be multiple MFEPs connecting any given starting and ending morphologies, and the MFEP that the string converges to will depend on the initial path. Various paths can be explored by specifying intermediate morphologies obtained, for example, from particle-based simulation. Here, we chose a simple linear interpolation between the two end morphologies. The MFEP is then calculated according to the following iterative, two-step procedure [S16]: First, each morphology along the string is evolved independently according to its steepest-descent dynamics. We use a version of the external potential dynamics (EPD) [S20, S21] that reformulates the dynamics of the collective densities in terms of the conjugate potential fields:

$$\frac{dW_\alpha(\mathbf{r})}{dt} = D \frac{\delta \mathcal{G}}{\delta W_\alpha(\mathbf{r})}, \quad (\text{S22})$$

where  $D$  is a scalar mobility coefficient. Thus the same numerical methods from EPD can be used to update the potential fields  $W_\alpha(\mathbf{r})$ , which are then used directly to update the volume fraction fields  $\phi_\alpha(\mathbf{r})$ . The updated  $\phi_\alpha(\mathbf{r})$  represent the average density of independent molecules in the presence of the new  $W_\alpha(\mathbf{r})$ . Second, a restoring step redistributes the morphologies in a direction tangent to the string. This is enforced by reparameterizing the

states to be equidistant along the string according to the collective volume fraction fields  $\phi_\alpha(\mathbf{r})$  of the monomeric species. The dynamics, followed by the reparameterization, are computed at every iteration step. Once converged, the string coincides with the MFEP. The contour parameter along the string is the optimal reaction coordinate.

### III. ATOMISTIC MD SIMULATION

#### A. Setup and parameters

Five membrane systems of 128 DLPC, DMPC, DPPC, POPC, DOPC, DLPG, DMPG, DPPG, POPG, or DOPG lipids and 40 water molecules per lipid were built with the MemGen web server [S22]. Systems of PG lipids were neutralized by adding 128 potassium ions. In previous work, we found that simulating 128 lipids is sufficient to avoid finite-size artifacts in simulations with a small pore [S23]. Each system was equilibrated until the box dimensions and the potential energy were fully converged. Parameters from CHARMM36 force field [S2] were used for the lipid molecules and water was modeled with the CHARMM-modified TIP3P model (with Lennard-Jones (LJ) interactions on hydrogen atoms) [S2]. All simulations were performed with the GROMACS 2016.2 software. Bonds and angles of water were constrained with the SETTLE algorithm [S24], and bonds of lipids between heavy atoms and hydrogen atoms were constrained with LINCS [S25]. The temperature of the simulations was controlled at 300 K for all lipids except DPPC. Since DPPC and DPPG form a gel phase at 300 K, the temperature of the DPPC and DPPG simulations was controlled at 323 K. The temperature was controlled with a stochastic dynamics integrator [S26] during umbrella sampling, and using velocity-rescaling during other simulations (with separate heat baths for solvent and lipid) [S27]. An integration time step of 2 fs was applied. The pressure was controlled at 1 bar using a semi-isotropic weak coupling scheme ( $\tau = 1$  ps) [S28]. Electrostatic interactions were calculated using the particle-mesh Ewald (PME) method [S29, S30]. Dispersion interactions and short-range repulsion were described by a switched Lennard-Jones (LJ) potential with cut-offs at 1 nm and 1.2 nm, respectively. Direct-space Coulomb interactions were truncated at 1.2 nm. Unbiased simulations were carried out with the Gromacs simulation software, version 2016 [S31]. Simulations with restraints along the reaction coordinate,  $\xi$ , were done with an in-house modification of Gromacs 2016 [S23], which is available from the authors upon request.

To exclude the possibility that the applied force field could qualitatively change the results, simulations using the Slipids [S1] instead of the CHARMM36 force field were also conducted. Here, all bonds were constrained with LINCS [S25], and a plain cut-off at 1.2 nm was applied for LJ interactions. Water was described with the

normal TIP3P model [S32]. All other parameters were identical to the CHARMM36 simulations. The PMFs computed with Slipids were found to be similar to the PMF computed with CHARMM36 (cf. Figs. 1A and S3).

