Quantifying membrane binding and diffusion with fluorescence correlation spectroscopy diffusion laws

Anita Mouttou, Erwan Bremaud, Julien Noero, Rayane Dibsy, Coline Arone, Johnson Mak, Delphine Muriaux, Hugues Berry, Cyril Favard

Manuscript submitted to **Biophysical***Journal*
Article

Quantifying membrane binding and diffusion with Fluorescence Correlation Spectroscopy diffusion laws

Anita Mouttou[1,eq], Erwan Bremaud[1,eq], Julien Noero[1], Rayane Dibsy[1], Coline Arone[1], Johnson Mak[3], Delphine Muriaux[1], Hugues Berry[2,*], and Cyril Favard[1,*]

[1]Membrane Domains and Viral Assembly, Montpellier Infectious Disease Research Institute, UMR CNRS 9004, 1919, route de Mende, 34293 Montpellier Cedex, France
[2]INRIA Rhone Alpes, Laboratoire d'InfoRmatique en Image et Systèmes d'information, UMR 5205 CNRS INSA Lyon, Université Claude Bernard Lyon 1, Université Lumière Lyon 2, École Centrale de Lyon, Campus de la Doua, 25 avenue Pierre de Coubertin, 69622 Villeurbanne Cedex, France
[3]Institute for Glycomics, Griffith University Gold Coast, QLD 4222, Australia
[*]Correspondence: cyril.favard@irim.cnrs.fr, hugues.berry@inria.fr
[eq]Equivalent Contribution: A. Mouttou, E. Bremaud

ABSTRACT Many transient processes in cells arise from the binding of cytosolic proteins to membranes. Quantifying this membrane binding and its associated diffusion in the living cell is therefore of primary importance. Dynamic photonic microscopies, e.g. single/multiple particle tracking, fluorescence recovery after photobleaching and fluorescence correlation spectroscopy (FCS) enable noninvasive measurement of molecular mobility in living cells and their plasma membranes. However, FCS with a single beam waist is of limited applicability with complex, non Brownian, motions. Recently, the development of FCS diffusion laws methods has given access to the characterization of these complex motions, although none of them is applicable to the membrane binding case at the moment. In this study, we combined computer simulations and FCS experiments to propose an FCS diffusion law for membrane binding. First, we generated computer simulations of spot-variation FCS (svFCS) measurements for a membrane binding process combined to 2D and 3D diffusion at the membrane and in the bulk/cytosol, respectively. Then, using these simulations as a learning set, we derived an empirical diffusion law with three free parameters: the apparent binding constant K_D, the diffusion coefficient on the membrane D_{2D} and the diffusion coefficient in the cytosol, D_{3D}. Finally, we monitored, using svFCS, the dynamics of retroviral Gag proteins and associated mutants during their binding to supported lipid bilayers of different lipid composition or at plasma membranes of living cells and we quantified K_D and D_{2D} in these conditions using our empirical diffusion law. Based on these experiments and numerical simulations, we conclude that this new approach enables correct estimation of membrane partitioning and membrane diffusion properties (K_D and D_{2D}) for peripheral membrane molecules.

WHY IT MATTERS Many processes in cells start with the binding of a cytosolic protein to the plasma membrane. Accurate determination of the protein binding efficiency as well as the protein diffusion coefficient at the plasma membrane of living cells is of utmost importance to decipher the different processes triggered by the initial binding event. Here we propose a non-invasive novel approach, based on spot-variation fluorescence correlation spectroscopy, to quantitatively determine these two parameters, directly in living cells expressing a fluorescent proteins of interest.

INTRODUCTION

The binding of cytosolic proteins to a membrane is often the starting point of dynamic processes occurring in the cell. This binding is frequently a key event in signal transduction, metabolism, membrane trafficking (endo/exocytosis) or enveloped virus assembly. Although much has been studied about these processes, quantifying the dynamics of the initial event, i.e. protein binding to and diffusion on the membrane, still remains a challenge in living cells.

Several methods, such as fluorescence recovery after photobleaching (FRAP), fluorescence correlation spectroscopy (FCS) and single(multiple) particle tracking (SPT), have been developed to monitor molecule motions in cells using fluorescence microscopy and have been extensively used in membrane biology (for review see (1)). Each of these dynamic microscopic techniques have pros and cons related to their timescale and statistics. FCS is sensitive on the millisecond to second timescale,

corresponding to the diffusion characteristic time of a fluorescently-labeled molecule in a lipid mixture, through an illumination focus with a waist of approximately 200nm. FCS has therefore provided a convenient way to investigate the motions of lipids and membrane proteins in living cells. FCS usually consists in computing the autocorrelogram function (ACF) from the temporal recording of fluorescence intensity fluctuations in the illumination beam. The ACF is then fitted with an analytical expression accounting for a given type of motion e.g., fractional or normal Brownian motion in 2D or 3D. This fit yields the mobility parameters as estimates of its free parameters. Interestingly, FCS was originally developed to determine the kinetic constants of chemical reactions at equilibrium (2). Recently, FCS has been proposed as an efficient method to monitor more complex dynamics such as reaction-diffusion processes occurring in the case of transcription factor binding to DNA in cells (3) or in the embryo (4). However, single-spot FCS, in which the beam waist is constant, is usually not applicable to discriminate and quantify complex motions. In the case of reaction-diffusion processes for example, there is no analytical solution available to retrieve the kinetic parameters from the fit of the autocorrelogram. Approximated expressions can be derived by simplifying the kinetics (e.g., neglecting diffusion of the bound molecule) or distinguishing between simplification regimes (reaction kinetic dominant, diffusion dominant...) (3). However even in the case of molecules diffusing in the membrane, single spot FCS measurements often results in inaccurate estimates of diffusion coefficient, mainly because the single spot size is not sufficient to probe the heterogeneity of the environment where the molecules diffuse.

One way to circumvent that issue is to perform FCS measurements at different beam waists. Plotting the half-time of the decorrelation ($\tau_{1/2}$) as a function of the surface probed by the illumination (w^2, where w is the beam waist, or spot size), one obtains so-called svFCS diffusion laws. Wawrezinieck et al. (5) showed that svFCS diffusion laws are a powerful tool for analyzing complex diffusions. They developed a svFCS experiment and, based on numerical simulations, managed to successfully associate svFCS diffusion laws with different types of complex, heterogeneous environments. The same approach has also been used experimentally to characterize molecular motions in membranes, below the diffraction limit (6–8) and recently in Imaging-FCS (Im-FCS) (9, 10). In terms of FCS diffusion laws $\tau_{1/2} = f(w^2)$, a pure Brownian motion (free diffusion) leads to $\tau_{1/2} \to 0$ at $w^2 \to 0$, i.e. the extrapolated value of $\tau_{1/2}$ for vanishing w^2 is expected to be zero. Moreover, one predicts a linear relationship between $\tau_{1/2}$ and w^2 with a slope that is inversely proportional to the free diffusion coefficient. Interestingly, the extrapolation of $\tau_{1/2}$ for $w^2 \to 0$, $\tau_{1/2}(0)$ gives information about the nature of the heterogeneity probed by the molecule. For example, molecules experiencing dynamic partitioning between liquid ordered and liquid disordered (Lo/Ld) lipid phases will have a positive $\tau_{1/2}(0)$ value, whereas molecules restricted in their diffusion by sub-membrane fences, such as cortical actin cytoskeleton for example, are expected to show a $\tau_{1/2}(0) < 0$ (5, 11). However, lipids exhibiting dynamic partition in solid/liquid disordered phase (S/Ld) also display $\tau_{1/2}(0) < 0$ (12), showing that the $\tau_{1/2}(0)$ value alone is not sufficient to correctly characterize the motion (13).

Recently, different types of svFCS diffusion laws have been derived and characterized for different types of heterogeneous environments based on numerical simulations (14, 15), mostly for complex motions related to anomalous sub-diffusion. In this article, we explored the applicability of svFCS diffusion laws to the quantification of membrane binding kinetics where the fluorescent molecule diffuses both at the 2D membrane and in the 3D bulk. Our main objective was to test the suitability of such experimental diffusion laws to quantify the main parameters: the apparent binding constant (K_D), the membrane bound diffusion coefficient (D_{2D}) and the cytosolic/bulk diffusion coefficient (D_{3D}) (see table 1 for the definitions of the different parameter use here).

We first performed extensive numerical simulations of svFCS diffusion laws for this system, exploring different regimes by varying K_D, D_{2D} and D_{3D} . Using part of these simulations as a learning set, we could derive an empirical svFCS diffusion law and use the other part of the simulations as a test set to confirm that the diffusion law can indeed be used to estimate K_D and D_{2D} of synthetic data with good accuracy. We then used this empirical diffusion law to fit experimental svFCS measurements of HIV-1 or HTLV-1 retroviral Gag proteins binding to either supported lipid membranes or living HEK 293T cells plasma membranes. HIV-1 and HTLV-1 Gag proteins are multi-domain proteins binding the plasma membrane of host cells and are essential for viral assembly. HIV-1 Gag for example contains three main domains, each having distinct roles during viral assembly. The matrix domain (MA) is responsible for membrane binding to the PI(4,5)P_2 lipids, present in the inner leaflet of the plasma membrane, thanks to a highly basic region (HBR). Membrane binding is reinforced by the N-terminal myristate MA substitution that inserts into the inner leaflet. The capsid domain (CA) is involved in Gag-Gag interactions during Gag self-assembly, the initial stage in the generation of a new virion. The nucleocapsid domain (NC) also contains enriched basic motifs that are involved in RNA binding to permit encapsidation of the RNA in the new virion. NMR data and coarsed grain molecular dynamics have shown that the HBR region of MA as well as other polybasic motifs bind to anionic lipids (PI(4,5)P_2, but also PS) (16–19). Conversely, the lack of myristate has been shown to decrease membrane binding of HIV-1 in living cells (20). Finally, Fogarty et al. (21) have shown that HTLV-1 Gag has a higher affinity for cell plasma membranes than HIV-1 Gag.

Although these results from the literature do not consist in a precise quantification of the binding affinities, they provide us with ordering relations between them, e.g. the binding affinity of HIV-1 in the presence of myristate is higher than in its absence. We used them to challenge the capacity of our svFCS diffusion law to retrieve the correct estimations from experimental svFCS

diffusion laws and to correctly estimate K_D and D_{2D} *in vitro*, using supported lipid bilayers (SLBs) of various composition and *in cellulo* by expressing Gag chimerae proteins in HEK-293T cells.

On the synthetic test dataset, we show that our svFCS diffusion law yields precise estimates for D_{2D} and K_D, with an accuracy around 12 %. The accuracy of our method on real experimental measurements cannot be assessed directly, but our results confirm the capacity of our 2D/3D diffusion and binding svFCS diffusion law to correctly estimate D_{2D} and K_D in the case of HIV-1 and HTLV-1 Gag proteins binding to model lipid SLB and to HEK-293T cells. Notably, our estimates confirm previous results obtained in the literature with other methods by different teams. Therefore, our results highlight the benefits of our empirical svFCS diffusion law for the determination of D_{2D} and K_D in the membrane binding problem with diffusion both at the membrane and in the cytosol.

MATERIALS AND METHODS

Lipids and Plasmids

Egg-PhosphatidylCholine (PC), Brain-PhosphatidylSerine (PS), Brain-Phosphatidyl Inositol (4,5)BisPhosphate$_2$ (PI(4,5)P_2), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-(Cyanine 5.5) (PE-Cy5.5) and 1-palmitoyl-2-(dipyrromethene boron difluoride)-undecanoyl-sn-glycero-3-phosphocholine (TopFluor®-PC) were purchased from Avanti Polar Lipids (Alabama,USA). Atto647N-PI(4,5)P_2 is a gift from Pr. Christian Eggeling (Jena, Germany). The plasmid encoding HIV-1-myr(-)Gag-GFP is a gift of Dr. H.de Rocquigny and described in (22). The plasmid encoding HTLV-1 Gag-YFP (named pHTLV-Gag-YFP) was obtained from Dr. D. Derse's lab and described previously in (23). The pHIV-1-Gag-mCherry was obtained from Dr. N. Jouvenet (24).

Gag purification and labelling

Purified Gag Proteins were produced and provided by Pr. J. Mak's lab (25). Protein stock concentration was measured at 1.2mg/mL using a NanoPhotometer (Implen), and 100μL were incubated overnight under agitation and at 4°C with 1μL of Alexa Fluor 488 C5-Maleimide (Invitrogen) at 20mM in DMSO. The reaction mix was transferred in Slide-A-Lyzer MINI Dialysis Device, 0.5mL (Thermo Scientific) and dialyzed 6h under agitation at 4°C in 15mL Buffer; TRIS (50mM), NaCl (1M), pH=8. Dialysing buffer was then renewed, and the reaction mix was dialysed again overnight under agitation at 4°C, collected, and stored at -20°C.

LUVs processing

Liposomes were processed by dissolving in chloroform the following lipids aiming at a 1mg/mL concentration: Egg-PC, Brain-PS, Brain-PI(4,5)P_2 and PE-Cy5.5 in the ratios described Figure 2a. Solvent was then evaporated during 20min in a rotary evaporator and 10min more in a desiccator. Lipids were rehydrated in 500μL Na Citrate Buffer; Na Citrate (10mM), NaCl (100mM), EGTA (0.5mM), pH 4.6. The mixture was then, 5 consecutive times: frozen 30s in liquid nitrogen, water bathed at 37°C for 30s and vortexed. MLVs processed as such can be stored at -20°C. LUVs are subsequently produced by diluting Liposomes 1:5 in Na Citrate Buffer. This solution is then passed through a 100nm Nucleopore Track-Etched Membranes in an extruder (Avanti Polar Lipids), and sonicated.

LUV binding experiments

Proteins at a final 40nM concentration were incubated 30min at room temperature with LUVs compositions described Figure 2a at a final 0.8mg/mL concentration. 100μL samples were then centrifuged at 42000rpm for 30min at 4°C (Beckman TLA 100 rotor). 2 fractions were isolated from these samples : 90μL were collected from the supernatants, and the pellets were resuspended with 80μL TRIS NaCl Buffer; TRIS HCl (10mM), NaCl (150mM), pH 7.4. Equal volumes from each fraction were loaded in SDS-PAGE gels and analysed by Western Blot. Gag proteins were detected using an anti-p24 as primary antibody and a secondary antibody coupled to HRP (Horse Radish Peroxydase). Membranes were imaged by enhanced chemiluminescence (ECL) in the ChemiDoc (Biorad Inc, USA) and bands relative intensities were quantified using ImageJ. K_D values were then calculated as in equation 1, ie the ratio of the bands signal corresponding to supernatants (S) over the bands signal corresponding to pellets (P) corrected by the pellet obtained with Gag alone P_{GA}, which account for spontaneous self-assembly. According to the mass action law, the amount of self-assembly was considered proportional to the total amount of Gag ($T = P + S$) and normalize to this total amount when Gag was alone (T_{GA}). To correctly estimate K_D we then performed the following calculation :

$$K_D = \frac{S}{P_{LUV_{Binding}}} = \frac{S}{(P - P_{GA}\frac{T}{T_{GA}})} \tag{1}$$

SLBs processing

Glass coverslips (25 mm No.1.5 H cover glass (Marienfeld)) were treated 30min with ozone in Ossila UV Ozone Cleaner and rinsed thoroughly with ultrapure water. The sample is delimited by a 7mm diameter plastic cylinder fixed to the coverslip using Twinsil (Picodent). 100μL of 0.2mg/mL SUVs of the different mix compositions are spread on the coverslip and incubated 40min at 37°C. Formed SLBs are washed 4 times with filtered TRIS NaCl Buffer to remove non-fused vesicles. Proteins at a final 10nM concentration are injected 30 min before the beginning of the experiment.

Living cell sv-FCS experiments

Human embryonic kidney cell line (HEK-293T) were seeded on 25 mm No.1.5 H cover glass (Marienfeld) coated with poly-L-lysine (Sigma). 0.15 million cells were maintained in 2ml Dulbecco's Modified Eagle's Medium (DMEM, GIBCO) for 24 hours at 37°C. Transfection was performed using $CaCl_2$ (250mM) and HBS2X, with 1 μg pHIV-Gag-mCherry, 0.3 μg pHIV-GagG2A-GFP and 1μg pHTLV-Gag-YFP. 15 hours post transfection, medium was replaced with phenol-red free medium L15 (GIBCO) supplemented with 20 mM Hepes, for sv-FCS experiments on live cells.

Acquisition and fit of the experimental diffusion laws

FCS set up

FCS was performed on a Zeiss LSM 780 microscope equipped with the variable pupil coverage system using water immersion (NA = 1.2). The underfilling of the objective back-aperture leads to the increasing of the beam waist.

Beam waist calibration

w^2 was calibrated using highly diluted Rhodamine-6G (R6G) and Tetramethylrhodamine (TMR, a gift from Dr. M.May) solutions excited at 488 nm and 561 nm respectively. At least 30 autocorrelogram functions (ACF) were obtained at T=37°C (for cell experiments) or 20°C (for SLB binding experiments) from 5s fluctuation intensity acquisition. Laser irradiance of the sample is adjusted for the different waists in order to get constant molecular brightness. The free diffusion time is obtained by fitting the autocorrelation function of the measured time - trace intensity with an analytical model of a 3D free diffusion including a triplet model (T + 3D).

$$G(\tau) = 1 + \left(1 + \frac{T}{1-T} e^{-t/\tau_T}\right) \frac{1}{N} \frac{1}{1+\tau/\tau_\mathrm{d}} \frac{1}{\left(1+\tau/(s^2\tau_\mathrm{d})\right)^{1/2}} \tag{2}$$

Rhodamine triplet time was measured from the smallest waist correlogram to be 4μs, and was therefore kept fixed to fit all the bigger waist calibration correlogram, only T, the triplet fraction was allowed to change. s is set to 5. w^2 is calculated for τ_d using $w^2 = 4D\tau_d$. The diffusion coefficient of R6G and TMR solutions were set to D = 360 $\mu m^2.s^{-1}$ for the SLB measurements and D = 550 $\mu m^2.s^{-1}$ (respective values of D for R6G at 20 and 37°C according to https://www.picoquant.com/images/uploads/page/files/7353/appnote_diffusioncoefficients.pdf, based on (26)). We measured excitation waists varying from 240 to 1120 nm (488nm excitation) or from 290 to 1220 nm (561 nm excitation) depending on objective's back-aperture coverage.

