Communication

Quantitative Super-Resolution Imaging Reveals Protein Stoichiometry and Nanoscale Morphology of Assembling HIV-Gag Virions

Julia Gunzenhäuser, Nicolas Olivier, Thomas Pengo, and Suliana Manley

Nano Lett., **Just Accepted Manuscript** • DOI: 10.1021/nl3021076 • Publication Date (Web): 21 Aug 2012

Quantitative Super-Resolution Imaging Reveals Protein Stoichiometry and Nanoscale Morphology of Assembling HIV-Gag Virions

Julia Gunzenhäuser, Nicolas Olivier, Thomas Pengo & Suliana Manley

Laboratory of Experimental Biophysics, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Abstract

The HIV structural protein Gag assembles to form spherical particles of radius ~70 nm. During the assembly process, the number of Gag proteins increases over several orders of magnitude, from a few at nucleation to thousands at completion. The challenge in studying protein assembly lies in the fact that current methods such as standard fluorescence or electron microscopy techniques cannot access all stages of the assembly process in a cellular context. Here, we demonstrate an approach using super-resolution fluorescence imaging that permits quantitative morphological and molecular counting analysis over a wide range of protein cluster sizes. We applied this technique to the analysis of hundreds of HIV-Gag clusters at the cellular plasma membrane, thus elucidating how different fluorescent labels can change the assembly of virions.

Keywords: Super-resolution imaging; protein assembly; protein counting; HIV-Gag

Viruses represent a major class of pathogens, whose assembly in the cellular context contains important information about the complex processes governing viral infection. Viruses are nanoscale objects that assemble from small nucleation complexes to ensembles containing thousands of molecules. In the case of human immunodeficiency virus (HIV), the viral components are targeted to the plasma membrane of infected cells where they assemble and eventually form spheres ~70 nm in radius. Viral assembly is widely studied using HIV-Gag, the main structural protein of HIV, which is sufficient to drive the assembly of virus-like particles (VLPs) in the absence of other viral components [1, 2]. Fluorescence imaging has been used to follow the time-dependent increase in the intensity of Gag clusters in living cells[3, 4], revealing the time scale of virion formation. Electron microscopy (EM) has elucidated the spatial arrangement of Gag in fully formed virions[5-8]. However, studying the complete assembly process requires nanoscale resolution over a large dynamic range, since the size of a cluster ranges from a few molecules to several thousand. This cannot be achieved with standard fluorescence imaging methods since they lack both the necessary resolution to determine cluster morphology, and the sensitivity to detect smaller clusters. EM-based methods in turn lack information on protein identity; thus, complexes composed of small numbers of Gag proteins are difficult to identify, precluding the first step toward quantitative analysis. As an alternative approach, super-resolution fluorescence imaging based on single molecule localization (SR)[9-11] offers nanoscale resolution of structures formed by specific proteins. Here, we use an SR-based approach to quantitatively image hundreds of forming HIV-Gag virions in different stages of cluster formation. With this information, we could extract protein stoichiometries as well as the nanoscale morphologies of Gag clusters, and detect differences in assembly for Gag proteins tagged with different fluorescent labels.

Gag proteins interact to form ordered assemblies of up to thousands of proteins, densely packed into nanoscale particles. Their high protein density and sub-diffraction limited sizes necessitate carefully adapted SR image acquisition procedures. This is because SR imaging relies on temporal separation of the fluorescence emission of each molecule within a diffraction-limited region, followed by molecular localization by fitting of the photon distribution, to finally yield over many frames the locations of multiple molecules per cluster. Thus, the fluorescence emission of single molecules within one cluster should be well separated temporally to guarantee maximal detection of molecules. Nanoscale clusters of proteins such as those formed by Gag must be imaged at a rate of less than one molecule per frame, since their size is smaller than a single diffraction-limited region. To fulfill these requirements, tight control of the fluorescence activation of single molecules is needed.

