

**Biophysical Journal, Volume 122**

**Supplemental information**

**Protein overexpression can induce the elongation of cell membrane  
nanodomains**

**Julie Cornet, Pascal Preira, Laurence Salomé, Frédéric Daumas, Bernard Lagane, Nicolas  
Destainville, Manoel Manghi, and Fabrice Dumas**

# Supplementary Material

## Protein overexpression can induce an elongation of cell membrane domains

Julie Cornet<sup>†</sup>, Pascal Preira<sup>\*</sup>, Laurence Salomé<sup>\*</sup>, Frédéric Daumas<sup>\*</sup>,  
 Bernard Lagane<sup>‡</sup>, Nicolas Destainville<sup>†</sup>, Manoel Manghi<sup>†</sup>, Fabrice Dumas<sup>\*</sup>  
<sup>†</sup> *Laboratoire de Physique Théorique, Université Toulouse III, CNRS, 31062 Toulouse, France*  
<sup>\*</sup> *Institut de Pharmacologie et de Biologie Structurale,*  
*Université Toulouse III, CNRS, 31062 Toulouse, France*  
<sup>‡</sup> *Infinity, Université de Toulouse, CNRS, Inserm, Toulouse, France*  
 (Dated: December 2, 2022)

### I. CELL LINES

To study HIV infection processes, Johnston et al. [1] have developed an inducible cell line in which two membrane proteins (CD4 and CCR5) can be simultaneously and independently regulated with a large range of surface expression with the help of variable concentrations of minocycline and ponasterone, respectively (Figure S1). We have stably transfected this cell line to express the third membrane protein (CXCR4) that can be involved in HIV infection. We have established two stable cell lines, one expressing a low number ( $15000 \pm 2000$  per cell) of CXCR4 and one expressing a high number ( $120000 \pm 7000$ ) of this protein (Figure S2). The CXCR4 copy number of each line was determined in triplicate using Quantibrite system described in the Mat. and Meth. section of the main text. The average of the minimum and maximum values of CD4, CCR5 and CXCR4 copy numbers are reported in Table 1 of the main text.

Figure S3 shows that the cell lines that we have established can be infected by X4- and R5-viruses (NL 4-3 and Bx08 respectively).

Based on these figures, we can also calculate a rough estimate of the protein density (or surface ratio) occupied by the proteins on a typical cell membrane in the case of overexpression. As shown in Table S1, the membrane surface ratio occupied by proteins is less than 1%. However, one should consider that a real cell is a far more complex system. The proteins do not cluster alone but also recruit certain lipids and partner proteins so that the surface ratio of the concerned domains is probably significantly higher than the value obtained by considering the sole tracked proteins.

### II. TRANSIENTLY CONFINED TRAJECTORIES

In Figure S4 is shown the distribution of the ratio between the trajectory time and the typical diffusion time to explore the domain of size  $L$  measured in the experiments,  $L^2/(4D)$  (where  $D$  is the measured lateral diffusion coefficient). For the trajectories studied, this ratio is always larger than 4, thus confirming that the full domain as indeed been explored. This ensures that elongated trajectories account for merely elongated domains and are

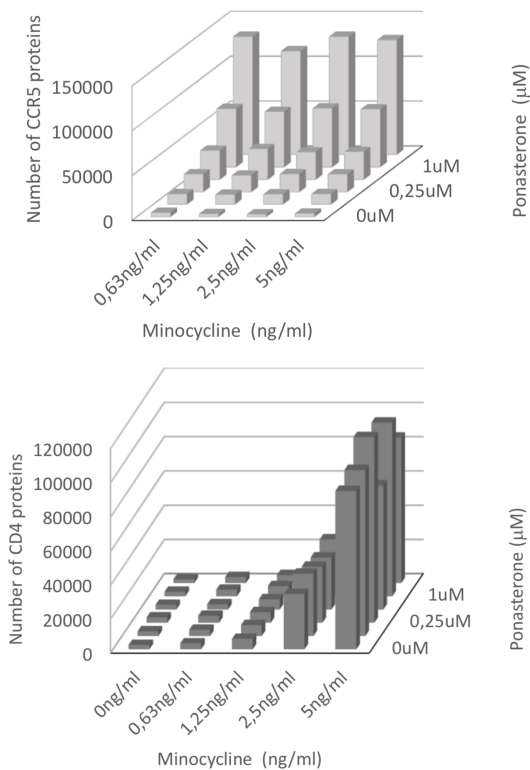


FIG. S1: CD4 and CCR5 can be independently and simultaneously regulated. Top: ponasterone induces the expression of CD4 proteins on the same manner whatever the minocycline concentration. Bottom: Minocycline induces the expression of CCR5 proteins on the same manner whatever the ponasterone concentration.

not due to a bias arising from a too short acquisition time leading to the partial exploration of a domain.

### III. QUANTIFICATION OF MEMBRANE DOMAIN SHAPE

In SPT, some diffusion modes can give rise to *apparently* elongated domains. Notably, directed diffusion trajectories might be confused with confined trajectories in an elongated domain. We put directed diffusion trajectories aside through the characterization of their MSD [3].