FCS diffusion law acquisition on SLBs and cell

The SLB z=0 axial position was retrieved thanks to the Cy5.5 PE fluorescence, ACF were recorded at 20°C. Cell intensity fluctuations and associated ACF correlograms were acquired at 37°C by parking the laser in the z-axis at the level of the bottom membrane, in order to avoid deformation of the laser shape due to index mismatch that will occur at the top membrane. The position of the laser spot within the cell was systematically controlled with a z-scan intensity profile before and after each waist of correlogram acquisition. In both cases, FCS measurements were done for 7 different beam waist and we recorded at least 50 series of time-trace intensity of 10s duration per waist by successive series of 10. Again,the laser irradiance is adjusted over the different waists to keep constant and independent of the waist probed a significant photon count per molecules in the range of 2-5 kHz, but also kept low enough to avoid important photobleaching. This is controlled *a posteriori* by plotting N, the number of molecules as a function of $V_{\mathrm{eff}} = \pi^{3/2} s w^3$ (see Fig.SI.3). In the case of GFP and YFP labelled proteins as well as for the Alexa-Fluor 488 labelled Gag, each ACF correlogram function are fitted (from 10μs to 1.7s) with an analytical model for Brownian 2D diffusion.

$$G(\tau) = 1 + \frac{1}{N}\left(\frac{1}{1+t/\tau_d}\right) \tag{3}$$

FCS measurements of mCherry labelled proteins were done with a 561 nm laser line. mCherry is known to exhibit flickering which lead to dark state population with typical ground state recovery kinetics (27). As we do not have an analytical model for this particular photophysics, we used a classical triplet state+2D diffusion model:

$$G(\tau) = 1 + \left(1 + \frac{T}{1-T} e^{-t/\tau_T}\right) \frac{1}{N} \frac{1}{1+\tau/\tau_{\mathrm{d}}} \tag{4}$$

with $\tau_T << \tau_d$. Average $\tau_{1/2}$ and their standard deviation (s.d.) values found for each waist in each conditions (either for the SLB experiments or for the live cells) as well as the corresponding number of ACF correlograms fitted to determine these average $\tau_{1/2}$ values are given in tables SI.1, SI.2, SI.3 and SI.4 .

Table 1: List of parameters used in this work

Variables	Parameters details
Confocal volume calibration parameters	
w_{xy}	Width of the observation area in x, y direction
$s = w_z/w_{xy}$	Shape factor of the 3D Gaussian excitation point spread function (PSF)
Parameters used in ACF fit	
$G(\tau)$	Autocorrelation function (ACF)
$\tau_{1/2}$	Characteristic time of half temporal decorrelation i.e. $G(\tau_{1/2}) = G(\infty) + (G(0) - G(\infty))/2$
τ_d	Characteristic diffusion time of particles in the confocal volume $\tau_d = \tau_{1/2}$ for 2D diffusion
N	Effective number of particles in confocal volume
T	Fraction of particles in triplet state
τ_T	Characteristic residence time in triplet state
Parameters of the 2D/3D diffusion and binding kinetics	
$D_{2\mathrm{D}}$	Diffusion coefficient of membrane-bound molecules
$D_{3\mathrm{D}}$	Diffusion coefficient of free molecules
k_{on}	Association constant of particles to the membrane
k_{off}	Dissociation constant of particles from the membrane
$K_D = k_{\mathrm{off}}/k_{\mathrm{on}}$	Binding affinity of the proteins to the membrane
$K_P = [F]_{\mathrm{eq}}/[B]_{\mathrm{eq}}$	Experimental apparent partition coefficient (with $[F]_{\mathrm{eq}}$ and $[B]_{\mathrm{eq}}$ the equilibrium concentrations of free and bound proteins, respectively).

NUMERICAL SIMULATIONS OF (SV)FCS EXPERIMENTS

Algorithm

We emulated FCS experiments by simulating the binding reaction and diffusion of 100 individual molecules in a cubic nucleus of size L^3 μm^3 during 11 seconds. Unless otherwise indicated, we use $L = 6\mu$m in this work. The plasma membrane was defined as the $z = 0$ face of the cell, with the cytoplasmic bulk as the $z > 0$ 3D volume. At the start of the simulation, a fraction of the individual molecules were positioned at random 3D positions (uniform distribution) within the cytoplasm and the reminder was located at random 2D positions (uniform distribution) on the plasma membrane (i.e. with coordinates $z = 0$) (Fig 1a1). The fraction of molecules initially free in the cytoplasm was set to its theoretical equilibrium fraction i.e., to $k_{\mathrm{off}}/(k_{\mathrm{on}} + k_{\mathrm{off}})$. Reaction-diffusion simulation processed by successive iterations of time step Δt (we used $\Delta t = 1\mu$s). The position $\mathbf{r}_i$ of every individual molecule $i = \{1, \ldots, 100\}$ was first moved according to independent Brownian dynamics: $\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \sqrt{2D_i\Delta t}\xi$ where ξ is a 3D vector of i.i.d. Normal random numbers and $D_i = D_{2\mathrm{D}}$ or $D_{3\mathrm{D}}$, for membrane-bound or free molecules in the bulk, respectively. Membrane-bound molecules were kept at the membrane at this step (i.e. we reset their coordinate z to zero). Classical boundary conditions e.g., periodic or reflective, introduce strong correlations in the simulations that are not present in

the experiments. We therefore opted for the following scheme. Molecules exiting the cell through the plasma membrane at $z = 0$ were changed into bound molecules (thus reset at $z = 0$), whereas the upper face ($z = L$) was considered reflective. Molecules exiting the cell through any of the 4 remaining faces were removed from the simulation. To keep the number of molecules constant in the cell, removed molecules were compensated by the addition of an equal number of new free molecules that were located in the bulk at a distance $\epsilon = 10$ nm of one of the 4 remaining faces, chosen at random (with uniform distribution).
We then simulated the reaction step. We considered that free bulk molecules located at a distance less than $\epsilon = 10$ nm from the plasma membrane (i.e., for which $z < 0.01\mu$m) were close enough to bind. Each of these molecules was therefore independently turned into a membrane-bound molecule with probability p_{on}. Conversely, every membrane-bound molecule was independently allowed to unbind i.e., was changed into a free bulk molecule, with probability p_{off}. Both reaction probabilities were set from the simulation parameters according to: $p_{\text{on}} = k_{\text{on}} L \Delta t / \epsilon$ and $p_{\text{off}} = k_{\text{off}} \Delta t$.
To emulate fluorescence emission, we used a 3D Gaussian illumination profile centered on the plasma membrane at $z = 0$. The probability that a molecule located at position (x, y, z) at time t emit one photon between t and $t + \Delta t$ was computed as $I(x, y, z) = I_0 \exp\left(\frac{-2(x^2+y^2)}{{w_{xy}}^2}\right) \exp\left(\frac{-2z^2}{{w_z}^2}\right) \Delta t$, with w_{xy} and w_z the beam width in the x, y or z direction respectively. We set I_0 to 15×10^4 to reproduce the experimentally-observed brightness. To achieve correct equilibrium in the molecule locations before recording photon emission, the collection of photon emission times was switched off during the first 500 ms of every simulation run. We counted the total number of photons emitted at each Δt time step during the simulation before computing the auto-correlation of the corresponding time series. For each parameter value, we ran 20 independent realisations of this 11 second-simulation, and averaged the 20 resulting auto-correlations to get the average auto-correlation $G(\tau)$.

Simulation of spot variation FCS

To emulate an svFCS experiment, we repeated the above simulations for a selection of 16 values of the beam width w_{xy} between 71 and 632 nm and obtained the auto-correlation $G(\tau)$ for each of these 16 values. The beam width in the z direction was set to $w_z = s w_{xy}$ with a constant value $s = 5$ (in agreement with experimental measurements).

Model selection

To determine what model is the best fit of an autocorrelation curve, we used the reduced Akaike Information criterion (AICc) (28). This criterion has the advantage to take into account the number of free parameters of the models, penalizing models with more free parameters. We note G_i the i^{th} value of autocorrelation measured at a delay τ_i and $m(\tau_i)$ the value predicted by model m for this delay. The residual sum of squares $RSS = \sum_{i=1}^{n} (G_i - m(\tau_i))^2$ is used to compute the criterion: $AICc = 2k + n \ln(RSS) + (2k^2 + 2k)/(n - k - 1)$ with k the number of free parameters of model m and n the number of autocorrelation values. When comparing two models, the model with the smaller AICc is the best, but the evidence is considered strong enough if the AICc difference is at least 6. In case several models are found to be the best-fit (i.e. the evidence for their difference is not strong enough), we selected the model with the smaller number of free parameters k.

Code availability

The computer code used in this article is freely available at https://gitlab.inria.fr/hberry/gag_svfcs.

RESULTS

Empirical determination of a phenomenological diffusion law.

Single spot size

A first possibility to interpret the results of FCS experiments is to derive an analytical expression for the auto-correlation function $G(\tau)$ and fit it to the data to estimate the parameters. In practice, this is possible only for the simplest cases. In particular, this is not feasible, to our knowledge, when the problem is not homogeneous nor isotropic as is our case here with a 2D membrane located in a 3D volume at $z = 0$. However, one can simplify the problem by considering that the membrane binding and diffusion sites are homogeneously distributed in the 3D bulk, and by neglecting the reduction of dimensionality. The result is a system where the bound molecules diffuse in the same 3D space as the free ones i.e., a spatially homogeneous reaction-diffusion problem with reaction $\text{F} \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} \text{B}$ with F the free molecules diffusing with diffusion coefficient $D_{3\text{D}}$ and B the bound molecules diffusing with diffusion coefficient $D_{2\text{D}}$. Textbook reviews and several articles have treated simplified versions of the problem. For instance, the correlation function for the case $D_{2\text{D}} \approx 0$ is derived in Michelman-Ribeiro et al. (3) whereas Krichevsky and Bonnet (29) derives it for $D_{3\text{D}} = D_{2\text{D}}$. However, we could not solve the general case $D_{3\text{D}} > D_{2\text{D}} > 0$ (see

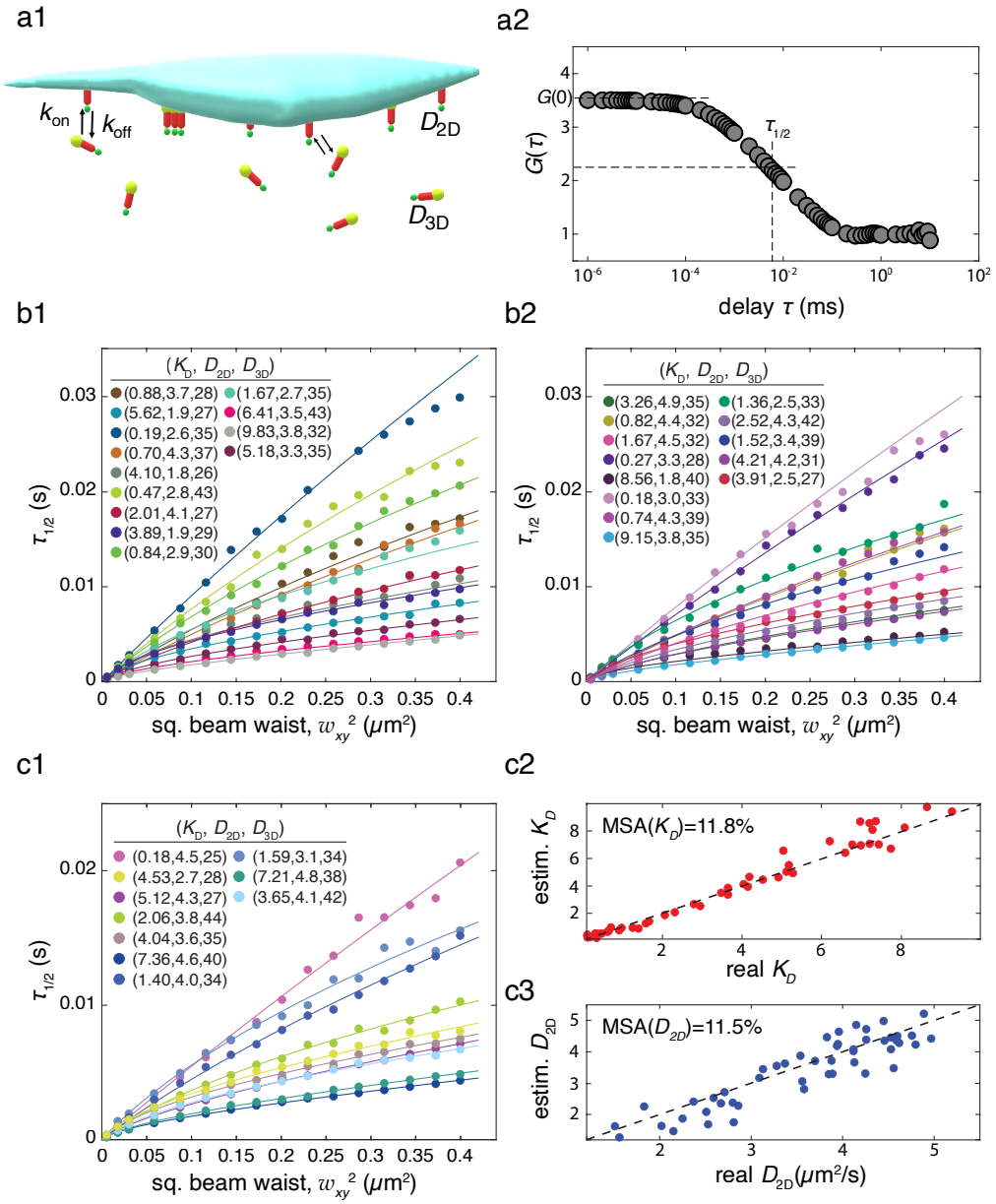

Figure 1: **Simulation-based derivation of a phenomenological diffusion law :** a1- We used the numerical simulations of svFCS described under Material and Methods to generate synthetic data of an svFCS experiment where HIV-1 gag molecules diffuse freely in 3D in the bulk / cytoplasm or in 2D as bound molecules at the plasma membrane, with binding and unbinding events driven by rate constants k_{on} and k_{off}, respectively. a2- The algorithm outputs the corresponding auto-correlation function, that we quantify by its FWHM (full width at half maximum), $\tau_{1/2}$. b- The diffusion laws, i.e. the changes of $\tau_{1/2}$ with beam width ${w_{xy}}^2$, are first fitted with the phenomenological diffusion law eq.(6) to the "training set" in order to estimate the six fit parameters of the law (see text). We show a selection of the synthetic diffusion laws (full circles) and their fits (full lines). The values of K_D, D_{2D}, D_{3D} for each simulation are given in the legend. c1- Fitting eq.(6) to the synthetic data of the test set is then used to estimate K_D and D_{2D} (D_{3D} is considered known, see text). The panel illustrates the obtained fits for a randomly chosen sample of simulations. c2 & c3- The estimated values of $K_D = k_{\text{off}}/k_{\text{on}}$ or D_{2D} are compared to their real values. The accuracy of these estimations is quantified using the Median Symmetric Accuracy (MSA).

Supporting Information SI1).

An alternative approach consists in fitting the auto-correlation curves $G(\tau) = f(\tau)$ with a selection of expressions derived for other problems and test on synthetic (simulation) data the accuracy or meaning of the parameters thus estimated. Supporting

Information Figure SI.1 shows the result of this approach. Here, we simulated diffusion and binding for a range of k_{on} and k_{off} values using the algorithm described above with $D_{2D} = 1.7$ and $D_{3D} = 30\,\mu\text{m}^2.\text{s}^{-1}$. For each parameter value, we fitted the obtained auto-correlation function with three models: a single population with 3D Brownian motion,

$$G(\tau) = 1 + \left[N\phi(\tau, D, w_{xy})\sqrt{\phi(\tau, D, w_z)}\right]^{-1}$$

with $\phi(\tau, D, w) = 1 + 4D\tau/w^2$, a mix of two populations with 3D Brownian motion,

$$G(\tau) = 1 + N^{-1}\left[\frac{K_D}{1+K_D}\phi(\tau, D_1, w_{xy})\sqrt{\phi(\tau, D_1, w_z)} + \frac{1}{1+K_D}\phi(\tau, D_2, w_{xy})\sqrt{\phi(\tau, D_2, w_z)}\right]^{-1}$$

or anomalous diffusion

$$G(\tau) = 1 + \left[N\left(1 + \left(4D\tau/{w_{xy}}^2\right)^{\alpha}\right)\right]^{-1}.$$

The best-fit model was then selected using corrected Akaike information criterias (see Methods) (Fig SI.1a). Note that, for the two-population mixture, one could use one population with 3D Brownian motion and the other with 2D. Using this 3D-2D mix in place of the 3D-3D mix used here does however not change our results.

We found that the best-fit model is the single Brownian model with low diffusion coefficient for $k_{on} \gg k_{off}$, or with large diffusion coefficient for $k_{off} \gg k_{on}$. Therefore, in these regimes, the approach correctly spots one of the two diffusion coefficients, but is blind to the other. Between these extreme regimes, the best-fit model was found to be the 2 population model with one small and one large diffusion coefficients (FigSI.1b1). Interestingly, the anomalous diffusion fit was never found to be the best-fit model except in one case. For most of the (k_{on}, k_{off}) values, this simple procedure is remarkably precise regarding the estimation of the diffusion coefficients (Fig SI.1b2). However, the estimation for K_D is very bad, with estimated values that can be one order of magnitude smaller than the real value (Fig SI.1b3). This was of course to be expected, given the simplifying hypotheses that support this approach. However, since our experimental interest here is mainly on the estimation of K_D, we opted for a strategy based on spot variation.

