Biophysical Journal, Volume *113*

Supplemental Information

Cholesterol Promotes Protein Binding by Affecting Membrane Electrostatics and Solvation Properties

Milka Doktorova, Frederick A. Heberle, Richard L. Kingston, George Khelashvili, Michel A. Cuendet, Yi Wen, John Katsaras, Gerald W. Feigenson, Volker M. Vogt, and Robert A. Dick

Supporting Material for:

Cholesterol promotes binding of retroviral matrix protein by indirectly affecting membrane electrostatics and solvation properties

Milka Doktorova,[1,#] Frederick A. Heberle,[2-4, *,#] Richard L. Kingston,[5] George Khelashvili,[6] Michel A. Cuendet,[6] Yi Wen,[7] John Katsaras,[2,4,8,9] Gerald W. Feigenson,[7] Volker M. Vogt,[7] and Robert A. Dick[7, *,#]

[1]Tri-Institutional PhD Program in Computational Biology and Medicine, Weill Cornell Medical College, New York, New York 10065, United States; [2]The Bredesen Center for Interdisciplinary Research and Graduate Education, University of Tennessee, Knoxville, Tennessee 37996, United States; [3]Joint Institute for Biological Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States; [4]Biology and Soft Matter Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States; [5]School of Biological Sciences, The University of Auckland, Auckland, New Zealand; [6]Department of Physiology and Biophysics, Weill Cornell Medical College, New York, New York 10065, United States; [7]Department of Biochemistry and Molecular Cell Biology, Cornell University, Ithaca, New York 14853, United States; [8]Department of Physics and Astronomy, University of Tennessee, Knoxville, Tennessee 37996, United States; [9]Shull Wollan Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States;

[#]Contributed equally
[*]Co-corresponding authors

DETAILED METHODS

SANS data analysis. SANS data were modeled with a heterogeneous core-shell (HCS) form factor (1) with modifications discussed here. This model is appropriate for describing scattering from a "patchy" spherical shell particle, such as a phase-separated or protein-bound unilamellar lipid bilayer vesicle, shown schematically in Fig. S2A. The coherent scattered intensity of such a particle contains three contributions:

$$I(q) = I_{hom}(q) + I_{intra}(q) + I_{inter}(q). \qquad (S1)$$

The first term in Eq. S1 accounts for the *homogeneous* contribution to the total scattering arising from structure normal to the plane of the bilayer. Differences in the atomic composition of lipid headgroups and chains generally results in different average neutron scattering length densities (NSLDs) for these layers, with the NSLD variation along the bilayer normal described quantitatively by a radial NSLD profile $\rho(r)$, where r is the radial distance from the center of a vesicle. In a phase-separated vesicle with two coexisting environments, the transverse structure from each phase contributes to the homogeneous scattering, which can be expressed as:

$$I_{hom}(q) = 4\pi\left[2\sqrt{\pi}M_0(q) + 2\sqrt{\pi}a_d W_0(q)\right]^2, \qquad (S2)$$

$$M_0(q) = \int_0^\infty [\rho_c(r) - \rho_s] r^2 j_0(qr) dr, \qquad (S3)$$

$$W_0(q) = \int_0^\infty [\rho_d(r) - \rho_c(r)] r^2 j_0(qr) dr. \qquad (S4)$$

Here, ρ_d, ρ_c, and ρ_s refer respectively to the NSLD of the domain phase, continuous phase, and surrounding aqueous solvent, a_d is the fraction of the vesicle surface area occupied by the domain phase, and j_0 is the zeroth order Bessel function. From Eqs. S2-4, it is clear that I_{hom} depends only on the radial (transverse) bilayer structure and relative amounts of the two phases, but not on the size or spatial organization of domains. The latter information is accounted for by the second and third terms in Eq. S1, I_{intra} and I_{inter}. Making use of a spherical harmonic expansion of the vesicle scattering amplitude, the intradomain scattering contribution is given by:

$$I_{intra}(q) = 4\pi N_d \sum_{l=1}^{\infty} \left|\widetilde{w}_l^0(\alpha_d)\right|^2 |W_l(q)|^2, \qquad (S5)$$

$$W_l(q) = \int_0^\infty [\rho_d(r) - \rho_c(r)] r^2 j_l(qr) dr, \qquad (S6)$$

$$\widetilde{w}_l^0(\alpha_d) = \frac{\sqrt{(2l+1)}}{2l}[\cos\alpha_d\, P_l(\cos\alpha_d) - P_{l+1}(\cos\alpha_d)], \qquad (S7)$$

where N_d is the number of domains, α_d is the angle formed by vectors pointing from the vesicle center to the domain center and edge, and P_l is the Legendre polynomial of degree l. Finally, the interdomain scattering arising from coherent interference between different domains is given by:

$$I_{inter}(q) = 4\pi \sum_{J \neq K} \sum_{l=1}^{\infty} \left|\widetilde{w}_l^0(\alpha_d)\right|^2 |W_l(q)|^2 P_l\left(\cos\theta_{JK}\right), \qquad (S8)$$

where θ_{JK} is the angle between the vesicle center and the centers of domains J and K. The effects of vesicle size polydispersity are included by averaging the monodisperse intensity $I(q, R)$ (i.e., Eqs. S1-8) over a Schulz distribution:

$$G(R) = \left(\frac{1}{R_m \sigma^2}\right)^{(1/\sigma^2)} \frac{R^{(1/\sigma^2 - 1)}}{\Gamma(1/\sigma^2)} exp\left[\frac{-R}{R_m \sigma^2}\right], \qquad (S9)$$

where R_m is the most probable vesicle radius, σ is the root mean square deviation from R_m, and Γ is the gamma function. The polydisperse intensity $I_P(q)$ is then given by:

$$I_P(q) = \int_0^{\infty} I(q, R) G(R) dR. \qquad (S10)$$

To summarize, the HCS model requires as input:

1. Radial SLD profiles for the domain and continuous phases, $\rho_d(r)$ and $\rho_c(r)$. Assuming that the radial bilayer structure does not depend on vesicle size R, then $\rho(r; R) = \rho(z + R)$ for all R, where $\rho_d(z)$ and $\rho_c(z)$ are transverse SLD profiles centered at $z = 0$. Diverse models for transverse SLD profiles can be found in the literature (reviewed in (2)); our analysis used a simple "slab" model described below.
2. The size and spatial arrangement of domains on the vesicle surface, given by the angle α_d and the distribution of domain center-center angles θ_{JK}. Our analysis assumed circular domains with a fixed area of 1375 Å^2 (corresponding to the cross-sectional area of an MA monomer), randomly arranged on the vesicle surface.
3. A vesicle size distribution $G(R; R_m, \sigma)$. We note that for vesicles larger than ~ 300 Å diameter, the precise values of R_m and σ do not affect $I(q)$ for q > 0.05 Å^{-1}. In our analysis, we fixed these parameters at 500 Å and 125 Å, respectively (i.e., a relative polydispersity of 0.25).

Transverse bilayer structure was modeled for each phase separately, using volume probability distributions for different lipid and protein "quasi-molecular fragments". The lipid headgroups and hydrocarbon chains were modeled as separate fragments with uniform probability distributions:

$$P_i(z) = \frac{N_i V_i}{A_L \sigma_i}[\Theta(\mathrm{z} - \mathrm{z}_i) - \Theta(\mathrm{z} - \mathrm{z}_i - \sigma_i)], \qquad (S11)$$

$$\Theta(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}, \qquad (S12)$$

where V_i is the fragment volume, σ_i is the fragment width along the bilayer normal, A_L is the unit cell area, z_i is the fragment's lower boundary ($z_i + \sigma_i$ is the upper boundary), and Θ is the unit step function. For the domain phase, externally-bound protein was modeled with a Gaussian probability distribution:

$$P_P(z) = \frac{2\chi_P V_P}{\sqrt{2\pi} A_L \sigma_p} exp\left[\frac{-(z - z_p)^2}{2{\sigma_p}^2}\right]. \qquad (S13)$$

In Eq. S13, χ_P is the protein mole fraction in the protein+lipid sample; because the bilayer unit cell by definition contains exactly two lipids, the (fractional) number of proteins per unit cell is given by $N_P = 2\chi_P$. All lipid and protein fragment volume probability functions satisfy the following relationships:

$$\int P_i(z)dz = \frac{N_i V_i}{A_L}, \qquad (S14)$$

$$P(z) = \sum_i P_i(z) + P_w(z) = 1, \qquad (S15)$$

where N_i is the number of fragment i contained in the unit cell, P_w is the water probability, and P is the total probability. These equations enforce local volume conservation: any volume not occupied by a lipid or protein fragment must be occupied by water. Equation S15 can be rearranged to define the water probability function in terms of lipid and protein fragment probabilities:

$$P_w(z) = 1 - \sum_i P_i(z). \qquad (S16)$$

The scattering length density profile is then given by a weighted sum of the lipid and protein fragment probabilities:

$$\rho(z) = \sum_i \rho_i P_i(z), \qquad (S17)$$

$$\rho_i = \frac{b_i}{V_i}, \qquad (S18)$$

where b_i is the fragment's coherent scattering length. In the case of mixtures of two or more lipids, the lipid fragments are composites whose properties represent average properties of the mixture, and are approximated as mole fraction-weighted sums of individual lipid properties, i.e.:

$$V_i = \sum_j \chi_j V_{ij}, \qquad (S19)$$

$$b_i = \sum_j \chi_j b_{ij}, \qquad (S20)$$

where V_{ij} and b_{ij} are the fragment i volume and scattering length of mixture component j, respectively, and χ_j is the component j mole fraction. Values for the volumes and scattering length densities of the different lipid and protein species are given in Table S3. The total bilayer (Luzzati) thickness D_B is calculated from the total lipid volume V_L and area per lipid:

$$D_B = V_L/A_L, \qquad (S21)$$

where V_C and V_H are the lipid chain and headgroup volumes, respectively, and $V_L = V_C + V_H$. Similarly, the hydrocarbon chain thickness $2D_C$ is calculated from the hydrocarbon chain volume and area per lipid:

$$2D_C = 2V_C/A_L. \qquad (S22)$$

Finally, to account for the smearing effects of thermal disorder, the NSLD profile was smoothed by convolution with a Gaussian function:

$$\tilde{\rho}(z) = \rho * g = \int_{-\infty}^{\infty} \rho(x) g(\sigma_s; z - x) dx, \qquad (S23)$$

with the width of the smoothing window σ_s fixed at 2 Å.

X-ray crystallography. All crystals were grown using the sitting-drop vapor diffusion technique. Crystallization conditions were identified using screening experiments based on orthogonal arrays (3). Details are given in Table S1. Prior to data collection crystals were transferred into cryo-protective solutions, and vitrified by direct immersion in liquid nitrogen. Diffraction data were collected by the oscillation method, using both laboratory and synchrotron radiation sources (Table S1), with crystals maintained at 100–110 K in a cold gas stream throughout. Data integration and scaling were performed with the program HKL2000 (4).

The structure of RSV MA was determined by the method of Multiple Isomorphous Replacement with Anomalous Scattering (MIRAS). To produce isomorphous derivatives, the crystals were soaked for 2-10 minutes in cryo-protective solutions incorporating either 1 M NaI or 1 M NaBr, prior to immersion in liquid nitrogen. Diffraction data on NaBr-soaked crystals were collected at several wavelengths near the Bromine K edge, while diffraction data on NaI-soaked crystals were collected at wavelengths of 1.033 and 1.653 Å. The program SHELXD (5) was used to identify the halide-binding sites. The NaI and NaBr-derivatized crystals shared a common site, with an additional unique site

for the NaI derivative. The program SHARP 2.0 (6) was used to refine site occupancies and calculate phases, producing a partially interpretable electron density map in which helices were clearly visible. Repeat rounds of model building and refinement using the programs Coot (7) and Refmac (8) allowed for the completion of a structural model for the N-terminal region of the molecule (amino acids 1-102). There was no interpretable electron density associated with the remainder of the sequence (amino acids 103-155). Subsequently structures of a truncated variant (MA2-102) were determined by the method of molecular replacement, using the program Phaser (9) to position the structural model where required. Statistics associated with the native data sets and refined structural models are shown in Table S1.

