Biophysical Journal, Volume *111*

Supplemental Information

Arbitrary-Region Raster Image Correlation Spectroscopy

Jelle Hendrix, Tomas Dekens, Waldemar Schrimpf, and Don C. Lamb

Supporting Materials and Methods

Intensity or intensity variance thresholding

Besides manually drawing the ROI or using an organelle-based ROI, the *R* mask can also be generated by thresholding images on the basis of intensity or the variance in intensity. Below, we discuss a few approaches to image thresholding.

Frame-averaged intensity: When structural heterogeneities in the imaged sample remain localized (static) over the course of the experiment, the frame-averaged intensity can be used as a threshold. The image of the average intensity over all frames, $\langle I \rangle_F$, is then compared at each pixel with given thresholds, I_{min} and I_{max}, the minimal and maximal allowed pixel intensities, respectively, to generate *R*:

$$\begin{aligned} &for \quad \langle I(x,y)\rangle_F < I_{min} \quad or \quad \langle I(x,y)\rangle_F > I_{max}, \quad R(x,y) = 0 \\ &for \quad I_{min} \le \langle I(x,y)\rangle_F \le I_{max}, \quad R(x,y) = 1 \end{aligned} \qquad \text{Eq.S1}$$

Frame-averaged intensity thresholding performs well for separating regions in the sample with clear differences in intensity, for example between intra- and extracellular regions or between the cytosol and the nucleus.

Rolling-window frame-averaged intensity: When structural heterogeneities in the imaged sample move slowly but remain localized (static) over at least $2\Delta F+1$ frames, averaging over a sliding window of $2\Delta F+1$ frames can be performed. In this case, the average intensity image is calculated between frames (f-ΔF) and (f+ΔF), $\langle I \rangle_{\Delta F}$, and the intensity at every pixel compared with the threshold intensities, I_{min} and I_{max}, to generate a frame-specific ROI mask of size $X \times Y \times (F-2\Delta F)$:

$$\begin{aligned} &for \quad \langle I(x,y)\rangle_{\Delta F} < I_{min} \quad or \quad \langle I(x,y)\rangle_{\Delta F} > I_{max}, \quad R(x,y) = 0 \\ &for \quad I_{min} \le \langle I(x,y)\rangle_{\Delta F} \le I_{max,} \quad R(x,y) = 1 \end{aligned} \qquad \text{Eq.S2}$$

Due to the window necessary for performing the rolling average, the first and last ΔF frames in the image are ignored for calculation of the SACF. The total I_{min} and I_{max} can be (*i*) absolute numbers, (*ii*) defined relatively to $\langle I \rangle_F$:

$$\begin{aligned} &for \quad \langle I(x,y)\rangle_{\Delta F} < \langle I(x,y)\rangle_F / c \quad or \quad \langle I(x,y)\rangle_{\Delta F} > c\langle I(x,y)\rangle_F, \quad R(x,y) = 0 \\ &for \quad \langle I\rangle(x,y)_F / c \le \langle I(x,y)\rangle_{\Delta F} \le c\langle I(x,y)\rangle_F, \quad R(x,y) = 1 \end{aligned} \qquad \text{Eq.S3}$$

where c defines the strictness of the threshold, (*iii*) defined relatively to $\langle \boldsymbol{I} \rangle_{\Delta F2}$, the average intensity image between frames (f-$\Delta F2$) and (f+$\Delta F2$) with $\Delta F2 > \Delta F$:

$$for \quad \langle I(x,y)\rangle_{\Delta F} < \langle I(x,y)\rangle_{\Delta F2}/c \quad or \ \langle I(x,y)\rangle_{\Delta F} > c\langle I(x,y)\rangle_{\Delta F2}, \quad R(x,y) = 0$$
$$for \quad \langle I(x,y)\rangle_{\Delta F2}/c \leq \langle I(x,y)\rangle_{\Delta F} \leq c\langle I(x,y)\rangle_{\Delta F2}, \quad R(x,y) = 1 \qquad \text{Eq.S4}$$

or (*iv*) defined relatively to the mean intensity in any other region surrounding the pixel. A rolling window-averaged intensity thresholding performs well for excluding slowly moving bright (*e.g.* aggregates or oligomers) or dim (*e.g.* non-fluorescent cell organelles) regions from ICS analysis.