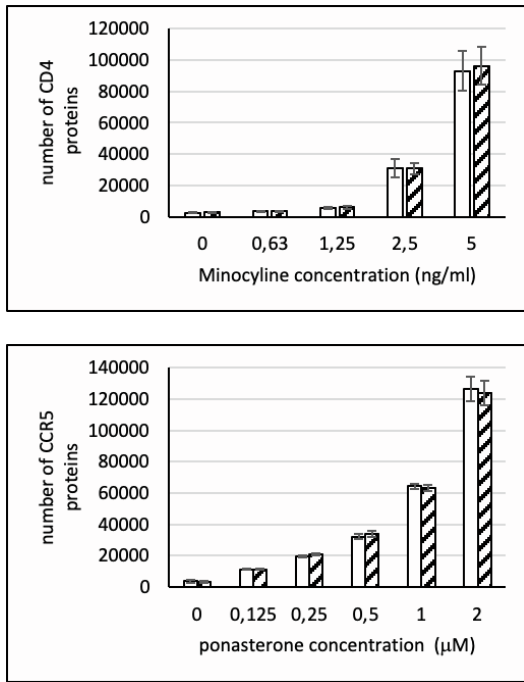


FIG. S2: The expression of CD4 and CCR5 can be controlled in the inducible 293-Affinofile cell lines that stably express high or low level of CXCR4 (dashed and white bars respectively). Minocycline (top) and ponasterone (bottom) induced CD4 and CCR5 expression, respectively, in a concentration dependent manner without affecting one another.

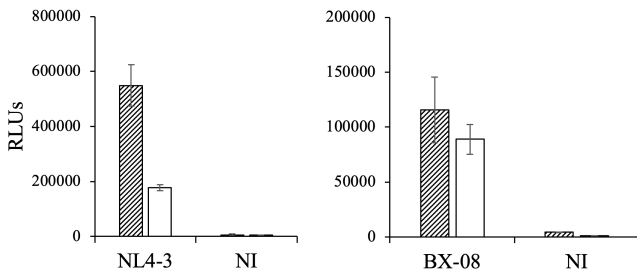


FIG. S3: Affinofile stable cells lines expressing high or low amounts of CXCR4 (dashed bars and white bars respectively) were induced in order to express CD4 and CCR5. They are successfully infected by R5 tropic (Bx08) or X4 tropic (NL4-3) viruses. Experiments have been carried out according to Ref. [2] as described in Section VI of SM. Results represent the luciferase activity in the cell lysates, expressed as relative light units (RLUs). Uninfected cells (NI) represent negative controls.

Note also that we did not observe any walking confined diffusion mode trajectories, where the tracked particle is confined in a domain that itself diffuses, which can also be confused with a trajectory in an elongated domain. Then we isolated the confined part of the trajectories thanks to the confinement index developed in Ref. [4] (see this article for more details about this detection method). The data collected give the  $(x, y)$  coordinates of the con-

	% of cell surface occupied by proteins	
	low expression	high expression
CCR5	0.0072 %	0.249 %
CXCR4	0.0356 %	0.285 %
CD4	0.0012 %	0.0509 %
Total	0.044 %	0.585 %

TABLE S1: Percentage of the cell surface occupied by each protein expressed at high or low concentration. Based on the structures of CCR5 (PDB ID: 4MBS), CXCR4 (PDB ID: 3OE9) and CD4 (PDB ID: 2KLU), we have estimated the surface occupied by each of these proteins in the plane of the cell membrane to be 9, 10.75 and 2.3 nm<sup>2</sup> respectively. We have considered that HEK cells have a radius of 6  $\mu$ m and taken into account the number of copies of each protein presented in Table 1 of the main text to calculate the percentage of cell surface occupied by each protein at low or high expression level.

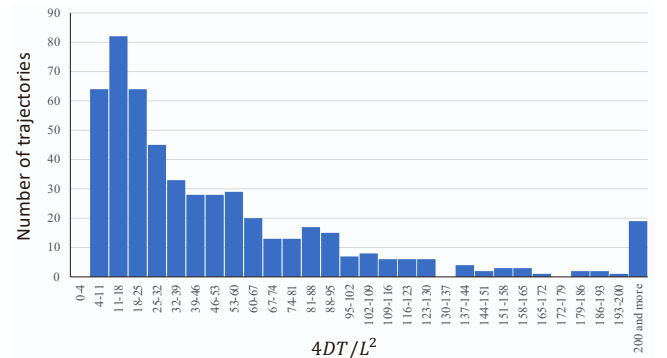


FIG. S4: Distribution of the ratio  $4DT/L^2$  where  $T$  is the confined trajectory duration of HIV receptors,  $D$  their measured lateral diffusion coefficient, and  $L$  the average size of the confinement zone.

fined parts of the trajectories of the tracked proteins every 40 ms. The localization precision has been measured to be of  $\leq 7$  nm in both  $x$  and  $y$  directions by recording and analyzing 80 s videos of quantum dots immobilized in 15 % polyacrylamide gel [3, 5] (see Figure S5). Examples of trajectories are displayed in Figure 2 in the main text. Such trajectories clearly give us information about domain sizes and shapes, thanks to the unrivalled accuracy of SPT. To quantify size and shape, we determine the main axes of the observed domains, from the eigenvalues of their covariance matrix:

$$V = \begin{pmatrix} \sigma_x^2 & \text{cov}(x, y) \\ \text{cov}(x, y) & \sigma_y^2 \end{pmatrix} \quad (\text{S1})$$