MD simulations. All MD simulations were performed with the NAMD software, versions 2.7-2.10 (10) and analyzed with VMD (11) and custom Tcl scripts. Protein secondary structure was calculated using DSSP (12).

The two bilayers, POPC/POPS 70/30 mol% and POPC/POPS/Chol 34/30/36 mol%, were constructed with CHARMM-GUI (13) and simulated as described in (14). The bilayers contained 70 and 100 lipids per leaflet, respectively, and were solvated with 45 waters/lipid and 50 mM NaCl. The POPC/POPS bilayer was simulated for a total of 191 ns and the POPC/POPS/Chol bilayer for a total of 270 ns. The last 100 ns of each simulation were used for subsequent analysis.

MA was placed on top of each bilayer using coordinates taken from the last frames of the two bilayer-only simulations. MA was oriented with respect to the membrane surface as previously done (15). The systems were energy minimized for 24000 steps and run for 1.2 ns with a 1 fs time step before the production runs. Three replica simulations were run for the POPC/POPS +MA system with total simulation times of 184 ns, 242 ns and 198 ns. The POPC/POPS/Chol +MA system was simulated for 74 ns, after which time two replica simulations were started and run for an additional 132 and 141 ns respectively. The last 100 ns of the trajectories were used for subsequent analysis.

To ensure that the applied periodic boundary conditions did not affect the interaction of the protein with the bilayer, an additional simulation was performed in which two MA proteins were placed as described above, but on each side of a POPC/POPS/Chol bilayer (i.e., one on the top leaflet and one on the bottom leaflet). Thus, any modes of interaction that could cause large leaflet deformations and accumulation of pressure if applied only on one side of the bilayer, would be counterbalanced and not suppressed. The simulation was run for a total of 204 ns and the interaction of each of the two MA proteins with their respective leaflets was analyzed separately. Since the results were the same as in the systems with a single MA, they were treated as two additional replicas of the POPC/POPS/Chol +MA system and were analyzed jointly with the other simulations.

Calculations of electrostatic potential and fraction of bound protein. The electrostatic potential on the membrane surface was calculated using the analytical solution to the non-linear Poisson-Boltzmann equation (Eq. 15 in (16)) with the membrane surface taken as $z = 0$. Surface charge density was calculated as the mole fraction of charge divided by

the average area per lipid of the bilayer, with units of $e^-/Å^2$. Unless otherwise noted, the areas per lipid calculated from MD simulations were used.

To generate the contour plots in Fig. 6B of the main text, the fraction of bound protein f from protein binding assays (i.e., the data in Fig. 5F of the main text) was modeled as a sigmoidal function of the membrane surface potential ψ:

$$f(\psi) = \frac{a}{1 + e^{b(\psi - \psi^*)}}, \qquad (S24)$$

where a is a scaling factor representing the maximum bound fraction, b is a stretching factor representing the width of the sigmoidal binding transition, and ψ^* is the surface potential at half-maximum binding. The three adjustable parameters were optimized with Mathematica's built-in NonlinearModelFit function using a Levenberg-Marquardt algorithm. The best-fit parameter values were: a = 70.9%, b = 0.285, and ψ^* = –55.9 mV (Fig. S10). This parameter set was used with Eq. S24 to map the calculated membrane surface potential (Fig. 6B, left-hand plot) to the percentage of bound protein, generating the right-hand plot in Fig. 6B (main text).

Monte Carlo simulations of equilibrium lipid distributions. Monte Carlo simulations of a 100 × 100 triangular lattice in the canonical ensemble were performed using custom code written in Mathematica and available from the authors upon request. Each lattice site represented either a PC or PS lipid, with the composition fixed at 70/30 mol% PC/PS. Neglecting multibody and long-range electrostatic interactions, the total energy of a lattice composed of a fixed number of PS and PC lipids (N_{PS} and N_{PC}, respectively) is given by the sum of unique nearest-neighbor pairwise interactions (17):

$$U^T = \frac{ZN_{PS}U_{PS-PS}}{2} + \frac{ZN_{PC}U_{PC-PC}}{2} + N_{PS-PC}\Delta E_m, \qquad (S25)$$

$$\Delta E_m = U_{PS-PC} - (U_{PS-PS} + U_{PC-PC})/2, \qquad (S26)$$

where U_{PS-PS} and U_{PC-PC} are the interaction energies for a neighboring pair of PS and PC lipids, respectively, N_{PS-PC} is the total number of PC/PS contacts, and Z is the number of nearest neighbors in a lattice site (6 for a triangular lattice). The sole adjustable parameter ΔE_m is the excess mixing energy of a PC/PS pair. The first two terms in Eq. S25 do not depend on the lipid distribution and therefore do not contribute to non-ideal mixing. As a result, only the third term was updated. For each proposed update, the position of two randomly chosen lipids was exchanged, generating a change in the lattice energy $\Delta U^T \equiv U^T_{new} - U^T_{init}$ that was either favorable/neutral ($\Delta U^T \leq 0$) or unfavorable ($\Delta U^T > 0$). Importance sampling was based on the Metropolis criterion, whereby a favorable move was always accepted, and an unfavorable move was accepted with probability $P = \exp(-\Delta U^T / k_B T)$ by first drawing a random number R from a uniform probability distribution $R \sim U[0,1]$ and then performing the exchange if $R \leq P$. Each simulation was equilibrated for a minimum of 10^3 MC cycles, where a cycle is defined as a number of proposed exchanges equal to the lattice size (here, 10^4

exchanges). Equilibrium was judged by convergence of the lattice energy upon starting from either (a) a completely random distribution of the lipids, or (b) a block distribution of the lipids. The number of MC cycles required to reach convergence increased with increasing ΔE_m, varying from ~ 300 cycles for ΔE_m = + 0.1 k_BT to ~ 10^4 cycles for ΔE_m = + 0.5 k_BT. For additional details on MC lattice simulations, we point the reader toward several studies relevant to lipid bilayers (17-27).

Calculation of relative surface potential from Monte Carlo snapshots. To calculate a surface potential map from MC snapshots, lattice sites were assigned relative real space coordinates $\boldsymbol{l} = (l_x, l_y, 0)$ using the relationship between the triangular lattice spacing λ and the unit cell area A:

$$\lambda = \left(\frac{2A}{\sqrt{3}}\right)^{1/2}, \qquad (S27)$$

and taking A to be the area per lipid (63 Å^2). The potential V at an arbitrary point $\boldsymbol{p} = (p_x, p_y, p_z)$ is then given by:

$$V = k_e \sum_j \frac{q_j}{r_j}, \qquad (S28)$$

where k_e is Coulomb's constant, $r_j = \|\boldsymbol{l_j} - \boldsymbol{p}\|$ is the distance between point $\boldsymbol{p}$ and lattice site j, q_j is the charge at lattice site j (i.e., 0 for a neutral PC lipid and the elementary charge e^- for an acidic PS lipid), and the sum is over all lattice sites j. Because the absolute potential depends strongly on the lattice size, V was normalized to a reference potential $\bar{v}$ arising from a uniformly charged lattice calculated as:

$$\bar{v} = k_e \chi e^- \sum_j r_j^{-1}, \qquad (S29)$$

where χ is the mole fraction of charged lipid in the mixture (here, 0.3). Finally the relative potential $\tilde{V}$ at point $\boldsymbol{p}$ was calculated as $\tilde{V} = V/\bar{v}$.

MM/GBSA calculations. The molecular mechanics-generalized Born and surface area (MM-GBSA) method (28, 29) is a so-called end-point free energy approach to estimate the binding free energy between two molecular binding partners, based on a sample of molecular conformations of the complex generated by all-atom molecular dynamics simulation. The MM-GBSA method has been used successfully to estimate the binding free energy of ligands to proteins, and to calculate single-residue contributions to binding free energies of large protein-protein complexes (28-30). For each trajectory frame, the solvent and ions are stripped away and only the coordinates of the binding partners are kept. In the one-trajectory MM-GBSA approach employed here, coordinates for each partner in isolation are extracted from the same trajectory frames of the complex, assuming that these are also acceptable conformations for the molecules in solution. As in

the preceding molecular mechanics-Poisson Boltzmann surface area (MM-PBSA) method (31), the solvation effects are approximated by immersing the molecules in a continuous medium with high relative dielectric constant $\varepsilon_r^{\mathrm{solv}} = 80$. Following the thermodynamic cycle shown in Fig. S11, the binding free energy between two binding partners, here membrane (M) and protein (P), is expressed as

$$\Delta G_{\mathrm{bind}} = \Delta E_{\mathrm{int}}^{\mathrm{vac}} + \Delta G_{\mathrm{desolv}}(M) + \Delta G_{\mathrm{desolv}}(P) - \Delta G_{\mathrm{desolv}}(MP) + \Delta S_{\mathrm{bind}}^{\mathrm{vib}} . \qquad (S30)$$

In general, $\Delta E_{\mathrm{int}}^{\mathrm{vac}}$ is the difference in internal bonded and non-bonded energies in M and P upon binding, calculated with the same Charmm36 parameters (32) as in the MD simulation. Here, a lot of terms cancel out since we use the same coordinates for bound and unbound molecules, such that $\Delta E_{\mathrm{int}}^{\mathrm{vac}}$ boils down to the Van der Waals and electrostatic interaction energies between M and P. $\Delta S_{\mathrm{bind}}^{\mathrm{vib}}$ is the difference in internal vibrational entropy upon binding, which we neglect in the present application. The desolvation penalty for molecular system X is composed of a polar and a non-polar term,

$$\Delta G_{\mathrm{desolv}}(X) = \Delta G_{\mathrm{P,desolv}}(X) + \Delta G_{\mathrm{NP,desolv}}(X). \qquad (S31)$$

The non-polar term accounts for energetic and entropic effects in the solvent related to creating the cavity occupied by X. This term is simply proportional to the SASA,

$$\Delta G_{\mathrm{NP,desolv}}(X) = \gamma\, SASA(X), \qquad (S32)$$

with $\gamma = -0.0072$ kcal/mol/Å^2 (33). The SASA is calculated by rolling a virtual 1.4 Å radius ball over the molecules. The polar solvation term is calculated using the generalized Born (GB) equation (28, 31) with an additional Debeye-Hückel correction to account for ionic screening (28):

$$\Delta G_{\mathrm{P,desolv}}(X) = \frac{c}{2}\left(\frac{1}{\varepsilon_r^{\mathrm{vac}}} - \frac{e^{-\kappa r_{ij}^{\mathrm{GB}}}}{\varepsilon_r^{\mathrm{solv}}}\right)\sum_{i,j\in X}\frac{q_i q_j}{r_{ij}^{\mathrm{GB}}} . \qquad (S33)$$

Here, the q_i are atomic partial charges, $c = 332.0672$ kcal/mol Å / u^2, and κ is the Debeye-Hückel screening constant expressed in Å^{-1} as $\kappa = 0.316\sqrt{[\mathrm{salt}]}$, where [salt] is the monovalent ion concentration in mol/L (34). The modified atomic distances entering the GB equation are given by

$$r_{ij}^{\mathrm{GB}} = \sqrt{r_{ij}^2 + \alpha_i \alpha_j \exp\left\{-\frac{r_{ij}^2}{8\alpha_i\alpha_j}\right\}}. \qquad (S34)$$

Critical quantities for the accuracy of the GB model are the Born radii α_i, which essentially express how far each atom is from the molecular surface. To calculate these, we use the GB-MV2 method (35, 36) implemented in the CHARMM software (37), which was shown to yield very good accuracy compared to Poisson-Boltzmann results.

Because all energies in the MM-GBSA framework are expressed in terms involving single atoms or pairs of atoms, ΔG_{bind} can be decomposed in contributions from separate groups of atoms (29, 30). For the Van der Waals, electrostatic, and GB pair terms, half of the interaction energy is attributed to each atom of the pair. When applied to amino acid side chains, this decomposition yields contributions comparable to those obtained by computational alanine scanning (38). In the present case, summing over all residues in M or P allows us to attribute contributions of each binding partner to ΔG_{bind}. These contributions can differ due to different desolvation penalties on each side. The MM-GBSA free energy decomposition were carried out using a custom set of scripts built upon the original implementation of V. Zoete (30, 38, 39) and the CHARMM version 37 software (37).