Spatial-averaged intensity: When structural heterogeneities in the imaged sample are dynamic from one frame to the next, the spatially averaged intensity image, $\langle \boldsymbol{I} \rangle_{\Delta L}$, where ΔL is the size of the spatial averaging window in each dimension in pixels, is useful. Again, every pixel is compared with $I_{\min}$ and $I_{\max}$, to generate R with size (X-ΔL)×(Y-ΔL)×F:

$$for \quad \langle I(x,y)\rangle_{\Delta L} < I_{min} \quad or \quad \langle I(x,y)\rangle_{\Delta L} > I_{max}, \quad R(x,y) = 0$$
$$for \quad I_{min} \leq \langle I(x,y)\rangle_{\Delta L} \leq I_{max}, \quad R(x,y) = 1 \qquad \text{Eq.S5}$$

We found it best to define $I_{\min}$ and $I_{\max}$ relatively to a larger spatially average image, $\langle \boldsymbol{I} \rangle_{\Delta L2}$, where $\Delta L2 > \Delta L$ is the averaging window size per dimension.

$$for \quad \langle I(x,y)\rangle_{\Delta L} < \langle I(x,y)\rangle_{\Delta L2}/c \quad or \ \langle I(x,y)\rangle_{\Delta L} > c\langle I(x,y)\rangle_{\Delta l2}, \quad R(x,y) = 0$$
$$for \quad \langle I(x,y)\rangle_{\Delta l2}/c \leq \langle I(x,y)\rangle_{\Delta l} \leq c\langle I(x,y)\rangle_{\Delta L2}, \quad R(x,y) = 1 \qquad \text{Eq.S6}$$

Typical values were $\Delta L2 = 30$, $\Delta L = 10$ and $c = 1.5$.

Space-averaged intensity variance: Lastly, the spatial variance image, $\boldsymbol{\sigma}^2_{\Delta L}$, where ΔL is the side length of the square in which the variance is calculated, can be compared with, the minimal and maximal allowed variance, $\sigma^2_{\min}$ and $\sigma^2_{\max}$ respectively, to generate R with a size of (X-ΔL)×(Y-ΔL)×F:

$$for \quad \sigma^2_{\Delta L} < \sigma^2_{min} \quad or \quad \sigma^2_{\Delta L} > \sigma^2_{max}, \quad R(x,y) = 0$$
$$for \quad \sigma^2_{min} \leq \sigma^2_{\Delta L} \leq \sigma^2_{max}, \quad R(x,y) = 1 \qquad \text{Eq.S7}$$

We found it best to define $\sigma^2_{\min}$ and $\sigma^2_{\max}$ relatively to $\boldsymbol{\sigma}^2_{\Delta L2}$, where $\Delta L2 > \Delta L$:

$$for \quad \sigma^2_{\Delta L} < \sigma^2_{\Delta L2}/c \quad or \quad \sigma^2_{\Delta L} > c\sigma^2_{\Delta L2}, \quad R(x,y) = 0$$
$$for \quad \sigma^2_{\Delta L2}/c \leq \sigma^2_{\Delta L} \leq c\sigma^2_{\Delta L2}, \quad R(x,y) = 1 \qquad \text{Eq.S8}$$

Typical values were $\Delta L2$ = 30, ΔL = 10 and c = 1.5. Since the $\boldsymbol{\sigma}^2_{\Delta L2}/\boldsymbol{\sigma}^2_{\Delta L}$ ratio is typically larger than the $\langle \boldsymbol{I} \rangle_{\Delta L2}/\langle \boldsymbol{I} \rangle_{\Delta L}$ ratio, variance-based thresholding can be slightly more sensitive. Intensity and variance thresholding allow removal of dynamic bright (*e.g.* fluorescent aggregates or oligomers) or dim (*e.g.* non-fluorescent vesicles) regions from ICS analysis.