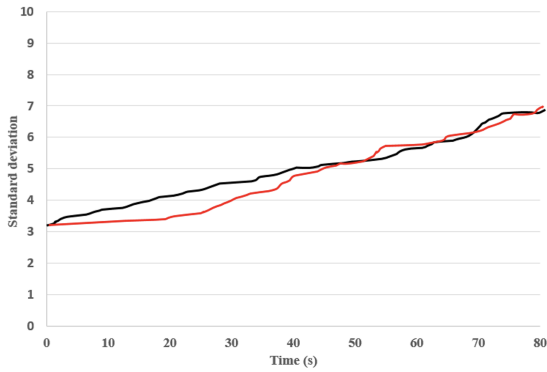


FIG. S5: Evolution of standard deviation of immobilized colloids as a function of time. To estimate the accuracy of our setup with respect to the determination of the particle position, we recorded a sequence of 80 s of several colloids deposited on a coverslip and immobilized in a 15% polyacrylamide gel. The standard deviations of the  $\approx 30$  colloids coordinates were found to increase with time indicating a limitation due to a thermomechanical instability of the microscope. After 80 s of measurements, the standard deviations  $\sigma_x$  (red) and  $\sigma_y$  (black) were found to be equal to 7 nm.

where  $\sigma_{x,y}^2$  are the variances and

$$\text{cov}(x, y) = \frac{1}{N} \sum_{i=1}^N x_i y_i - \left( \frac{1}{N} \sum_{i=1}^N x_i \right) \left( \frac{1}{N} \sum_{i=1}^N y_i \right), \quad (\text{S2})$$

is the covariance for each confined trajectory. Let  $\lambda_1 > \lambda_2$  be the real eigenvalues of the symmetric matrix  $V$ . We approximate the domain area by

$$\mathcal{A} \simeq 4\pi \sqrt{\lambda_1 \lambda_2} \quad (\text{S3})$$

so that in the case where the domain is a disc in which trajectory points are uniformly distributed, one recovers the area of the disc. This also amounts to set that the semi-major and minor axes are defined by  $r_1 = 2\sqrt{\lambda_1}$  and  $r_2 = 2\sqrt{\lambda_2}$  when the domain is elliptic. Experimental values are given in Table S2.

In this Table we use the following notations for the experimental conditions: when a protein is expressed at a low expression level, its name is written in lowercase letters, when a protein is overexpressed, its name is written in uppercase letters. The protein that has been tracked is underlined.

The aspect ratio of the domain is then naturally defined as

$$\text{AR} = \sqrt{\frac{\lambda_1}{\lambda_2}} \quad (\text{S4})$$

We set the threshold value above which domains are considered elongated to  $\text{AR}_0 = 2$ . It comes from the numerical distributions showing a “hinge point” for this value when concentration is increased and domains start to become more elongated, as observed in Fig. 5 in the main

	semi-minor axis (nm)	semi-major axis (nm)
cd4 ccr5 cxcr4	$120 \pm 14$	$204 \pm 24$
CD4 CCR5 CXCR4	$156 \pm 10$	$280 \pm 18$
cd4 <u>ccr5</u> cxcr4	$127 \pm 13$	$218 \pm 26$
CD4 <u>CCR5</u> CXCR4	$150 \pm 8$	$288 \pm 14$
cd4 ccr5 <u>cxcr4</u>	$130 \pm 8$	$236 \pm 14$
CD4 CCR5 <u>CXCR4</u>	$146 \pm 12$	$312 \pm 32$

TABLE S2: Average semi-major ( $r_1$ ) and minor ( $r_2$ ) axes in various experimental conditions as inferred from the eigenvalues of the covariance matrix (see main text for the convention on experimental condition notations). Error bars are standard errors on means (s.e.m.). See text for notations.

text and Fig. S10 below.

In Figure S6 we compare the aspect ratio distributions of the cumulated data of all the proteins having low expression level on the one hand, and the cumulated data of all the proteins overexpressed on the other hand. We can already notice that the domains in overexpressed conditions show an aspect ratio distribution shifted towards higher values.

However, we have also noticed that for some domains, the aspect ratio was not a sufficient parameter to characterize their shape. Indeed some ring-shaped, spiral-shaped, or horseshoe-shaped domains lead to  $\text{AR} \simeq 1$ , whereas they are in fact coiled elongated domains (Figure S7). We cannot numerically discriminate these domains through our sole AR analysis. In the experimental trajectories however we noticed that these domains were rather numerous ( $> 10\%$ ), in majority with a horseshoe shape, which motivated us to implement a second verification, carried out on the domains that are classified by their aspect ratio as roundish ( $\text{AR} < 2$ ): we calculate the domain barycenter and divide the domain into regular angular sectors emerging from this barycenter. We then compute the number of points per sector, as illustrated in Figure S7 (top).