To perform the MM-GBSA analysis on each of the two MA/membrane systems, we first created a single trajectory file by concatenating the last ~150 ns from the respective replica simulations of a single MA and the bilayer. The trajectories of the –Chol and +Chol systems consisted of a total of 5929 and 3450 frames, respectively, all output with a stride of 80 ps. Only the top membrane leaflet was considered in the reported energy analysis (taking the bottom leaflet into account had an insignificant effect on the energies). The calculation was set up and run with 50 mM salt and a 20 Å cutoff for VdW and electrostatics.

ESR measurements. ESR was performed as previously described (40) with the following changes. Multilamellar vesicles were extruded to form 100 nm large unilamellar vesicles (LUVs). The LUVs were incubated with MA protein at a ratio of 0.31 mg protein to 1 mg lipid (the same ratio used for SANS measurements) in buffer (20 mM Tris HCl pH 8, 50 mM NaCl, 2 mM TCEP) for no less than 30 minutes at ambient room temperature (~22 °C) prior to loading into glass ESR capillary tubes. The final protein concentration in the ESR experiment was ~ 68 μM, or approximately three times the MA binding constant measured by SPR (41) for similar binding reactions. The model-free order parameters for POPC/POPS LUVs with and without MA were each 0.11, and the model-free order parameters for POPC/POPS/Chol LUVs with and without MA were each 0.22 (42).

Continuum mean-field modeling of MA protein-membrane interactions. To quantify the extent of lipid segregation around the MA protein adsorbed to a membrane, we used a previously developed continuum mean-field (CMF) computational approach (43, 44). The CMF method quantifies essential components of the energetics of protein-lipid interactions and describes the combined kinetic effect of many lipid species interacting with the membrane-adsorbed protein. The protocol defines the steady state of the system consisting of the membrane-associated protein, and includes all important degrees of freedom (electrostatics, mixing entropy of lipids and solution ions), as described in detail in our publications (43-47). To this end, a hybrid representation of the computational model is constructed in which membrane-associated proteins are treated at detailed atomic level in three dimensions, and the lipid membrane is considered as a continuum

elastic medium comprised of two-dimensional smooth charged surfaces representing the lipid polar headgroups, and a low-dielectric hydrocarbon core volume.

This system is subjected to a self-consistent minimization of the governing mean-field-based free energy functional F that depends on local lipid component densities $\varphi(x,y)$ and mobile ion concentrations c_+ and c_- (for positive and negative ionic species, respectively) in the solution. In particular, as detailed in (43, 44), F can be written as the sum of electrostatic energy (F_{el}), lipid mixing entropy (F_{lip}), and salt ion translational entropy (F_{ion}) contributions:

$$F = F_{el} + F_{lip} + F_{ion}, \quad (S35)$$

where,

$$F_{el} = \frac{1}{2}\epsilon_0\epsilon_w\left(\frac{k_BT}{e^2}\right)\int_V (\nabla\Phi)^2 dv, \quad (S36)$$

$$F_{lip} = \frac{k_BT}{a}\int_A \left[\varphi\, ln\frac{\varphi}{\varphi_0} + (1-\varphi)\, ln\frac{(1-\varphi)}{(1-\varphi_0)}\right] dA, \quad (S37)$$

$$F_{ion} = k_BT\int_V \left[c_+\, ln\frac{c_+}{c_0} + c_-\, ln\frac{c_-}{c_0} - (c_+ + c_- - 2c_0)\right] dv. \quad (S38)$$

In the above, k_B is the Boltzmann constant, T is the temperature, e is the elementary charge, ϵ_0 is the permeability of free space, and $\epsilon_w = 80$ is the dielectric constant of the aqueous solution. Φ represents reduced (dimensionless) electrostatic potential in space, a denotes the area per lipid, c_0 is the salt concentration in the bulk, and φ_0 represents bulk concentration of a charged lipid species. The $\varphi(x,y)$ local field relates to the surface charge densities $\sigma(x,y)$ through $\sigma(x,y) = (e/a)\varphi(x,y)z(x,y)$, where $z(x,y)$ denotes valency of the lipid at (x,y). Minimization of F with respect to c_+ and c_- leads to the non-linear Poisson-Boltzmann (NLPB) equation (48):

$$\nabla\Phi = \lambda^{-2}\, sinh\,\Phi, \quad (S39)$$

which is solved to obtain Φ in space (λ being the Debye length of the electrolyte solution). As seen from Eqs. S35-38, this electrostatic potential is self-consistently dependent on the local lipid concentrations through the entropic penalty (F_{lip}) due to lipid segregation or de-mixing. Thus, a self-consistent search for the free energy minimum is conveniently carried by linking Φ (obtained from the NLPB equation) and spatial charged-lipid compositions φ on each leaflet of the membrane to the respective electrochemical potentials μ through the Cahn-Hilliard (CH) equation (49):

$$\frac{\partial \varphi(\vec{r},t)}{\partial t} = D_{lip}\nabla^2\mu(\vec{r},t). \qquad (S40)$$

To quantify interactions between MA protein and the membrane composed of 70/30 mol% mixture of POPC/POPS lipids with the CMF approach, we took one of the top 5 simulation frames with the lowest MA-bilayer total interaction energy (as calculated from MM/GBSA). Then, by applying only global transformations to the protein, we positioned MA in a manner where we maximized the exposure of its lysine residues to a flat lipid surface of $\sigma(x,y) \sim 4.93\times10^{-3}$ e^- charge density, corresponding to ~30 mol% PS lipid content (assuming a = 60.9 $Å^2$).

The self-consistent minimization of F was then carried out for the protein by solving the NLPB equation using the multigrid solver of the Adaptive Poisson-Boltzmann Solver (APBS) suite (50) on 1 Å-spaced cubic 256 $Å^3$ mesh as described previously (43). This was done using a 0.05 M ionic solution of monovalent counterions (corresponding to λ = 13.49 Å Debye length), and a dielectric constant of 2 for the membrane interior and protein, and 80 for the solution. The protein models were positioned so that the minimum distance between the protein and the lipid surface was 2 Å (43, 47).

The CMF calculation shows the adsorption free energy of MA protein, $\Delta F = F-(F_{prot}+F_{memb})$ (F_{prot} and F_{memb} representing the free energy of protein and membrane system in separation) onto the POPC/POPS 70/30 mol% lipid membrane is ~ –9 k_BT (see Fig. S4A). Importantly, lipid de-mixing had an insignificant effect on ΔF (the minimization procedure resulted in a change in ΔF of < 1 k_BT, Fig. S4A) as the local concentration of PS under the adsorbed protein calculated by integrating the charge map of the smallest rectangle enclosing the protein shadow, was ~ 31%, or only slightly higher than the bulk PS concentration of 30% (see Fig. S4B). The minor extent of lipid segregation, together with the concomitant insignificant change in the adsorption free energy, suggests that MA protein adsorbs onto the PC/PS membrane mainly through the electrostatic sensing of PS lipid domains.

SUPPORTING TABLES

Table S1. Protein crystallization conditions, and statistics associated with the X-ray diffraction data and atomic models.

Protein	**RSV MA**	**RSV MA_{2-102}**	**RSV MA_{2-102}**
Crystallization Conditions			
Protein concentration (μM)	680	540	540
Reservoir Solution	2.60 M Ammonium formate 0.20 M β-Alanine/KOH pH 10.3	18%(w/v) PEG 8000 0.2 M Succinic acid/KOH pH 5.5 1.0 M Ammonium nitrate	0.6 M Malonic acid /KOH pH 9.1 0.1 M Boric acid /KOH pH 9.1
Temperature (°C)	Ambient	18	18
X-ray diffraction data			
Cryoprotectant	4.00 M Ammonium formate 0.20 M β-Alanine/KOH pH 10.3 30 %(v/v) Ethylene glycol	20%(w/v) PEG 8000 0.1 M Succinic acid/KOH pH 5.5 1.0 M Ammonium nitrate 20 %(v/v) Ethylene glycol	0.6 M Malonic acid /KOH pH 9.1 0.1 M Boric acid /KOH pH 9.1 20 %(v/v) Ethylene glycol
Space group	$I4_122$	$I4_1$	$I4_122$
Unit cell lengths (Å)	a=b= 66.2, c= 218.8	a=b= 79.0, c= 27.8	a=b= 67.4, c=220.8
X-ray source	ALS Beamline 8.2.1	Rigaku MicroMax-007 HF Rotating Copper Anode	Rigaku MicroMax-007 HF Rotating Copper Anode
X-ray wavelength (Å)	1.03320	1.54179	1.54179
Sample Temperature (K)	100	110	110
Data resolution limits (Å) [a]	49.0 - 2.85 (2.95 - 2.85)	39.5 - 1.86 (1.93 - 1.86)	49.7 - 3.20 (3.31 - 3.20)
Number of unique observations [a]	6104 (582)	7363 (681)	4535 (429)
Mean Redundancy [a]	8.7 (9.0)	12.4 (11.4)	10.1 (10.5)
Completeness (%) [a]	99.9 (100.0)	99.5 (94.6)	99.9 (100.0)
$R_{measure}$ [a]	ND (ND)	0.058 (0.289)	0.139 (0.703)
R_{merge} [a]	0.070 (0.455)	0.055 (0.277)	0.132 (0.678)
Mean I / σI [a]	34.0 (4.1)	56.0 (10.0)	22.2 (4.5)
Crystallographic models			
Number of protein molecules in the asymmetric unit	1	1	1
R_{work} / R_{free} [b]	0.195 / 0.239	0.164 / 0.210	0.202 / 0.249
Total number of protein atoms	767	741	791
Number of water molecules	7	54	6
Other ligands	-	Ethylene Glycol NO_3^-	-
Disorder Model	Individual Isotropic B-factors	Individual Isotropic B-factors	Individual Isotropic B-factors
Mean total isotropic B-factor, all protein atoms ($Å^2$):	67.4	23.1	74.7
Bulk Solvent model	Mask	Mask	Mask
RMSD from ideal geometry : Bond lengths (Å) / Bond angles (°)	0.008 / 1.325	0.011 / 1.392	0.013 / 1.695
Residues in Favoured / Allowed regions of Ramachandran plot (%)[c]	100.0 / 100.0	97.1 / 100.0	95.3 / 100.0
PDB ID code	5KZ9	5KZA	5KZB

[a]Numbers in parentheses are for the highest resolution shell. [b]Calculated from a randomly selected 5% of observations omitted from all model refinement. [c]Defined by the MolProbity web-server (51).

Table S2. Bilayer structural parameters obtained from the refinement of SANS data. Italicized values indicate constrained parameters, and asterisks indicate parameters that were varied during the fitting routine (all remaining structural and compositional values are obtained through mathematical relationships).