Simulations for confined diffusion

Random motion of particles was simulated using a pseudo-random number generator based on the commonly used Mersenne Twister algorithm (1). The smallest simulation interval was set to 1 µs. Random photon emission was also ensured by the pseudo-random number generator, based on the particle's brightness and its position relative to the focus. The focus shape was approximated by a 2D Gaussian with a focal size ω_r of 200 nm. The restriction to 2D movement was done to increase the local concentration and thus reduce the total calculation time of the simulation. The molecular brightness was 100 kHz. During the simulation, the focus position was moved to emulate the raster scanning. The total area scanned was 250×250 pixels or 12.5 µm × 12.5 µm, with pixel, line and frame times of 12 µs, 3 ms and 750 ms, respectively. Consequently, the pixel size was 50 nm. In total, 2000 particles per condition were simulated for a total of 100 frames. Upon exiting the simulation box, particles were reentered on the opposite side. Since this might lead to artifacts in the correlation function, the simulation box was increased to 15 µm per dimension, thus limiting the border effects. For each simulation two different particle types were simulated. The first type of particles was confined to small round compartments and moved with a diffusion coefficient of 2 µm^2/s. The rest of the particles diffused freely around these confinements with D = 10 µm^2/s.

Simulations for diffusion pseudo-maps

The same algorithms as in the previous section were employed, with the exception that the diffusion coefficient or, more precisely, the step sizes of the particles depended on their spatial position, defined by an input diffusion map. This spatial

diffusion distribution was generated using a map of random numbers that was smoothed with a Gaussian blur and rescaled to an appropriate diffusion range (Fig. S4*A*). The resolution of the map was 5 nm. The molecular brightness of the particles was 10 kHz for the channel used for correlation analysis and 100 kHz for the one used for generating the ROIs.

Supporting Results

Diffusion in confinements

In order to test the minimal size of compartments in/around which diffusion can still be accurately quantified with ARICS, simulations with known parameters were employed. Four different diameters of circular confinements were simulated: 200 nm, 400 nm, 800 nm, and 1600 nm (Fig. S3*A*, *from top to bottom*), corresponding to 1, 2 4, and 8 times the focus size. The confinements covered approximately 10% of the total image area. Concentration inside and outside were chosen such that the brightness inside the confinements was 4-5 times higher than in the surrounding space. The average count rate of all simulations was in the range of 75-80 kHz. Pixels with an average count rate of 70 kHz and less were attributed to the free component (Fig. S3*A*, *red regions*), while all pixels with more than 100 kHz were assumed to belong to the confinements (Fig. S3*A*, *blue regions*). The intermediate border regions with 70-100 kHz were omitted from further analysis.

SACFs for diffusion outside the confinements showed a similar shape for all conditions (Fig. S3*B*). Quantitatively, a deviation of *D* from the simulated value was more pronounced at smaller confinement sizes, but always less than 10% (Fig. S3*D*). Additionally, the reflective nature of the confinements might further contribute to the decreased diffusion.

SACFs for confined diffusion, on the other hand, were more affected by the confinement size (Fig. S3*C*). The SACFs exhibited a 300-500-nm dip along the fast scanning axis, with the depth decreasing for larger sizes and almost vanishing at confinement diameters that were 8 times the focus size. This dip is most likely caused by the confinement and the reflective borders and not by the ARICS algorithm itself. Quantitatively, for confinement sizes up to 400 nm, *D* was larger than the set value (Fig S3*D*). This is again caused by problems in correctly assigning the pixels so that the correlation functions still contain contributions from the freely diffusion component. Interestingly, in the simulations with confinements that were four time larger than the focus (800 nm), the individual compartments were large enough that miss-assignment no longer contributed to the correlation function, but small enough such that the confinement still significantly affected the mobility. This lead to an apparent *D* that was even lower than expected. For the largest sizes (1600 nm), the diffusion coefficient can be recovered to within less than 5% of the expected value.