We are interested in characterizing the fact that some sectors might contain no or a few points, much less than in the other sectors, corresponding to a hole in the angular distributions. We use the criterion described in Figure S7 (bottom): after computing the mean number of points per sector, we set a threshold  $t$  below which the domain is considered to have a hole if at least one of its sectors contains less points than this value  $t$ , and then considered as coiled elongated. We tried different threshold values and chose to focus on two of them,  $t = 25\%$  and  $30\%$  of the average number of points per sector. The distributions given in the figures are computed with  $t = 25\%$ . We checked visually that it discriminated most of the hidden elongated domains and included very few real “roundish” domains in this category. The modifications generated by this correction were on the order of few percents on the percentages of elongated domains for

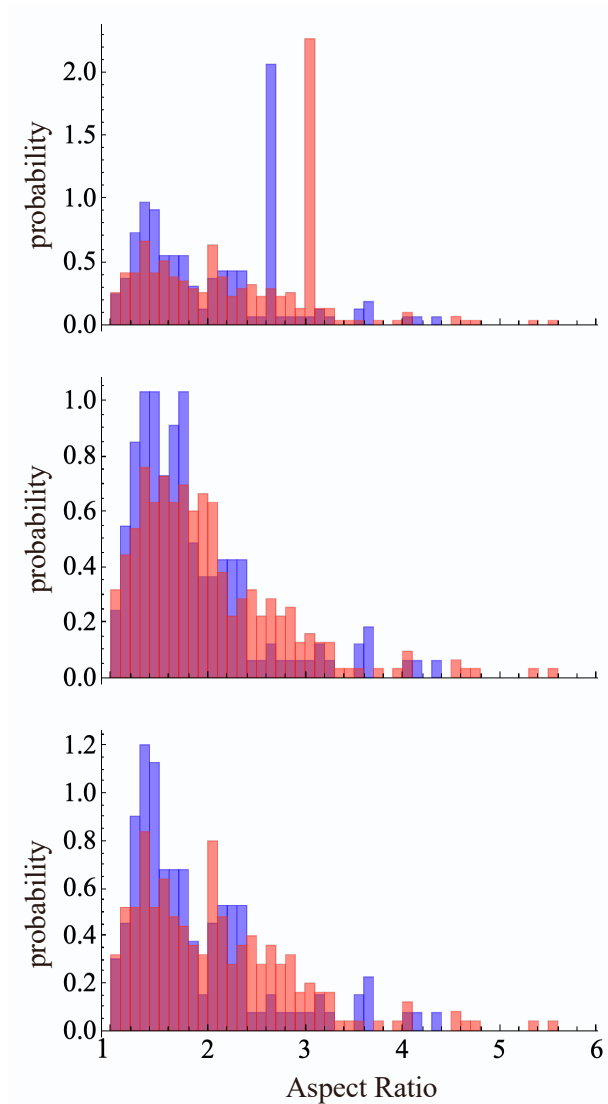


FIG. S6: Experimental aspect ratio cumulated probability distributions for all the conditions where proteins have low expression level and all the conditions where all proteins CD4, CCR5 and CXCR4 are over-expressed. Top: same as Figure 6 in the main text, except that the two peaks corresponding to the corrected AR of curled or coiled elongated nanodomains are represented here with their full heights. Middle: same distribution except that curled or coiled elongated nanodomains appear here with their raw AR, prior to correction. Bottom: same distribution except that curled or coiled elongated nanodomains do not appear at all. Only nanodomains not detected as curled or coiled have been pooled in these histograms. The  $p$ -values (KS statistical test, see below) are respectively  $< 10^{-7}$ , 0.00145 and 0.00143.

each experimental condition.

Since the sole AR is not sufficient for characterizing these domain shapes but we still wanted to take these “hidden” elongated domains into account, we assigned to them the average aspect ratio value of the initially detected elongated domains (verifying  $AR > 2$ ). Figure 6