The necessary control of fluorescence activation can be achieved by using photoactivatable fluorescent proteins (PA-FPs). We labeled Gag with two widely used versions of the PA-FP Eos (EosFP), the monomeric mEos2[12] and the larger, tandem-dimeric tdEos[13]. EosFP irreversibly photoconverts from a green fluorescent to a red fluorescent state upon activation with UV light and the intensity of UV light controls the density of activated molecules in a sample. When expressing Gag-mEos2 or Gag-tdEos in cells, Gag-enriched clusters appear in the green fluorescence channel as diffraction limited puncta at the plasma membrane[14] (Figure 1A). We imaged Gag in the activated red fluorescence channel under intermediate-angle total internal reflection (TIR) illumination and applied extremely low UV intensities to activate less than one molecule per cluster per raw image (Supporting Figure 1A, B). We used continuous wave (CW) activation to minimize long-lived fluorescent dark states[15]. The individual molecules were subsequently localized and their positions

rendered as the envelope of the molecular probability distribution to create an SR image of Gag (Figure 1B). We rendered our data this way for visual representation only. Indeed, any image processing performed on rendered images can be biased by the rendering procedure itself[16]. Moreover, rendering obscures information on the number of molecules as well as the internal molecular organization of clusters. Therefore we performed quantitative analysis using the molecular positions themselves.

We extracted HIV-Gag clusters, representing assembling virions, from the list of molecular positions based on the proximity of molecules using the Hoshen-Kopelman algorithm[17]. Neighboring Gag molecules were defined as belonging to the same cluster if their intermolecular distance was less than 50 nm. This cutoff, at least 5 times smaller than the average intermolecular distance measured in non-clustered regions of the SR images, serves to prevent false cluster identifications. Using this empirically determined parameter, all Gag clusters observed in the diffraction-limited as well as the SR image were correctly identified by an automated algorithm (Figure 1C).

In quantifying the number of Gag proteins per cluster, one must account for the impact of fluorescent protein photophysics on molecular counting. Reversible photoconversion and photoblinking can result in overcounting. We used irreversibly photoactivatable fluorophores to avoid overcounting due to multiple photoconversions of a single dye. Photoblinking of the converted mEos2 fluorophore has previously been characterized, and we used the 2 sigma value of the empirical dark-state lifetime distribution[18] of mEos2 as an input parameter to our software to identify and group signals from single blinking molecules (Supporting Information, Supporting Figure 1). Simultaneous activation of multiple dyes in turn can result in undercounting. We use very low UV intensities to avoid simultaneous activation, but the tandem-dimeric version of Eos contains two chromophores, which due to their linkage and close proximity could still be correlated in activation. If this were the case, the number of identified tdEos molecules would correspond to the number of Gag-tdEos proteins. In the uncorrelated scenario, each tdEos molecule would be detected twice leading to double counting of the number of Gag proteins. Correlated activation of a single label should result in an increase in fluorescence upon activation followed by two photobleaching steps. Molecular traces of activated Gag-tdEos molecules were collected and analyzed (Figure 2A, Supporting Information), and individual activation and bleaching steps were observed. Of a total of 2,000 traces, 95% were imaged as molecules that bleached in a single step; thus the contribution from simultaneous activations of the two chromophores composing tdEos is negligible. As a result, we estimate that each Gag-tdEos molecule can be detected twice, corresponding to a pair of dyes.

Another requirement for quantitative SR imaging is an estimation of the fraction of molecules that has been detected. This is important when directly comparing objects of potentially different sizes or labeling densities. Because the proteins in a single nascent virion must be imaged and bleached literally one at a time, the number of molecules composing it directly determines the minimum number of raw images needed for complete imaging. To quantify the molecular counting process, cells were imaged until all fluorophores had been activated and bleached. For EosFPs the complete imaging was evidenced by a lack of signal in both the unconverted green and activated red channels. Plots of the cumulative number of molecules in individual HIV-Gag assembling sites (Figure 2B, C) show an initial rise that flattens to reach an apparent final plateau value.