Spot variation FCS

Our motivation for svFCS is based on the following argument. Let F and B be the concentration (or density) of molecules in the cytoplasm or bound at the membrane, respectively. With the $z = 0$ plan of the beam located at the membrane, for increasing values of the beam waist w_{xy}, the surface of illuminated membrane will grow more slowly than the volume of illuminated cytoplasm, because the former scales as ${w_{xy}}^2$ and the later as ${w_{xy}}^3$. Hence, intuitively, one expects the FCS measurement to be more sensitive to the bound fraction at small beam waists and more sensitive to the free fraction for large beam waists. A back- of-the-envelop calculation is following: the number of bound molecules found inside the focal volume is $N_B \propto B{w_{xy}}^2$, whereas the number of free molecules found in the focal volume is $N_F \propto F{w_{xy}}^2 w_z = Fs{w_{xy}}^3$. The ratio between both is thus $N_B/N_F = B/F \times 1/(sw_{xy})$. The concentrations B and F are constant. The ratio $s = w_z/w_{xy}$ is as well roughly constant in our experimental setup. Therefore we expect from this simple analysis that the svFCS signal at very small beam widths w_{xy} will mostly be dominated by the bound molecules whereas free molecules should dominate at very large spot sizes.

In the svFCS literature, the evolution of $\tau_{1/2}$, the FWHM of $G(\tau)$ (full width at half maximum i.e., the value of τ were $G(\tau)$ is half its maximum value, see Fig 1a2) is referred to as a diffusion law. A first phenomenological diffusion law for our case therefore consists in expressing $\tau_{1/2}$ as the sum of a 2D and a 3D contribution, adding the constraint that the 2D bound fraction should dominate for $B/F \to \infty$ or $w_{xy} \to 0$, i.e. $\tau_{1/2}$ should converge to ${w_{xy}}^2/(4D_{2D})$ at these limits, whereas the 3D free fraction should dominate for $B/F \to 0$ and $w_{xy} \to \infty$ i.e. $\tau_{1/2}$ should converge to ${w_{xy}}^2/(4D_{3D})$ at these limits. A simple ansatz that respects these constraints is:

$$\tau_{1/2} = \left[\frac{1}{D_{2D}} + \frac{4K_D s w_{xy}}{3D_{3D}}\right]\frac{{w_{xy}}^2}{4\left(1 + 4/3K_D s w_{xy}\right)} \tag{5}$$

Unfortunately, eq.(5) was not found to yield correct estimates of D_{2D} or K_D on synthetic data. For instance we found that the prefactor of the sw_{xy} terms in eq.(5) could not be constant since it was decreasing with increasing D_{2D} and decreasing D_{3D}, and was increasing with K_D. Likewise, we found that the dependence of the second addend in the first term of the RHS of eq.(5) was actually non-linear in sw_{xy}. All this information was included into eq.(5), yielding a generalized, more complex phenomenological expression:

$$\tau_{1/2} = \left[\frac{1}{D_{2D}} + \psi(D_{2D}, D_{3D}, K_D)\frac{(sw_{xy})^{n_3}}{D_{3D}}\right]\frac{{w_{xy}}^2}{4\left(1 + \psi(D_{2D}, D_{3D}, K_D) \times (sw_{xy})^{n_3}\right)} \tag{6}$$

with

$$\psi(D_{2D}, D_{3D}, K_D) = \frac{\alpha}{\kappa_{2D} + D_{2D}} \frac{D_{3D}}{\kappa_{3D} + D_{3D}} {K_D}^{n_1} \exp\left((n_2 \ln(K_D))^2\right). \quad (7)$$

In addition to the simulation parameters (D_{2D}, D_{3D}, K_D, s), eq.(6) comprises six free fit parameters: $\alpha, n_1, n_2, n_3, \kappa_{2D}$ and κ_{3D}. We next estimated their values using a simulation-based approach.

We ran a total of 200 svFCS simulations, each with different values for D_{2D}, D_{3D}, k_{on} and k_{off}, thus obtaining 200 diffusion laws $\tau_{1/2} = f({w_{xy}}^2)$ (see supporting figure SI.2 for examples of simulated fluorescence intensity traces and ACF at different waists). For each simulation, K_D was set by random sampling with uniform distribution in the $[0.1, 1]$ interval for 30 % of the simulations (60/200), and in $[1, 10]$ for the rest. Likewise, k_{on} was set by random sampling (uniform distribution) in $[8.3, 167]$. k_{off} was set using $k_{off} = K_D k_{on}$. The diffusion coefficients for each simulation were also chosen at random with uniform distribution in $[1.5, 5]\ \mu m^2.s^{-1}$ for D_{2D} and in $[25, 45]\ \mu m^2.s^{-1}$ for D_{3D}. These 200 simulations were then distributed at random (uniformly) into two groups: 150 simulations for the "training set" and 50 for the "test set". No quadruplet of parameter values ($k_{on}, k_{off}, D_{2D}, D_{3D}$) were common between the two sets.

In a first stage ("training"), we considered the values of D_{2D}, D_{3D}, and K_D as known, and fitted the 150 diffusion laws of the training set with $\alpha, n_1, n_2, n_3, \kappa_{2D}$ and κ_{3D} as free parameters. Figure 1b gives a random selection of the diffusion laws and the fits that were obtained. To improve robustness of the parameter estimation and to obtain variation ranges, we used a "bootstrap aggregation" (bagging) process (30): we generated 100 variants of the training set by randomly sampling, with replacement, the initial training set. Each variant contained 150 simulations, like the original training set. For each variant, we then fitted eq. (6) to the 150 simulations of the variant with $\alpha, n_1, n_2, n_3, \kappa_{2D}$ and κ_{3D} as free parameters and considering D_{2D}, D_{3D}, and K_D as known. This yielded 100 estimations for free parameters, that we averaged. The resulting values were (mean ± s.d.): $\alpha = 1.162 \pm 0.110$, $\kappa_{2D} = 0.723 \pm 0.103$, $\kappa_{3D} = 16.84 \pm 4.18$, $n_1 = 1.104 \pm 0.017$, $n_2 = 0.202 \pm 0.031$ and $n_3 = 1.372 \pm 0.008$.

In a second stage, we fixed the values of $\alpha, n_1, n_2, n_3, \kappa_{2D}$ and κ_{3D} to their mean values above and fitted the phenomenological diffusion law eq.(6) to the 50 diffusion laws of the test set using D_{2D} and K_D as free fit parameters. We considered that D_{3D} can easily be determined from independent experiments (see below) and therefore considered it as known. Note again that the values of D_{2D}, D_{3D}, k_{on} and k_{off} used for this test set have never been used in the training set. A random choice of the obtained fits is shown in Figure 1c1. For each of these 50 simulations, we compare the estimates of $K_D = k_{off}/k_{on}$ and D_{2D} thus obtained with their real values on Fig. 1c2 and c3, respectively. In these panels, the main diagonals (red line) represent perfect estimations. In both cases, the estimation is most of the time close to the diagonal, indicating a good estimation. Although K_D was varied over two orders of magnitude, its estimation is correct over the whole range. We quantified the accuracy of these estimates by the median symmetric accuracy (31): $\mathrm{MSA} = 100 \exp\left(M\left(\left|\ln(y^{pred}/y^{real})\right|\right) - 1\right)$, where M designates the median function and y^{pred} or y^{real} are the predicted or real values of the parameters. We found an estimation error of 11.8 % and 11.5 % for K_D and D_{2D}, respectively. We therefore considered that the phenomenological diffusion law eq.(6) provides us with estimates that are correct enough to be tested using experimental data.

Monitoring HIV-1 Gag partition using FCS diffusion laws on model membranes

In the following, we use equivalently and interchangeably the equilibrium constant $K_D = k_{off}/k_{on}$ and the partition coefficient $K_P = [F]_{eq}/[B]_{eq}$, since they are equivalent for the simple binding equilibrium $\mathrm{F} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \mathrm{B}$. To assess the ability of our empirical analytical expression of FCS diffusion law (eq.6) to correctly estimate the K_P and the membrane bound diffusion coefficient D_{2D} in the case of a 2D/3D binding unbinding kinetics, we performed svFCS of HIV-1-myr(-)-Gag protein kinetics in the presence of supported lipid bilayers (SLBs) as depicted in figure 2a. HIV-1-Gag is known to interact with lipid membranes thanks to a bipartite motif consisting of a myristate and a polybasic domain (Highly basic region, HBR) (for review see (32)). Electrostatic interactions of HIV1-Gag with negatively charged lipids (such as phosphatidylserine (PS) and phosphatidylinositolbisphosphate (PIP_2) have been shown using NMR (17, 18).

We monitored differences occurring in membrane binding using four different lipid compositions containing decreasing amount of negatively charged lipids (namely mix 1 to 4, see the table in figure 2a for the composition and the molar proportion of each lipid). Interaction with the plasma membrane is a key step in the initiation of HIV-1 Gag self-assembly, occurring during the generation of a new virion (33). However, HIV-1 Gag self assembly only occurs above a critical concentration of HIV-1 Gag. (34) shown that for HIV-1 Gag concentration far below 50 nM, no self-assembly was observed on SLB with composition equivalent to our mix1. To avoid self-assembly, in order to stay in the limit of our kinetic model, fluorescent HIV-1 Gag was injected in the bulk phase, above the SLB, to a final concentration of 10 nM. In parallel, fluorescent PE-Cy5.5 was added to a negligible molar proportion into the lipid composition in order to precisely locate the SLB in the axial direction and correctly focus the laser to ensure maximal w_{xy} value at the SLB. Correlograms were then collected at different waists, as illustrated in figure 2b and figure SI.2b to determine the mean value of $\tau_{1/2}$ for each waist, in order to establish the svFCS

a

lipid mol%	PIP_2	PS	PC	PE Cy5.5	Color Code
mix 1	2	28	69.9	0.1	
mix 2	0.5	28	71.4	0.1	
mix 3	0	15	84.9	0.1	
mix 4	0	0	99.9	0.1	

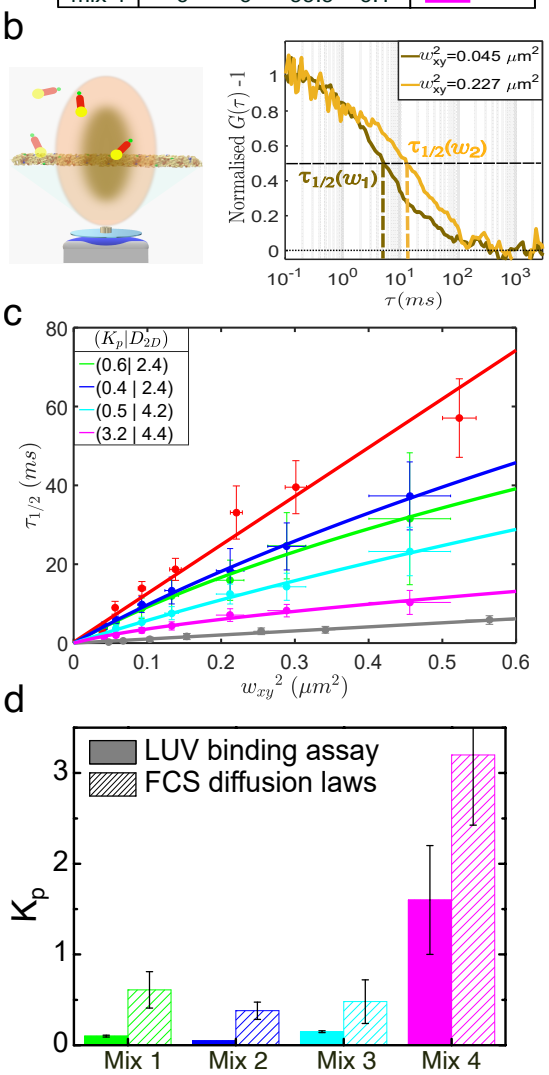

Figure 2: **FCS diffusion laws on model membranes :** a- Different composition of the SLB used here. b- These SLBs were spread on glass coverslips and set in the microscope to perform sv-FCS measurements. Fluorescent myr(-)Gag was introduced in the buffer at the beginning of the experiment. Example of normalized correlograms recorded during an experiment at two different laser waists (210 nm and 480 nm) showing increasing $\tau_{1/2}$ values with increasing waists. In order to establish the diffusion laws, the $\tau_{1/2}$ values are obtained at different waists by fitting the correlogram with equation 3 . c- Experimental FCS diffusion laws of myr(-)Gag obtained at the surface of the different SLB composition (dots) and their fits (lines) using eq.6 with K_P and D_{2D} as free parameters (see table for colour code). In grey, FCS diffusion law of myr(-)Gag in the buffer (dots) and its fit using a linear model (line). In red, FCS diffusion law of Atto647N-PIP_2 in the supported lipid bilayer (dots) and its fit with a linear model (line). d- Comparison of the K_P values obtained from LUV binding assay (full bars) and sv-FCS experiments fits (dashed bars). Values are mean±sd obtained from the fit in the case of sv-FCS and mean±sd of n=2 LUV binding experiments.

diffusion laws. Figure 2c shows the svFCS diffusion laws ($\tau_{1/2} = f({w_{xy}}^2)$) obtained for the four different lipid composition, from the highest charged one (mix 1, blue dots) to the neutral one (mix4, pink dots). Each $\tau_{1/2}$ value plotted on these diffusion laws represent the mean±sd of $30 < n < 110$ correlograms measured on $2 < n < 3$ different SLBs for each ${w_{xy}}^2$. Each ${w_{xy}}^2$

value is itself a mean±sd of 60 different measurements on rhodamine6G standard (see table tablesup for details). Interestingly, from figure2c, it can be seen that each of these experimental diffusion laws are flanked by two different diffusion laws. One, the red dots, is the diffusion law of Atto647N-PI(4,5)P_2 inserted in the mix1 SLB with its linear fit for free diffusion model (red line). The other, in gray dots, was obtained by measuring $\tau_{1/2}$ decorrelation times of labelled myr(-)Gag in the buffer, far away (z= 15μm) from the mix1 SLB, with its linear fit for free diffusion (gray line). These linear fits lead to the following diffusion coefficients: in the case of Atto647N-PI(4,5)P_2, D_{2D}=2.0 ± 0.2 $\mu m^2.s^{-1}$ (mean±sem) and in the case of labelled myr(-)Gag, D_{3D}=24.3 ± 1.5 $\mu m.^2.s^{-1}$. In comparison, each of the diffusion laws obtained for labelled myr(-)Gag when focusing the laser at the SLB does not seem to have a linear tendency. Depending on the lipid composition, they exhibit different curvature as in our simulated diffusion laws (see fig.1b1,b2,c1), suggesting a 2D/3D diffusion plus binding/unbinding process. These diffusion laws were fitted with eq.(6), in order to extract D_{2D}, K_P and D_{3D} values. Fitting the experimental diffusion laws leaving the three parameters free, systematically led to highly erroneous D_{3D} (see figSI.3 for details). For this reason, we fixed D_{3D}=24 $\mu m^2.s^{-1}$ in our fit, as experimentally determined. In this case, we measured respectively D_{2D} = 2.4 ± 0.3 $\mu m^2.s^{-1}$ and D_{2D} = 2.4 ± 0.1 $\mu m^2.s^{-1}$for myr(-)Gag in the case of PI(4,5)P_2 containing SLBs (mix 1& 2). Interestingly this value is similar to the diffusion coefficient value measured for Atto647N-PI(4,5)P_2 in mix 1 using Brownian svFCS diffusion laws. These D_{2D} then have higher values in SLBs lacking of PI(4,5)P_2 and increase with decreasing surface charges from 4.2± 0.4 to 4.4 ± 1.4 $\mu m^2.s^{-1}$. These determinations therefore are in perfect agreement with the previous literature on the effect of PI(4,5)P_2 on myr(-)Gag. Our main goal here was to assess the ability to measure K_P using svFCS. To compare our quantification with results independently obtained with a standard method, we performed LUV binding experiments with the 4 different lipid mixtures of figure2 (see also supplemental figSI.3). Figure 2d shows the K_P values obtained respectively by LUV binding experiments (full bars) and by fitting the FCS diffusion laws (dashed bars) with our empirical expression eq.(6). For every lipid mixture, the K_P determined from our sv-FCS measurements were systematically larger than those obtained by LUV binding assays. As K_P depends on the number of accessible lipids, the different experimental procedures can lead to the systematic deviation observed here. Indeed, although our experiments were designed in such a way that the protein to lipid (P/L) ratio is similar in both approach (LUV binding and sv-FCS on SLBs), SLB preparation necessitates additional steps for small unilamellar vesicules preparation and fusion, where a lot of lipids could be lost, which could lead to systematic overestimate of the K_P values.

However, overall, our estimation of K_P with the empirical svFCS diffusion law follows exactly the same trend as those obtained with LUV binding assay. Indeed, only mix4, that has no negatively charged lipids, exhibits a $K_P > 1$, i.e. a partition of myr(-)Gag in favor of the bulk phase, whereas all the other composition (mix1 to mix3) are found to have myr(-)Gag mainly partitioning at the lipid membrane ($K_P < 1$). Fradin *et al.* (35) previously shown that, close to a fluctuating membrane, parallel to the z-axis of measurement, the experimental D_{3D} value obtained from FCS could be underestimated by 40%. Using D_{3D} values measured far from the membrane might therefore lead to wrong estimation of D_{2D} and K_P. However, in our experimental conditions (membrane perpendicular to the z-axis) we found here D_{2D} and K_P to be very robust to D_{3D} variation (see fig. SI.2)

Quantifying different retroviral Gag proteins binding at the plasma membrane of HEK-293T living cells

We then explored the capability of our svFCS experiments to measure K_P and D_{2D} values in living cells. To this aim, we transfected HEK-293T cells with plasmids expressing either HIV-1-Gag-mCherry or HIV-1-myr(-)GagGFP or HTLV1-GagYFP chimera proteins. Figure 3a shows representative confocal fluorescence microscopy images of the HEK-293T cells expressing these three proteins. As can be seen from the fluorescence intensities in the images, fluorescent protein expression amongst the cells was heterogeneous. We selected the cells with the lowest fluorescence intensity to perform the sv-FCS experiments as HTLV-1 (poorly) as well as HIV-1 Gag (mostly) proteins might self-aggregating at concentration higher than 500nM (36). As for model membranes experiments, this self-aggregation will strongly impact our measurements and lead to kinetics reaction scheme totally different from the one we used to develop our numerical simulations and our empirical analytical solution (eq.6). Based on our ACF fits with eq.3 or eq.4 (see figure SI.5a & b for example of intensity traces and their associated ACF) , we found the average number of molecules (N) in the smallest waist to be between 3 and 10 which corresponds to apparent concentrations ranging between 35 to 120 nM (the average number of molecules found in the different confocal volumes are shown in supplemental figure SI.6a).