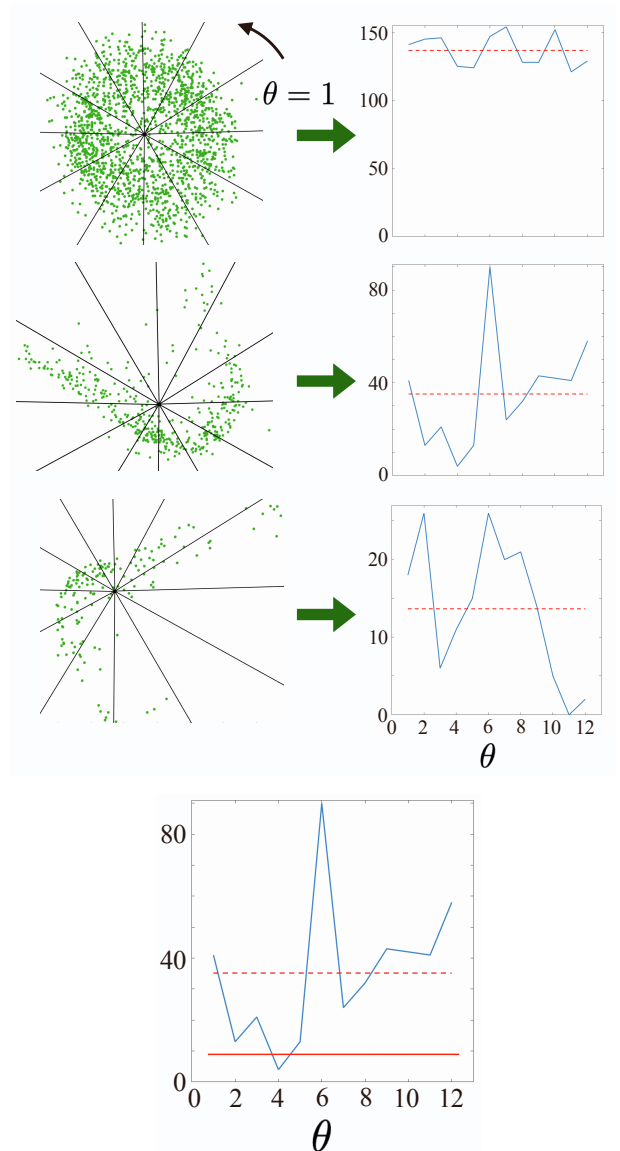


FIG. S7: Illustrative examples of the method used to identify the “false roundish” domains. Top: Analysis of the number of points per angular sector. The plane is divided into 12 regular angular sectors originating from the point-set barycenter, indexed by  $\theta = 1$  to 12 (left); the number of points per sector is computed and plotted as a histogram (right). The average value of the number of points per sector is depicted with a red dashed line. Bottom: Domains for which at least one sector contains less than  $t = 25\%$  of the average number of points per sector are considered to have a hole and then to be coiled elongated. The average value is still represented with a red dashed line, whereas the threshold  $t$  is depicted with a red full line.

in the main text shows the modification of the AR distributions shown in Figure S6, after applying the correction described above. The correction is captured through two new peaks, around 2.6 in the distribution associated with a low level of expression, and 3.1 in the distribution associated with protein overexpression.

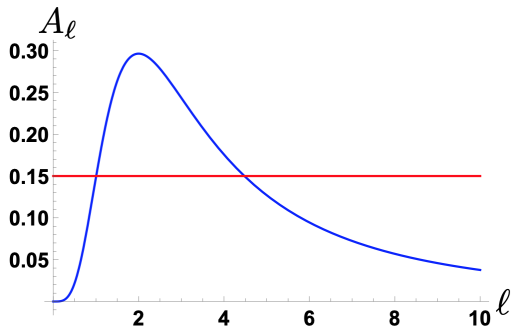


FIG. S8: Determination of the stability of elongated domains. In blue, plot of  $A_\ell$  in function of the repulsion range  $\ell$ . The maximum coordinates are ( $\ell^* = 2, A^* = \frac{8}{27}$ ). The red horizontal line represents  $\frac{3}{2\pi\epsilon}$ . Ellipses are more stable than discs when the red line is below the blue curve.

We tested the two different threshold values,  $t = 25\%$  and  $30\%$ , to detect the elongated domains, and checked that our conclusions were robust to modifications of the threshold value. We also performed the same analysis with an aspect ratio threshold  $AR_0$  of 1.8 instead of 2 and also checked that the same conclusions hold with marginal influence on the trajectory classification.

#### IV. STABILITY OF ELLIPSES WITH RESPECT TO DISCS

We discuss in the main text that ellipses are more stable than discs when

$$A_\ell > \frac{3}{2\pi\epsilon} \quad (\text{S5})$$

As illustrated in Figure S8,  $A_\ell$  has a maximum  $A^* = \frac{8}{27}$  at  $\ell^* = 2$ . Hence there exists an interval  $[\ell_i^*, \ell_s^*]$  of values of  $\ell$  where ellipses are stable if and only if  $A^* > 3/(2\pi\epsilon)$ .

#### V. SUPPLEMENTARY INFORMATION ON NUMERICAL SIMULATIONS

##### A. Triangulated-vesicle model

We present briefly the numerical model developed and validated in Ref. [6]. It is a lattice-gas model, equivalent to the 2D Ising model with Ising interaction parameter  $J_I$ , describing the binary mixture of two species A and B. The species A can be considered as forming a separated phase containing some membrane proteins and/or particular lipids. The lattice-gas model is coupled to a discretized Helfrich model [7] accounting for the membrane elasticity, with bending modulus  $\kappa$  and surface tension  $\sigma$ . The membrane has the topology of a spherical vesicle, and is modeled as a triangulated surface. A com-

position variable, equal to A or B, is attached to each of the  $N = 2562$  vertices of the triangulation. The Ising model is defined on this 2562-vertex lattice, which is a triangular lattice except for 12 vertices having 5 neighbors only. The local spontaneous curvature  $C$  depends on the composition. The species A has a spontaneous curvature  $C_A$  different from the majority phase (species B with  $C_B = 2/R < C_A$  where  $R$  is the average radius of the vesicle). We introduce the dimensionless parameters  $c_1 = R(C_A - C_B) = RC_A - 2$ ,  $\tilde{\sigma} = R^2\sigma/(k_B T)$  ( $\sigma$  is the surface tension) and  $\tilde{J}_I = J_I/(k_B T)$ . Together with the reduced bending rigidity  $\kappa/(k_B T)$  and the fraction  $\bar{\phi}$  of species A vertices, they fully characterize the system.

The system undergoes two kinds of Monte Carlo (MC) moves: small radial displacements of the vertices and spin-exchange Kawasaki moves as far as composition is concerned. A total of more than  $10^{10}$  MC steps is performed for each condition. The equilibration times have been measured in Ref. [6], and estimated to be shorter than  $10^8$  MC moves for the regime of parameters studied in the present work. In general, simulations are run after starting from a perfect sphere on which species A and B are randomly distributed. To ascertain that the system under study are in equilibrium, we have run one simulation after starting from a single macro-cluster (Figure S9). We observe that the system has rapidly forgotten its initial state. The unstable macro-cluster splits into smaller fragments. Accordingly, the measured distributions of cluster sizes are essentially insensitive to the initial configuration. This demonstrates that small clusters are not just the fruit of slow coarsening when starting from a random configuration.