Parameter	– Cholesterol		+ Cholesterol	
	– RSV-MA	+ RSV-MA	– RSV-MA	+ RSV-MA
global				
χ_{POPC}	*0.7*	*0.7*	*0.34*	*0.34*
χ_{POPS}	*0.3*	*0.3*	*0.3*	*0.3*
χ_{CHOL}	--	--	*0.36*	*0.36*
POPC-rich domain				
χ_{POPC}	0.9 ± 0.1*	0.9 ± 0.1*	0.46 ± 0.02*	0.45 ± 0.07*
χ_{POPS}	0.1 ± 0.1	0.1 ± 0.1	0.18 ± 0.02	0.19 ± 0.07
χ_{CHOL}	--	--	*0.36*	*0.36*
V_L [Å^3] [a]	1234 ± 2	1236 ± 2	1012 ± 1	1012 ± 2
A_L [Å^2] [b]	63.1 ± 1.4*	61.5 ± 2.0*	48.2 ± 0.6*	47.7 ± 0.9*
D_B [Å] [c]	39.2 ± 0.9	40.3 ± 1.2	42.0 ± 0.5	42.4 ± 0.8
$2D_C$ [Å] [d]	29.0 ± 0.7	29.9 ± 0.9	33.8 ± 0.4	34.1 ± 0.6
D_H [Å] [e]	6.9 ± 1.5*	6.2 ± 0.6*	7.2 ± 1.4*	6.6 ± 1.3*
POPS-rich domain				
χ_{POPC}	0.3 ± 0.3	0.3 ± 0.3	0.08 ± 0.03	0.19 ± 0.09
χ_{POPS}	0.7 ± 0.3	0.7 ± 0.3	0.56 ± 0.03	0.45 ± 0.09
χ_{CHOL}	--	--	*0.36*	*0.36*
V_L [Å^3] [a]	1222 ± 5	1222 ± 5	1003 ± 2	1005 ± 3
A_L [Å^2] [b]	63.0 ± 4.3*	65.2 ± 2.3*	48.8 ± 4.4*	47.4 ± 2.9*
D_B [Å] [c]	38.9 ± 1.4	38.6 ± 0.9	41.7 ± 1.9	42.2 ± 1.3
$2D_C$ [Å] [d]	29.1 ± 1.1	29.0 ± 0.6	33.8 ± 1.5	34.2 ± 1.0
D_H [Å] [e]	9.2 ± 1.0*	6.4 ± 1.3*	7.8 ± 1.7*	7.6 ± 1.8*
Protein bound to POPS-rich domain				
A_P [Å^2] [f]		*1375*		*1375*
z_P [Å] [g]	--	35.9 ± 0.8*	--	38.0 ± 0.6*
σ_P [Å] [h]	--	7.2 ± 0.6*	--	6.6 ± 0.4*
N_L [i]	--	21 ± 1	--	29 ± 2
N_{POPS} [j]	--	15 ± 5	--	13 ± 3
f_P [k]	--	0.47 ± 0.04*	--	0.50 ± 0.05*
a_P [l]	--	0.35 ± 0.04	--	0.40 ± 0.04

[a]average lipid volume calculated as mole fraction-weighted sum of lipid and chol molecular volumes [b]average area per lipid [c]total bilayer (Luzzati) thickness [d]bilayer hydrocarbon thickness [e]headgroup thickness [f]area per protein [g]center of Gaussian protein volume distribution [h]width of Gaussian protein volume distribution [i]average number of outer leaflet lipids in protein shadow [j]average number of outer leaflet POPS in protein shadow [k]fraction of protein bound to vesicle [l]vesicle area fraction bound by protein

Table S3. Molecular volume V, neutron scattering length b, and neutron scattering length density ρ of different species.

	V [Å^3]		b [fm]		ρ [fm Å^{-3}]	
D_2O	30.1		19.145		0.636	
RSV-MA in D_2O	20280[a]		6325[b]		0.312	
	Head	*Chains*	*Head*	*Chains*	*Head*	*Chains*
POPC[c]	331	916	60.072	-26.624	0.181	-0.029
POPC-D31[c]	331	916	60.072	296.086	0.181	0.323
POPS[d]	278	917	115.789	-26.624	0.417	-0.029
POPS-D31[d]	278	917	115.789	296.086	0.417	0.323

[a]sum of residue volumes taken from ref. (52) [b]accounts for exchangeable protons following ref. (53) [c]volume data from ref. (54) [d]volume data from ref. (55)

Table S4. Estimated composition (56-58) and surface charge density (calculated from the mole fraction-weighted charge and area) of the mammalian plasma membrane inner leaflet.

lipid	mole fraction χ	charge [e$^-$]		area [Å^2]	
		per lipid	χ-weighted	per lipid	χ-weighted
Chol	0.4	0	0	27[a]	10.8
(PO)PE	0.28	0	0	59[b]	16.5
(PO)PC	0.05	0	0	65[c]	3.3
(PO)PS	0.18	-1	0.18	64[d]	11.5
(SA)PI	0.08	-1	0.08	68[e]	5.4
(SA)PIP2	0.02	-3	0.06	68[e]	1.4
Total	1.0		0.32		48.9

[a]ref. (59) [b]ref. (60) [c]ref. (54) [a]ref. (55) [e]ref. (61)

Table S5. Decomposition of the interaction energy of MA and the lipids calculated with the MM-GBSA method from the simulation trajectories as described in the SM text. Shown are the Van der Waals (VdW) and electrostatic (Elec) energies calculated in vacuum and their sum, $\Delta E_{\mathrm{int}}^{\mathrm{vac}}$; the polar (P) and nonpolar (NP) desolvation penalties and their sum, $\Delta\Delta G_{solv}$; and the total binding free energy approximated by $\Delta E_{\mathrm{int}}^{\mathrm{vac}}+\Delta\Delta G_{solv}$. All energy units are in kcal/mol.

system	component	Interaction energy in vacuum			Desolvation penalty			total
		VdW	Elec	ΔG_{in}^{vac}	P	NP	$\Delta\Delta G_{solv}$	
MA and PC/PS	MA	-9.3	-607.8	-617.1	620.9	-4.8	616.1	-1.0
	lipids	-9.3	-607.8	-617.1	620.6	-4.2	616.4	-0.7
MA and PC/PS/Chol	MA	-8.5	-765.6	-774.1	779.1	-5.1	774.0	-0.1
	lipids	-8.5	-765.7	-774.1	773.1	-4.8	768.3	-5.8

SUPPORTING FIGURES

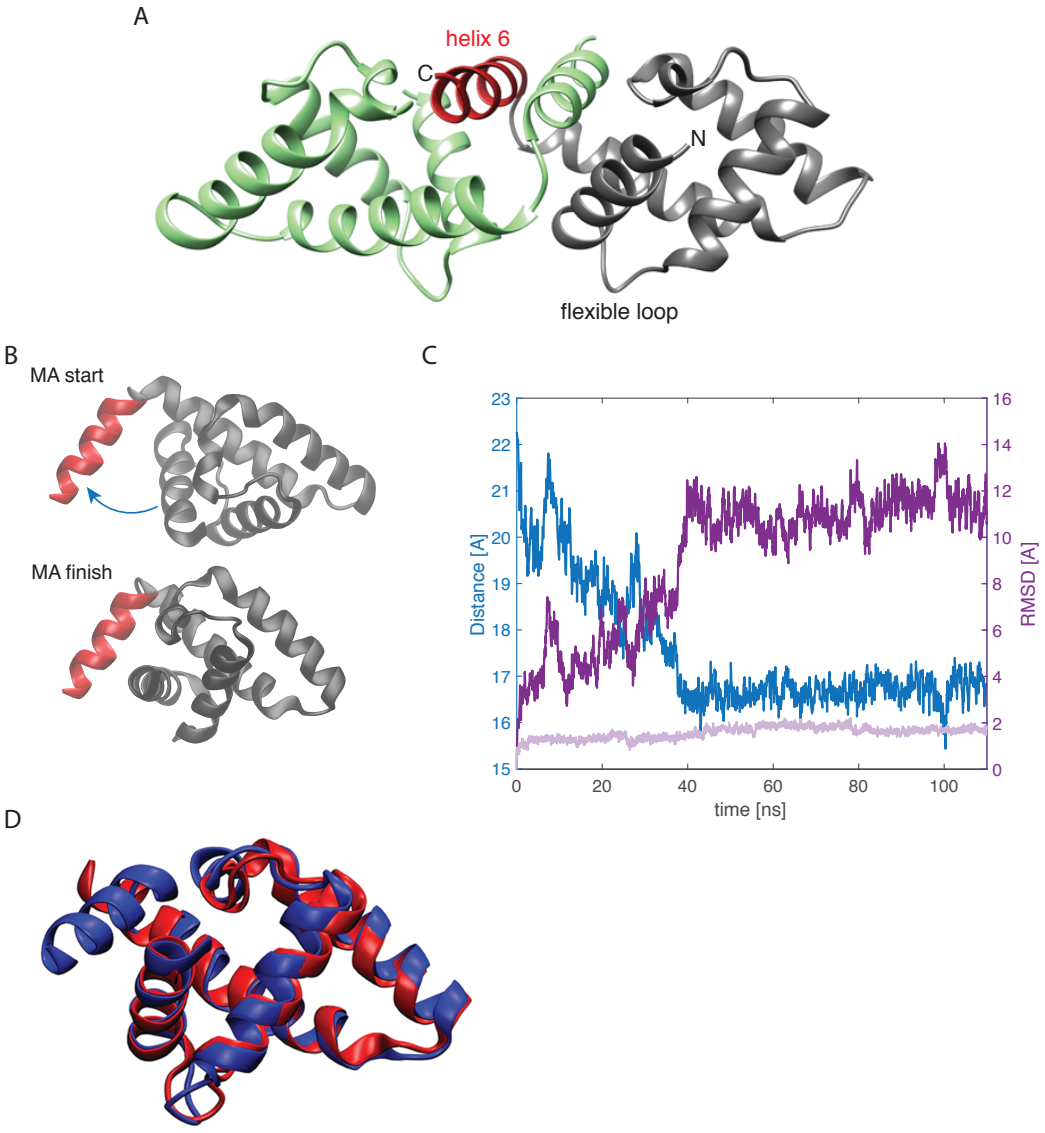

Figure S1. (A) Crystal structure of the resolved MA dimer, with monomers colored in green and grey. Helix 6 of the grey-colored monomer is shown in red. (B) Snapshots of the protein at the start and finish of the water box simulation. Helix 6 (red) is fixed during the simulation and is displayed in the same plane while the rest of the protein (referred to protein *body* from here on) rotates ~ 30 degrees. An arrow denotes the observed structural change. (C) Time evolution of the distance between the centers of mass of helix 6 and the protein body (residues 1-90, blue). Also plotted are the *full protein backbone* RMSD with respect to the starting structure after alignment on helix 6 (dark purple) and the *protein body* RMSD after alignment on the protein body (light purple). (D) Overlay of the monomer crystal structure of MA (red) and the structure of MA used in the simulations (blue). The backbone RMSD between the two structures (excluding helix 6 which is not present in the monomer crystal structure) is 2.8 Å.

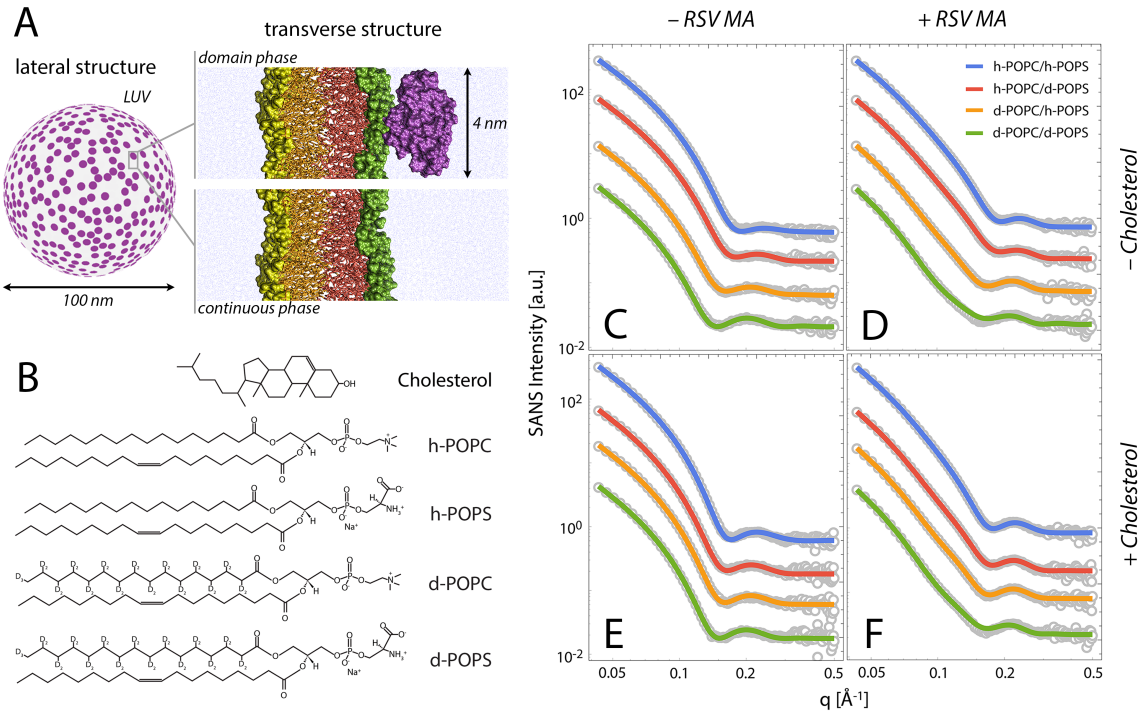

Figure S2. Bilayer structure from analysis of SANS data. (A) Schematic illustration of a 100 nm LUV and the two environments used in the model to analyze the SANS data. The structural model accounts for both transverse and lateral structure. Transverse structure arises from the layered distribution of matter projected onto the bilayer normal, and is mathematically described by parameters related to the volume probability distributions of inner and outer leaflet lipid headgroups, and hydrocarbon chains, in addition to bound protein. Lateral structure can arise from lipid clustering or phase separation, as well as partial surface coverage of bound protein monomers; it is mathematically described by parameters related to the domain size, shape and spatial arrangement (here, domains were modeled as circular disks, randomly arranged on the vesicle surface). Two distinct coexisting environments were modeled—the domain and the surrounding continuous phase—with the lipid composition and transverse structure of each allowed to vary as described in the SM text. (B) Structures of lipids used in SANS experiments. Palmitoyl chain-perdeuterated variants of POPC and POPS (i.e., d-POPC and d-POPS) were used to provide a scattering length density contrast in order to highlight lateral structure. (C-F) Scattering data (open circles) and fits (solid lines) for four different neutron contrast data sets, with different sample compositions: POPC/POPS 70/30 mol% (C); POPC/POPS 70/30 mol% + RSV MA (D); POPC/POPS/Chol 34/30/36 mol% (E); POPC/POPS/Chol 34/30/36 mol% + RSV MA (F).