As we did for model membranes, we first fitted our svFCS data with eq.6 leaving the three parameters K_P, D_{2D} and D_{3D} free. As found previously, these fits systematically led to erroneous values of D_{3D} (D_{3D} = 10^4 $\mu m^2.s^{-1}$ in the case of HIV-1-Gag-mCherry). To circumvent that pitfall, we therefore measured D_{3D} directly in the cell, by performing svFCS with the laser focused in the cytosol, far from the plasma membrane (see black line and dots in figure 3b and SI.6 for examples). Linear fit of the $\tau_{1/2}$ values led to an estimated average D_{3D} value of 37.2 ± 3 $\mu m^2.s^{-1}$ (mean±sem, n=4 cells) D_{3D} was therefore fixed to 37 $\mu m^2.s^{-1}$ and the data were fitted with K_P and D_{2D} parameters left free.

Figure 3b shows the experimental diffusion laws obtained in 4 cells expressing HTLV-1-GagYFP (redish squares), 12 cells

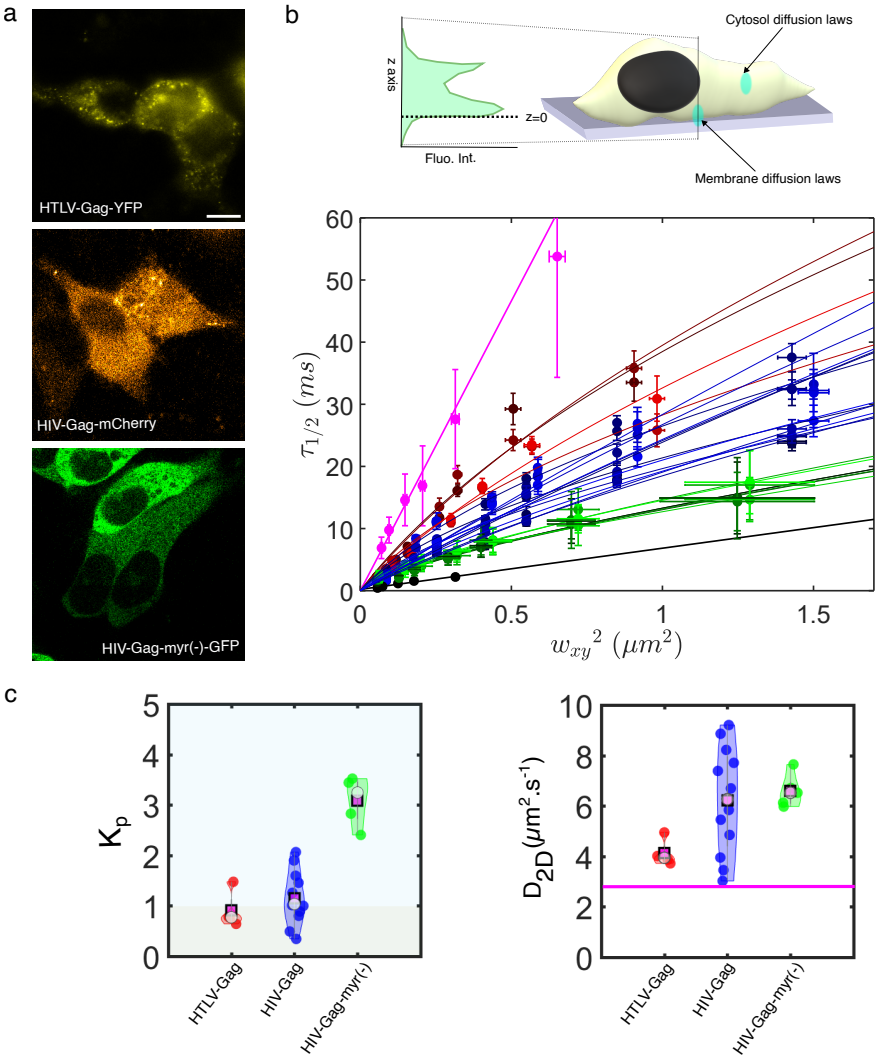

Figure 3: **svFCS diffusion of different viral Gag proteins in HEK-293T cells.** a- Typical confocal images obtained in HEK 293-T cells expressing, from top to bottom, HTLV-1-GagYFP, HIV-1-GagmCherry and HIV-1-myr(-)GagGFP from top to bottom. Scale bar is 10 μm for the three images. b- Experimental diffusion laws obtained in cells expressing the different fluorescent viral Gag proteins. Red squares are HTLV-1-GagYFP data, blue circles are HIV-1GagmCherry data and green diamonds are HIV-1-myr(-)GagGFP data with their respective fits (dots, dashed, dot-dashed lines) using eq.6 with K_P and D_{2D} as free parameters. Pink dots and line represent svFCS data and their fit using a linear model obtained at the plasma membrane of TopFluor®-PC lipid labelled HEK-293T cells (n=2). Black dots and line represent svFCS data and their fit using a linear model of a cytosolic HTLV-1-GagYFP. c- Comparison of the K_P and D_{2D} values obtained from the fit of the svFCS diffusion laws for the three different viral Gag expressed in HEK-293T cells. On the left panel (K_P values for the three proteins) grey area is the area were $[B]_{eq}$>$[F]_{eq}$ and blue the opposite. On the right panel (D_{2D} values for the three proteins), pink line depicts the value of the TopFluor®-PC 2D-diffusion coefficient.

expressing HIV-1-GagmCherry (blueish circles) and 5 cells expressing HIV-1-myr(-)GagGFP (greenish diamonds) with their respective fits as described above. As for model membranes, each of these diffusion laws are flanked by the protein cytosolic free diffusion laws (black dots and line) and by TopFluor®PC labelled plasma membrane (pink dots and line). In these plots, one sees three distinct sets of curves, corresponding to the three different proteins. These differences are confirmed when looking at the K_P and D_{2D} values obtained from the fits (Fig. 3c).

Using a linear fit of the diffusion law of TopFluor®PC in the membrane of living HEK-293T cells, we found D_{2D}=2.7±0.3 $\mu m^2.s^{-1}$. Since this value can appear relatively high for a membrane lipid in cells, we compared it to the value obtained in ref (11), after replotting of their initial diffusion law with correction of the waist using the most recent rhodamine6G diffusion

coefficient value (see methods for details). In this case, we found D_{2D}=2.1±0.2 $\mu m^2.s^{-1}$ for the TopFluor®PC inserted in Cos-7 cells membrane (fig. SI.7), showing the consistency of our value. Contrary to our SLB experiments results, each of the three proteins exhibit membrane diffusion coefficients that are 1.5 to 2 times higher than the diffusion coefficient of TopFluor®PC. HTLV-1-GagYFP was found to have the lower diffusion coefficient of the three, D_{2D}=4.1±0.3 $\mu m^2.s^{-1}$ (mean ± sem, n=4 cells), whereas HIV-1-GagmCherry and HIV-1-myr(-)GagGFP exhibited roughly the same value: 6.2±0.6 (mean ± sem, n=12 cells) and 6.6±0.3 $\mu m^2.s^{-1}$ (mean ± sem, n=5 cells), respectively. Different reasons could account for this systematic overestimation of Gag D_{2D} in cells, compared to the values obtained in SLBs or to that for TopFluor®PC, including physiological ones. Unlike our SLB experiments, the plasma membrane of the cells was not labelled with a fluorescent lipid. The z=0 position (interception of the laser with the plane of the plasma membrane) was therefore roughly determined by inspecting the fluorescence intensity distribution along the z-axis (fig.3b and SI.9). Performing numerical simulations for different K_P, with fixed D_{3D} and D_{2D} , but shifting the focus position above the plasma membrane inside the cells (z>0), we observed that the shift in focus induced a systematic overestimation of D_{2D} (see fig SI.8b). The higher the displacement, the strongest the overestimation. Moreover, as the beam needs to be positioned again for each waist tested, it is always prone to higher uncertainty at large waists since the fluorescence intensity profile does not depict sharp peaks anymore (see fig.SI.9 for an experimental fluorescence intensity z-profile).

The partition of HIV-1-myr(-)GagGFP between cytosol and membrane was found much larger than the other two proteins. We found K_P=3.1±0.2 (mean ± sem, n=5) for HIV-1-myr(-)Gag, which implies that, on average, only 24 % ([21,28]%, (mean±95%CI)) of the HIV-1-myr(-)Gag is bound to the plasma membrane of HEK293T cells. By contrast, the K_P of HIV-1-GagmCherry was found to be 1.2±0.2 (mean ± sem, n=12), corresponding to 45 % of the total HIV-1-GagmCherry bound to the membrane ([38,57] %, (mean±95%CI)). HTLV-1-GagYFP was the most heavily bound protein as on average 52% ([40,79] %, (mean±95%CI)) of the HTLV-1-GagYFP was foudn bound to HEK-293T cells plasma membrane (K_P=0.9±0.2). Interestingly, unlike D_{2D}, we found that our estimation method for K_P was very robust to inaccurate $z = 0$ plan positioning in computer simulations (fig. SI.8c).

DISCUSSION

Monitoring and quantifying molecular motions using FCS mainly rely on the ability to derive analytical solutions in order to fit the autocorrelograms obtained from single spot FCS measurements. Except in simple cases, such as flow or free diffusion, single point FCS often fails to correctly determine and quantify molecular motions in heterogeneous and non-isotropic environment. This failure is frequently (and sometimes abusively) circumvented by the addition of an α exponent, signing for anomalous sub-diffusion motion in the heterogeneous media. However, anomalous sub-diffusion may occur from many different processes which theoretically cannot be uniquely identified or quantified by the sole value of α (see (37, 38) and references therein). In the case of reaction-diffusion dynamics, it has been shown that no simple analytical solutions could be derived to fit single spot FCS experiments (3). This is also what we found here. Deriving an expression systematically required to simplify the dynamics process with different hypotheses, as was also observed in (3). In the 2D/3D diffusion and binding dynamics, we expected to have two limit regimes, namely pure 3D free diffusion (membrane unbound molecule, $k_{off} \gg k_{on}$) and pure 2D diffusion (membrane bound molecule, $k_{off} \ll k_{on}$) as well as different intermediate regimes where the system dynamically equilibrates. Interestingly, the fit of the autocorrelograms with different analytical expressions (3D free diffusion, 2D free diffusion, 2-components diffusion and anomalous diffusion) show that this type of dynamics is hardly correctly fitted by anomalous diffusion if, in order to select the best-fitting expression (as is the case with information-theory derived criteria), the number of free parameters to fit is accounted for. In addition, although K_D could be extracted from single point FCS, we found its estimated value to be at least ten times lower than its real value over a large range of k_{off}& k_{on} values.

A powerful method to correctly probe and characterize the dynamics of molecules in complex media is the spot size variation method (sv-FCS) (5, 11) that establishes FCS diffusion laws. We therefore used computer simulations to generate synthetic FCS diffusion laws for 2D/3D diffusion + binding in different regimes. We first controlled that these simulated diffusion laws were flanked by the two expected limit regimes, namely 2D and 3D diffusion, that both lead to linear FCS diffusion laws. We then generated a set of synthetic diffusion laws in the intermediate regime, by varying the values of D_{2D}, D_{3D} and K_D in our simulations. In these intermediate regimes, we systematically obtained non-linear synthetic FCS diffusion laws. We therefore derived an empirical non-linear analytical expression (eq.6) that, when fitted to our synthetic diffusion laws, provides a a quantitative estimation of K_D and D_{2D}, with relative precision around 12% in the range $0.1 < K_D < 10$ and $1.5 < D_{2D} < 5\,\mu m^2.s^{-1}$.

Since this precision was acceptable, we challenged the capacity of our empirical expression to estimate the change of the membrane apparent partition coefficient K_P and the membrane diffusion D_{2D} occurring when HIV-1 and HTLV-1 Gag proteins and a HIV-1 derivative (myr(-)) binds either to model lipid membranes, with controlled composition, or at the plasma membrane of living HEK-293T cells. In both cases, fitting the experimental FCS diffusion law with (eq.6) did not provide correct estimates

when we tried determine the three parameters of the model simultaneously, namely D_{3D}, D_{2D} and K_P. The simplest explanation, is that the number of beam waists to fit is not sufficient to achieve a good estimation of the three parameters at the same time.

Increasing the number of waists monitored should help in having better fits, but while it is easy to achieve numerically, it remains illusive experimentally. Indeed, correct determination of $\tau_{1/2}$ needs avoiding photobleaching during successive experiments, which cannot be done if we drastically increase the number of waists probed in the same cell. To circumvent that issue, we directly measured the D_{3D} in the bulk (in the case of model membranes) or in the cytosol (in the case of HEK293T cells). Once D_{3D} determined, we could successfully fit the experimental svFCS diffusion laws the two free parameters remaining (K_P and D_{2D}). Interestingly, although Fradin *et al.* (35) shown that the apparent D_{3D} close to a membrane could differ from the D_{3D} measured far from it (as we do here), we found that our estimation of D_{2D} and to a lesser extent K_P, were robust over a large range of D_{3D} values. Note that, as small beam waist values are more sensitive to diffusion of the proteins bound on the membrane, STED-FCS, that address much smaller beam waists than classical confocal FCS, could be another way to circumvent this issue by estimating D_{2D} instead of D_{3D} as the fixed parameter. Combination of the two approaches will certainly help in increasing the accuracy of the method.

We first used SLBs model membranes made of PC, PS and PI(4,5)P_2 with decreasing surface negative charges by tuning the molar ratio of PS and PI(4,5)P_2. We found the membrane diffusion D_{2D} of HIV-1-myr-Gag to be equivalent to that of Atto-647N-PI(4,5)P_2 in the SLB. PI(4,5)P_2 is known to be the specific target lipid of HIV-1-Gag association to the plasma membrane (33) and has been shown to be trapped by HIV-1-Gag in model membranes (34) as well as in living T-cells (39). However, PS is also involved in binding of Gag to the membrane (18, 19). This could explain why we only measured K_P>1 for SLBs lacking negatively charged lipids. Although surprising, as we expected to obtain increasing values of K_P with decreasing amount of PI(4,5)P_2, the K_P values obtained with svFCS diffusion laws exhibited exactly the same trend than those obtained by LUV binding experiments (i.e., K_P>1 only observed with neutral lipids). We also previously showed that lack of myristate strongly decreases the specificity for PI(4,5)P_2 (19). In addition, in a 2:1 PC:PS mol:mol lipid composition, (40) reported no significant change in membrane binding for molar proportion of PI(4,5)P_2 varying from 0 to 2%, in agreement with our own measurements.

Finally, we examined the ability of our method to quantify membrane binding and diffusion of retroviral Gag proteins and its derivatives in HEK-293T cells. In this case, we found D_{2D} to be 4.1±0.3 $\mu m^2.s^{-1}$ and 6.2±0.6$\mu m^2.s^{-1}$ for HTLV-1 and HIV-1 Gag, respectively, questioning the origin of such high D_{2D} values. Unlike in SLB experiments where the lipid membrane is labelled, easying the z-axis laser parking, the cell plasma membrane is not labelled and its location is defined thanks to z-axis fluorescence intensity profile of the protein. This can lead to a high inaccuracy in the laser z-positioning, which in turns strongly impacts the estimation of D_{2D} in the sense of a strong overestimation. Fortunately, we found that this inaccuracy has low impact in the determination of K_P. However, if one aims at correctly estimating D_{2D} values of membrane binding proteins in cells, a stable fluorescent labelling of the plasma membrane (using MemBright for example (41)) will definitely help.

Using a combination of TIRF and fluorescence fluctuation spectroscopy, (21) showed that HTLV-1-Gag has a higher affinity than HIV-1-Gag for the plasma membrane of HeLa cells. Using svFCS diffusion laws, we also measured a lower K_P for HTLV-1-Gag compared to HIV-1-Gag. This reflects the higher affinity of HTLV-1-Gag vs HIV-1-Gag for the plasma membrane of HEK-293T cells and illustrates again the ability of our method to correctly determine this parameter. We also showed that removing myristate from HIV-1-Gag resulted in the re-localization of the Gag protein towards the cytosol, with only 24±3% of the total HIV-1-myr(-)Gag proteins bound to the membrane. Again this value is in good agreement with the results previously obtained by (21). Importantly, the results we obtained here are perfect in line with previous results obtained by different groups using different approaches/techniques.

It was not the main aim of this study to establish a molecular mechanism for retroviral Gag proteins by accurately quantifying their binding to lipid membranes in different conditions. However, we believe it is worth stressing out that not only we have strong differences in the K_P values obtained for HIV-1-myr(-)Gag in model vs cells lipid membranes but also a lower K_P in the case of HIV-1-Gag in cells vs. HIV-1-myr(-)Gag on negatively charged SLBs. Several reasons could account for such discrepancies between model lipid membranes and living cell plasma membranes :

- The accessibility of PI(4,5)P_2 in the plasma membrane of the cells, and more generally of negatively charged lipids, might be strongly decreased compared to model membranes, as many other proteins present in the cells are also known to interact with these lipids. This would result in a screening of the targeted lipids and, consequently an increase of apparent K_P values as the total concentration of available binding sites decreases.

- In the presence of RNA (which are highly present in cells), HIV-1 Gag has been shown to adopt a horseshoe configuration where both MA and NC domain binds to the RNA (42). The membrane binding process is then mediated by the HBR domain of the MA interaction with PI(4,5)P_2, inducing the release of the RNA which stays bound to the NC domain (43). This screening of HIV-1-Gag membrane binding domain in cells might be more important than in our model membrane experiments, where RNA is hardly present, resulting again in an increase of K_D.

- In the absence of RNA, the NC domain has been shown to exhibit significant affinity for negatively charged lipids (44). In our model membranes, where lipids are in large excess compared to RNA, this second lipid binding motif in the NC domain of HIV-1-Gag could compensate the lack of myristate and favor stronger binding to the SLB (lower K_P), while it will be screened in cells by the cytosolic/cellular RNAs.