We modeled the cumulative number of molecules counted by considering that SR relies on the stochastic activation of molecules. This stochasticity implies that the probability of activating a given molecule has a constant value, *c*, over time. Then, on average the number of activated molecules at any given time *t* is this constant *c* multiplied by the number of non-activated molecules. This gives rise to a relationship between the total number of molecules detected at time *t* and the number detected in the subsequent time step. Consequently, for a structure with a total number of molecules N_{total}, the number of detected molecules N_{det} as a function of the total imaging time t and the exposure time α is given by an exponential form

(see Supporting Information for more details). This model, applied with *c* and N_{total} as free parameters, captures the data well (Figure 2B, C) and allows us to estimate the number of molecules in the structure without requiring complete imaging. For nascent HIV-Gag virions, this analysis shows that by acquiring 10'000 frames we imaged on average 90% of detectable molecules. This model precludes a determination of the absolute number of molecules because it cannot account for molecules that escape detection due to misfolding, failed photoconversion or premature photobleaching. However, under the assumption that the fluorescent proteins used have similar folding efficiency it allows us to make direct, quantitative comparisons between different objects imaged using different fluorophores.

Since we measured similar levels of detection for Gag-mEos2 and Gag-tdEos fluorophores, distributions of the number of Gag molecules per cluster can be directly compared with no further consideration of differences in dye behavior. This corresponds to the number of mEos2 fluorophores detected and half the number of tdEos fluorophores detected. We observed that the distribution of the number of molecules per assembling virion was shifted to lower values for Gag-tdEos relative to Gag-mEos2, with the mean value reduced by a factor of 0.5 (Figure 2D). We note that we can identify for both labels small nucleation complexes composed of a few molecules as well as dense clusters corresponding to HIV-Gag virions (Figure 2E), likely near the final state of assembly. Intriguingly, 43% (41%) of the detected Gag-mEos2 (Gag-tdEos) clusters contain between 32 (our minimum cutoff number for cluster identification) and 100 molecules. These small clusters would not be reliably identified or detected with standard fluorescence or EM imaging. Their presence indicates that the nucleation process that initiates virion formation may represent a considerable fraction of the total time for virion assembly, as the observed population likely represents a snapshot of the steady-state assembly process.

Gag clusters were further analyzed to extract quantitative measures of their size and morphological features. To extract the shape characteristics of each cluster, we study the distribution of its molecules in 16 angular sectors centered in the cluster's center of mass. To have sufficient statistics on the smallest clusters, we require on average two points per sector, which translates into a minimum of 32 Gag molecules per cluster. We used this approach to extract quantitative data such as the mean radius and its coefficient of variation and the aspect ratio (Supporting Information, Supporting Figure 2A, B). We defined the radius of each cluster by the average over all sectors of the mean distance of a molecule from the center of mass (Supporting Information, Supporting Figure 2A). To judge how well this estimator reflects the cluster size, we performed our sector-based analysis on simulated clusters of points with known localization precision and radius. We note that

the performance of our estimator is comparable to standard image based size extraction methods such as Gaussian profile fitting to histograms of molecular positions (Supporting Information, Supporting Figure 3). However, the information on molecular distances from the center of mass per sector thus obtained also allowed us to extract more subtle morphological features such as the aspect ratio and the standard deviation from the mean radius for each sector (Figure 3D, E). The coefficient of variation of the radius reflects how isotropic clusters are, with higher values corresponding to more anisotropic morphologies (Figure 3E, inset). Working directly with the molecular position list enables this morphological analysis, which would be obscured if it were rendered as an image.

Interestingly, the distribution of the coefficient of variation (Figure 3E) showed a double-peaked distribution for both Gag-FP fusions. The first peak is on the order of the anisotropy induced by the imprecise localization of single molecules observed for simulated data (Supporting Figure 3). The second peak, however, corresponds to clusters diverging from a circular shape (Figure 3E, insets). Performing our sector based analysis on HIV-Gag clusters also revealed striking differences between the sizes of assembling virions formed with Gag-mEos2 or Gag-tdEos (Figure 3A, B, C). The average radius was 53 +/- 12 nm for Gag-mEos2[5, 19] versus 86 +/- 26 nm for Gag-tdEos (Figure 3C), differing by a factor of 1.6. For comparison with previous measurements of fully formed virions, we examined the largest 10% of clusters, and found an average diameter of 166 +/- 26 nm for Gag-mEos2 and 302 +/- 90 nm for Gag-tdEos. For Gag-mEos2 this is in good agreement with EM of budded VLPs, reported in different studies to be 100-200 nm or 145 +/- 25 nm in size[5, 19]. It moreover indicates that the Gag-tdEos particles are unusually large.