Together, these results show that ARICS can reliably probe diffusion in compartments down to a factor of 4-8 times the focus size, which in practice means 600-1200 nm for a realistic focus size of 150 nm. This limit can be further decreased by decreasing the focus size, *e.g.* by using STED-RICS (2). However, another limit to the quantitative analysis, not specific to the ARICS algorithm, is the actual influence of confinements of the diffusion itself.

Diffusion pseudo-maps

In order to test the validity of the diffusion pseudo-map generation procedure, simulations with known parameters were employed. The differences in the diffusion coefficients in the different parts of the simulated image resulted in corresponding differences in concentration and, therefore, also in count rate (Fig. S4*A*). Just as described in the Results part of the main text, these differences in signal intensity were used to sort the pixels into five ROIs. A separate channel with higher brightness was used for this segmentation, to limit crosstalk between the selection procedure and the correlations. The segmentation was based on the count rate in a 5×5 pixel and 5-frame moving average (Fig. S4*B*). The additional averaging over multiple frames was needed due to the low concentrations and strong concentration fluctuations. The original diffusion map and the pseudo diffusion map generated using ARICS are shown in Fig. S4*C* and *D*. There is excellent agreement between the two images. However, as the map has gradual diffusion coefficients but the algorithm only divides the images into only five ROIs, the complete gradient was not recovered. Especially for the highest and the lowest intensity regions, the gradient was flattened by this undersampling. This shows that for each system the number and threshold of the ROIs need to be adjusted individually to ensure a good recovery of the underlying features, while at the same time limit the change of miss-assignment of the individual pixels (especially important when the diffusion and the signal intensity are not strictly correlated).