In this study, we have demonstrated the ability of svFCS diffusion laws to better estimate (apparent) membrane binding coefficients and membrane diffusion coefficients than single spot FCS can. Using a numerical simulation-based approach, we have derived an empirical analytical expression that we then used to fit experimental svFCS data obtained on retroviral Gag proteins binding either to model membranes or to plasma membrane of living cells. Overall, the results we obtained with our method are in perfect line with the ones reported in previous literature, obtained with different methods. This method has been shown to be stable over a large range of K_P values, covering cases where the bound fraction at equilibrium varies from 90% down to 5 % of the total protein . We conclude that our results provide a non-invasive and direct way to fairly estimate, in the same experiment, membrane binding and diffusion coefficients (with a correct z-positioning) in living cells.

AUTHOR CONTRIBUTIONS

CF, HB & DM designed the project and the experiments. HB developed the computer code, performed the numerical simulations and derived the empirical expression. AM, EB and JN performed the FCS experiments. AM, EB and JN fitted experimental data. RD & CA prepared the cell samples, EB the SLBs. JM purified and provided the myr(-)Gag protein. HB & CF draft the manuscript assisted by AM, EB, CA & DM.

ACKNOWLEDGMENTS

The authors acknowledge the Imabio CNRS (GdR Imabio) consortium for their continuous support of the project and for inititally granting EB. AM Ph.D. fellowship is granted by CNRS Prime 80. EB was then granted by CNRS. CA is a recipient of Université Montpellier Ph.D. fellowship. RD Ph.D. fellowship is granted by Sidaction. The project was initially granted by ANR Fluobuds and then by CNRS. Authors acknowledge Montpellier RIO Imaging (MRI, Biocampus, UAR CNRS) microscopy facility. Authors acknowledge C. Eggeling, H.de Rocquigny, M. May, D. Derse and N. Jouvenet for the gift of the different fluorescent lipids, dyes and plasmids.

DECLARATION OF INTEREST

The authors declare no competing interests.

REFERENCES

1. Sezgin, E., and P. Schwille, 2011. Fluorescence Techniques to Study Lipid Dynamics. *Cold Spring Harbor Perspectives in Biology* 3:a009803–a009803.

2. Magde, D., E. Elson, and W. W. Webb, 1972. Thermodynamic Fluctuations in a Reacting System—Measurement by Fluorescence Correlation Spectroscopy. *Physical Review Letters* 29:705–708.

3. Michelman-Ribeiro, A., D. Mazza, T. Rosales, T. J. Stasevich, H. Boukari, V. Rishi, C. Vinson, J. R. Knutson, and J. G. McNally, 2009. Direct measurement of association and dissociation rates of DNA binding in live cells by fluorescence correlation spectroscopy. *Biophys J* 97:337–346.

4. Dufourt, J., A. Trullo, J. Hunter, C. Fernandez, J. Lazaro, M. Dejean, L. Morales, S. Nait-Amer, K. N. Schulz, M. M. Harrison, C. Favard, O. Radulescu, and M. Lagha, 2018. Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs. *Nature Communications* 9.

5. Wawrezinieck, L., H. Rigneault, D. Marguet, and P. Lenne, 2005. Fluorescence correlation spectroscopy diffusion laws to probe the submicron cell membrane organization. *Biophys. J.* 89:4029–4042.

6. Eggeling, C., C. Ringemann, R. Medda, G. Schwarzmann, K. Sandhoff, S. Polyakova, V. N. Belov, B. Hein, C. von Middendorff, A. Schoenle, and S. W. Hell, 2009. Direct observation of the nanoscale dynamics of membrane lipids in a living cell. *Nature* 457:1159–1162.

7. Wenger, J., F. Conchonaud, J. Dintinger, L. Wawrezinieck, T. Ebbesen, H. Rigneault, D. Marguet, and P.-F. Lenne, 2007. Diffusion analysis within single nanometric apertures reveals the ultrafine cell membrane organization. *Biophys. J.* 92:913–919.

8. Winkler, P. M., R. Regmi, V. Flauraud, J. Brugger, H. Rigneault, J. Wenger, and M. F. García-Parajo, 2017. Transient Nanoscopic Phase Separation in Biological Lipid Membranes Resolved by Planar Plasmonic Antennas. *ACS Nano* 11:7241–7250.

9. Bag, N., D. H. X. Yap, and T. Wohland, 2014. Temperature dependence of diffusion in model and live cell membranes characterized by imaging fluorescence correlation spectroscopy. *Biochimica et biophysica acta* 1838:802–813.

10. Veerapathiran, S., and T. Wohland, 2018. The imaging FCS diffusion law in the presence of multiple diffusive modes. *Methods* 140-141:140–150.

11. Lenne, P.-F., L. Wawrezinieck, F. Conchonaud, O. Wurtz, A. Boned, X.-J. Guo, H. Rigneault, H.-T. He, and D. Marguet, 2006. Dynamic molecular confinement in the plasma membrane by microdomains and the cytoskeleton meshwork. *EMBO J.* 25:3245–3256.

12. Favard, C., J. Wenger, P.-F. Lenne, and H. Rigneault, 2011. FCS diffusion laws in two-phase lipid membranes: determination of domain mean size by experiments and Monte Carlo simulations. *Biophysical journal* 100:1242–1251.

13. Gupta, A., I. Y. Phang, and T. Wohland, 2020. To Hop or not to Hop: Exceptions in the FCS Diffusion Law. *Biophysical Journal* 118:2434–2447.

14. Burns, M. C., M. Nouri, and S. L. Veatch, 2016. Spot size variation FCS in simulations of the 2D Ising model. *Journal of Physics D: Applied Physics* 49:214001.

15. Stolle, M. D., and C. Fradin, 2019. Anomalous Diffusion in Inverted Variable-Lengthscale Fluorescence Correlation Spectroscopy. *Biophysical Journal* 116:791–806.

16. Saad, J. S., J. Miller, J. Tai, A. Kim, R. H. Ghanam, and M. F. Summers, 2006. Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proceedings of the National Academy of Sciences* 103:11364–11369.

17. Mercredi, P. Y., N. Bucca, B. Loeliger, C. R. Gaines, M. Mehta, P. Bhargava, P. R. Tedbury, L. Charlier, N. Floquet, D. Muriaux, C. Favard, C. R. Sanders, E. O. Freed, J. Marchant, and M. F. Summers, 2016. Structural and Molecular Determinants of Membrane Binding by the HIV-1 Matrix Protein. *Journal of Molecular Biology* 428:1637–1655.

18. Vlach, J., and J. S. Saad, 2013. Trio engagement via plasma membrane phospholipids and the myristoyl moiety governs HIV-1 matrix binding to bilayers. *Proceedings of the National Academy of Sciences* 110:3525–3530.

19. Charlier, L., M. Louet, L. Chaloin, P. Fuchs, J. Martinez, D. Muriaux, C. Favard, and N. Floquet, 2014. Coarse-Grained Simulations of the HIV-1 Matrix Protein Anchoring: Revisiting Its Assembly on Membrane Domains. *Biophysical Journal* 106:577–585.

20. Bryant, M., and L. Ratner, 1990. Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proceedings of the National Academy of Sciences* 87:523–527.

21. Fogarty, K. H., S. Berk, I. F. Grigsby, Y. Chen, L. M. Mansky, and J. D. Mueller, 2014. Interrelationship between Cytoplasmic Retroviral Gag Concentration and Gag–Membrane Association. *Journal of Molecular Biology* 426:1611–1624.

22. Boutant, E., J. Bonzi, H. Anton, M. B. Nasim, R. Cathagne, E. Réal, D. Dujardin, P. Carl, P. Didier, J.-C. Paillart, R. Marquet, Y. Mély, H. de Rocquigny, and S. Bernacchi, 2020. Zinc Fingers in HIV-1 Gag Precursor Are Not Equivalent for gRNA Recruitment at the Plasma Membrane. *Biophysical Journal* 119:419–433.

23. Mazurov, D., G. Heidecker, and D. Derse, 2006. HTLV-1 Gag protein associates with CD82 tetraspanin microdomains at the plasma membrane. *Virology* 346:194–204.

24. Jouvenet, N., P. D. Bieniasz, and S. M. Simon, 2008. Imaging the biogenesis of individual HIV-1 virions in live cells. *Nature* 454:236–240.

25. Tanwar, H. S., K. K. Khoo, M. Garvey, L. Waddington, A. Leis, M. Hijnen, T. Velkov, G. J. Dumsday, W. J. McKinstry, and J. Mak, 2017. The thermodynamics of Pr55Gag-RNA interaction regulate the assembly of HIV. *PLOS Pathogens* 13:e1006221.

26. Müller, C. B., A. Loman, V. Pacheco, F. Koberling, D. Willbold, W. Richtering, and J. Enderlein, 2008. Precise measurement of diffusion by multi-color dual-focus fluorescence correlation spectroscopy. *EPL (Europhysics Letters)* 83:46001.

27. Hendrix, J., C. Flors, P. Dedecker, J. Hofkens, and Y. Engelborghs, 2008. Dark States in Monomeric Red Fluorescent Proteins Studied by Fluorescence Correlation and Single Molecule Spectroscopy. *Biophysical Journal* 94:4103–4113.

28. Cavanaugh, J. E., 1997. Unifying the derivations for the Akaike and corrected Akaike information criteria. *Statistics & Probability Letters* 33:201–208. https://www.sciencedirect.com/science/article/pii/S0167715296001289.

29. Krichevsky, O., and G. Bonnet, 2002. Fluorescence correlation spectroscopy: the technique and its applications. *Reports on Progress in Physics* 65:251.

30. Opitz, D., and R. Maclin, 1999. Popular Ensemble Methods: An Empirical Study. *J. Artif. Int. Res.* 11:169–198.

31. Morley, S. K., T. V. Brito, and D. T. Welling, 2018. Measures of Model Performance Based On the Log Accuracy Ratio. *Space Weather* 16:69–88. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017SW001669.

32. Freed, E. O., 2015. HIV-1 assembly, release and maturation. *Nature Reviews Microbiology* 13:484–496.

33. Ono, A., S. D. Ablan, S. J. Lockett, K. Nagashima, and E. O. Freed, 2004. Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. *Proceedings of the National Academy of Sciences of the United States of America* 101:14889–14894.

34. Yandrapalli, N., Q. Lubart, H. S. Tanwar, C. Picart, J. Mak, D. Muriaux, and C. Favard, 2016. Self assembly of HIV-1 Gag protein on lipid membranes generates PI(4,5)P2 /Cholesterol nanoclusters. *Scientific reports* 6:39332.

35. Fradin, C., A. Abu-Arish, R. Granek, and M. Elbaum, 2003. Fluorescence Correlation Spectroscopy Close to a Fluctuating Membrane. *Biophysical Journal* 84:2005–2020.

36. Fogarty, K. H., Y. Chen, I. F. Grigsby, P. J. Macdonald, E. M. Smith, J. L. Johnson, J. M. Rawson, L. M. Mansky, and J. D. Mueller, 2011. Characterization of Cytoplasmic Gag-Gag Interactions by Dual-Color Z-Scan Fluorescence Fluctuation Spectroscopy. *Biophysical Journal* 100:1587–1595.

37. Höfling, F., and T. Franosch, 2013. Anomalous transport in the crowded world of biological cells. *Reports on Progress in Physics* 76:046602. https://doi.org/10.1088/0034-4885/76/4/046602.

38. Metzler, R., J.-H. Jeon, A. G. Cherstvy, and E. Barkai, 2014. Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Phys. Chem. Chem. Phys.* 16:24128–24164.

39. Favard, C., J. Chojnacki, P. Merida, N. Yandrapalli, J. Mak, C. Eggeling, and D. Muriaux, 2019. HIV-1 Gag specifically restricts PI(4,5)P2 and cholesterol mobility in living cells creating a nanodomain platform for virus assembly. *Science Advances* 5.

40. Chukkapalli, V., I. B. Hogue, V. Boyko, W.-S. Hu, and A. Ono, 2008. Interaction between the Human Immunodeficiency Virus Type 1 Gag Matrix Domain and Phosphatidylinositol-(4,5)-Bisphosphate Is Essential for Efficient Gag Membrane Binding. *Journal of Virology* 82:2405–2417.

41. Collot, M., P. Ashokkumar, H. Anton, E. Boutant, O. Faklaris, T. Galli, Y. Mély, L. Danglot, and A. S. Klymchenko, 2019. MemBright: A Family of Fluorescent Membrane Probes for Advanced Cellular Imaging and Neuroscience. *Cell Chemical Biology* 26:600–614.e7.

42. Munro, J. B., A. Nath, M. Färber, S. A. K. Datta, A. Rein, E. Rhoades, and W. Mothes, 2014. A Conformational Transition Observed in Single HIV-1 Gag Molecules during in Vitro Assembly of Virus-Like Particles. *Journal of Virology* 88:3577–3585.

618 43. Chukkapalli, V., S. J. Oh, and A. Ono, 2010. Opposing mechanisms involving RNA and lipids regulate HIV-1 Gag
619 membrane binding through the highly basic region of the matrix domain. *Proceedings of the National Academy of Sciences*
620 107:1600–1605.

621 44. Kempf, N., V. Postupalenko, S. Bora, P. Didier, Y. Arntz, H. de Rocquigny, and Y. Mély, 2015. The HIV-1 Nucleocapsid
622 Protein Recruits Negatively Charged Lipids To Ensure Its Optimal Binding to Lipid Membranes. *Journal of Virology*
623 89:1756–1767.

SUPPORTING INFORMATION

SI1. Derivation of the auto-correlation function for the single-spot case

Main Hypotheses Here we simplify the problem by considering that the membrane binding and diffusion sites are not restricted to at $z = 0$ but are homogeneously distributed in the 3D bulk. Moreover, we neglect the reduction of dimensionality and consider that the bound molecules diffuse in the same 3D space as the free ones, but with a different diffusion coefficient. Under these assumptions, the problem is now that of a protein that reversibly switches between two forms: (F)ree and (B)ound, and where the two forms have non-identical but non-zero diffusion coefficients: F $\underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}}$ B, with $D_{3D} > D_{2D} > 0$.

Assuming Brownian motion in 3D (for *both free and membrane-bound* Gag), yields the following reaction diffusion system:

$$\begin{aligned} \partial_t F &= D_{3D}\nabla^2 F - k_{\text{on}} F + k_{\text{off}} B \\ \partial_t B &= D_{2D}\nabla^2 B + k_{\text{on}} F - k_{\text{off}} B \end{aligned} \tag{SI.1}$$

where F and B refer to the local concentration of the two species.

General case Noting f and b the fluctuations of F and B, respectively (i.e. $F(\vec{r},t) = f(\vec{r},t) + \langle F \rangle$ and $G(\vec{r},t) = g(\vec{r},t) + \langle G \rangle$ where $\langle F \rangle$ and $\langle G \rangle$ are the average concentrations over the volume), Eq. SI.1 yields:

$$\begin{aligned} \partial_t f &= D_{3D}\nabla^2 f - k_{\text{on}} f + k_{\text{off}} b \\ \partial_t b &= D_{2D}\nabla^2 b + k_{\text{on}} f - k_{\text{off}} b \end{aligned} \tag{SI.2}$$

The coefficient matrix of the Fourier transform of Eq. SI.2, $M_{ij} = K_{ij} - D_i q^2 \delta_{ij}$ (where K is the stoechiometry matrix and $q^2 = {q_x}^2 + {q_y}^2 + {q_z}^2$ is the square norm of the Fourier variable) is given by:

$$M = \begin{bmatrix} -k_{\text{on}} - q^2 D_{3D} & k_{\text{off}} \\ k_{\text{on}} & -k_{\text{off}} - q^2 D_{2D} \end{bmatrix} \tag{SI.3}$$

whose eigenvalues are:

$$\begin{aligned} \lambda_1 &= \left(-k_{\text{on}} - k_{\text{off}} - (D_{3D} + D_{2D})q^2 - \sqrt{\Delta}\right)/2 \\ \lambda_2 &= \left(-k_{\text{on}} - k_{\text{off}} - (D_{3D} + D_{2D})q^2 + \sqrt{\Delta}\right)/2 \end{aligned} \tag{SI.4}$$

with $\Delta = q^4(D_{3D} - D_{2D})^2 + (k_{\text{on}} + k_{\text{off}})^2 + 2q^2(D_{3D} - D_{2D})(k_{\text{on}} - k_{\text{off}})$.

The corresponding eigenvectors are

$$v_1 = \begin{pmatrix} 1 \\ \frac{k_{\text{on}} - k_{\text{off}} + q^2(D_{3D} - D_{2D}) - \sqrt{\Delta}}{2k_{\text{off}}} \end{pmatrix}$$

and

$$v_2 = \begin{pmatrix} 1 \\ \frac{k_{\text{on}} - k_{\text{off}} + q^2(D_{3D} - D_{2D}) + \sqrt{\Delta}}{2k_{\text{off}}} \end{pmatrix}$$

Their inverse eigenvectors are

$$v_1^{(-1)} = \begin{pmatrix} k_{\text{on}} - k_{\text{off}} + q^2(D_{3D} - D_{2D}) + \sqrt{\Delta} \\ -2k_{\text{off}} \end{pmatrix} \left(2\sqrt{\Delta}\right)^{-1}$$

and

$$v_2^{(-1)} = \begin{pmatrix} k_{\text{off}} - k_{\text{on}} - q^2(D_{3D} - D_{2D}) + \sqrt{\Delta} \\ 2k_{\text{off}} \end{pmatrix} \left(2\sqrt{\Delta}\right)^{-1}$$

.

Assuming a Gaussian point-spread function:

$$I(x,y,z) = I_0 \exp\left(-\frac{-2(x^2+y^2)}{{w_{xy}}^2} - \frac{-2z^2}{{w_z}^2}\right) \tag{SI.5}$$

the autocorrelation of the fluctuations $G(\tau)$ can be generically computed (29) as :

$$\begin{aligned} G(\tau) &= \frac{1}{(2\pi)^3 \left(\sum_i Q_i \bar{C}_i\right)^2} \int d^3\vec{q} \exp\left(-\frac{{w_{xy}}^2({q_x}^2+{q_y}^2)}{4\mu} - \frac{{w_z}^2{q_z}^2}{4\mu}\right) \\ &\times \sum_{j,l=1}^{2} Q_j Q_l \bar{C}_j \left(v_1(l) e^{\lambda_1 \tau} v_1^{(-1)}(j) + v_2(l) e^{\lambda_2 \tau} v_2^{(-1)}(j)\right) \end{aligned} \tag{SI.6}$$

where $\mu = 1$ or 2 for one- or two-photon fluorescence, respectively, $Q_i, i = 1,2$ are the fluorescence cross-sections of F and B, respectively, $\bar{C}_i, i = 1,2$ are their average distribution in the sample, i.e. $\bar{C}_1 = \langle F \rangle$ and $\bar{C}_2 = \langle B \rangle$, and $u(i)$ represents the ith element of vector u.