The dramatic change in nascent virion size distribution together with the difference in the mean number of Gag proteins per cluster indicates a change in the nanoscale organization within clusters for different fluorescent labels. HIV-Gag assembles into a hexagonal lattice as observed with cryoEM[20, 21]. It is proposed that this hexameric lattice grows from a nucleation point by incorporating more Gag proteins during assembly, with an inherent curvature set by protein-protein interactions. But on a sphere thus formed, hexamers further from the nucleation point must pack at increasing density[20] (Supporting Figure 4), with an energetic cost. As a consequence, at some point during assembly it should become more energetically favorable to leave gaps in the lattice than to pack proteins at higher densities; consistent with this, the Gag lattice covers only 60% of the virion surface[20]. This is similar to the coverage values of 40-70% that we estimate based on size and protein number for the largest 10% of Gag-mEos2 clusters, which likely correspond to fully formed VLPs. In this context, our data suggest how the fusion of a fluorescent protein to Gag can interfere with assembly. Gag-tdEos (the larger fluorescent label) may disrupt hexameric ordering closer to the nucleation point due to its increased steric hindrance; this would translate into smaller ordered domains. In support of this, we estimate based on our measurements of nascent virion size and protein number that the Gag-tdEos forms a patchy lattice covering only ~10-20% of the virion surface, based on the lattice spacing measured by EM. As further evidence, a separate study of elongated Gag proteins revealed an increase in virion size and a discontinuous density of Gag[19]. However, our data allow us to go further, in showing that these differences in packing exist throughout most of the assembly process.

The assembly of proteins into functional nanoscale structures is a ubiquitous process in cellular systems. We have demonstrated that SR imaging combined with the quantitative measurements

described here allow us to characterize virion assembly at many intermediate stages, revealing differences in protein packing and cluster morphologies. In principle, the same procedure can be applied to other biological processes. In particular, morphological analysis from molecular positions can help to quantify changes in the spatial arrangement of proteins. Complete imaging of structures coupled with morphological and molecular counting analysis can allow one to detect whether, and quantify how the nanoscale organization of proteins is affected by protein structure.

Acknowledgement

We wish to thank Harald Hess for the PeakSelector software, George Patterson for the kind gift of plasmid mEos2, Joerg Wiedenmann for the gift of tdEos plasmids and purified proteins, Daniel Blair for the clustering algorithm and Vinoth Sundar Rajan for additional experiments during the review process. The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme / ERC grant agreement n° 243016-PALMassembly. TP also received support from the Brazilian Swiss Joint Research Program.

Supporting information available

This material is available free of charge via the Internet at http://pubs.acs.org.

Figure 1: SR imaging and identification of HIV-Gag virions at different stages of assembly