Supporting Figures

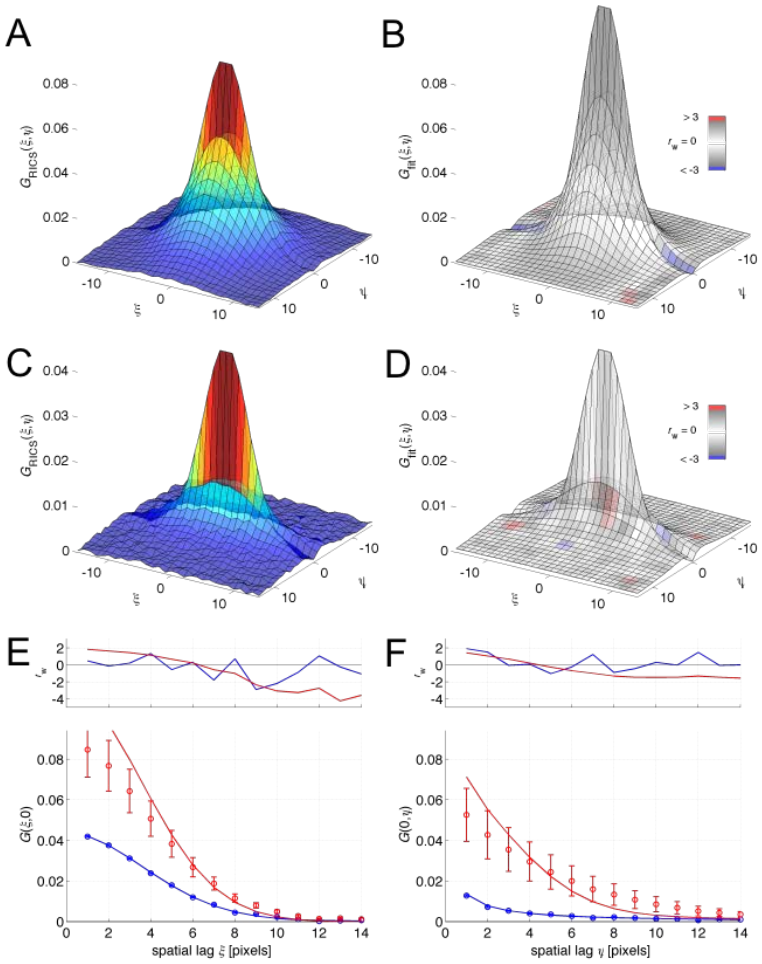

Fig. S1 | Further analysis of the data from Fig 2. (*A*) The average raw-data SACF, color-coded according to the correlation value. (*B*) Best fit of the data from panel *A* to Eq. 14, color-coded using the weighted residuals goodness-of-fit parameter r_w. The best fit does not provide reliable results (see *red data* panels *E* and *F*) because the error on the average SACF is very large. Hence, r_w and the reduced $\chi^2 = 1.39$ are still relatively small. (*C*) The average arbitrary-ROI SACF, color-coded according to the correlation value. (*D*) Best fit of the data from panel *C* to Eq. 14, color-coded using the weighted residuals goodness-of-fit parameter r_w (reduced $\chi^2 = 1.02$). (*E* and *F*) Experimental $G(\xi,0)$ (*symbols* in panel *E*) and $G(0,\psi)$ (*symbols* in panel *F*), standard deviation (*error bars*) and $G_{fit}(\xi,0)$ (*solid lines* in panel *E*) and $G_{fit}(0,\psi)$ (*solid lines* in panel *F*) corresponding to the data in panels *A* and *B* (*red*) and panels *C* and *D* (*blue*). The average arbitrary-ROI SACF has a much lower error and is described excellently by the fit model. (*Plot on the upper panel*) Weighted residuals r_w.

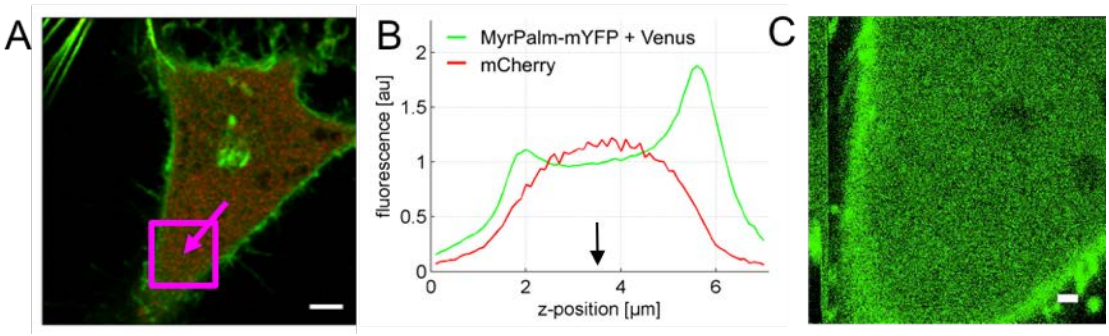

Fig. S2 | Further information on Fig. 3. (*A*) Dual-color confocal image of the HeLa cell co-expressing Venus, MyrPalm-mYFP and mCherry. The *pink arrow* points towards the location where the *z*-scan shown in panel *B* was performed. The *pink square* marks the region shown in panel *C* used for the RICS analysis. The scale bar is 5 µm. (*B*) Axial *z*-scan, clearly illustrating the enrichment of MyrPalm-mYFP at the bottom and top cell membrane. The *black arrow* points to the middle of the cell, where imaging for RICS was performed. (*C*) Zoom-in of the *pink square* in panel *A*. Scale bar is 1 µm.