Simplifications *Fast free diffusion*. Eq.(SI.6) cannot be solved analytically without further simplifications. In particular, one can assume that the diffusion of F in the bulk is too fast to contribute a signal in FCS. In this case $Q_1 = Q_F \approx 0$ and the only non-vanishing terms in the sum of the rhs of eq.(SI.6) are for $j = l = 2$. Therefore in the specific case studied here, eq.(SI.6) becomes

$$\begin{aligned} G(\tau) &= \frac{1}{16\pi^3 \langle B \rangle^2} \int d^3\vec{q} \, \frac{\left(k_{\text{off}} - k_{\text{on}} - q^2(D_{3D} - D_{2D}) + \sqrt{\Delta}\right) \exp\left(\lambda_1 \tau - \frac{{w_{xy}}^2({q_x}^2+{q_y}^2)+{w_z}^2{q_z}^2}{4\mu}\right)}{\sqrt{\Delta}} \\ &+ \frac{\left(k_{\text{on}} - k_{\text{off}} + q^2(D_{3D} - D_{2D}) + \sqrt{\Delta}\right) \exp\left(\lambda_2 \tau - \frac{{w_{xy}}^2({q_x}^2+{q_y}^2)+{w_z}^2{q_z}^2}{4\mu}\right)}{\sqrt{\Delta}} \end{aligned} \tag{SI.7}$$

where the eigenvalues λ_i are given by eq.(SI.4). Eq.(SI.7) greatly simplifies when e.g. $D_{3D} = D_{2D}$ since in this case, $\sqrt{\Delta} = k_{\text{off}} + k_{\text{on}}$ so that all the integral variables q_i are found inside the exponentials. The integral eq.(SI.7) can be analytically computed in this case and leads to the formula given in (29). In our case $D_{3D} > D_{2D} > 0$, though, we have not been able to compute the integral analytically.

In (3), the authors explore a series of parameter regimes where the expression for $G(\tau)$ can be further simplified.
The Hybrid Regime. This corresponds to the case where binding is much larger than unbinding, $k_{\text{on}} \gg k_{\text{off}}$. In particular, this can be used to support the approximation $k_{\text{on}} - k_{\text{off}} \approx k_{\text{on}} + k_{\text{off}}$. In this case, Δ simplifies according to

$$\sqrt{\Delta} \approx q^2(D_{3D} - D_{2D}) + k_{\text{on}} + k_{\text{off}}$$

and

$$\begin{aligned} G(\tau) &= \frac{1}{16\pi^3 \langle B \rangle^2} \int d^3\vec{q} \, \frac{2k_{\text{off}} \exp\left(-\tau(k_{\text{on}} + k_{\text{off}} + q^2 D_{3D}) - \frac{{w_{xy}}^2({q_x}^2+{q_y}^2)+{w_z}^2{q_z}^2}{4\mu}\right)}{q^2(D_{3D} - D_{2D}) + k_{\text{on}} + k_{\text{off}}} \\ &+ \frac{\left(2k_{\text{on}} + 2q^2(D_{3D} - D_{2D})\right) \exp\left(-D_{2D} q^2 \tau - \frac{{w_{xy}}^2({q_x}^2+{q_y}^2)+{w_z}^2{q_z}^2}{4\mu}\right)}{q^2(D_{3D} - D_{2D}) + k_{\text{on}} + k_{\text{off}}} \end{aligned} \tag{SI.8}$$

Unfortunately, this simpler expression is still too complex to permit the calculation of an analytical expression.
Other Regimes None of the other regimes explored in (3) are applicable when $D_{2D} > 0$.

- The "pure diffusion" regime ($k_{\text{on}} \ll k_{\text{off}}$) implies that all the Gag molecules are free. The case however is trivial, and corresponds to classical 3D diffusion in the bulk.
- The "effective diffusion" regime ($\tau_D = {w_{xy}}^2/(4\mu D) \gg 1/k_{\text{on}}$) assumes that diffusion is slow enough that the binding reaction is at equilibrium (locally) everywhere and allows to reduce the reaction-diffusion system eq.(SI.2) to a single diffusion equation (no coupling to reaction). This reduction is however not possible when $D_{2D} > 0$
- the "reaction dominant" regime ($\tau_D \ll 1/k_{\text{on}}$) assumes that the characteristic time for a Gag molecule to diffuse across the focal volume is much shorter than the characteristic time to bind the membrane. Here again, the simplification is critically based on the fact that one of the species does not diffuse and cannot be applied in our case.

SI2. Supporting Figure

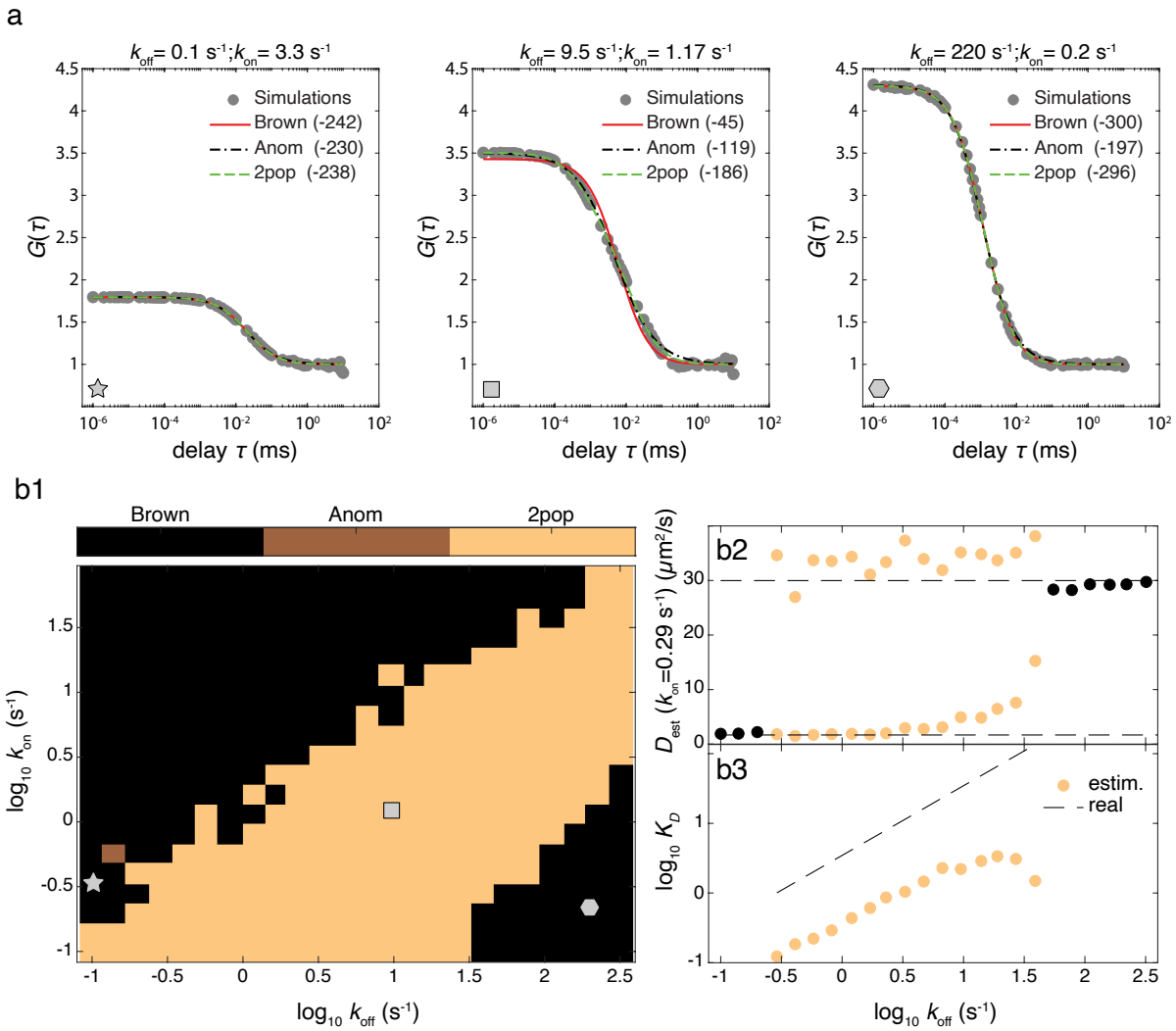

Figure SI.1: **Fitting the auto-correlation function with adhoc expressions**. Molecular diffusion in the cytoplasm and binding to the plasma membrane was simulated as described in Materials and Methods for a range of k_{off} and k_{on} values. We fixed the other parameters to constant values, in particular $w_{xy} = 0.379\,\mu$m, $D_{2D} = 1.7$ and $D_{3D} = 30\,\mu\text{m}^2/\text{s}$. Illustration of some of the obtained auto-correlations are shown in (a) (full gray circles). Every auto-correlation functions were fitted by three models: a single population of pure 3D Brownian motion (full red line), a mix of two populations with pure 3D Brownian motion (dashed green line) or anomalous diffusion (dashed-dotted black line). The best model was selected by comparison of the corrected Akaike information criterias (28) (AICc, values shown in parenthesis, see Materials and Methods). The colormap (b1) shows what is the best model for the range of k_{off} and k_{on} values explored. The gray star, rectangle and hexagon locate the illustrations shown in (a). With k_{on} fixed to 0.29 s^{-1}, we also show the resulting estimation by the best model of the diffusion coefficient(s) (b2) and equilibrium constant $K_D = k_{\text{off}}/k_{\text{on}}$ (b3). The real values of diffusion and the equilibrium constants are shown as dashed lines.

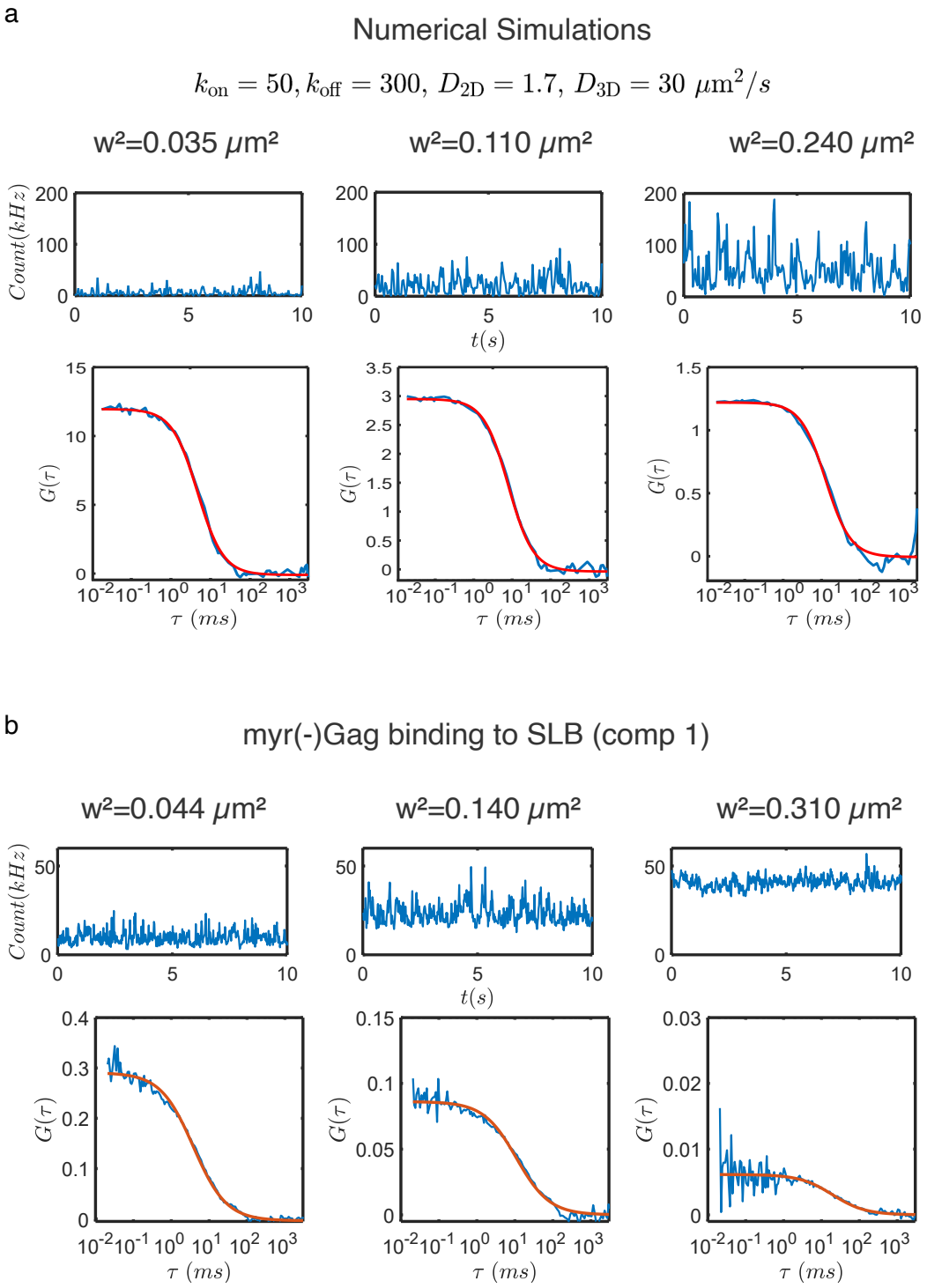

Figure SI.2: **Numerical and Experimental fluorescence intensity time trace and auto-correlation functions** a- Example of time traces and ACF obtained by numerical simulations for increasing waists. As expected in time traces, the average photon count (in kHz) increases with increasing waists (from left to right). Below are depicted the corresponding ACFs (in blue) and their fit using equation 3 (in red). From these ACF it can be seen that G(0) decreases (reflecting an increase in molecules number in the observed volume) while $\tau_{1/2}$ increases with increasing waists. b- Example of time traces and ACF acquired in myr(-)Gag binding to SLBs experiments. ACF are also fitted using eq.3. As for the numerical simulations, average fluorescence intensity, G(0) and $\tau_{1/2}$ increase with increasing waists.

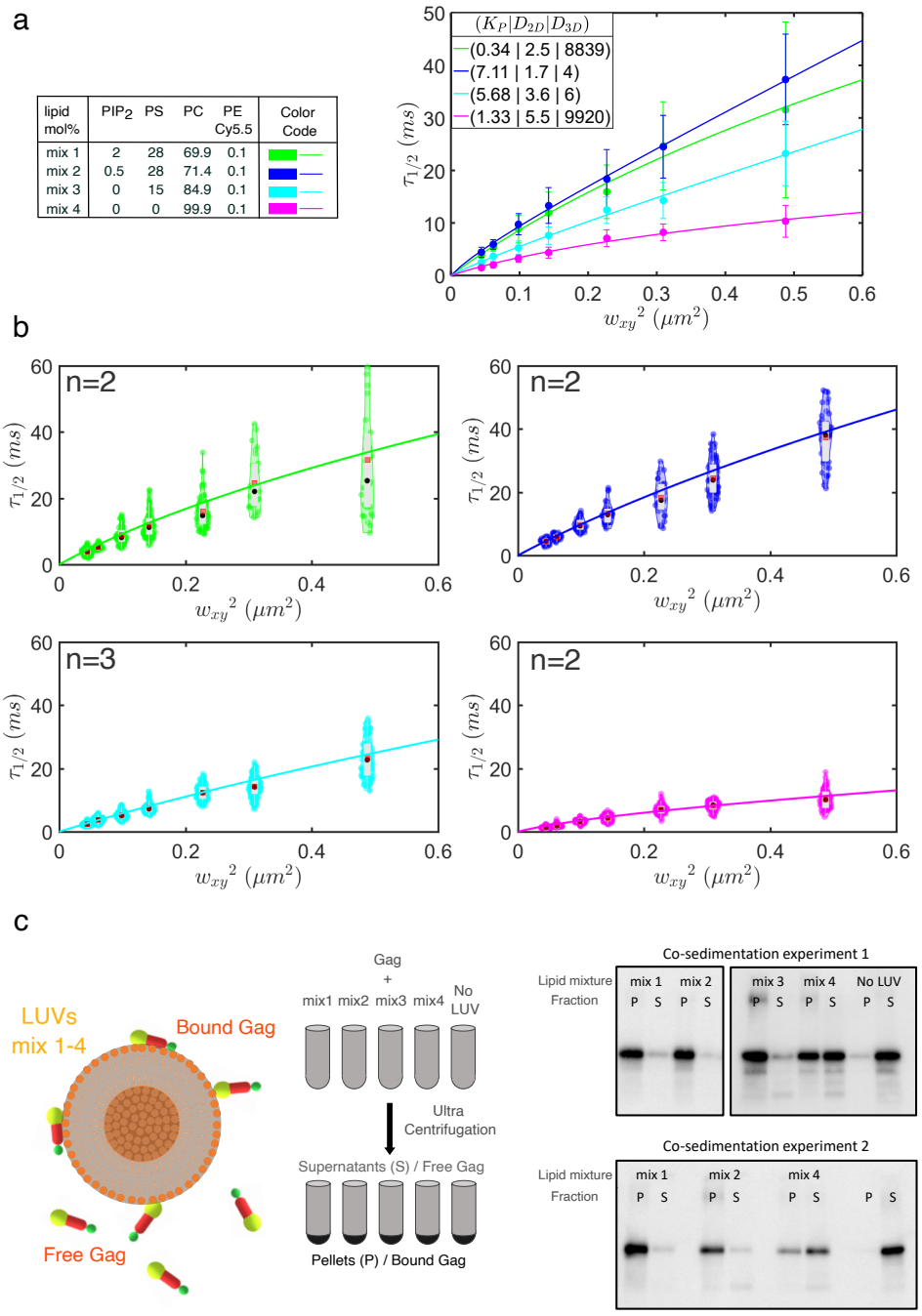

Figure SI.3: **Supplemental results for Gag binding to model membranes** a- Fit of the FCS diffusion laws obtained with the different lipid mixes (see table on the left for lipid composition) using a 3 free parameters model (K_P, D_{2D}, D_{3D}). Although K_P shows increasing values with decreasing negative surface charge and PI(4,5)P_2, the D_{3D} values obtained were totally incoherent. b- Diffusion laws represented with violin plots of the whole set of $\tau_{1/2}$ (also display in table SI.1) obtained by fitting the experimental ACF with eq.3 for the four different lipid mixture (n=2 to 3 different experiments, pooled together in the violin plot). For each waist, the median value is represented in black, the mean value in red. Except for the two last waists data of mix 1, where mean differs from median value, the distributions of the $\tau_{1/2}$ exhibited mean values similar to median values. No clear bimodal distribution was observed, demonstrating reproducibility of the experiment. The fits of the diffusion laws using eq. 6 are plotted as full lines. c- Principle of the LUV binding experiment (left) and images of the 2 western blots obtained in two different experiments and used to quantify the K_P (see methods for details)