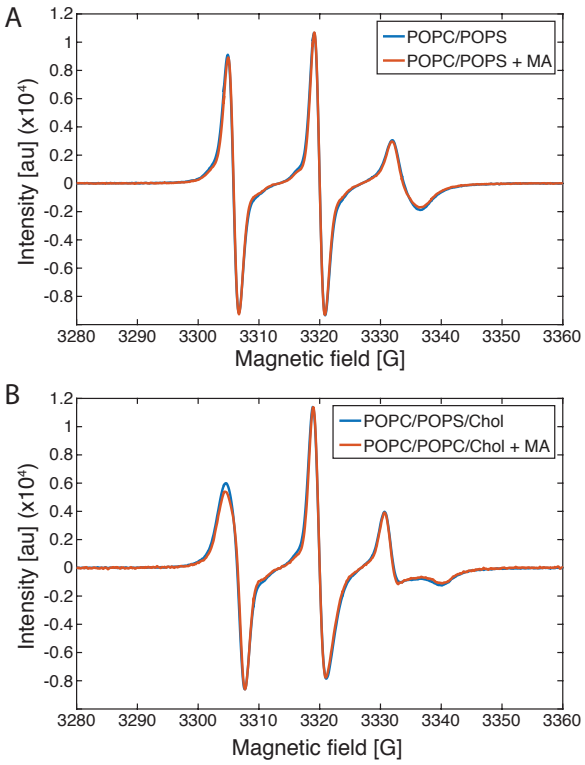

Figure S3. ESR spectra of (A) POPC/POPS with and without MA, and (B) POPC/POPS/Chol with and without MA. All LUVs contained 0.5 mol% of 16:0-16 Doxyl PC.

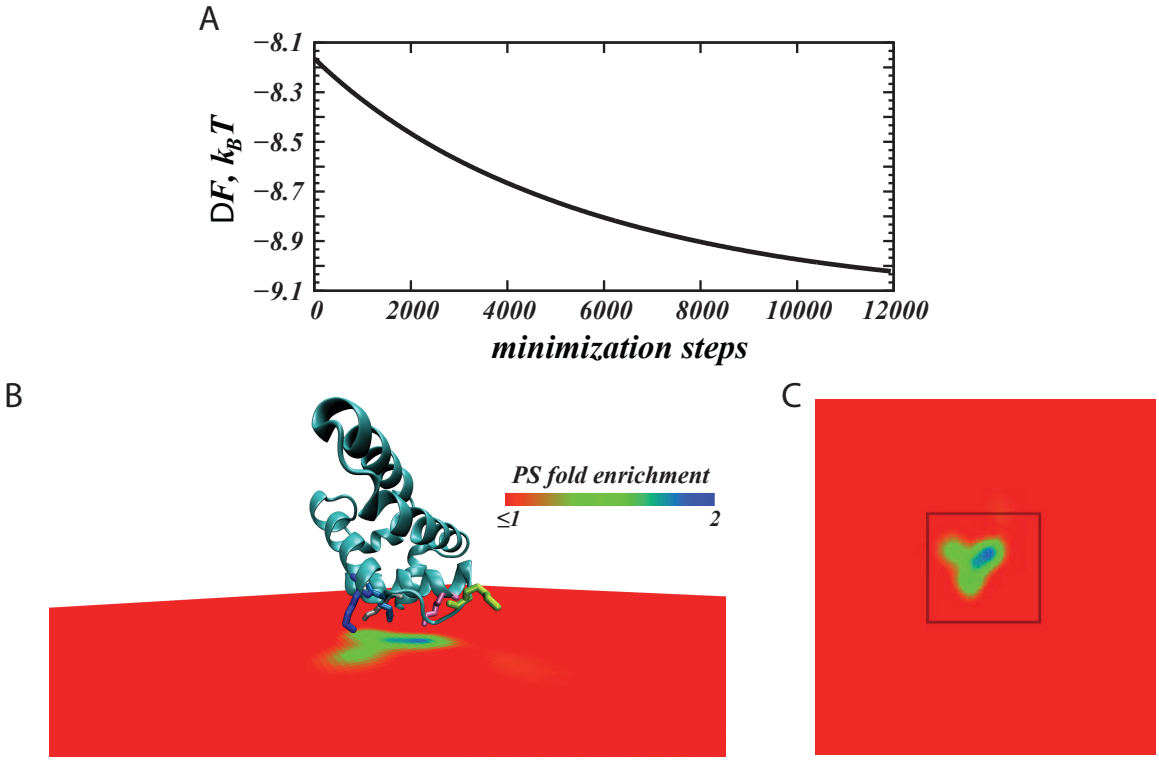

Figure S4: (A) Convergence of the adsorption free energy (ΔF) in the CMF calculations. Shown is the change in ΔF as the mean-field free energy functional is minimized. Value of ΔF at 0^{th} minimization step corresponds to the adsorption free energy onto homogeneous membrane composed of charged and neutral lipids generating a surface charge density of 4.93×10^{-3} e^- (corresponding to a PC/PS lipid mixture with ~30 mol% PS). (B) View of the MA protein (cartoon) adsorbing on the lipid membrane (as seen from the side). The level of PS lipid segregation by the protein, calculated with CMF approach, is illustrated (as ratios of local and bulk lipid fraction values) in color code. Lysine residues close to the surface are shown in licorice and color according to the following: K6 – gray, K13 – purple, K18 – green, K23 – light blue, K24 – dark blue. (C) View of the lipid membrane (as seen from above). Highlighted with a black box is the smallest rectangle enclosing the protein shadow , whose charge density was integrated to calculate the local concentration of PS under MA (~31 mol%).

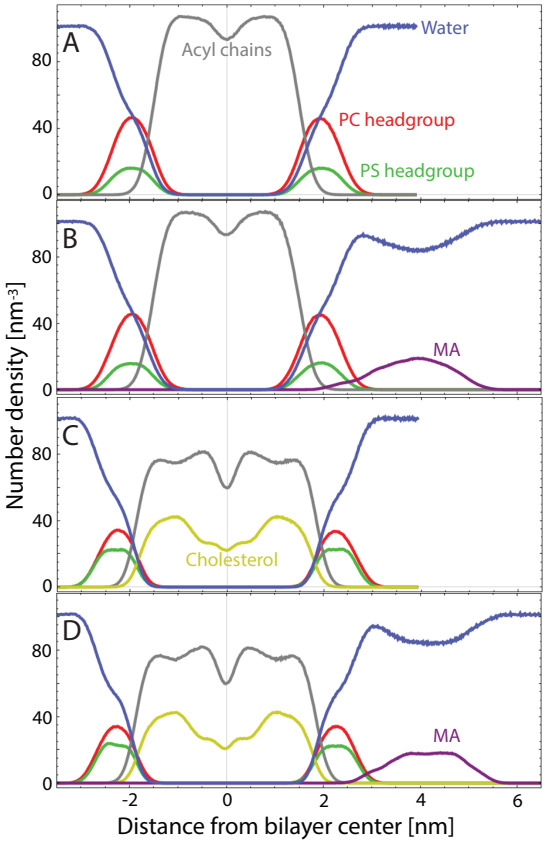

Figure S5. Number density profiles calculated from four different MD simulations. Color coded are the profiles of lipid acyl chains (grey), water (blue), POPC headgroups (red), POPS headgroups (green), Chol (yellow) and MA (purple) calculated from the last 100 ns of each simulation: (A) POPC/POPS 70/30 mol%; (B) POPC/POPS 70/30 mol% with MA; (C) POPC/POPS/Chol 34/30/36 mol%; and (D) POPC/POPS/Chol 34/30/36 mol% with MA.

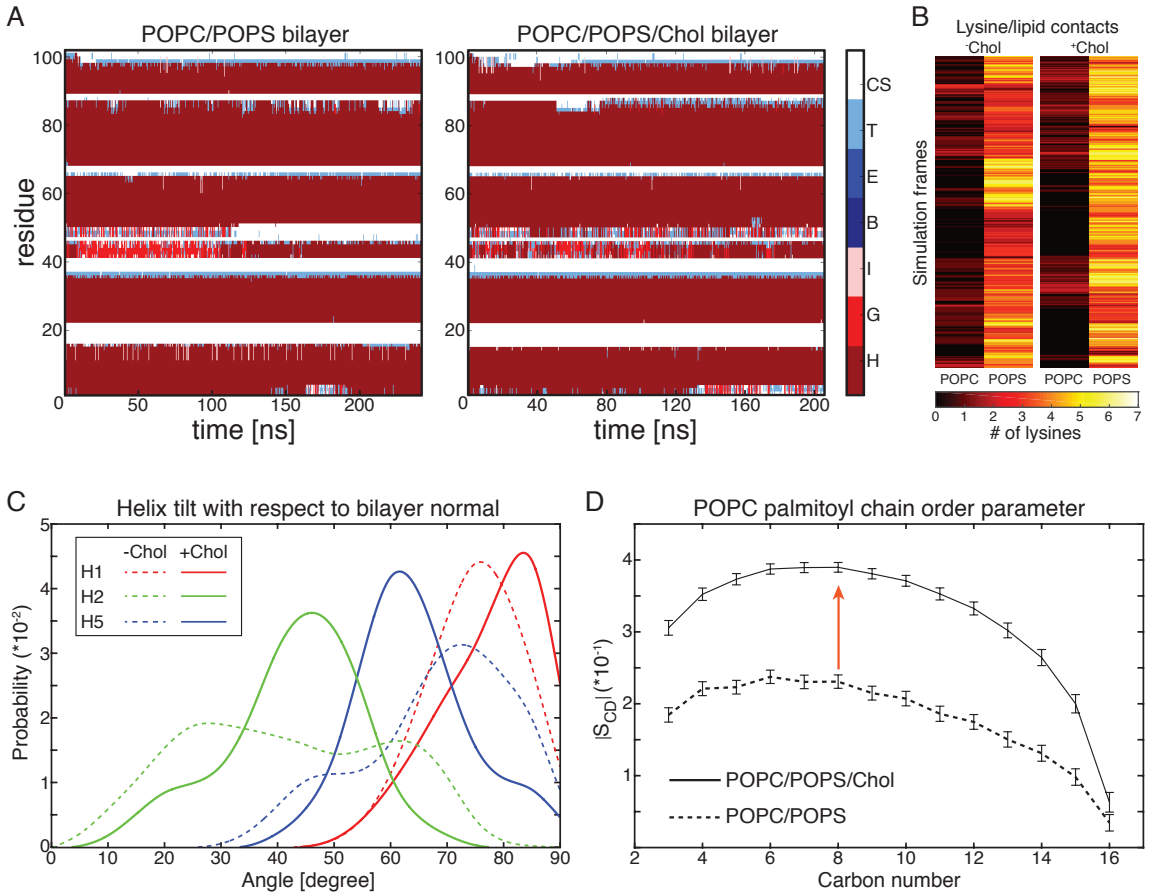

Figure S6. (A) Time evolution of MA's secondary structure in bilayers composed of POPC/POPS 70/30 mol% (left) or POPC/POPS/Chol 34/30/36 mol% (right). The secondary structure assignment of each residue was calculated with the DSSP software. They are as follows: H, alpha helix; G, 3-turn helix; I, pi helix; B, residue in isolated beta-bridge; E, extended beta sheet; T, hydrogen bonded turn; and CS, coil and bend. (B) Heat map of lysine-lipid contacts defined as having no more than 4 Å distance between the centers of mass of the NH_3 Lysine group and either the serine or phosphate groups on the lipid headgroups. Each row represents a single frame and the color denotes the number of instantaneous lysine-POPC or lysine-POPS contacts. Shown is data from the last 100 ns of three replica simulations for each system. (C) Tilt distributions of helices 1, 2 and 5 with respect to bilayer normal. The helices direction vectors are defined by the Cα atoms of residues E2 and K13, K23 and Q34, and E70 and A84, respectively, and the bilayer normal is the z-dimension of the simulation cell. The broad distributions in the absence of Chol indicate the dynamic nature of MA orientation with respect to the membrane (dashed lines, Movie S1A), which becomes more stable upon the addition of Chol (solid lines, Movie S1B). (D) The acyl chain order parameter of the palmitoyl chain of POPC, S_{CD}, increases by more than 65% in the presence of Chol as denoted by the red arrow.