#### B. Reaction coordinate for atomistic description

Umbrella sampling simulations were used to calculate PMFs for pore nucleation in the atomistic membrane models. For umbrella sampling, we used a reaction coordinate,  $\xi$ , that is capable of probing the formation or rupture of a continuous polar defect in a lipid membrane [S23].  $\xi$  is defined using a membrane-spanning cylinder that is decomposed into  $N_s$  slices taken symmetrically around the center of mass of the hydrophobic membrane atoms along the membrane normal:

$$\xi = N_s^{-1} \sum_{s=0}^{N_s-1} \delta_s(N_s^{(p)}). \quad (\text{S23})$$

Technical details about  $\xi$  and its performance were discussed previously [S23]. In brief,  $\xi$  defines the fraction of slices that are occupied by polar atoms, where  $N_s^{(p)}$  denotes the number of polar heavy atoms in slice  $s$ , and  $\delta_s$  is a continuous indicator function ( $0 \leq \delta_s < 1$ ) that equals zero if  $N_s^{(p)} = 0$  and a value close to unity if  $N_s^{(p)} \geq 1$ . Note that  $\delta_s$  remains smaller than unity, irrespective of how many polar atoms are located in slice  $s$ . Functions  $\delta_s$  and  $N_s^{(p)}$  are further defined as differentiable switch functions using Cartesian coordinates of atoms as arguments.

By design,  $\xi$  differentiates between (i) polar defects partly penetrating the membrane and (ii) a continuous defect spanning the entire membrane and, hence, it is quite suitable to model the transition region of pore formation. We have previously shown that a pulling system along  $\xi$  efficiently introduced a pore into the membrane and that restraints along  $\xi$  are capable of restraining the system close to the transition state of pore formation. Further, we showed that PMF calculations along  $\xi$  do not suffer from hysteresis effects [S23].

#### C. PMF calculations for pore nucleation

PMFs were computed along the reaction coordinate  $\xi$  defined in Eq. (S23) using the technique of umbrella sampling [S33]. Starting frames for umbrella sampling were taken from a “slow-growth” pore-opening simulation. Accordingly, starting in the equilibrium state, the pore was opened by pulling the system along  $\xi$  with a harmonic potential (force constant  $2000 \text{ kJ mol}^{-1}$ ). The minimum of the harmonic potential was moved with constant velocity from  $\xi = 0$  at time 0 to  $\xi = 1$  at 45 ns (20 ns for PG membranes). Umbrella sampling was performed using 24 umbrella windows. To ensure that the transition state of pore formation is well sampled, we used

Table I: Summary of parameters used for umbrella sampling simulations and used to define the reaction coordinate. Simulation time per umbrella window  $t$ , temperature  $T$ , and four parameters required to define the reaction coordinate: the parameter  $\zeta$ , cylinder radius  $R_{\text{cyl}}$ , slice thickness  $d_s$ , and number of slices  $N_s$ .

Lipid	$t$ (ns)	$T$ (K)	$\zeta$	$R_{\text{cyl}}$ (nm)	$d_s$ (nm)	$N_s$
DLPC	150	300	0.75	1.2	0.1	19
DMPC	150	300	0.75	1.2	0.1	23
DPPC	150	323	0.75	1.2	0.1	29
POPC	150	300	0.75	1.2	0.1	29
DOPC	150	300	0.75	1.2	0.1	30
DLPG	100	300	0.75	0.8	0.1	17
DMPG	100	300	0.75	0.8	0.1	19
DPPG	100	323	0.75	0.8	0.1	23
POPG	100	300	0.75	0.8	0.1	28
DOPG	100	300	0.75	0.8	0.1	26