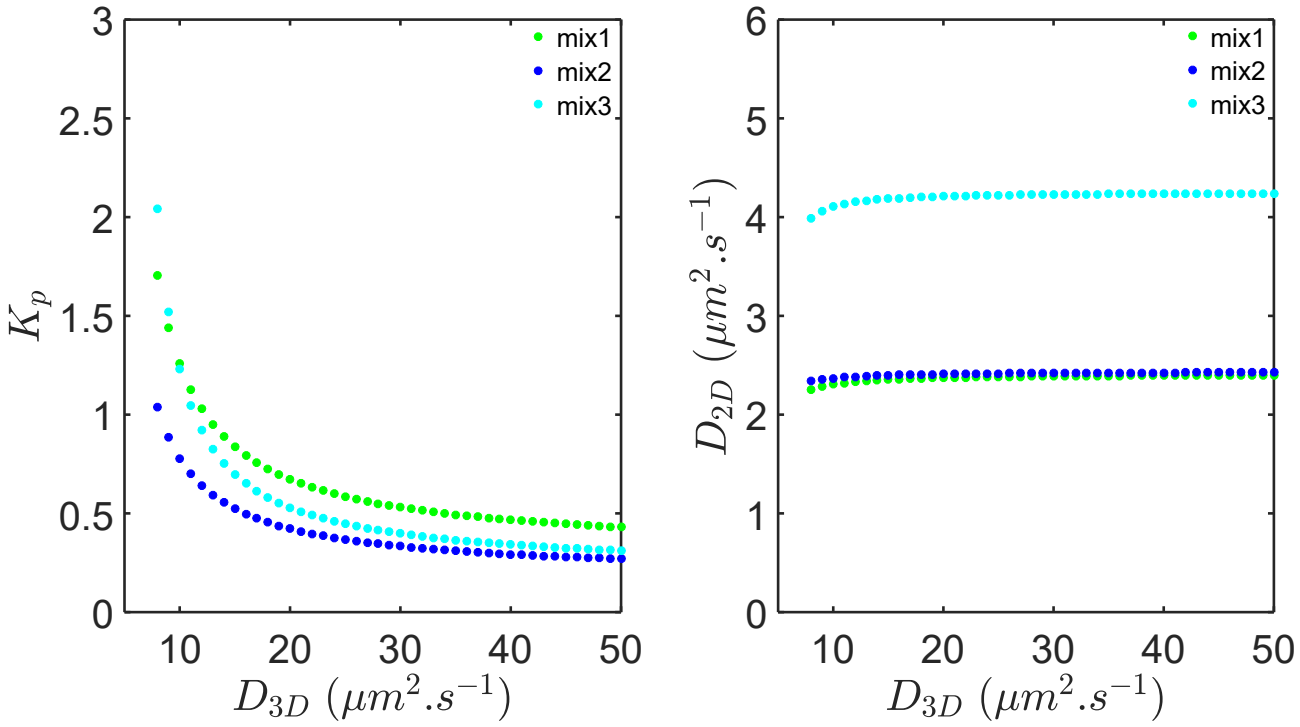

Figure SI.4: **Changes in K_P and D_{2D} values as a function of D_{3D}** The estimation of K_P and D_{2D} are obtained by fitting diffusion laws with fixed values of D_{3D}, this value being determined by independent experiments in the cytosol/bulk, far from the membrane. As it has been suggested that the value of D_{3D} determined by FCS in the vicinity of a membrane could differ by 40% from its bulk value (35), we quantified here how our estimation of K_P and D_{2D} depends of the estimated value of D_{3D}. The left panel shows the changes in estimated K_P as a function of different estimated D_{3D} for the 3 first lipid mixes (where bound fraction of myr(-)Gag is superior to the unbound fraction). The right panel shows the changes in estimated D_{2D} as a function of imposed D_{3D} values.

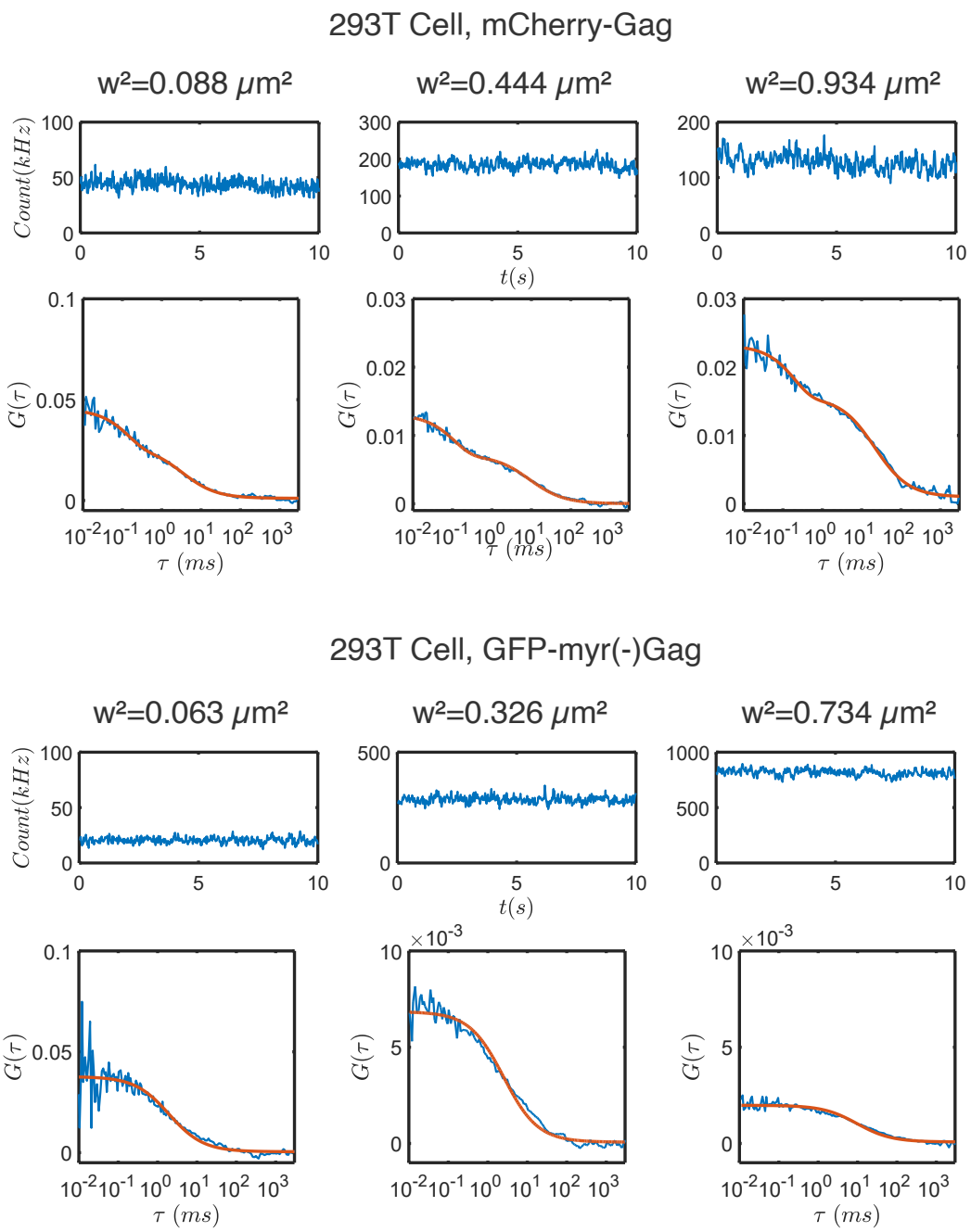

Figure SI.5: **Experimental fluorescence intensity time trace and auto-correlation functions observed in live cells.** a- Example of time traces and ACF obtained in HEK-293T cells expressing HIV-1-Gag-mCherry for increasing waists (from left to right). The bottom line shows the corresponding ACFs (in blue) and their fit (in red) using the equation 4. These plots evidence the presence of a first component in the ACF corresponding to the blinking of the mCherry. b- Examples of time traces and ACF obtained in HEK 293T cells expressing myr(-)Gag GFP. ACF are fit using eq.3.

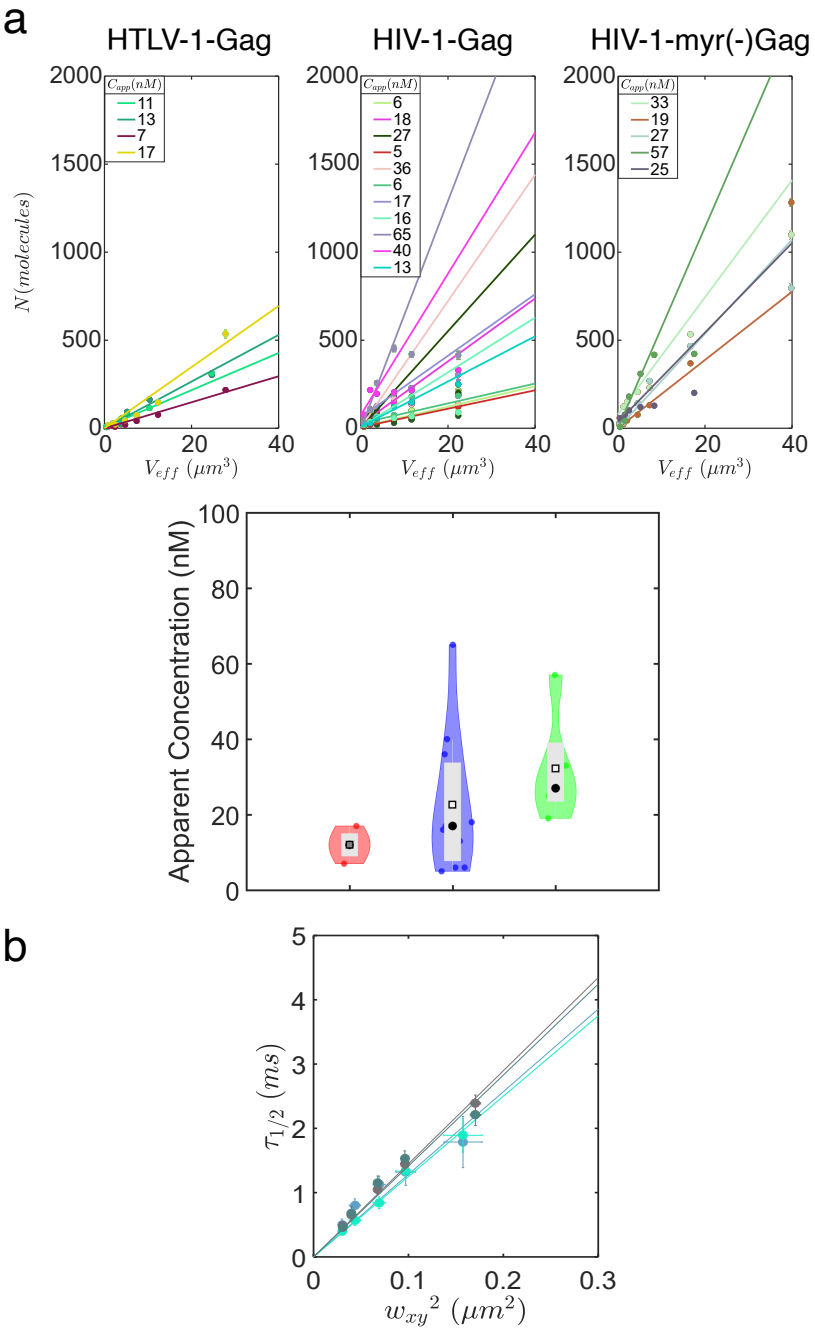

Figure SI.6: **Supplemental results for Gag binding to cell plasma membranes** a- Evolution of the apparent number of molecules (N) obtained from the fit of the, ACF with eq.3 or eq. 4 (see methods), as a function of the effective volume (V_{eff}) for all the cells analyzed here. The slope of each curve is the apparent concentration, constantly increasing from HIV-1-Gag-mCherry to HIV-1-myr(-)Gag-GFP, but still remaining (far) below the threshold concentration of self-assembly (see text for detailed explanations). Apparent concentration is systematically found smaller at larger waists compared to the smaller ones. This is mainly due to the fact that at big waists, the apparent number of molecules N is larger, and thus difficult to be correctly estimated. This leads to and underestimation of apparent concentrations. A violin plot of the different concentrations obtained from the fit is shown on the bottom. b- Example of FCS diffusion laws obtained in the cytosol of 4 different cells, in the case of HTLV-1-Gag-YFP (shades of grey dots) and HIV-1-myr(-)-Gag-GFP (shades of cyan dots) and their linear fit (lines with same colors) assuming 3D free diffusion. Slopes of the fits are used to determine D_{3D}.

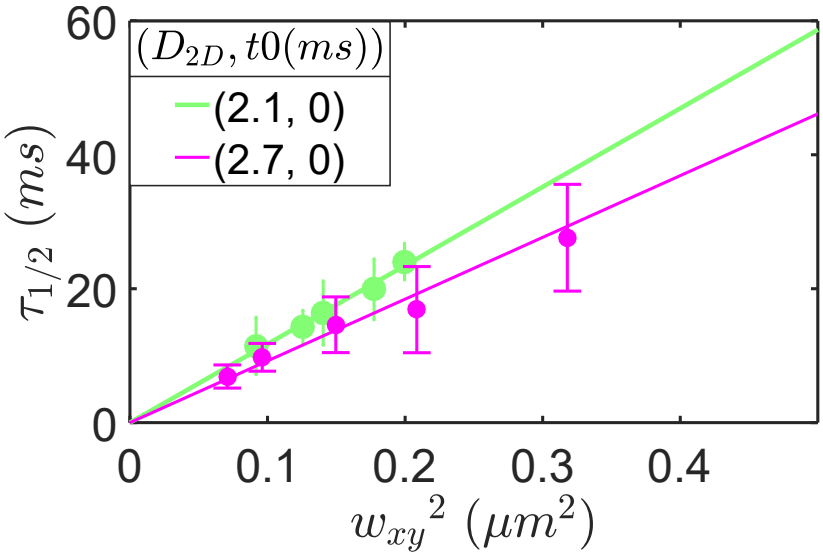

Figure SI.7: **TopFluor®PC FCS diffusion laws.** In our study, FCS diffusion laws were established with waists calibrated at 37°C using D_{Rhod}=550μm^2.s^{-1} as the rhodamine diffusion coefficient (see Material and Methods for details). Previous confocal FCS diffusion laws of lipids were established using D_{Rhod}=280μm^2.s^{-1} as the rhodamine diffusion coefficient. This approximately half D_{Rhod} value leads to half w^2 calibrated values, which of course changes the diffusion law (since $\tau_{1/2} = f(w^2)$), and therefore the apparent diffusion coefficient of the molecule of interest. Based on the diffusion law published in Lenne *et al.* (11), using D_{Rhod}=280μm^2.s^{-1}, we calculated the apparent diffusion coefficient of TopFluor®PC to be D=1.4±0.3μm^2.s^{-1}. We then recalibrated the waists using the most recent published value of rhodamine diffusion coefficient at 25°C, D_{Rhod}=414μm^2.s^{-1} (26), simply multiplying w^2 by 414/280 (since we considered that the original value of the waists were calibrated at 25°C not 37°C (not indicated in the original paper)). The recalibrated TopFluor®PC diffusion law, obtained in Cos-7 cells by Lenne *et al.* is plotted in green, and our TopFluor®PC diffusion law, determined here, in HEK 293-T cells in pink. The apparent diffusion coefficient using a linear fit is found to be 2.1±0.2 μm^2.s^{-1} in the case of Cos-7 cells and 2.7±0.3 μm^2.s^{-1} in our case (mean±sem) which is in the same range according the different cell types and the uncertainty on the temperature at which the calibration was made in Lenne *et al.*