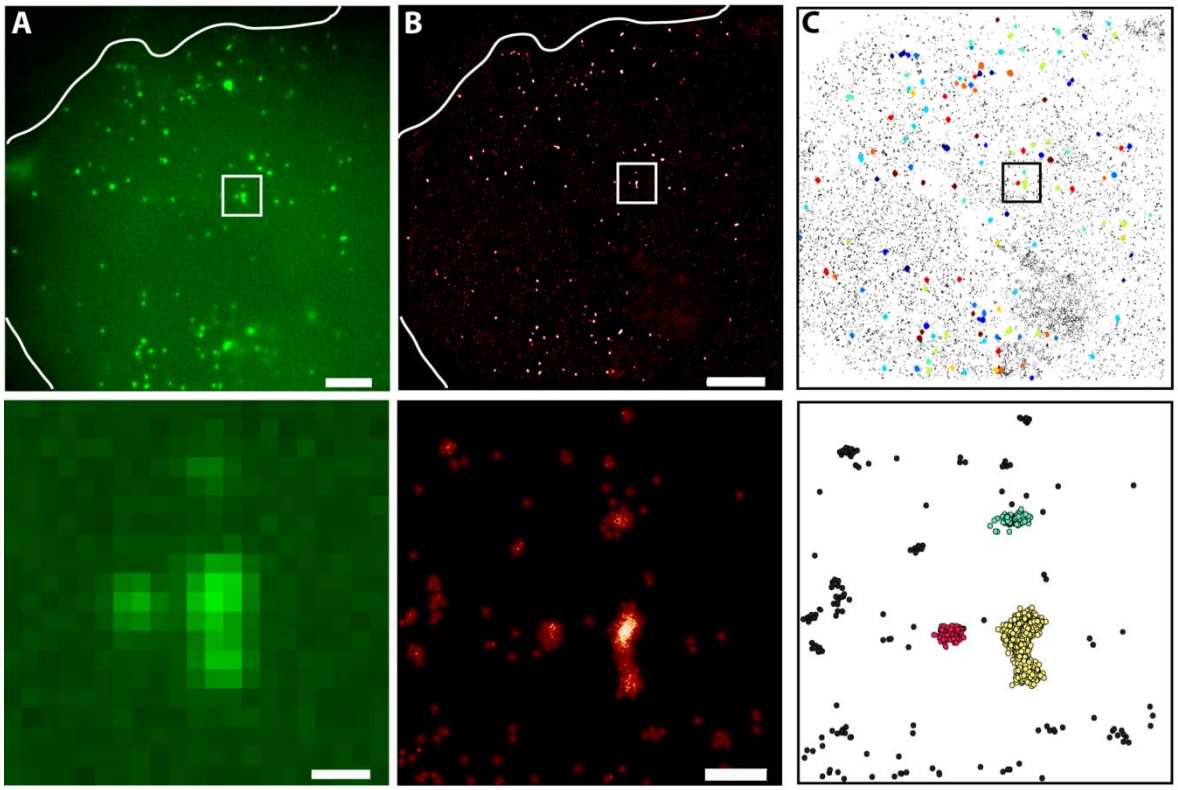

Figure 1: SR imaging and identification of HIV-Gag virions at different stages of assembly. (**A**) Cos7 cells transfected with Gag-mEos2, fixed and imaged using wide-field epi-fluorescence imaging (top) and zoom of the boxed region (bottom). (**B**) The same cell imaged with SR (10'000 images, total recording time 5 min) rendered as the envelope of the molecular probability distribution to emphasize molecular locations (top). Higher magnification view of the boxed region (bottom). (**C**) HIV-Gag clusters identified by the clustering algorithm (top) and zoom of the boxed region (bottom), colors are used to distinguish individual clusters. Scale bars: 5 µm (**A, B,** top), 500 nm (**A, B,** bottom)

Figure 2: Complete molecular imaging and counting.