##### B. Interacting-protein model

We also performed simulations on planar membranes with a different model of interacting “proteins” [8]. The mid-range repulsion between the proteins in this case is not explicitly promoted by the difference of spontaneous curvature as in the previous model, but comes from an effective pairwise potential, as if the membrane degrees of freedom were integrated out. Indeed, as explained in the main text, those two approaches are equivalent from a statistical physics perspective [9]. Both mechanisms make large domain unstable because the energy of a domain grows faster than its area, and promote mesophases. This also enables us to extend our analysis and results beyond the case of effective repulsion mediated by composition-shape coupling, e.g. interactions of electrostatic origin in soft matter physics [10, 11]. The pairwise potential reads [8, 10]:

$$U(r) = -\epsilon_a e^{-\gamma_a r} + \epsilon_r e^{-\gamma_r r} \quad (\text{S6})$$

$r$  being the distance between the molecules. The parameters used for the simulations are  $\epsilon_a = 21.3k_B T$ ,  $\epsilon_r = 0.31k_B T$ ,  $\gamma_a = 2$  and  $\gamma_r = 0.25$  in inverse units of



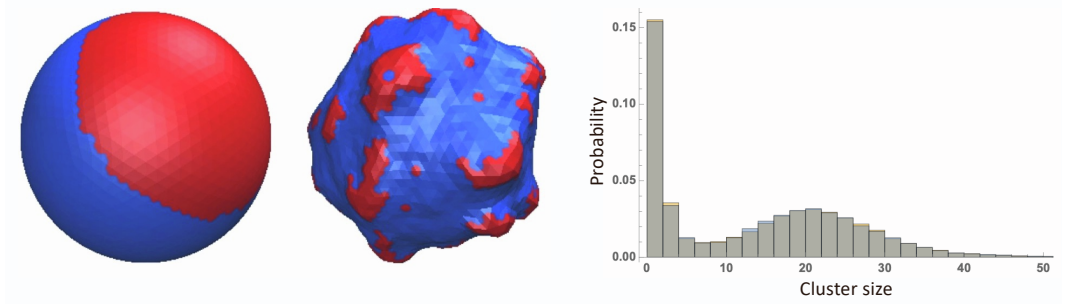


FIG. S9: If the simulations starts from a macro-cluster (left), one rapidly arrives (middle, example after  $10^9$  MC moves) to the same kind of configuration as if one had started from a random distribution of species A and B. Right: Probability distributions of cluster sizes when starting from a random configuration (orange) or the macro-cluster (blue). They are almost super-imposed, indicating that the system is in equilibrium. Simulation parameters are the same as in Figure 5 in the main text, with  $\bar{\phi} = 0.2$ .

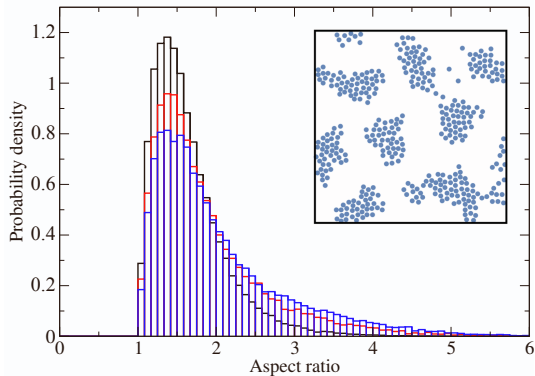


FIG. S10: Aspect ratio probability distributions for the domains observed in the in-plane simulations with explicit short-range attraction and long-range repulsion [8] as defined in Eq. (S6) with  $\bar{\phi} = 0.18$  (black),  $\bar{\phi} = 0.25$  (red) and  $\bar{\phi} = 0.27$  (blue). The inset is a snapshot for  $\bar{\phi} \simeq 0.25$  showing elongated domains (400 particles, periodic boundary conditions).

molecule diameter <sup>1</sup>. Figure S10 shows that this model also leads to the formation of more elongated domains when the concentration is high enough, whereas the domains are roundish at lower concentrations. Again the “hinge point” in the aspect ratio distributions is close to 2, which strengthens further our choice  $AR_0 = 2$ .

We now estimate the parameter  $\varepsilon$  for this interacting-protein model. As compared to the main text, we have to estimate the parameters  $\lambda$  and  $E_0$  in a different way. We know the short-range attraction strength  $\varepsilon_a$  and the mid-range repulsion strength  $\varepsilon_r$  of the pairwise interaction, and we denote by  $d$  their typical diameter (i.e. their typical separation). The line tension is then estimated to be  $\lambda \approx \varepsilon_a/d$  and the repulsion parameter to be  $E_0 \approx$

$\varepsilon_r/d^4$  <sup>2</sup>. In addition, we introduce the reduced diameter  $\tilde{r}_0 = r_0/d$  (estimated to be close to 4 in the simulation outputs). It follows that  $\lambda/(E_0 r_0^3) \approx \varepsilon_a/(\varepsilon_r \tilde{r}_0^3) \approx 1$  with the parameter values of our simulations. Again, we are close to the threshold, even though we do know the exact numerical prefactor in the stability condition for ellipses for the pairwise potential used in these simulation and the real density profile inside the clusters.