a tighter spacing of higher umbrella force constants at  $\xi \geq 0.7$  as compared to  $\xi < 0.7$ . Accordingly, 11 windows were distributed between  $\xi = 0$  and  $\xi = 0.65$  in steps of 0.065, using a force constant of  $5000 \text{ kJ mol}^{-1}$ . An additional 13 windows were distributed between  $\xi = 0.7$  and  $\xi = 1$  in steps of 0.025, using a force constant of  $10000 \text{ kJ mol}^{-1}$ . Each umbrella window was simulated for 150 ns for PC, and for 100 ns for PG membranes (Table I). For PG membranes, we defined  $\xi$  using a narrower transmembrane cylinder ( $R_{\text{cyl}} = 0.8 \text{ nm}$ ) as compared to PC membranes ( $R_{\text{cyl}} = 1.2 \text{ nm}$ ). This choice is rationalized by more pronounced membrane undulations of the soft PG membranes, which could lead to undesirable hysteresis effects when pulling the system along  $\xi$ . The first 10 ns were omitted for equilibration, if not stated otherwise. Longer equilibration changed the PMFs only marginally. The PMFs were constructed from the umbrella histograms with the weighted histogram analysis method (WHAM), as implemented in the `g-wham` software [S34, S35]. The parameters required for defining the reaction coordinate are summarized in Table I. Here, the thickness  $d_s = 0.1 \text{ nm}$  of the slices was chosen such that pairs of polar atoms in neighboring slices are capable of forming stable hydrogen bonds, even in the presence of some disorder and fluctuations. The parameter  $\zeta = 0.75$  defines the fraction to which a slice is filled by the addition of the first polar atom.  $N_s$  was chosen such that  $\xi \approx 0.25$  for the flat unperturbed membrane; hence larger  $N_s$  were chosen for thicker membranes.

#### D. Head group volume fraction of atomistic lipids

The head group volume fraction  $f_H$  of CG lipids is simply given via the number of head group and tail beads. For atomistic models,  $f_H$  was computed using three-dimensional (3D) Voronoi tessellation, as implemented in the `Voro++` software (Fig. S7) [S36]. For each lipid type, the volume of an equilibrated, unperturbed mem-

brane system was decomposed into atomic volumes of water, head group, and tail atoms. Atoms of the choline, phosphate, glycerol, and ester groups were taken as “head group”, and the aliphatic chains as “tails”. Water atoms were not considered for  $f_H$ .

#### E. Free-energy calculations of pore growth

The growth of a fully formed polar defect was simulated as follows. A simulation system of 200 DMPC lipids plus 60 water molecules per lipid was set up with MemGen [S22] and subsequently fully equilibrated. A prepore was induced using a slow-growth pulling simulation along the reaction coordinate,  $\xi$ , within 40 ns and up to  $\xi = 1$ . Next, the cross-sectional area,  $A$ , of the box was increased in 200 steps. Within each step, the box dimensions and atomic coordinates were scaled in the  $x$ - and  $y$ -directions, corresponding to the membrane plane, by a factor of  $1.2^{1/200}$ . After the scaling, the energy was carefully minimized, and the system was equilibrated for 2250 ps. During equilibration, pressure coupling was allowed in the  $z$ -direction but the area,  $A$ , was fixed. In addition, the system was restrained to  $\xi = 1$  to ensure that the pore did not close. The equilibrated system was used as input for the next step of scaling of the  $x$ - $y$  coordinates. This procedure generated 200 frames with increasing membrane area and hence increasing pore radius.