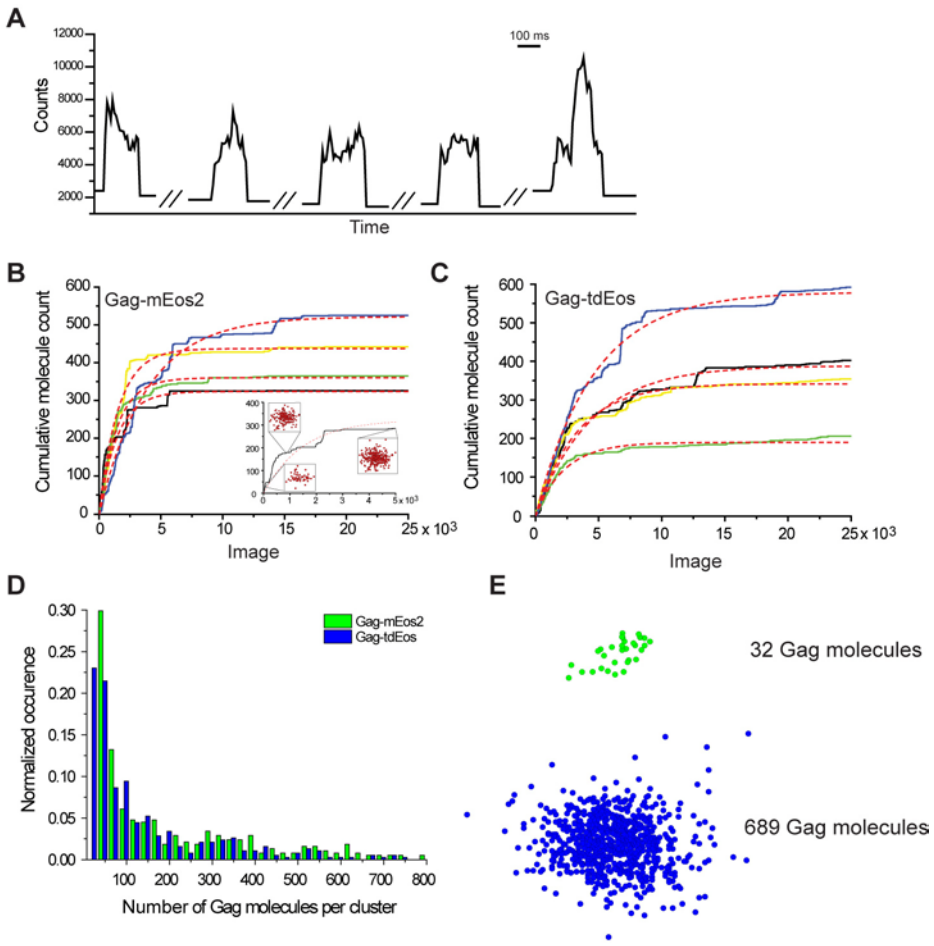

Figure 2: Complete molecular imaging and counting. (**A**) Single molecule time-intensity traces for Gag-tdEos. (**B**) Cumulative molecule count for four Gag-mEos2 virions in assembly (solid lines) and fits to the model (dashed red lines) as a function of the number of acquired images. The inset shows the cumulative molecule count during the first 5'000 images for the black Gag-mEos2 curve, the fit to our model (dashed red line) and the spatial maps of the detected molecules after 100, 1'000 and 5'000 images. (**C**) Cumulative molecule count for four Gag-tdEos clusters (solid lines) and fits to the model (dashed red lines). (**D**) Normalized histogram of molecules detected per cluster for Gag-mEos2 and Gag-tdEos. (**E**) Examples of clusters of different sizes.

Figure 3: Morphological and statistical analysis of individual assembling virions labeled with Gag-mEos2 or Gag-tdEos.