### C. Supplementary analyses

Regarding the triangulated-vesicle model, Figure S11 shows the cluster-size distribution for the two different concentrations  $\bar{\phi}$  of the minority phase A studied in this work. Indeed, we implemented the cluster detection algorithm described in Ref. [6] to identify each domain and measure its size on independent system configurations throughout the simulation.

Figure S12 shows the spatial auto-correlation function, and the associated structure factor in the spherical harmonics basis, for same two concentrations in species A. In both cases, spatial wavelengths are very close. The oscillations of the spatial correlation functions are almost in phase and the two systems have a maximum in their structure factor for the same mode (order  $l = 8$ ). The differences in amplitudes for the correlation functions and in the peak widths in the structure factors arise from the different input concentrations. Although the correlation function is computed from an average measurement in every direction, we can qualitatively interpret these observations as the fact that the domains elongate in one direction only and their typical spacing and thus width remain roughly constant. This has also been observed in the experimental data were the domain minor axis is weakly varying and their major axis is varying more significantly with concentration, as shown in Table S2. For

<sup>1</sup> The density  $\bar{\phi}$  is computed at null temperature, with all the particles considered to be condensed in a perfect triangular lattice. They experience hard-core repulsion.

<sup>2</sup> Here we approximate the area occupied by each protein by  $d^2$ .

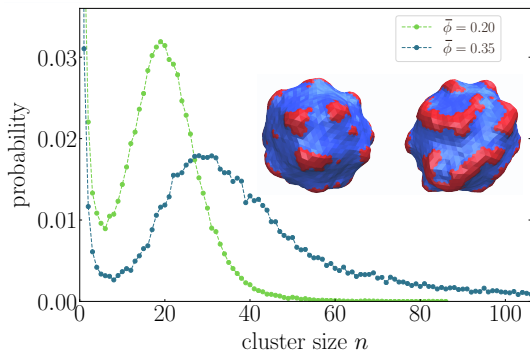


FIG. S11: Domain size probability distribution for vesicles with  $\bar{\phi} = 0.20$  and  $\bar{\phi} = 0.35$  in terms of number of sites. The typical domain size for a system corresponds to the position of the peak. The system at higher concentration ( $\bar{\phi} = 0.35$ ) shows domains with a higher typical size (by  $\approx 50\%$ ) than the one at lower concentration ( $\bar{\phi} = 0.20$ ). Its largest domains are also bigger than the ones at lower concentration and can reach  $\approx 400$  sites. Same parameters as in Figure 5 in the main text. Inset: simulations snapshots of the corresponding vesicles with  $\bar{\phi} = 0.20$  (left) and  $\bar{\phi} = 0.35$  (right); same snapshots as in Figure 5 in the main text.

example, in the condition where CCR5 is tracked and all protein expression levels are low, the minor axis of the domains was in average equal to  $64 \pm 7$  nm, and it slightly shifted to  $75 \pm 4$  nm in the case where all the proteins are overexpressed. In comparison, the major axis mean value shifted from  $109 \pm 13$  nm to  $144 \pm 7$  nm, a significantly more pronounced increase.

Note also that the analysis based upon the triangulated-vesicle model was additionally run for the same parameter values except a slightly different curvature coupling value as compared to the value  $c_1 = 10$  examined in the main text. It led to very similar results:  $c_1 = 9$  led to 24.5% of elongated domains at  $\bar{\phi} = 0.2$  and 37.0% at  $\bar{\phi} = 0.35$ , whereas  $c_1 = 11.0$  led to 31.1% of elongated domains at  $\bar{\phi} = 0.2$  and 40.8% at  $\bar{\phi} = 0.35$ . This reinforces our conclusions on the effect of concentration on domain shape.

Even though coiled elongated structures as presented in section III above, that could not be detected by the sole aspect ratio criterion, were observed in the simulations for both numerical models (see Figure S13), this was limited to very high concentrations. They were exceptional in the concentrations used in Figures 5 in the main text and S10. For this reason, the classification correction of domain shapes based on angular distribution analysis was not applied there.

## VI. ABOUT $p$ -VALUES

We have used several approaches to test whether two unpaired datasets arise from the same probability distribution through the  $p$ -value, as follows.

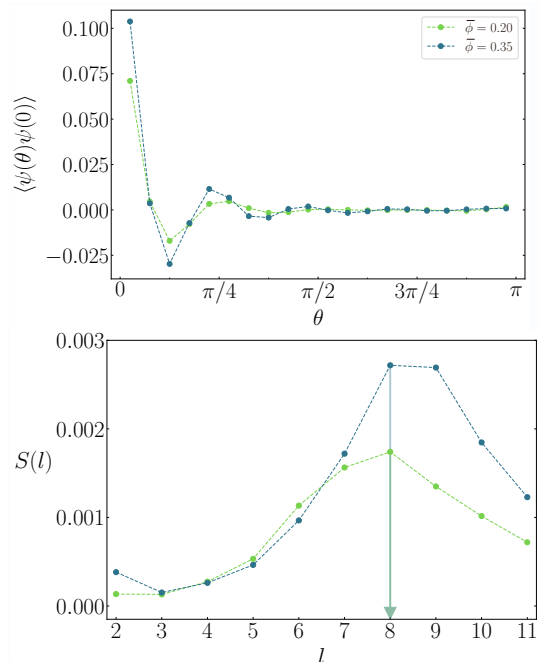


FIG. S12: Correlation functions (top) and associated structure factor (bottom) for the triangulated-vesicle numerical model, with  $\bar{\phi} = 0.20$  and  $\bar{\phi} = 0.35$ . The parameter  $l$  is the order of the spherical harmonics. The two systems are characterized by very close spatial wavelengths. Other simulation parameters are  $c_1 = 8.0$ ,  $\tilde{\sigma} = 300$  and  $\tilde{J}_I = 0.5$ .