30 frames of increasing area were picked for long simulations with constant area but without restraints along  $\xi$ . 17 systems with smaller area were simulated for 300 ns; 13 systems with larger area were simulated for 200 ns. After removing the first 5 ns of equilibration, the tension was computed from the anisotropy of the pressure tensor,

$$\Sigma(A) = \left\langle L_z \left[ P_{zz}(t) - \frac{P_{xx}(t) + P_{yy}(t)}{2} \right] \right\rangle_A. \quad (\text{S24})$$

Here,  $P$  is the pressure tensor, and  $\langle \cdot \rangle_A$  is the average over simulation time,  $t$ , at constant box area,  $A$ . The free energy was computed following

$$\Delta F(A) = F_0 + \int_{A_0}^A \Sigma(A') dA', \quad (\text{S25})$$

where the offset,  $F_0$ , was adjusted so that the PMF for pore growth (Fig. 3D, right) matched the PMF for pore nucleation near  $\xi \approx 1$  (Fig. 3D, left). The area,  $A$ , was translated to the increase in area relative to a flat unperturbed membrane,  $\Delta A = A - A_0$ , where we obtained  $A_0 = 62.60 \text{ nm}^2$  under the simulation conditions.

The radius of the pore was computed from the average water density within an 8 Å-layer parallel to the membrane and centered in the  $z$ -direction by the membrane center-of-mass. The water density was computed as a function of the radial distance from the axis of the pore. Here, the axis of the pore was taken as the cylinder axis

defined by our reaction coordinate, as described previously [S23]. The radius of the pore  $R_{\text{pore}}$  was defined as the radial distance from the pore axis where the water density drops below half the bulk water density. The pore radius versus  $\Delta A$  is shown in Fig. S10. Following Tolpekina *et al.*, the line tension of the pore rim was computed using  $\sigma(A) = R_{\text{pore}}(A)\Sigma(A)$  [S37]. Statistical errors of  $\sigma(A)$  were computed by binning analysis, as implemented in the Gromacs module “gmx analyze” [S38].

To reduce the computational cost, the simulations of

pore growth were conducted with lipid parameters taken from Berger *et al.* [S39], using a 4 fs time step, PME electrostatics [S29, S30], and a cut-off at 1 nm for direct-space interactions. Water parameters were taken from the SPC model [S40]. The pressure was controlled with the Parrinello-Rahman barostat ( $\tau = 5$  ps) [S28], using a compressibility of zero in the  $x$ - $y$  plane, thereby fixing the box area. All lipid bonds were constrained with LINCS [S25]. Other parameters were chosen as described above.

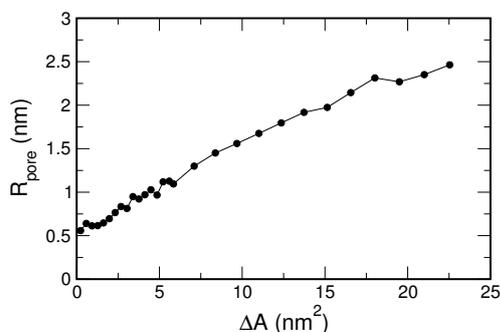


Figure S10: Pore radius versus increase in box area during simulations of pore growth, as used for computing the line tension,  $\sigma$ , of the pore edge (Fig. 3D, right).