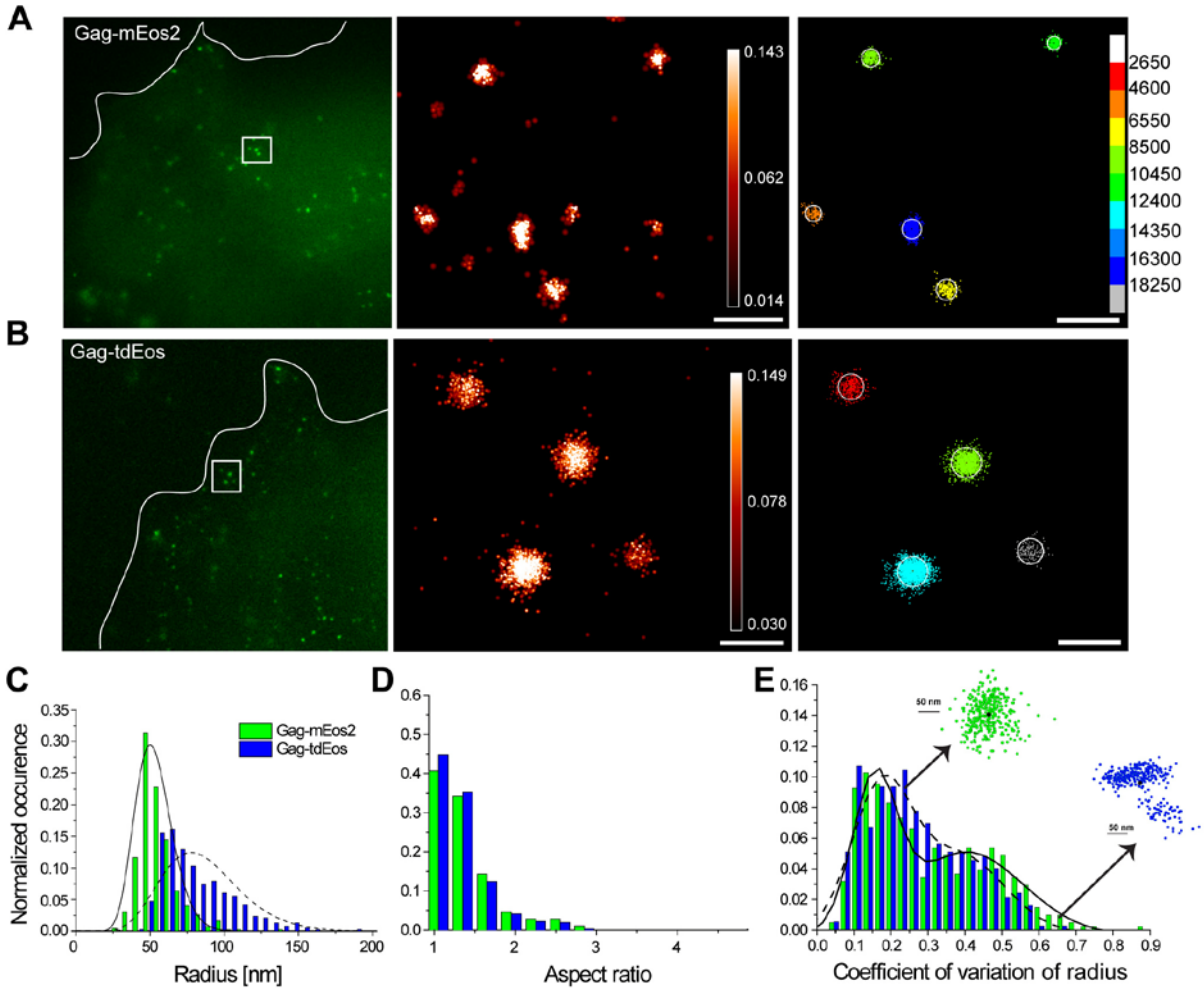

Figure 3: Morphological and statistical analysis of individual assembling virions labeled with Gag-mEos2 or Gag-tdEos. (**A**) Gag-mEos2 clusters at the plasma membrane. Diffraction-limited image with the white line corresponding to the cell edge (left), SR image of the boxed region (center) where color indicates the local molecular probability density (x100/nm^2) as indicated by the color bar. Cluster map with diameter overlaid in white (right), the color indicates the overall cluster density in molecules/µm^2 as indicated by the color scale. (**B**) Corresponding images for Gag-tdEos. Normalized histograms of (**C**) the radius, (**D**) the aspect ratio, and (**E**) the coefficient of variation of the mean radius. The inset shows typical clusters with different coefficients of variation. Curves show the fit to gamma distributions for (**C**) and a double peak distributions for (**E**). Gag-mEos2, n > 400 clusters and Gag-tdEos, n > 400 clusters. Scale bar: 500 nm (**A, B**)