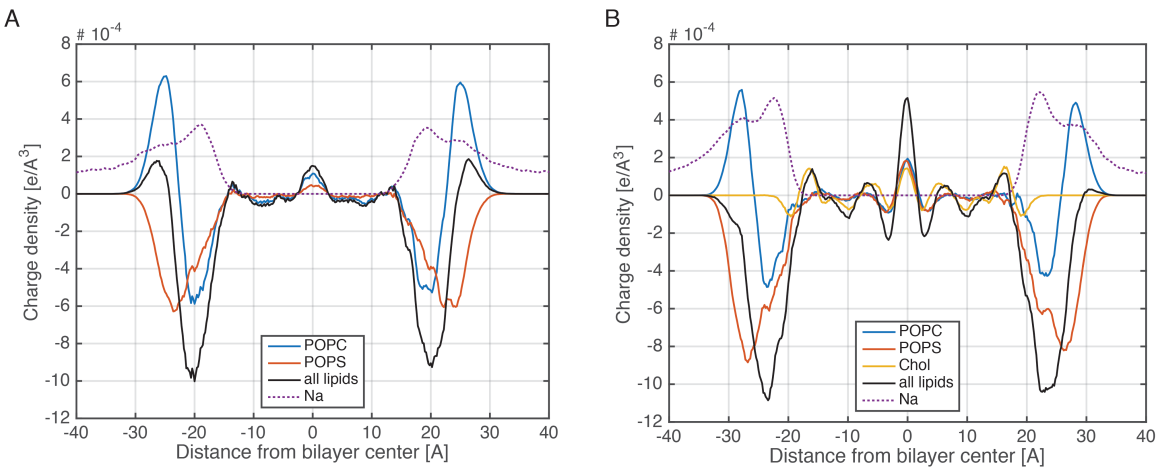

Figure S7. Charge density of POPC (blue), POPS (red), Chol (yellow), all lipids (black) and sodium (dashed purple line) from the two bilayer-only simulations of POPC/POPS (A) and POPC/POPS/Chol (B). The positively charged choline group on the POPC headgroups counteracts the negatively charged serine group on the POPS headgroups, and depending on the PC/PS ratio, the net effect is the accumulation of a higher (A) or lower (B) positive charge density on the bilayer surface (~ 27 and 30 Å, respectively). Note that since the charge density calculation is performed at high resolution (slabs in z with thickness of 0.2 Å) on individual atoms with assigned partial charges, relative imbalances in the distributions of hydrogen and carbon atoms result in a peak of positive charge density at the interface between the two leaflets. If the calculation is instead performed on the neutral chemical atomic groups (e.g. methylene, methyl), instead of the individual atoms, the positive density at the midplane would disappear.

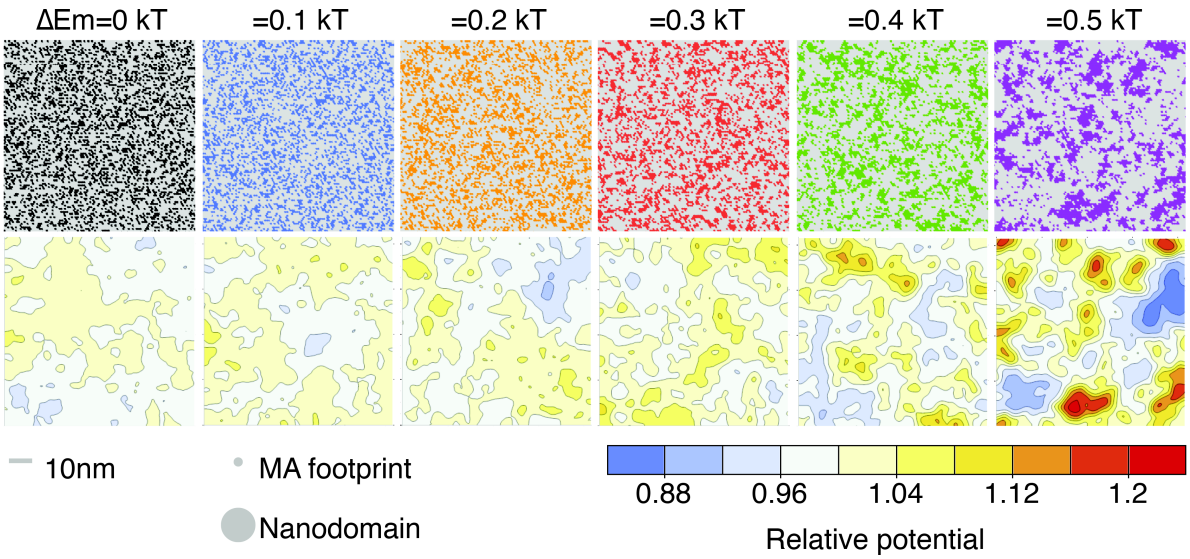

Figure S8. Non-ideal mixing does not induce large changes in membrane surface potential but may influence the spatial organization of bound protein. *Upper*, plots of the lateral distribution of a binary lipid mixture (70/30 mol% neutral/charged) obtained from Monte Carlo lattice simulations. Increasing the unfavorable pairwise interaction energy ΔEm results in larger clusters of the charged lipid (colored regions) within the neutral lipid matrix (gray regions). *Lower*, corresponding maps of the relative electrostatic surface potential calculated 3 Å above the bilayer surface and normalized to the potential of a uniformly mixed bilayer having the same average charge density. Also displayed for reference are the relative sizes of the MA protein (4 nm diameter) and a lipid nanodomain (15 nm diameter (62)). Scale bar 10 nm.

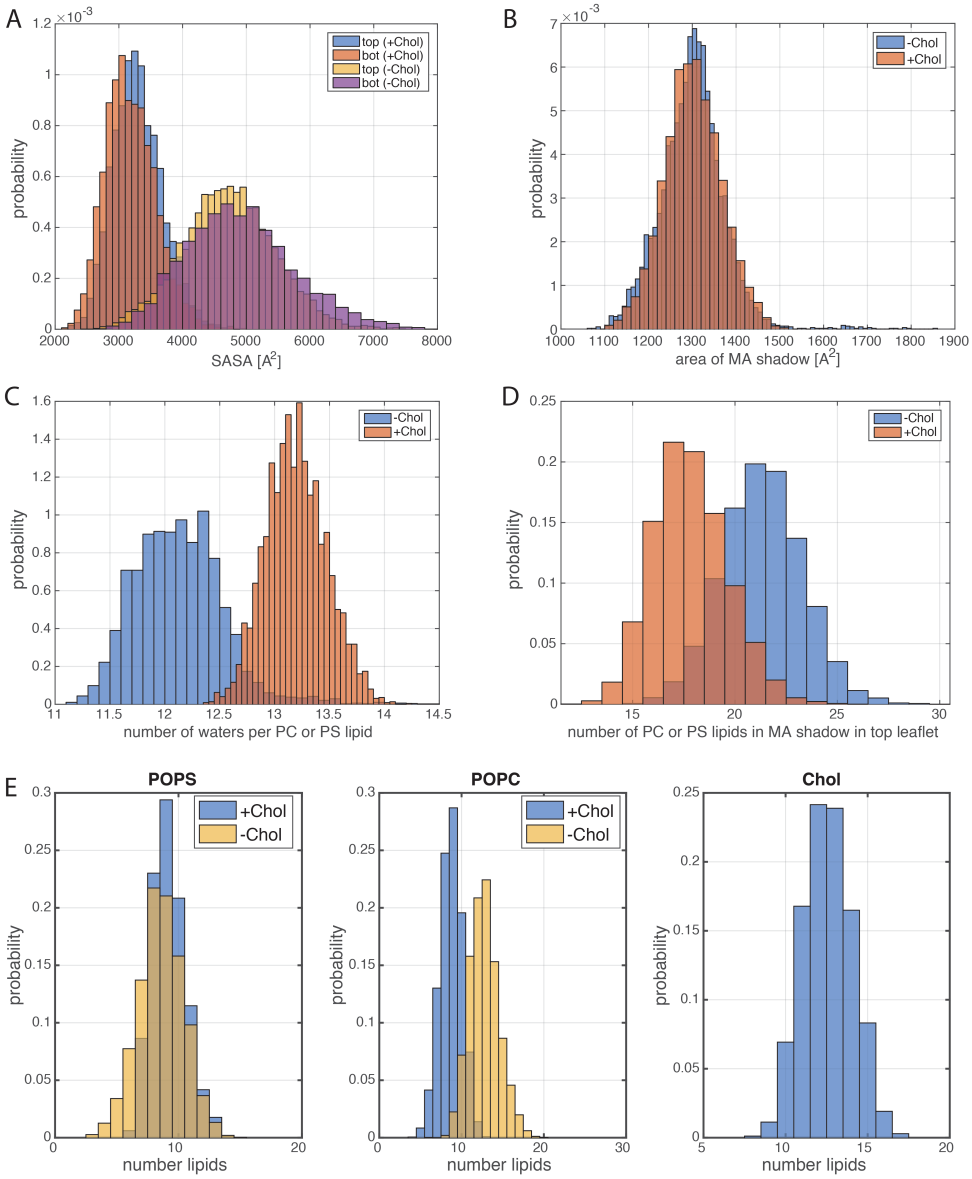

Figure S9. (A) Solvent exposed surface area (SASA) of each leaflet of the bilayer patch in the MA shadow. The MA shadow is defined as the set of atoms whose x and y coordinates are within 3 Å of the x and y coordinates of any protein atom. MA is interacting directly with the top leaflet. (B) Area of the MA shadow calculated as the area of the convex hull containing all 2D atomic coordinates in the MA shadow. (C) Distribution of the number of water molecules per POPC or POPS headgroup calculated from the bilayer-only simulations. A water per headgroup is defined as a water molecule within 3 Å of any lipid atom of a POPC or POPS lipid. (D) Distribution of the number of POPC or POPS lipids in the top leaflet from the MA-bilayer trajectories, whose phosphate atoms are within the MA shadow. (E) Number of different lipid counts in the top leaflet of the bilayer patch in the MA shadow in the two systems. POPS and POPC are represented with their phosphate atoms and Chol with its O3 atom. All distributions in this figure were constructed from calculations performed across all frames of the respective simulation trajectories.