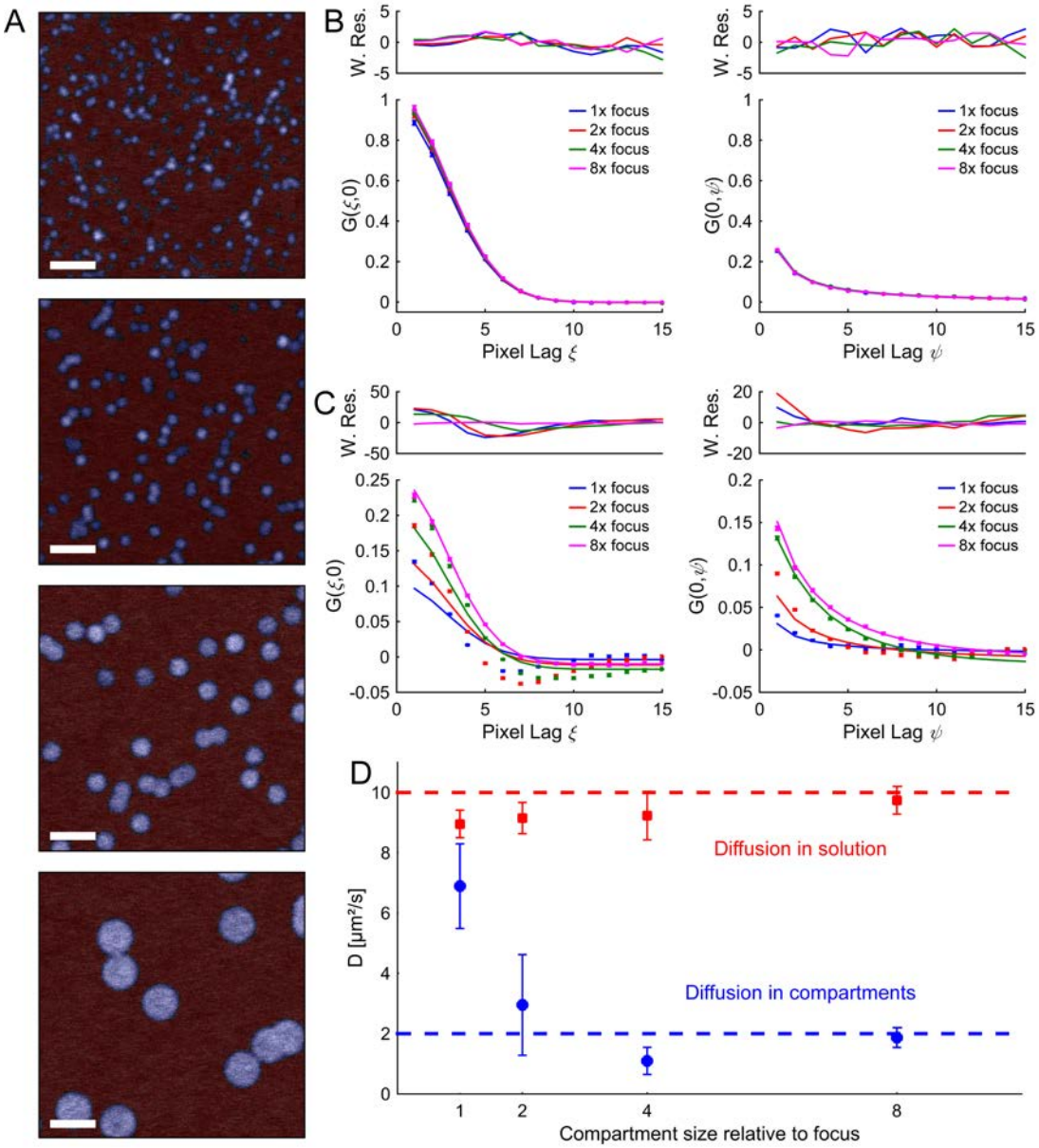

Fig. S3 | Simulations for different confinement sizes. (*A*) Images of samples containing confinements of varying diameter: (*from top to bottom*) 200 nm, 400 nm, 800 nm and 1600 nm, corresponding to 1, 2, 4 and 8 times the focus size, respectively. Blue regions contain the pixels in the confinements, red regions encompass the surrounding area. Gray pixels at the borders were omitted from further analysis. Scale bars are 2 µm. (*B*) SACFs calculated only using pixels from the surrounding area. (*C*) SACFs calculated with pixels in the confinements. In panels *B* and *C*, the fast scanning axis is shown in on the left, the slow scanning axis on the right. Points show the data, including the s.e.m., solid lines depict the fit. (*D*) Extracted *D* values for different confinement sizes for the free (*red squares*) and confined (*blue circles*) components. Error bars indicate the s.d. of five measurements. Dashed lines show the values used for the simulation.

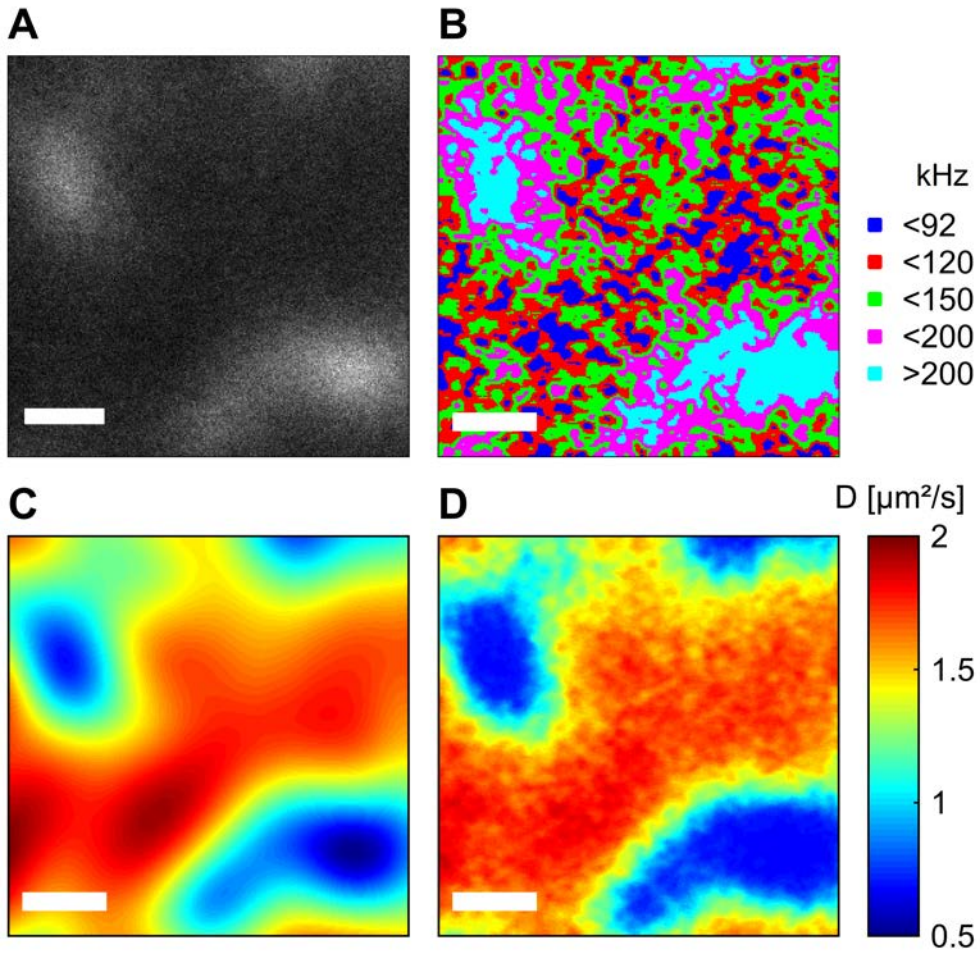

Fig. S4 | Simulations for diffusion pseudo-map analysis. (*A*) Average fluorescence intensity image of the simulated experiment with 5000 particles with a molecular brightness of 10 kHz. (*B*) Single frame segmented into 5 ROIs. Maximum intensities for selection of the respective ROIs are given in the legend. (*C*) Diffusion map used for the simulations. (*D*) A pseudo-diffusion map reconstructed from the ROI segmentation procedure. The scale bar corresponds to 2 µm. The color table used for panels *C* and *D* is shown to the right.