References

(1) Freed, E. O., *Virology* **1998,** *251* (1), 1-15.
(2) Gheysen, D.; Jacobs, E.; De Foresta, F.; Thiriart, C.; Francotte, M.; Thines, D.; De Wilde, M., *Cell* **1989,** *59* (1), 103-112.
(3) Ivanchenko, S.; Godinez, W. J.; Lampe, M.; Kräusslich, H. G.; Eils, R.; Rohr, K.; Bräuchle, C.; Müller, B.; Lamb, D. C., *PLoS Path* **2009,** *5* (11).
(4) Jouvenet, N.; Bieniasz, P. D.; Simon, S. M., *Nature* **2008,** *454* (7201), 236-240.
(5) Briggs, J. A. G.; Johnson, M. C.; Simon, M. N.; Fuller, S. D.; Vogt, V. M., *J Mol Biol* **2006,** *355* (1), 157-168.
(6) Briggs, J. A. G.; Simon, M. N.; Gross, I.; Kräusslich, H. G.; Fuller, S. D.; Vogt, V. M.; Johnson, M. C., *Nat Struct Mol Biol* **2004,** *11* (7), 672-675.
(7) Fuller, S. D.; Wilk, T.; Gowen, B. E.; Kräusslich, H. G.; Vogt, V. M., *Curr Biol* **1997,** *7* (10), 729-738.
(8) Wright, E. R.; Schooler, J. B.; Ding, H. J.; Kieffer, C.; Fillmore, C.; Sundquist, W. I.; Jensen, G. J., *EMBO J* **2007,** *26* (8), 2218-2226.
(9) Betzig, E.; Patterson, G. H.; Sougrat, R.; Lindwasser, O. W.; Olenych, S.; Bonifacino, J. S.; Davidson, M. W.; Lippincott-Schwartz, J.; Hess, H. F., *Science* **2006,** *313* (5793), 1642-1645.
(10) Rust, M. J.; Bates, M.; Zhuang, X., *Nat Methods* **2006,** *3* (10), 793-5.
(11) Hess, S. T.; Girirajan, T. P. K.; Mason, M. D., *Biophys J* **2006,** *91* (11), 4258-4272.
(12) McKinney, S. A.; Murphy, C. S.; Hazelwood, K. L.; Davidson, M. W.; Looger, L. L., *Nat Methods* **2009,** *6* (2), 131-133.
(13) Wiedenmann, J.; Ivanchenko, S.; Oswald, F.; Schmitt, F.; Röcker, C.; Salih, A.; Spindler, K. D.; Nienhaus, G. U., *Proc Natl Acad Sci USA* **2004,** *101* (45), 15905-15910.
(14) Jouvenet, N.; Neil, S. J.; Bess, C.; Johnson, M. C.; Virgen, C. A.; Simon, S. M.; Bieniasz, P. D., *PLoS Biol* **2006,** *4* (12), 2296-2310.
(15) Annibale, P.; Scarselli, M.; Kodiyan, A.; Radenovic, A., *J Phys Chem Lett* **2010,** *1* (9), 1506-1510.
(16) Baddeley, D.; Cannell, M. B.; Soeller, C., *Microsc Microanal* **2010,** *16* (1), 64-72.
(17) Hoshen, J.; Kopelman, R., *Phys Rev B* **1976,** *14*, 3438-3445.
(18) Annibale, P.; Vanni, S.; Scarselli, M.; Rothlisberger, U.; Radenovic, A., *PLoS ONE* **2011,** *6* (7).
(19) Pornillos, O.; Higginson, D. S.; Stray, K. M.; Fisher, R. D.; Garrus, J. E.; Payne, M.; He, G. P.; Wang, H. E.; Morham, S. G.; Sundquist, W. I., *J Cell Biol* **2003,** *162* (3), 425-34.
(20) Briggs, J. A. G.; Riches, J. D.; Glass, B.; Bartonova, V.; Zanetti, G.; Kräusslich, H. G., *Proc Natl Acad Sci USA* **2009,** *106* (27), 11090-11095.
(21) Carlson, L. A.; Briggs, J. A. G.; Glass, B.; Riches, J. D.; Simon, M. N.; Johnson, M. C.; Müller, B.; Grünewald, K.; Kräusslich, H. G., *Cell Host and Microbe* **2008,** *4* (6), 592-599.

TOC graphic