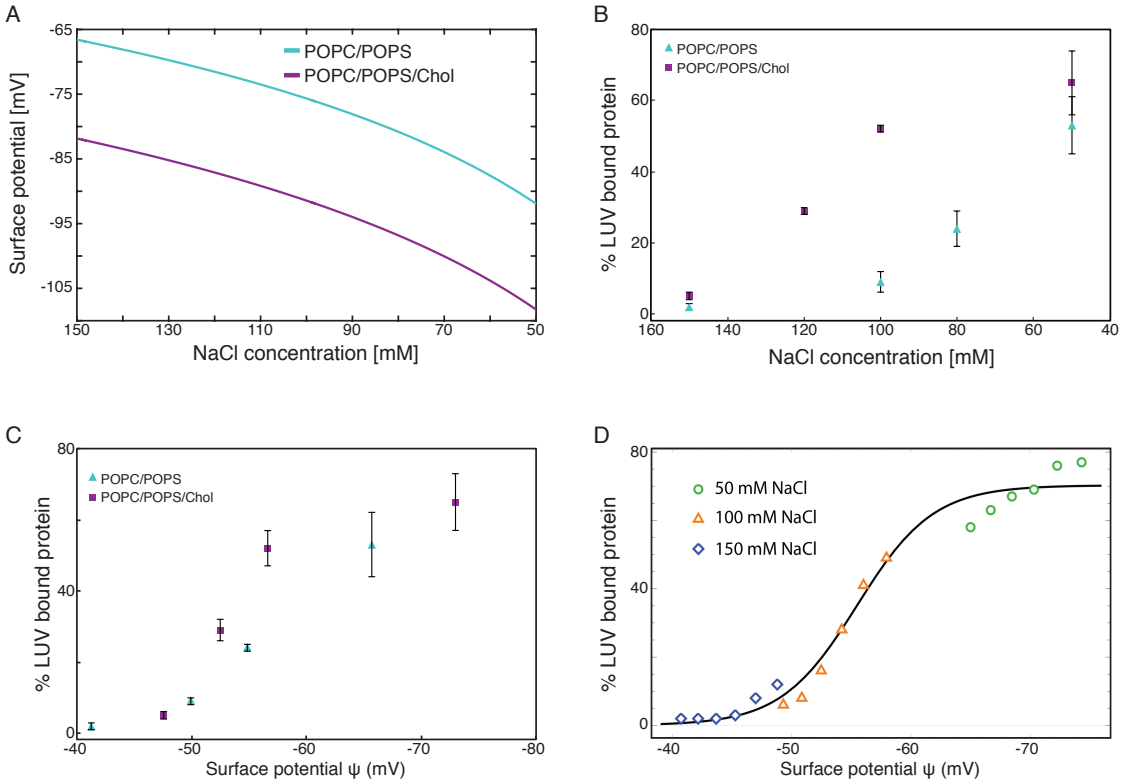

Figure S10. RSV MA membrane association as a function of NaCl concentration. (A) Example of a calculated membrane surface potential as a function of increasing NaCl concentrations for POPC/POPS (70/30 mol%) (light blue) and POPC/POPS/Chol (34/30/36 mol%) (purple) bilayers. (B) % of LUV-bound MA plotted against decreasing NaCl concentration. (C) Binding data from B plotted against calculated membrane surface potential. (D) Best sigmoidal fit to the MA binding data versus surface potential from Fig. 5F (see SM text for more details).

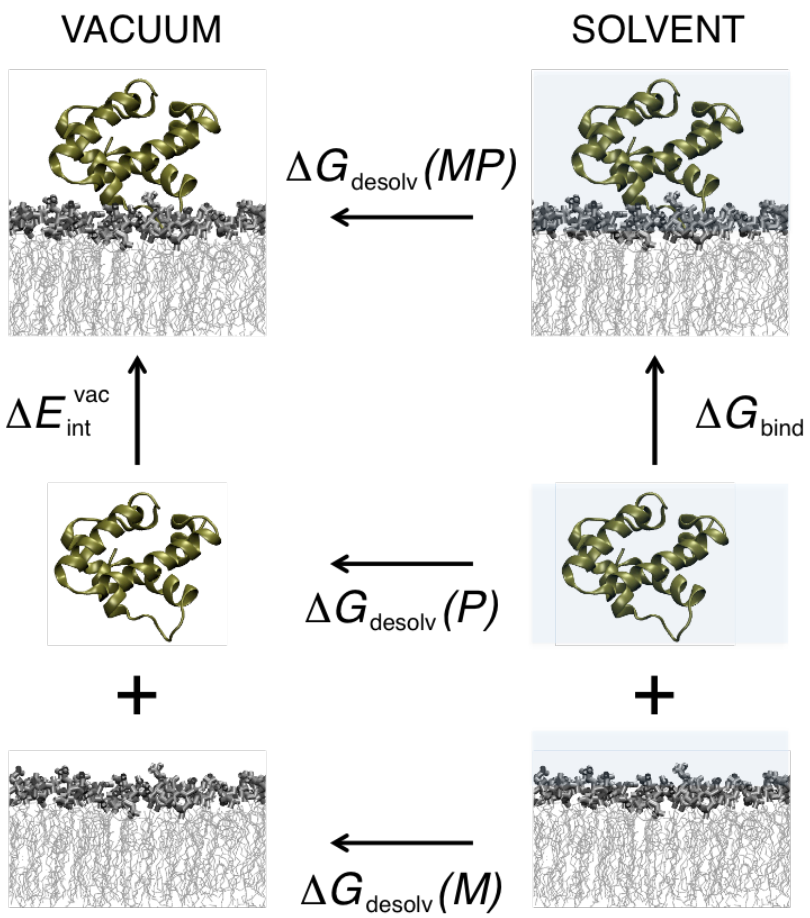

Figure S11: Thermodynamic cycle underlying the MM-GBSA method. For each frame of the trajectory, the binding free energy ΔG_{bind} is estimated from the vacuum interaction energy and the desolvation penalties for each binding partner and the complex.

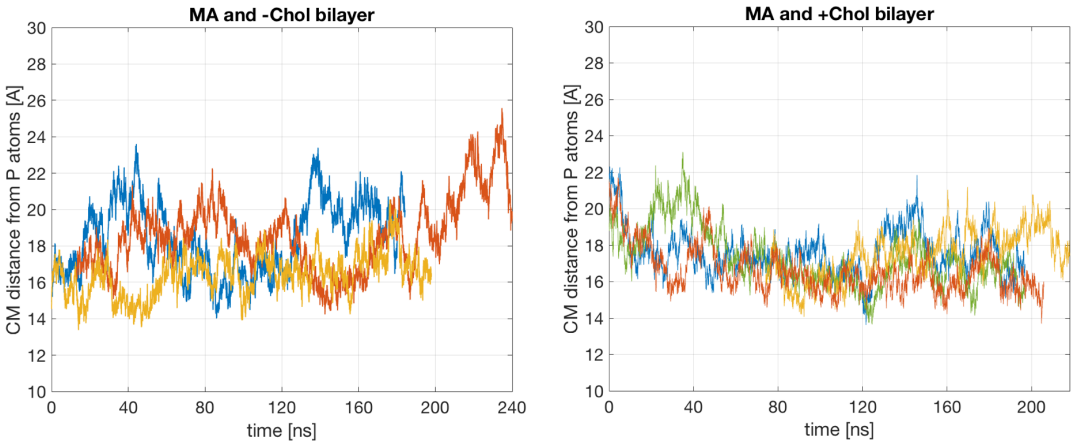

Figure S12. Fluctuations of the distance between MA and bilayer surface. Time evolution of the MA-bilayer distance defined as the distance in z between the center of mass of the protein and the center of mass of the phosphate atoms in the MA-proximal leaflet. Data is shown for the –Chol (left) and +Chol (right) systems. Colors denote different replica simulations (or in the latter also proteins interacting with different leaflets).

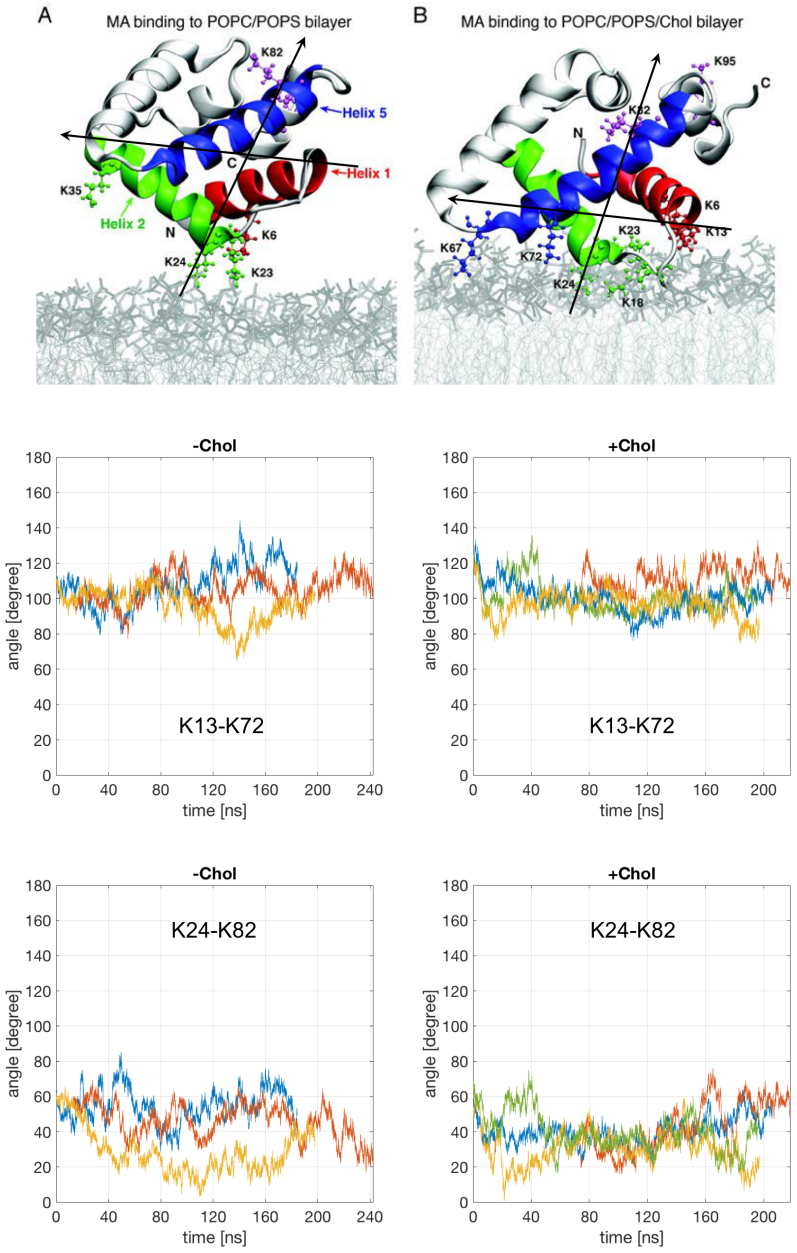

Figure S13. Orientation of MA with respect to the bilayer surface. *Top* Fig. 2AB with two axes (director vectors) shown, defining the orientation of MA relative to the bilayer surface: one, connecting the Cα atoms of K13 and K72, and another one connecting the Cα atoms of K24 and K82. Time evolution of the tilt angle of each axis (middle and bottom rows) with respect to the bilayer normal (z dimension of the simulation box) is shown for the –Chol (left) and +Chol (right) systems. Colors denote different replica simulations (or in the latter also proteins interacting with different leaflets). After about 80 ns of initial relaxation, the tilts fluctuate ~103º for K13-K72 and ~41º for K24-K82 with standard deviations within 10 degrees for individual replicas, and within 16 or 10 degrees across different replicas for –Chol and +Chol systems respectively.

Supporting Movies

SI Movie 1. Cholesterol enhances MA-membrane contacts. Simulation segment showing 40 ns of MA interaction with POPC/POPS .7/.3 (A) and POPC/POPS/Chol .34/.3/.36) (B) membranes. Lipid acyl chains and head groups are in light and dark gray, respectively. MA coloring is as described in main text. All lysine residues are shown in ball-and-stick representation and colored according to the helix to which they belong, except for K18 (gold) which is positioned in a loop, and K95 (silver) which is positioned in helix 6. Water and ion atoms are omitted for viewing clarity.