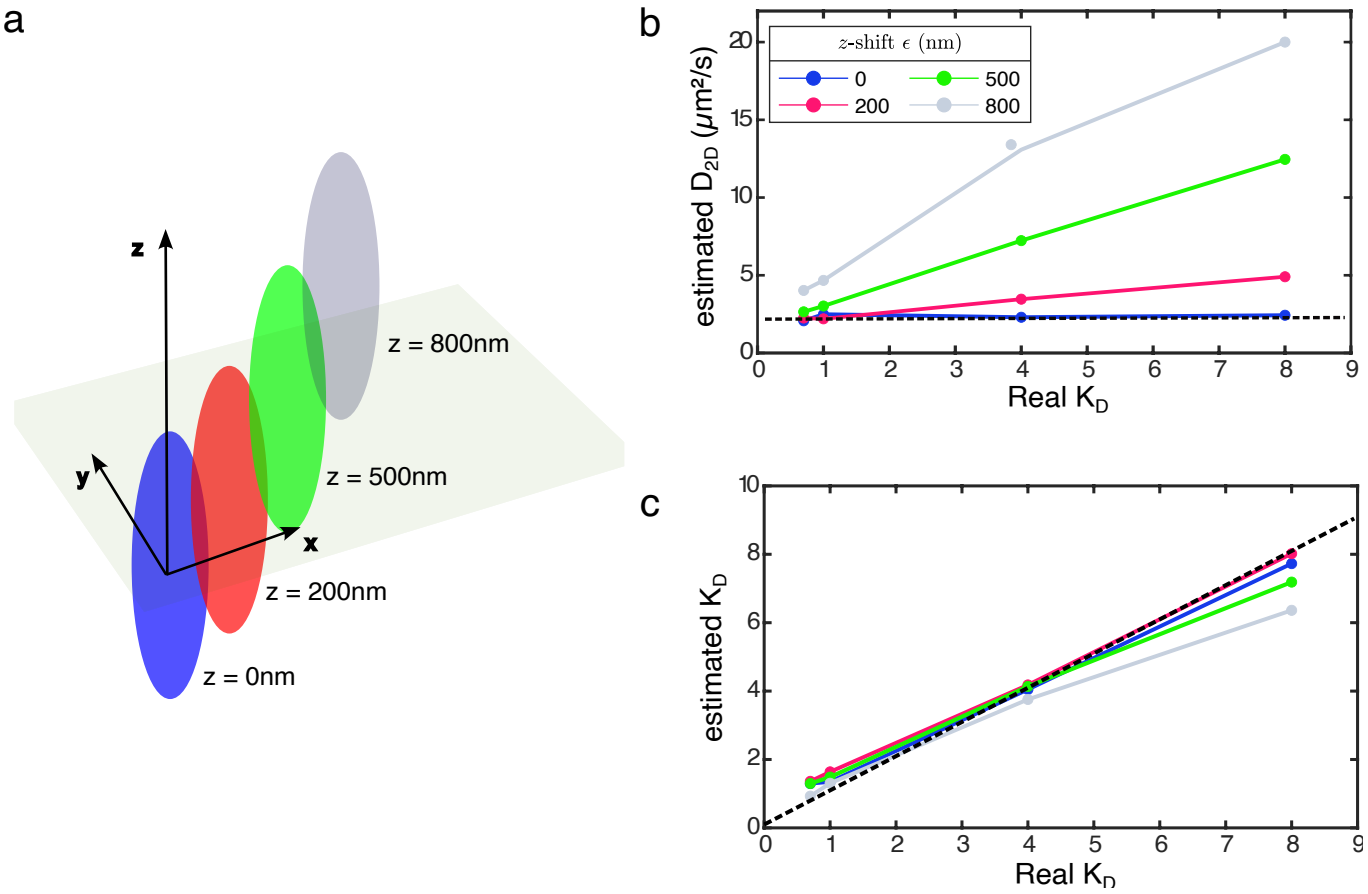

Figure SI.8: **Estimations of K_P and D_{2D} for different shifts in the z-focus.** Numerical simulations were performed for different K_D values (real K_D) with D_{2D}=2.7μm^2.s^{-1} and D_{3D}=37μm^2.s^{-1}. The simulation algorithm was the same as described in the main text, except that the $z = 0$ plan of the 3-D Gaussian illumination was shifted towards the cytosol by the indicated values. Estimations of D_{2D} and K_D were then made as described in the main text. a- The laser spot was shifted along the z-axis by 0 (blue), 0.2 (red), 0.5 (green) or 0.8 (gray) μm above the membrane, into the cytosol, keeping kinetics parameters of the simulation constant. b- Plot of the estimated D_{2D} as a function of the real K_D values for the different shifts in z (colour code same as a- panel). The dashed line correspond to the real D_{2D} value. c- Plot of the estimated K_D as a function of real K_D for the different shifts in z. The estimation of D_{2D} from FCS diffusion law fit is seen to be highly sensitive to the shift in focus, while estimation of K_D is not.

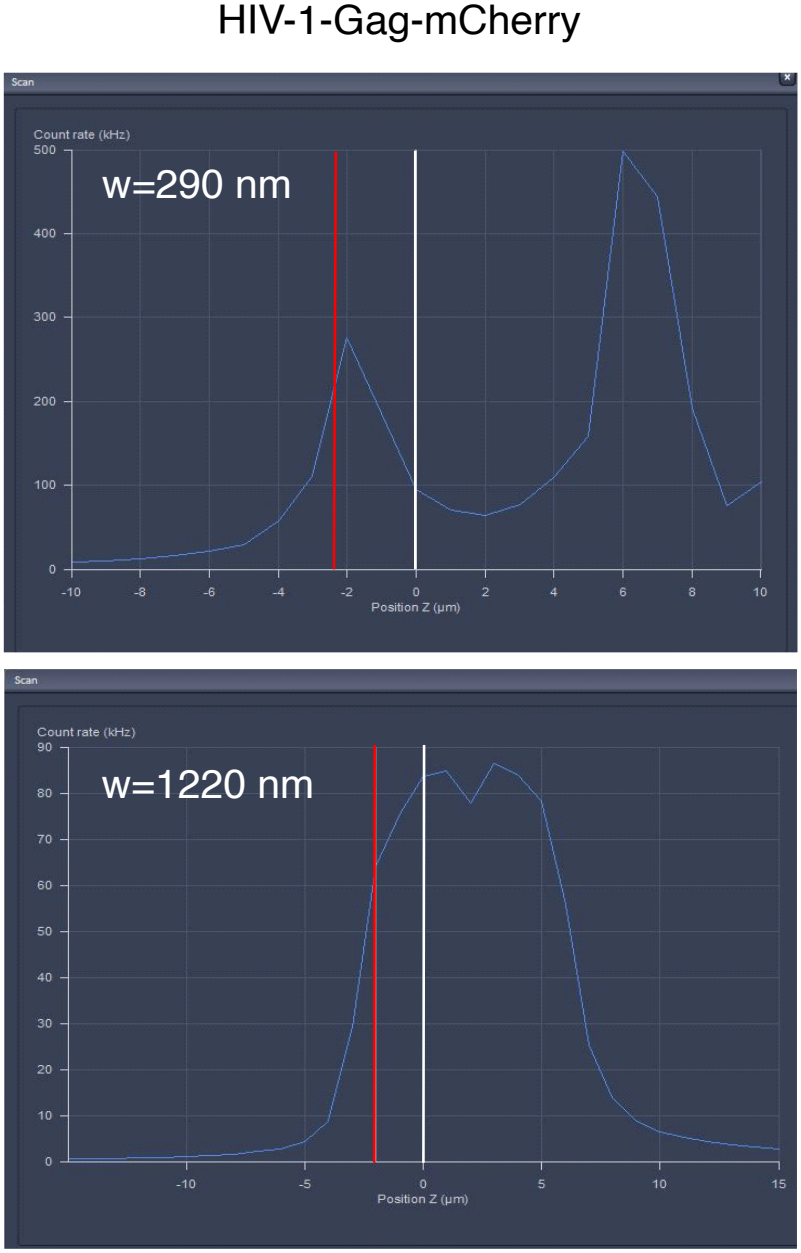

Figure SI.9: **Z-axis fluorescence intensity distribution observed in cells at small and large waists** Experimental fluorescence intensity distribution along the z-axis of an HEK 293 T cell expressing HIV-1-Gag-mCherry is exhibiting two distinct peaks with small waists (upper part), corresponding to the membrane and the cytosol below and above the unlabelled nucleus of the cell (valley). These two peaks tends to become indistinguishable at the larger waist (lower part) making it difficult to correctly focus on the membrane of the cell. Red lines represent the z-position where the focus is set to acquire FCS data for establishing diffusion laws.

SI3.Supporting Tables

Table SI.1: Example of experimental values of the $\tau_{1/2}$ obtained for myr(-)Gag in different lipid mixtures

Rhodamine			mix1			mix2			mix3			mix4		
$w^2(\mu m^2)$	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n
0.045	0.004	60	3.8	1.1	68	4.4	1	30	2.4	0.5	90	1.4	0.4	67
0.062	0.005	60	5.3	0.7	62	5.9	0.9	29	3.7	0.9	90	2	0.5	71
0.099	0.006	60	8.8	2.6	66	9.8	2	32	5.3	1.3	99	3.2	0.7	69
0.142	0.01	60	11.9	3.9	69	13.3	3.4	30	7.6	1.7	90	4.4	1	70
0.227	0.02	60	15.9	5.1	62	18.4	5.6	31	12.4	2.5	92	7.1	1.6	70
0.309	0.028	60	24.7	8.4	27	24.5	6	66	14.3	3.4	110	8.2	1.6	73
0.488	0.059	60	31.5	16.7	35	37.3	8.6	45	23.2	6.1	106	10.3	3	71

Table SI.2: Example of experimental values of the $\tau_{1/2}$ obtained for HIV-1-myr(-)GagGFP for the different cells

Rhodamine			cell1			cell2			cell3		
$w^2(\mu m^2)$	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n
0.055	0.007	29	1.4	0.4	31	1.8	0.4	19	1.8	0.3	16
0.128	0.012	30	2.5	0.4	32	2.3	0.5	34	2.1	0.4	27
0.179	0.019	30	3.3	0.4	37	3.7	0.7	31	3.2	0.8	30
0.291	0.037	30	5.5	1	35	5.2	1.1	40	5.2	1.1	41
0.4	0.036	30	7.2	1.8	36	6.9	1.5	28	8.1	2.4	30
0.699	0.079	29	11.2	3.6	35	10.7	2.4	39	11.4	4.6	16
1.248	0.256	30	14.9	5.8	29	14.9	6.5	27	14.4	4.7	22

Rhodamine			cell4			cell5		
$w^2(\mu m^2)$	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n
0.062	0.009	30	1.8	0.5	26	2.4	0.9	24
0.149	0.026	30	3.4	0.7	28	3.3	1.1	21
0.202	0.036	30	3.9	0.9	44	4.5	1.3	30
0.32	0.044	31	5.6	1.5	43	6.5	1.5	27
0.439	0.062	30	7.9	2.1	50	8	2.1	32
0.721	0.069	30	13.1	3.3	38	11.6	2.3	49
1.289	0.214	46	17	5.6	36	17.5	5.1	41

Table SI.3: Example of experimental values of the $\tau_{1/2}$ obtained for HTLV-1-Gag-YFP for the different cells

Rhodamine			cell1			cell2		
$w^2(\mu m^2)$	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n
0.054	0.002	29	1.7	0.3	34	2.2	0.2	41
0.118	0.003	30	5	0.3	53	4.2	0.2	40
0.156	0.005	30	7.1	0.6	44	6	0.2	36
0.262	0.009	30	11.8	0.7	28	13.5	1.4	41
0.322	0.014	30	16.1	0.9	25	18.7	1.4	38
0.506	0.026	29	29.2	2.5	31	24.2	1.7	38
0.906	0.025	30	33.5	3	32	35.8	2.8	28
Rhodamine			cell3			cell4		
$w^2(\mu m^2)$	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n
0.055	0.055	30	1.7	0.1	15	1.2	0.1	34
0.106	0.106	30	4.8	0.2	23	3.6	0.2	21
0.17	0.17	30	5.9	0.3	25	6	0.3	25
0.301	0.301	31	10.9	0.6	25	11.5	0.9	25
0.404	0.404	30	16.5	0.7	21	16.8	1.2	50
0.568	0.568	30	23.2	1.2	27	23.5	1.3	23
0.982	0.982	46	25.8	2.6	25	30.9	3.6	63

Table SI.4: Example of experimental values of the $\tau_{1/2}$ obtained for HIV-1-Gag-mCherry for the different cells

Rhodamine			cell1			cell2			cell3			cell4		
$w^2(\mu m^2)$	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n
0.086	0.001	30	2.9	0.1	45	2.1	0.1	44	2.9	1.1	25	2	2.1	29
0.183	0.003	30	7.3	0.4	32	4	0.1	49	9.4	0.6	24	4	0.1	30
0.252	0.004	30	9.9	0.4	42	6.2	0.3	53	11.5	0.9	27	6.9	0.3	23
0.436	0.006	29	11.6	0.5	58	12.4	0.9	44	14	1.1	20	14.1	0.9	39
0.588	0.008	30	15	0.8	46	19.3	2.3	43	19.9	1.4	30	17.5	1.5	29
0.917	0.011	25	21.2	1	50	24.8	3	32	26	1.5	32	21.9	1.8	31
1.5	0.051	29	27	1.1	54	30.2	4.7	30	32.7	1.7	21	33.2	2.3	48
Rhodamine			cell5			cell6			cell7			cell8		
$w^2(\mu m^2)$	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n
0.085	0.001	30	2.2	0.1	29	3.5	0.1	56	1.6	1.1	30	2	2.1	37
0.177	0.003	30	4.7	0.3	30	5.3	0.2	28	4.4	0.2	27	4.2	0.2	30
0.253	0.004	30	5.8	0.4	35	7.2	0.4	35	5.7	0.3	28	7.9	0.5	25
0.413	0.006	29	7.2	0.6	45	10.7	0.4	50	7.8	0.5	24	13.9	0.9	26
0.55	0.008	30	11	0.6	38	15.2	0.7	51	12.3	1.3	35	18	0.9	27
0.849	0.011	25	17.9	0.9	36	17.7	0.8	35	18.6	1.6	29	25.7	1.5	48
1.429	0.047	29	24	1	45	23.8	1.2	40	24.8	1.6	43	32.5	2.7	29
Rhodamine			cell9			cell10			cell11			cell12		
$w^2(\mu m^2)$	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n	$\tau_{1/2}$(ms)	s.d.	n
0.085	0.001	30	2.7	0.1	18	2.1	0.1	22	2	1.1	30	2.8	2.1	19
0.177	0.003	30	5	0.2	28	4.9	0.2	23	4.8	0.2	28	5.1	0.2	57
0.253	0.004	30	6.9	0.3	35	6.5	0.3	29	6.9	0.2	29	8.3	0.4	29
0.413	0.006	29	11.5	0.6	48	9.6	0.4	26	9.4	0.4	27	10.4	0.5	21
0.55	0.008	30	16.7	0.7	46	15.5	0.4	34	11.2	0.6	23	15.6	1.2	16
0.849	0.011	25	27	1.1	38	22.2	1.4	27	17.5	0.9	41	19.2	1.3	20
1.429	0.047	29	37.5	2.2	58	32.4	1.6	37	25	1.2	32	26	1.5	23

Table SI.5: Example of experimental values of the proportion of molecules in the dark state (0<T<1) and the average blinking time τ_T (μs) obtained by fitting the HIV-1-Gag-mCherry ACF of the different cells with equation 4. The number of measurements n for each waist are equivalent to the one given in table SI.4.

Rhodamine	cell1		cell2		cell3		cell4	
$w^2(\mu m^2)$±s.d.	T±s.d.	τ_T± s.d.(μs)	T±s.d.	τ_T± s.d.(μs)	T±s.d.	τ_T± s.d.(μs)	T±s.d.	τ_T± s.d.(μs)
0.086±0.001	0.44±0.01	154±15	0.28±0.01	171±19	0.19±0.02	134±32	0.29±0.05	298±95
0.183±0.003	0.4±0.01	276±18	0.3±0.01	187±13	0.29±0.03	214±52	0.15±0.01	178±25
0.252±0.004	0.43±0	261±15	0.38±0.01	205±8	0.29±0.04	829±149	0.19±0.01	220±28
0.436±0.006	0.41±0	232±5	0.41±0.01	230±20	0.22±0.04	525±170	0.24±0.03	372±64
0.588±0.008	0.41±0	269±9	0.35±0.02	631±90	0.19±0.01	329±37	0.24±0.04	438±71
0.917±0.011	0.46±0	214±4	0.28±0.02	613±134	0.21±0.02	443±49	0.27±0.02	281±48
1.5±0.051	0.39±0.01	397±14	0.28±0.02	406±96	0.31±0.06	212±38	0.25±0.02	229±38
Rhodamine	cell5		cell6		cell7		cell8	
$w^2(\mu m^2)$±s.d.	T±s.d.	τ_T± s.d.(μs)	T±s.d.	τ_T± s.d.(μs)	T±s.d.	τ_T± s.d.(μs)	T±s.d.	τ_T± s.d.(μs)
0.085± 0.001	0.38±0.11	114±52	0.3±0.01	152±12	0.23±0.01	121±12	0.48±0.01	78±5
0.177±0.003	0.33±0.01	213±20	0.24±0.01	132±15	0.27±0.01	207±29	0.31±0	122±10
0.253±0.004	0.27±0.05	213±37	0.23±0.01	140±15	0.26±0.01	163±14	0.28±0	136±9
0.413±0.006	0.25±0.01	210±14	0.24±0	168±7	0.32±0.01	240±17	0.26±0.01	233±15
0.55±0.008	0.21±0.01	202±16	0.25±0.01	211±15	0.25±0.01	221±25	0.28±0.01	141±11
0.849±0.011	0.22±0.01	251±27	0.23±0	232±20	0.25±0.01	323±45	0.31±0	175±11
1.429±0.047	0.23±0.01	218±27	0.3±0	272±16	0.25±0.01	236±23	0.34±0.01	268±36
Rhodamine	cell9		cell10		cell11		cell12	
$w^2(\mu m^2)$±s.d.	T±s.d.	τ_T± s.d.(μs)	T±s.d.	τ_T± s.d.(μs)	T±s.d.	τ_T± s.d.(μs)	T±s.d.	τ_T± s.d.(μs)
0.085±0.001	0.48±0.01	88±4	0.47±0.01	89±5	0.41±0.01	90±5	0.47±0.01	103±12
0.177±0.003	0.47±0	147±9	0.45±0.01	133±4	0.42±0	154±7	0.45±0.01	132±8
0.253±0.004	0.44±0	174±9	0.43±0	140±4	0.26±0	146±11	0.41±0.01	188±10
0.413±0.006	0.45±0	171±6	0.42±0	160±5	0.25±0	170±14	0.42±0	209±10
0.55±0.008	0.46±0	155±5	0.46±0	144±4	0.29±0.01	139±12	0.46±0	186±7
0.849±0.011	0.46±0	154±5	0.43±0	227±10	0.3±0	198±14	0.46±0.01	185±10
1.429±0.047	0.44±0.01	209±9	0.42±0	264±14	0.32±0.01	165±12	0.43±0.01	254±18