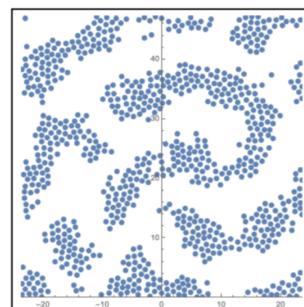


FIG. S13: Example of curved and curled domains obtained in the interaction-protein model for higher concentrations  $\bar{\phi} = 0.31$  (800 particles, periodic boundary conditions).

Every transient confinement zone (TCZ) is considered a sample. Indeed, if several TCZs are extracted from the same cell, they do not in general correspond to the same nanodomain, and are thus uncorrelated in terms of cell surface topography.

We first computed the  $p$ -values based on the aspect ratio lists of two samples with the help of the Student t-test (TT) provided by Wolfram Mathematica (unpaired two-tailed t-test). However, the algorithm systematically displayed a warning because the distribution are not sufficiently close to Gaussian ones to ensure the validity of the Student test. We therefore switched to a more



Test	$t$	CD4	CCR5	CXCR4
TT	0.25	0.012	$5 \times 10^{-5}$	0.005
TT	0.3	0.002	$1 \times 10^{-4}$	0.001
KS	0.25	$2 \times 10^{-4}$	0.002	0.001
KS	0.3	$2 \times 10^{-6}$	$8 \times 10^{-4}$	$5 \times 10^{-4}$

TABLE S3:  $p$ -values obtained with different statistical tests and different values of the threshold  $t$  in the case where the three proteins are overexpressed and one of them is tracked, as compared to the case where no protein is overexpressed and the same one is tracked, as in Figure 7 of the main text.

adapted non-parametric test, namely the Kolmogorov-Smirnov (KS) test also provided by the Wolfram Mathematica software. In addition, we tested two numerical values of the threshold  $t$  used to detect false roundish domains. In the main text figures, a value of  $t = 25\%$  is used. We have also calculated the  $p$ -values for an alternative value of 30%, as displayed in Table S3. A  $p$ -value  $< 0.05$  is generally considered to be statistically signifi-

cant [12]. Our KS  $p$ -values are on the order of, or lower than 0.001.

In all the figures, \* means  $p \leq 0.05$ , \*\* means  $p \leq 0.01$ , and \*\*\* means  $p \leq 0.001$ . When indicated on top of histograms displayed in insets, the  $p$ -values are calculated on the basis of the full distributions, not on the sole histograms that contain less information.

## VII. HIV-1 INFECTION ASSAYS

Infection assay were carried out as described in [2]. Briefly, the induced 293-Affinofile cells ( $2 \times 10^5$  cells per well) in 96-well plates were inoculated with 100 ng p24 of the pNL4-3-derived viral clones expressing Renilla luciferase and gp160 from the JR-CSF or JR-FL strains. Infected cells were further incubated for 30 h at 37°C before being lysed. Viral replication was then determined by measuring luciferase activity in the cell lysates using a Glomax 96-well plate luminometer (Renilla Luciferase Assay, Promega, Madison, WI, USA) [13, 14].

- 
- [1] S.H. Johnston *et al.*, *J. Virol.* **83**, 11016–11026 (2009).
  - [2] P. Colin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 9475–9480 (2013).
  - [3] F. Daumas *et al.*, *Biophys. J.* **84**, 356–366 (2003).
  - [4] N. Meilhac, L. Le Guyader, L. Salome, N. Destainville, *Phys. Rev. E* **73**, 011915 (2006).
  - [5] P. Mascalchi, Thesis of the Université Toulouse III - Paul Sabatier (2012).
  - [6] J. Cornet, N. Destainville, M. Manghi, *J. Chem. Phys.* **152**, 244705 (2020).
  - [7] G. Gueguen, N. Destainville, M. Manghi, *Soft Matter* **13**, 6100 (2017).
  - [8] N. Destainville, *Phys. Rev. E* **77**, 011905 (2008).
  - [9] S. Weitz, N. Destainville, *Soft Matter* **9**, 7804–7816 (2013).
  - [10] R. P. Sear *et al.*, *Phys. Rev. E* **59**, R6255–6258 (1999).
  - [11] T. M. Konyakhina *et al.*, *Biophys. J.* **101**, L8 (2011).
  - [12] M. Krzywinski, N. Altman, *Nat. Methods* **10**, 1041–1042 (2013).
  - [13] P. Colin *et al.*, *PLoS Pathog.* **14**, e1007432 (2018).
  - [14] M. Armani-Tourret *et al.*, *PLoS Pathog.* **17**, e1009526 (2021).