REFERENCES

1. Heberle, F. A., V. N. P. Anghel, and J. Katsaras. 2015. Scattering from phase-separated vesicles. I. An analytical form factor for multiple static domains. J Appl Crystallogr 48:1391-1404.
2. Heberle, F. A., J. Pan, R. F. Standaert, P. Drazba, N. Kucerka, and J. Katsaras. 2012. Model-based approaches for the determination of lipid bilayer structure from small-angle neutron and X-ray scattering data. European biophysics journal : EBJ 41:875-890.
3. Kingston, R. L., H. M. Baker, and E. N. Baker. 1994. Search designs for protein crystallization based on orthogonal arrays. Acta Crystallogr D Biol Crystallogr 50:429-440.
4. Otwinowski, Z., and W. Minor. 1997. Processing of X-ray diffraction data collected in oscillation mode. Method Enzymol 276:307-326.
5. Schneider, T. R., and G. M. Sheldrick. 2002. Substructure solution with SHELXD. Acta Crystallogr D Biol Crystallogr 58:1772-1779.
6. Bricogne, G., C. Vonrhein, C. Flensburg, M. Schiltz, and W. Paciorek. 2003. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. Acta Crystallogr D Biol Crystallogr 59:2023-2030.
7. Emsley, P., B. Lohkamp, W. G. Scott, and K. Cowtan. 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66:486-501.
8. Murshudov, G. N., P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, and A. A. Vagin. 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr 67:355-367.
9. McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read. 2007. Phaser crystallographic software. J Appl Crystallogr 40:658-674.
10. Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. 2005. Scalable molecular dynamics with NAMD. J Comput Chem 26:1781-1802.
11. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. J Mol Graph 14:33-38, 27-38.

12. Touw, W. G., C. Baakman, J. Black, T. A. te Beek, E. Krieger, R. P. Joosten, and G. Vriend. 2015. A series of PDB-related databanks for everyday needs. Nucleic Acids Res 43:D364-368.
13. Jo, S., T. Kim, V. G. Iyer, and W. Im. 2008. CHARMM-GUI: a web-based graphical user interface for CHARMM. J Comput Chem 29:1859-1865.
14. Doktorova, M., D. Harries, and G. Khelashvili. 2017. Determination of bending rigidity and tilt modulus of lipid membranes from real-space fluctuation analysis of molecular dynamics simulations. Physical chemistry chemical physics : PCCP.
15. Dalton, A., P. Murray, D. Murray, and V. Vogt. 2005. Biochemical characterization of rous sarcoma virus MA protein interaction with membranes. J Virol 79:6227-6238.
16. Andelman, D. 2006. Introduction to electrostatics in soft and biological matter. Scot Grad Ser:97-122.
17. Huang, J., and G. W. Feigenson. 1993. Monte Carlo simulation of lipid mixtures: finding phase separation. Biophys J 65:1788-1794.
18. Huang, J., J. E. Swanson, A. R. Dibble, A. K. Hinderliter, and G. W. Feigenson. 1993. Nonideal mixing of phosphatidylserine and phosphatidylcholine in the fluid lamellar phase. Biophys J 64:413-425.
19. Jorgensen, K., M. M. Sperotto, O. G. Mouritsen, J. H. Ipsen, and M. J. Zuckermann. 1993. Phase-Equilibria and Local-Structure in Binary Lipid Bilayers. Biochimica Et Biophysica Acta 1152:135-145.
20. Risbo, J., M. M. Sperotto, and O. G. Mouritsen. 1995. Theory of Phase-Equilibria and Critical Mixing Points in Binary Lipid Bilayers. J Chem Phys 103:3643-3656.
21. Sugar, I. P., T. E. Thompson, and R. L. Biltonen. 1999. Monte Carlo simulation of two-component bilayers: DMPC/DSPC mixtures. Biophysical Journal 76:2099-2110.
22. Frazier, M. L., J. R. Wright, A. Pokorny, and P. F. F. Almeida. 2007. Investigation of domain formation in sphingomyelin/cholesterol/POPC mixtures by fluorescence resonance energy transfer and Monte Carlo simulations. Biophysical Journal 92:2422-2433.
23. Honerkamp-Smith, A. R., P. Cicuta, M. D. Collins, S. L. Veatch, M. den Nijs, M. Schick, and S. L. Keller. 2008. Line tensions, correlation lengths, and critical exponents in lipid membranes near critical points. Biophysical Journal 95:236-246.
24. Heberle, F. A., and G. W. Feigenson. 2011. Phase separation in lipid membranes. Cold Spring Harb Perspect Biol 3.
25. Dai, J. A., M. Alwarawrah, M. R. Ali, G. W. Feigenson, and J. Y. Huang. 2011. Simulation of the l(o)-l(d) Phase Boundary in DSPC/DOPC/Cholesterol Ternary Mixtures Using Pairwise Interactions. Journal of Physical Chemistry B 115:1662-1671.
26. Lis, M., L. Pintal, J. Swiatek, and L. Cwiklik. 2012. GPU-Based Massive Parallel Kawasaki Kinetics in the Dynamic Monte Carlo Simulations of Lipid Nanodomains. Journal of Chemical Theory and Computation 8:4758-4765.
27. Meerschaert, R. L., and C. V. Kelly. 2015. Trace membrane additives affect lipid phases with distinct mechanisms: a modified Ising model. Eur Biophys J Biophy 44:227-233.

28. Srinivasan, J., M. W. Trevathan, P. Beroza, and D. A. Case. 1999. Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects. Theor Chem Acc 101:426-434.
29. Gohlke, H., C. Kiel, and D. A. Case. 2003. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RaIGDS complexes. Journal of Molecular Biology 330:891-913.
30. Zoete, V., M. Meuwly, and M. Karplus. 2005. Study of the insulin dimerization: binding free energy calculations and per-residue free energy decomposition. Proteins 61:79-93.
31. Srinivasan, J., T. E. Cheatham, P. Cieplak, P. A. Kollman, and D. A. Case. 1998. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate - DNA helices. Journal of the American Chemical Society 120:9401-9409.
32. Best, R. B., X. Zhu, J. Shim, P. E. Lopes, J. Mittal, M. Feig, and A. D. Mackerell, Jr. 2012. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. J Chem Theory Comput 8:3257-3273.
33. Still, W. C., A. Tempczyk, R. C. Hawley, and T. Hendrickson. 1990. Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics. Journal of the American Chemical Society 112:6127-6129.
34. Onufriev, A., D. Bashford, and D. A. Case. 2004. Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins-Structure Function and Bioinformatics 55:383-394.
35. Lee, M. S., M. Feig, F. R. Salsbury, and C. L. Brooks. 2003. New analytic approximation to the standard molecular volume definition and its application to generalized born calculations. Journal of Computational Chemistry 24:1348-1356.
36. Lee, M. S., F. R. Salsbury, and C. L. Brooks. 2002. Novel generalized Born methods. J Chem Phys 116:10606-10614.
37. Brooks, B. R., C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. 2009. CHARMM: The Biomolecular Simulation Program. Journal of Computational Chemistry 30:1545-1614.
38. Zoete, V., and O. Michielin. 2007. Comparison between computational alanine scanning and per-residue binding free energy decomposition for protein-protein association using MM-GBSA: Application to the TCR-p-MHC complex. Proteins-Structure Function and Bioinformatics 67:1026-1047.
39. Zoete, V., M. B. Irving, and O. Michielin. 2010. MM-GBSA binding free energy decomposition and T cell receptor engineering. J Mol Recognit 23:142-152.
40. Dick, R. A., S. L. Goh, G. W. Feigenson, and V. M. Vogt. 2012. HIV-1 Gag protein can sense the cholesterol and acyl chain environment in model membranes. Proc Natl Acad Sci U S A 109:18761-18766.
41. Dick, R. A., M. Barros, D. Jin, M. Losche, and V. M. Vogt. 2016. Membrane Binding of the Rous Sarcoma Virus Gag Protein Is Cooperative and Dependent on the Spacer Peptide Assembly Domain. J Virol 90:2473-2485.

42. Gaffney, B. J. 1976. The chemistry of spin labels. Spin Labeling: Theory and Applications L. J. Berliner, ed.:182-238.
43. Khelashvili, G., H. Weinstein, and D. Harries. 2008. Protein diffusion on charged membranes: A dynamic mean-field model describes time evolution and lipid reorganization. Biophysical Journal 94:2580-2597.
44. Khelashvili, G., and D. Harries. 2010. Modeling Signaling Processes across Cellular Membranes Using a Mesoscopic Approach. Ann Rep Comp Chem 6:237-261.
45. Khelashvili, G., M. Doktorova, M. A. Sahai, N. Johner, L. Shi, and H. Weinstein. 2015. Computational modeling of the N-terminus of the human dopamine transporter and its interaction with PIP2 -containing membranes. Proteins 83:952-969.
46. Khelashvili, G., A. Galli, and H. Weinstein. 2012. Phosphatidylinositol 4,5-biphosphate (PIP(2)) lipids regulate the phosphorylation of syntaxin N-terminus by modulating both its position and local structure. Biochemistry 51:7685-7698.
47. Khelashvili, G., D. Harries, and H. Weinstein. 2009. Modeling membrane deformations and lipid demixing upon protein-membrane interaction: the BAR dimer adsorption. Biophys J 97:1626-1635.
48. Andelman, D. 1995. Electrostatic properties of membranes: The Poisson-Boltzmann theory. . Lipowsky, R., Sackmann, E., editors. Amsterdam: Elsevier Science B.V.
49. Chaikin, P., and T. Lubensky. 2000. Principles of Condensed Matter Physics. Cambrige University Press.
50. Baker, N. A., D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A 98:10037-10041.
51. Davis, I. W., A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall, 3rd, J. Snoeyink, J. S. Richardson, and D. C. Richardson. 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res 35:W375-383.
52. Perkins, S. J. 1986. Protein volumes and hydration effects. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. Eur J Biochem 157:169-180.
53. Efimova, Y. M., A. A. van Well, U. Hanefeld, B. Wierczinski, and W. G. Bouwman. 2004. On the neutron scattering length density of proteins in H(2)O/D(2)O. Physica B 350:E877-E880.
54. Kucerka, N., M. P. Nieh, and J. Katsaras. 2011. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. Biochim Biophys Acta 1808:2761-2771.
55. Pan, J., X. Cheng, L. Monticelli, F. A. Heberle, N. Kucerka, D. P. Tieleman, and J. Katsaras. 2014. The molecular structure of a phosphatidylserine bilayer determined by scattering and molecular dynamics simulations. Soft Matter 10:3716-3725.
56. Lorizate, M., T. Sachsenheimer, B. Glass, A. Habermann, M. J. Gerl, H. G. Krausslich, and B. Brugger. 2013. Comparative lipidomics analysis of HIV-1

particles and their producer cell membrane in different cell lines. Cell Microbiol 15:292-304.
57. van Meer, G., D. R. Voelker, and G. W. Feigenson. 2008. Membrane lipids: where they are and how they behave. Nat Rev Mol Cell Biol 9:112-124.
58. Murate, M., and T. Kobayashi. 2016. Revisiting transbilayer distribution of lipids in the plasma membrane. Chem Phys Lipids 194:58-71.
59. Edholm, O., and J. F. Nagle. 2005. Areas of molecules in membranes consisting of mixtures. Biophys J 89:1827-1832.
60. Kucerka, N., B. van Oosten, J. Pan, F. A. Heberle, T. A. Harroun, and J. Katsaras. 2015. Molecular structures of fluid phosphatidylethanolamine bilayers obtained from simulation-to-experiment comparisons and experimental scattering density profiles. J Phys Chem B 119:1947-1956.
61. Eldho, N. V., S. E. Feller, S. Tristram-Nagle, I. V. Polozov, and K. Gawrisch. 2003. Polyunsaturated docosahexaenoic vs docosapentaenoic acid-differences in lipid matrix properties from the loss of one double bond. J Am Chem Soc 125:6409-6421.
62. Heberle, F. A., R. S. Petruzielo, J. Pan, P. Drazba, N. Kucerka, R. F. Standaert, G. W. Feigenson, and J. Katsaras. 2013. Bilayer thickness mismatch controls domain size in model membranes. J Am Chem Soc 135:6853-6859.