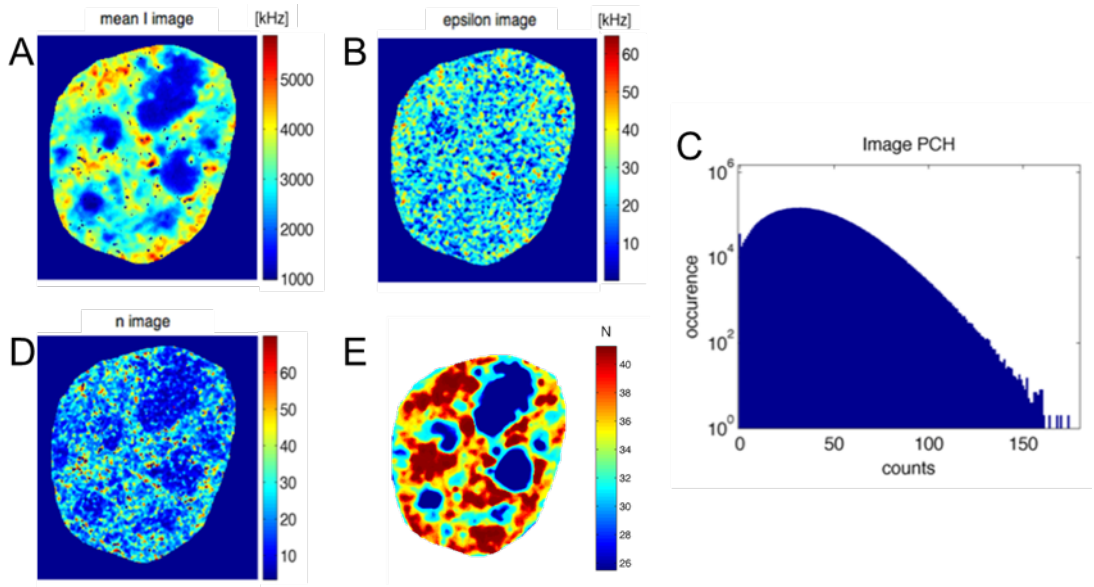

Fig. S5 | Comparative analysis of the eGFP-LEDGF/p75 imaging data in Fig. 5 with the Number & Brightness method. (*A*) pixel intensity image averaged over all frames. (*B*) Epsilon (brightness) image. (*C*) The pixel intensity distribution histogram. (*D*) Number image, illustrating similar concentrations as with the ROI segmentation method (*E*). The pointillism-like appearance of the n image is because of low statistics. For the N&B analysis, pixel intensities were pre-processed using a space-time moving average (3-pixel radius and $\Delta F = 3$) as reported before (3). n and epsilon images were median-filtered (3-by-3).

Supporting References

1. Matsumoto, M., and T. Nishimura. 1998. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. ACM Trans. Model. Comput. Simul. 8:3-30.
2. Hedde, P. N., R. M. Dorlich, R. Blomley, D. Gradl, E. Oppong, A. C. Cato, and G. U. Nienhaus. 2013. Stimulated emission depletion-based raster image correlation spectroscopy reveals biomolecular dynamics in live cells. Nature communications 4:2093.
3. Hendrix, J., V. Baumgartel, W. Schrimpf, S. Ivanchenko, M. A. Digman, E. Gratton, H. G. Krausslich, B. Muller, and D. C. Lamb. 2015. Live-cell observation of cytosolic HIV-1 assembly onset reveals RNA-interacting Gag oligomers. J Cell Biol 210:629-646.