- 
- [S1] J. P. M. Jämbeck and A. P. Lyubartsev, *J. Chem. Theory Comput.* **8**, 2938 (2012).
- [S2] R. Pastor and A. MacKerell Jr, *J. Phys. Chem. Lett.* **2**, 1526 (2011).
- [S3] A. Ben-Shaul, I. Szleifer, and W. Gelbart, *Proc. Natl. Acad. Sci. USA* **81**, 4601 (1984).
- [S4] A. Ben-Shaul and I. Szleifer, *J. Chem. Phys.* **83**, 3597 (1985).
- [S5] I. Szleifer, A. Ben-Shaul, and W. Gelbart, *J. Chem. Phys.* **83**, 3621 (1985).
- [S6] T. Zemb and C. Chachaty, *Chem. Phys. Lett.* **88**, 68 (1982).
- [S7] J. Charvolin, *J. Chim. Phys. Phys. Chim. Biol.* **80**, 15 (1983).
- [S8] P. van der Ploeg and H. Berendsen, *J. Chem. Phys.* **76**, 3271 (1982).
- [S9] P. van der Ploeg and H. Berendsen, *Mol. Phys.* **49**, 233 (1983).
- [S10] O. Edholm, H. Berendsen, and P. van der Ploeg, *Mol. Phys.* **48**, 379 (1983).
- [S11] M. Müller, K. Katsov, and M. Schick, *J. Polym. Sci. B: Polymer Physics* **41**, 1441 (2003).
- [S12] C. Ting and Z.-G. Wang, *Biophys. J.* **100**, 1288 (2011).
- [S13] G. H. Fredrickson, *The Equilibrium Theory of Inhomogeneous Polymers* (Oxford University Press, Oxford, 2006).
- [S14] K. Hong and J. Noolandi, *Macromolecules* **13**, 964 (1980).
- [S15] J. M. H. M. Scheutjens and G. J. Fleer, *J. Chem. Phys.* **83**, 1619 (1979).
- [S16] W. E, W. Ren, and E. Vanden-Eijnden, *J. Chem. Phys.* **126**, 164103 (2007).
- [S17] X. Y. Cheng, L. Lin, W. E, P. W. Zhang, and A. C. Shi, *Phys. Rev. Lett.* **104**, 148301 (2010).
- [S18] C. Ting, D. Appelö, and Z.-G. Wang, *Phys. Rev. Lett.* **106**, 168101 (2011).
- [S19] W. H. Li, P. F. Nealey, J. J. de Pablo, and M. Müller, *Phys. Rev. Lett.* **113**, 168301 (2014).
- [S20] N. M. Maurits and J. G. E. M. Fraaije, *J. Chem. Phys.* **107**, 5879 (1997).
- [S21] E. Reister, M. Müller, and K. Binder, *Phys. Rev. E* **64**, 041804 (2001).
- [S22] C. J. Knight and J. S. Hub, *Bioinformatics* **31**, 2897 (2015).
- [S23] J. S. Hub and N. Awasthi, *J. Chem. Theory Comput.* (2017).
- [S24] S. Miyamoto and P. A. Kollman, *J. Comp. Chem.* **13**, 952 (1992).
- [S25] B. Hess, *J. Chem. Theory Comput.* **4**, 116 (2008).
- [S26] W. F. van Gunsteren and H. J. C. Berendsen, *Mol. Sim.* **1**, 173 (1988).
- [S27] G. Bussi, D. Donadio, and M. Parrinello, *J. Chem. Phys.* **126**, 014101 (2007).
- [S28] H. J. C. Berendsen, J. P. M. Postma, A. DiNola, and J. R. Haak, *J. Chem. Phys.* **81**, 3684 (1984).
- [S29] T. Darden, D. York, and L. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).
- [S30] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, *J. Comp. Phys.* **103**, 8577

- (1995).
- [S31] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, *SoftwareX* **1**, 19 (2015).
- [S32] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Comp. Phys.* **79**, 926 (1983).
- [S33] G. M. Torrie and J. P. Valleau, *Chem. Phys. Lett.* **28**, 578 (1974).
- [S34] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, *J. Comput. Chem.* **13**, 1011 (1992).
- [S35] J. S. Hub, B. L. de Groot, and D. van der Spoel, *J. Chem. Theory Comput.* **6**, 3713 (2010).
- [S36] C. Rycroft, *Chaos* **19**, 041111 (2009).
- [S37] T. Tolpekina, W. Den Otter, and W. Briels, *J. Chem. Phys.* **121**, 8014 (2004).
- [S38] B. Hess, *J. Chem. Phys.* **116**, 209 (2002).
- [S39] O. Berger, O. Edholm, and F. Jähnig, *Biophys. J.* **72**, 2002 (1997).
- [S40] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, in *Intermolecular Forces*, edited by B. Pullman (D. Reidel Publishing Company, Dordrecht, 1981), pp. 331–342.